Analysis of protein expression in Chinese hamster ovary cells and breast cancer.

A thesis submitted for the degree of PhD
by Mark Gallagher, BSc.

September 2016

The experimental work in this thesis was performed
under the supervision of

Dr. Paula Meleady
&
Dr. Annemarie Larkin

School of Biotechnology

Dublin City University

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____ (Candidate) ID No.: _____ Date: _____

Acknowledgments

It has always been known that I do somethings slowly but...this...bears never to be repeated. Half a decade of peaks and troughs. Peaks mostly created by others, troughs by me and the data.

Those high points are namely Paula, Martin and Annemarie. Paula for when I'd need a pep talk even though my madness may have been incomprehensible, Martin for philosophical conversations (at least that's what I'll call them from my perspective) and Annemarie when I was at my most pessimistic with some of the breast cancer data. Their patience, tested to the absolute limit, I am extremely grateful for and humbled by.

Never forgetting the wacky peer group (frequently refereed by Niall as "A strange bunch") I'd like to thank Shane and Deirdre for their proteomic solidarity. Those times when that experiment went wrong and tears would happen. In particular Shane for taking the post of secretary when I started to fade out of existence toward the end. Kevin the source of food, all things German and positive outlooks. Gemma for translating my whale whisper voice at lunch time and always listening. Paul for his jingles and impressions, because you need to laugh doing a PhD. Andrew and Edel as the Annemarie holy trinity. Andrew always willing to help with work and Edel one of the few people to make me cry with laughter. Orla and Alan the next generation to which I say good luck with the data we leave you to inherit.

After that I'd like to thank all the behind the scenes but just as influential parties. Mairead always having very substantial chats with and sorting out most if not all of the administration for which we'd be lost otherwise. Carol for doing the same and always willing to make time for you. Gillian for her indispensible role in managing all our experiments gone wrong. Then there is Mick...he just about does everything already mentioned. An absolute asset professionally and at a personal level to proteomics and the centre.

I'd also like to thank my external and internal examiner Professor Sean Doyle and Dr Phil Cummins who were extremely generous in correcting my soft bound "thesis".

Paula, Martin and Shane deserve the special final mention. Without the alignment of those three people I would definitely have faltered, fell and failed. I required more than I ought to be given in help from them and they gave me more than I could ask for.

Contents

List of figures		**vi**
List of tables		**xi**
Abbreviations		**xiv**
Publications		**xvii**
Scientific talks and presentations		**xviii**
Abstract		**1**
CHAPTER 1 Background		**2**
1.1	**General Introduction - Chinese hamster ovary and bioprocessing**	**3**
1.2	**Chinese hamster ovary cells**	**4**
1.2.1	Biopharmaceutical uses	5
1.2.2	Cell factory optimisation	7
1.2.2.1	CHO cell productivity optimization	7
1.2.2.2	Media formulation and waste management	7
1.2.2.3	Bioprocess optimisation	9
1.3	**Temperature shift**	**13**
1.3.1	Temperature shift mechanism	13
1.3.2	Temperature shift and CHO	15
1.4	**microRNA interference**	**17**
1.4.1	microRNA formation	17
1.4.2	microRNA utility and target prediction	19
1.4.3	CHO and microRNA	21
1.4.4	microRNA-7 for CHO cell engineering	22
1.5	**Proteomics**	**23**
1.5.1	Sample preparation	23
1.5.2	Sample separation	24
1.5.2.1	Gel based separation	25
1.5.2.2	Fractionation and enrichment	26
1.5.2.3	Liquid Chromatography	28
1.5.3	Mass spectrometry	28
1.5.3.1	Mass spectrometry sample preparation	29
1.5.3.2	Nano-liquid chromatography coupled to mass spectrometry	29
1.5.1.3	Mass spectrometry data output	33
1.5.1.4	Bioinformatics processing	35
1.5.1.5	CHO database	36
1.5.1.6	Quantitative label-free LC-MS/MS	37
1.6	**General introduction - Novel breast cancer proteins**	**39**
1.7	**Breast cancer subtypes**	**40**
1.7.1	ER positive	40
1.7.2	PR positive	41
1.7.3	HER2 positive	42
1.7.4	TNBC and Basal-like	43
1.8	**Breast cancer treatment**	**44**
1.8.1	Conventional therapy	44
1.8.2	Chemotherapy	44
1.8.3	Targeted therapy	45
1.9	**Aims of thesis**	**49**

CHAPTER 2 Materials and Methods **51**
2.1 Cell culture techniques **52**
2.1.1 Preparation of culture media 52
2.1.2 Suspension culture 52
2.1.3 Adherent culture 53
2.1.4 Cryopreservation of cells 54
2.1.5 Thawing cells 54
2.1.6 Mycoplasm testing 55
2.2 Cell counting and viability **55**
2.2.1 Trypan blue 55
2.2.2 Cedex automated cell counter 56
2.2.3 Flow cytometry 56
2.3 Molecular techniques **57**
2.3.1 Transfecion with miRNA/siRNA 57
2.3.2 RNA extraction 60
2.3.3 RNA quantitation 60
2.3.4 RNA to cDNA synthesis 61
2.3.5 Real-time PCR 61
2.4 Proteomic techniques **63**
2.4.1 Cell lysate preparation 63
2.4.2 Nuclear/Cytoplasmic enrichment 64
2.4.3 Membrane enrichment 65
2.4.4 Protein quantitation 67
2.4.5 Brilliant Blue G Colloidal Coomassie staining 68
2.5 LC-MS/MS protein identification **68**
2.5.1 Protein sample cleanup 68
2.5.2 In solution digest for mass spectrometry 69
2.5.3 C18 peptide cleanup 70
2.5.4 LC/MS analaysis 71
2.5.5 Mass Spectrometry generated protein identifications 72
2.5.6 Label-free LC-MS data analysis 73
2.6 Western blot analysis **73**
2.6.1 Gel electrophoresis 73
2.6.2 Western blot transfer 74
2.6.3 Enhanced chemiluminescent detection 75
2.7 Bioinformatics and statistical analysis **76**
2.7.1 Pathway analysis 76
2.7.2 Statistical significance 77
2.7.3 Bioinformatic dataset analysis 77
2.8 Immunohistochemistry **78**
2.8.1 Scoring guidelines 79

CHAPTER 3 Effect of miR7 on CHO cells **81**
3.1 Impact of miR-7 on CHO cells **82**
3.2 Quantitative proteomic profiling of pre-miR-7 transfected CHO-K1-SEAP **85**
3.2.1 Label-free analysis 85
3.2.2 Mammalian vs CHO databases 97
3.2.3 Western blot identification 102
3.3 Predicted miR-7 target analysis using miRWalk 109
3.4 Pathway analysis 113

3.4.1	Protein annotation	113
3.4.2	DAVID analysis	115
3.4.3	PANTHER analysis	122
3.4.4	KEGG analysis	127
3.5	Proteomic and microarray overlap of targets associated with miR-7	132
3.6	Results summary	134
CHAPTER 4 Effect of temperature shift on CHO-K1-SEAP cells using subcellular fractionation		**135**
4.1	**Effect of temperature shift and deep proteome target discovery**	**136**
4.2	**Fractionation validation**	**137**
4.2.1	Western blot validation of enrichment	137
4.2.2	Mass spectrometry validation of enrichment	139
4.2.2.1	Protein identification overlap	140
4.2.2.2	Annotated enrichment analysis	141
4.3	**Quantitative label-free LC-MS/MS analysis**	**144**
4.3.1	Database comparison BBCHO vs NCBI	146
4.3.2	Differentially expressed/abundant proteins at 31^{o}C temperature shift	147
4.3.3	Pathway analysis of temperature shift	150
4.3.3.1	DAVID temperature shift	150
4.3.3.2	PANTHER temperature shift	152
4.3.3.3	KEGG temperature shift	154
4.3.4	Differentially expressed/abundant proteins over time at 37 and 31 ^{o}C	158
4.3.5	Pathway analysis of temperature shift over time at 31 and 37 ^{o}C	163
4.3.5.1	DAVID analysis time course	163
4.3.5.2	PANTHER analysis time course	165
4.3.5.3	KEGG analysis time course	167
4.4	**Functional Validation**	**175**
4.5	**Results summary**	**181**
CHAPTER 5 Identification of novel membrane protein targets in breast cancer		**182**
5.1	**Identifying potential membrane proteins over expressed in breast cancer**	**183**
5.2	**Validation of expression in cell lines**	**187**
5.3	**Preliminary IHC analysis**	**190**
5.3.1	Normal breast vs breast tumor	190
5.3.2	Expression in proliferative tissue	193
5.3.3	Normal tissue expression	195
5.4	**IGSF9**	**200**
5.4.1	IGSF9 expression in HER2+ and TNBC tumors	201
5.4.2	IGSF9 expression in other cancer tissues	203
5.4.3	TMA IHC analysis of IGSF9	205
5.5	**KLRG2**	**211**
5.5.1	KLRG2 expression in HER2+ and TNBC tumprs	212
5.5.2	KLRG2 expression in other cancer tissues	214
5.5.3	TMA IHC analysis of KLRG2	216
5.5.4	Survival outcome	217

CHAPTER 6 Discussion **220**
6.1 Effect of miR-7 on CHO cells introduction **221**
6.1.1 Quantitative label-free LC-MS/MS 222
6.1.2 Species homology identifications 222
6.1.3 Chinese hamster ovary identifications 223
6.1.4 Predicted direct targets of miR-7 223
6.1.4.1 Catalase 224
6.1.4.2 Stathmin 225
6.1.4.3 CCT3 226
6.1.4.4 PA2G4 227
6.1.4.5 RAN 228
6.1.4.6 EEF1A1 228
6.1.4.7 RPL15 229
6.1.5 Pathway analysis of differentially expressed proteins in miR-7 over expressing CHO cells 230
6.1.5.1 Global effect of miR-7 on protein in CHO cells 230
6.1.5.2 miR-7 negatively regulates proteins associated with translation 231
6.1.5.3 miR-7 negatively regulates proteins associated with cellular structure 232
6.1.5.4 miR-7 positively regulates homeostasis and anti-apoptotic proteins 233
6.1.5.5 miR-7 positively regulates proteins involved in glutathione metabolism 233
6.1.6 Transcriptomic and proteomic overlap of differentially expressed targets in miR-7 over expressing CHO cells 235
6.1.6.1 HDAC1 235
6.1.7 Proposed effect of miR-7 on CHO cell proteome 237
6.2 Effect of temperature shift on CHO cells introduction **239**
6.2.1 Subcellular fractionation of CHO-K1-SEAP cells undergoing temperature shift 239
6.2.1.1 Overlap between fractionated samples 240
6.2.1.2 Pathway analysis validation of enrichment 240
6.2.1.3 Western blot validation of enrichment 241
6.2.2 CHO database annotation 241
6.2.3 Quantitative Label-free LC-MS/MS analysis of temperature shifted sub cellular enriched CHO-SEAP cells 242
6.2.3.1 Differentially regulated proteins at 31 ^{o}C, 8 and 24 hr after temperature shift enriched CHO-K1-SEAP cells 242
6.2.3.2 Pathway analysis of differentially regulated proteins at 31 ^{o}C, 8 and 24 hr after temperature shift 244
6.2.4 Differentially regulated proteins over 16hr at 31 and 37 ^{o}C 245
6.2.4.1 Pathway analysis of differentially regulated proteins over 16hr at 31 ^{o}C and 37 ^{o}C 246
6.2.5 Selection of targets for functional validation 247
6.2.5.1 Ezrin (EZR) and Moesin (MSN) 247
6.2.5.2 Cyclon 248
6.2.5.3 Lamin A/C 249
6.2.6 Conclusion 250
6.3 Breast Cancer target discovery **251**
6.3.1 SLAMF8 251

6.3.2	LPR8	252
6.3.3	TSPAN13	253
6.3.4	IGSF9	254
6.3.4.1	IGSF9 expression in cell lines	255
6.3.4.2	IGSF9 has low expression in normal tissue	256
6.3.4.3	IGSF9 expression is not specifically associated with HER2+ breast tumor tissue	256
6.3.5	KLRG 2	257
6.3.5.1	KLRG2 expression in cell lines	257
6.3.5.2	KLRG2 is lowly expressed in normal tissue	258
6.3.5.3	KLRG2 expression is appears highest in TNBC tissue samples	258

CHAPTER 7 Conclusion **260**

7.1	**Effect of miR-7 on the CHO cell proteome**	**261**
7.2	**Effect of temperature shift on CHO-K1-SEAP cells using sub cellular fractionation**	**263**
7.3	**Identification of novel membrane targets in breast cancer**	**265**

Future work **267**

Bibliography **270**

Appendix **305**

List of figures

Figure 1.2.1	Top 8 biopharmaceutical product approval categories from 1982-2014 (cumulative) and 2010 to 2014.	6
Figure 1.2.2	Bioreactor schematic	11
Figure 1.2.3	Bioprocess schematic including upstream and downstream of the bioreactor	12
Figure 1.3.1	Temperature shift mechanism.	14
Figure 1.4.1	The biogenesis and processing of microRNA in eukaryotes.	18
Figure 1.5.1	Proteome complexity in humans.	25
Figure 1.5.2	Quadrupole mass analyser schematic.	31
Figure 1.5.3	Thermo Hybrid LTQ-Orbitrap XL Schematic.	32
Figure 1.5.4	Work flow of tandem mass spectrometry (MS/MS) analysis.	33
Figure 1.5.5	MS/MS spectrum of peptide fragments for Albumin.	35
Figure 1.5.6	Schematic of label-free LC-MS/MS peptide and protein quantitation.	38
Figure 1.8.1	Delivery of cytotoxic drugs conjugated to a monoclonal antibody (ADC).	46
Figure 2.2.1	Haemocytometer grid.	55
Figure 2.3.1	Real-time PCR using miRNA-specific primer.	62
Figure 2.3.2	Diagram of 5' nuclease reporter process.	63
Figure 2.4.1	Schematic of membrane enrichment methodology.	66
Figure 2.8.1	Immunoreactivityguidlines used for staining intensity	80
Figure 3.1.1	Impact on growth and viability in CHO K1 SEAP cells at 48 and 96 hr after transfection with pre-miR-7.	83
Figure 3.1.2	Relative expression of miR-7 at 48 and 96 hr after transfection with pre-miR-7 compared to endogenous miR-7 in pre-miR negative control (PM-Neg).	84
Figure 3.2.1	Principal component analysis (PCA) of differentially regulated proteins with statistical cut off of ANOVA <0.05 and features with +1, +2 and +3 charge state using the CHO database.	86
Figure 3.2.2	Overlap between gene name identifications associated with differentially expressed proteins in CHO-K1 SEAP cells over expressing miR-7 derived from a mammalian (human, mouse and rat) and Chinese hamster (BBCHO and NCBI non redundant) databases.	97
Figure 3.2.3	Graph of number of peptides found for each protein at 48 hr up-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database.	98
Figure 3.2.4	Graph of number of peptides found for each protein at 48 hr down-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database.	99
Figure 3.2.5	Graph of number of peptides found for each protein at 96 hr up-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database.	100
Figure 3.2.6	Graph of number of peptides found for each protein at 96 hr down-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database.	101
Figure 3.2.7	Western blot analysis showing down regulation of Histone H3 in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg).	103

Figure 3.2.8	Western blot analysis showing down regulation of Histone H4 in CHO- SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg).	104
Figure 3.2.9	Western blot showing up regulation of PDIA6 in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg).	105
Figure 3.2.10	Western blot showing up regulation of GRP78 in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg).	106
Figure 3.2.11	Western blot analysis showing down regulation of HSPA8 at 48 hr in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg).	107
Figure 3.2.12	Western blot analysis (left) showing up regulation of 14-3-3 epsilon in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg).	108
Figure 3.2.13	Western blot analysis for GAPDH showing equal loading for Figure 3.2.7 to 3.2.12.	108
Figure 3.3.1	Normalised abundance of peptides associated with predicted direct target of miR-7 Catalase from the label-free LC-MS/MS data using the CHO database.	111
Figure 3.3.2	Normalised abundance of peptides associated with predicted direct target of miR-7 Stathmin from the label-free LC-MS/MS data using the CHO database.	112
Figure 3.4.1	Examples of typical text strings from the BBCHO in house CHO database (top) and the NCBI non redundant CHO database (bottom) requiring annotation.	114
Figure 3.4.2	Diagram representing protein up-regulation in the Glutathione Metabolism pathway using DAVID and KEGG analysis.	128
Figure 3.4.3	Diagram representing protein down-regulation in the ribosome pathway using DAVID and KEGG analysis.	129
Figure 3.4.4	Diagram representing protein up-regulation in the Glutathione Metabolism pathway using DAVID and KEGG analysis.	130
Figure 3.4.5	Diagram representing protein down-regulation in the ribosome pathway using DAVID and KEGG analysis.	131
Figure 3.5.1	Transcriptional expression of HDAC1 (A) and Western blot analysis of Histone/Acetyl Histone (B) and of HDAC1 expression (C).	133
Figure 4.2.1	Validation of membrane, cytoplasmic and nuclear enrichment in CHO-K1 SEAP cells at 72 hr using commercially available benchtop fraction kits.	138
Figure 4.2.2	Venn diagram showing the overlap of protein IDs using gene index (GI) identifiers from the BBCHO database using CHO-K1-SEAP cells grown for 72 hr at 37 ^{o}C.	141
Figure 4.2.3	Representation of the top ten significant (Bonferroni adjusted $p \leq 0.05$) Cellular Component GO terms associated with each fraction from Table 4.2.3.	143
Figure 4.3.1	Overview of the experimental workflow of temperature shift study with subcellular fractionation.	145
Figure 4.3.2	Venn diagram showing the overlap in CHO IDs between the up and down-regulated IDs of the BBCHO and non redundant NCBI databases using the 16hr time course of CHO-K1-SEAP cells grown at 37 ^{o}C and 31 ^{o}C.	146

Figure 4.3.3	Up-regulated proteins (A) and down-regulated proteins (B) in CHO-K1 SEAP cells at 31 ^{o}C after 8 and 24 hr in membrane, cytoplasmic and nuclear fractions.	148
Figure 4.3.4	Overlapping up-regulated proteins in CHO-K1 SEAP cells at 31 ^{o}C after 8 (A) and 24 hr (B) between membrane, cytoplasmic and nuclear fractions.	149
Figure 4.3.5	Overlapping down-regulated proteins in CHO-K1 SEAP cells at 31 ^{o}C after 8 (A) and 24 hr (B) between membrane, cytoplasmic and nuclear fractions.	149
Figure 4.3.6	KEGG pathway diagram representing nuclear enriched protein down-regulation in the ribosome pathway 8 hr after temperature shift.	155
Figure 4.3.7	KEGG pathway diagram representing cytoplasm enriched protein up-regulation in the DNA replication pathway 24 hr after temperature shift.	156
Figure 4.3.8	KEGG pathway diagram representing nuclear enriched protein down-regulation in the ribosome pathway 24 hr after temperature shift.	157
Figure 4.3.9	Up (A) and down (B) regulated proteins in CHO-K1 SEAP cells at 24 hr compared to 8 hr time point at 31 and 37 ^{o}C in membrane, cytoplasmic and nuclear fractions.	159
Figure 4.3.10	Up-regulated proteins overlap in CHO-K1 SEAP cells over 16hr at 24 hr compared to 8 hr time point at 31 (A) and 37 ^{o}C (B) between membrane, cytoplasmic and nuclear fractions.	160
Figure 4.3.11	Down-regulated proteins overlap in CHO-K1 SEAP cells over 16hr at 24 hr compared to 8 hr time point at 31 (A) and 37 ^{o}C (B) between membrane, cytoplasmic and nuclear fractions.	161
Figure 4.3.12	Overlapping identifications up or down-regulated over time between 31 and 37 ^{o}C.	162
Figure 4.3.13	KEGG pathway diagram representing membrane enriched protein up-regulation in the spliceosome pathway over 16 hr at 37 ^{o}C.	168
Figure 4.3.14	KEGG pathway diagram representing membrane enriched protein down-regulation in the ribosome pathway over 16hr at 37 ^{o}C.	169
Figure 4.3.15	KEGG pathway diagram representing cytoplasm enriched protein down-regulation in the ribosome pathway over 16hr at 37 ^{o}C.	170
Figure 4.3.16	KEGG pathway diagram representing nuclear enriched protein up-regulation in the ribosome pathway over 16hr at 37 ^{o}C.	171
Figure 4.3.17	KEGG pathway diagram representing membrane enriched protein up-regulation in the spliceosome pathway over 16 hr at 37 ^{o}C.	172
Figure 4.3.18	KEGG pathway diagram representing cytoplasm enriched protein down-regulation in the ribosome pathway over 16 hr at 31 ^{o}C.	173
Figure 4.3.19	KEGG pathway diagram representing nuclear enriched protein up-regulation in the ribosome pathway over 16 hr at 31 ^{o}C.	174
Figure 4.4.1	Western blot analysis showing the expression intensity of targets in CHO-K1-SEAP cells grown over 24 hr (A**)** with corresponding label-free LC-MS/MS data (B).	175
Figure 4.4.2	Effect of Cyclon knockdown after 72 hr on target expression	177

using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F).
Figure 4.4.3 Effect of Ezrin knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). 178
Figure 4.4.4 Effect of Moesin knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). 179
Figure 4.4.5 Effect of Lamin A/C knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). 180
Figure 5.2.1 Western blot analysis showing differential protein expression of IGSF9 and KLRG2 across a panel different breast cancer cell lines (A) whole cell preparation and (B) membrane enriched fractions. 189
Figure 5.3.1 Representative normal breast tissue (A, B) and invasive breast cancer tissue (C, D) (subtype unknown) images of IHC analysis of IGSF9. 191
Figure 5.3.2 Representative normal breast tissue (A, B) and invasive breast cancer tissue (C,D) (subtype unknown) images of IHC analysis of KLRG2. 192
Figure 5.3.3 Normal colon images showing IHC analysis of IGSF9 (A) and KLRG2 (B) in normal colon tissue sections stain with Ki-67 proliferation marker. 194
Figure 5.3.4 IGSF9 immunoreactivity across representative normal tissues. 196
Figure 5.3.5 KLRG2 immunoreactivity across representative normal tissues. 198
Figure 5.4.1 IGSF9 staining of HER2+ breast tumors. 202
Figure 5.4.2 IGSF9 staining of metastatic (A) and primary melanoma (B) and also pancreatic cancer sections (C, D). 204
Figure 5.4.3 HER2+ TMA tumor core (n=19) stained for IGSF9 expression. 205
Figure 5.4.4 TNBC TMA tumor core (n=16) stained for IGSF9 expression. 205
Figure 5.4.5 ER+ TMA tumor core (n=59) stained for IGSF9 expression. 207
Figure 5.4.6 PR+ TMA tumor core (n=48) stained for IGSF9 expression. 207
Figure 5.4.7 HER2+ TMA tumor core (n=69) stained for IGSF9 expression. 208
Figure 5.4.8 HER2+, ER+ TMA tumor tissue (n=43) stained using IGSF9. 208
Figure 5.4.9 HER2+, ER- TMA tumor tissue (n=26) stained using IGSF9. 209
Figure 5.4.10 HER2+, LN+ TMA tumor tissue (n=33) stained using IGSF9. 209
Figure 5.4.11 HER2+, LN- TMA tumor tissue (n=35) stained using IGSF9. 210
Figure 5.5.1 KLRG2 expression in HER2+ breast cancer tissue. 213
Figure 5.5.2 KLRG2 staining of metastatic (A) and primary melanoma (B) and pancreatic cancer sections (C, D). 215
Figure 5.5.3 HER2+ TMA tumor cores (n=18) stained for KLRG2 expression. 217
Figure 5.5.4 TNBC TMA tumor cores (n=16) stained for KLRG2 expression 217

Figure 5.5.5 ER+ TMA tumor cores (n=63) stained for KLRG2 expression 218
Figure 5.5.6 PR+ TMA tumor cores (n=63) stained for KLRG2 expression 218
Figure 5.5.7 Kaplan-Meier cumulative survival curve showing poor patient outcome associated with KLRG2 (3466689) over expression in basal type breast cancer. 219
Figure 6.1.1 Proposed effect of miR-7 on the CHO cell proteome 238

List of tables

Table 1.2.1	Selection of biopharmaceutical products produced from 1987 to 2010 in Chinese hamster ovary cells.	6
Table 1.3.1	Increased recombinant protein yields in temperature shift CHO cells.	15
Table 1.4.1	List of online prediction tools for miRNA targets.	18
Table 1.4.2	List of miRNA shown to produce industrially desirable phenotype changes in CHO cells.	21
Table 1.5.1	Amino acids and their average mass in daltons (Da) as determined by MS/MS analysis	34
Table 1.8.1	ADCs approved for therapeutoc use as of January 2016.	47
Table 1.8.2	ADCs for solid tumors in clinical trial phase I, II and III.	48
Table 2.1.1	Breast cancer cell lines used in Chapter 6.	53
Table 2.3.1	Sequence of miRNA mimic and control used in Chapter 3.	58
Table 2.3.2	Sequences of custom CHO double stranded siRNA.	59
Table 2.3.3	Reverse transcription master mix volumes for a single reaction.	61
Table 2.3.4	Real-time PCR single reaction mix.	62
Table 2.4.1	Volumes required for Nuclear/cytoplasmic extraction of 2 x 10^7 cells.	65
Table 2.6.1	List of primary antibodies used for Western blot analysis.	75
Table 2.7.1	Number of genes ≥1.3 fold increased in ER+, HER2+, LN+ or triple negative breast cancer compared to normal breast tissue.	78
Table 3.2.1	Overview of the number of down-regulated (↓) and up-regulated (↑) proteins 48 and 96 hr after miR-7 over-expression compared to scramble negative control in CHO-K1-SEAP cells using a mammalian database search and a custom Chinese hamster ovary database search.	87
Table 3.2.2	List of proteins derived from human, mouse and rat UniProt-SwissProt databases with decreased expression following transient tansfection of miR-7 in CHO compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 and >1 peptide between experimental groups.	89
Table 3.2.3	List of proteins derived from human, mouse and rat UniProt-SwissProt databases with increased expression following transient tansfection of miR-7 in CHO compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 and >1 peptide between experimental groups.	91
Table 3.2.4	List of proteins derived from CHO database with decreased expression following transient tansfection of miR-7 in CHO compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 between experimental groups.	94
Table 3.2.5	List of proteins derived from CHO database with increased expression following transient tansfection of miR-7 in CHO compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 peptide between experimental groups.	96
Table 3.2.6	Label free LC-MS fold change of proteins chosen for Western blot validation.	102
Table 3.3.1	Catalase and stathmin down regulation (↓) in miR-7 over expressing CHO-K1-SEAP cells according to label free LC-MS	103

data from mammalian and CHO (BBCHO/NCBI) databases.
Table 3.3.2 Predicted direct targets of mir-7 using miRWALK across mouse, rat and human species. In total 5 search algorithms were used to search the down-regulated label-free LC-MS mammalian protein data associated with miR-7 over expression. 104
Table 3.4.1 GO biological process (BP) analysis through DAVID of up or down-regulated proteins 48 hr after transient over expression of miR-7 in CHO-K1-SEAP cells. 118
Table 3.4.2 GO biological process (BP) analysis through DAVID of up or down-regulated proteins 96 hr after transient over expression of miR-7 in CHO-K1-SEAP cells. 119
Table 3.4.3 GO molecular function (MF) analysis through DAVID of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. 120
Table 3.4.4 GO cellular component (CC) analysis through DAVID of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. 121
Table 3.4.5 PANTHER biological process (BP) enrichment of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. 124
Table 3.4.6 PANTHER molecular function (MF) enrichment of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. 125
Table 3.4.7 PANTHER cellular component (CC) enrichment in up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. 126
Table 3.5.1 Gene names associated with both transcriptional and proteomic differential regulation in response to miR-7 over expression compared to a miR Scramble control. 132
Table 4.2.1 Summary of the number of qualitative IDs obtained from the BBCHO database and the relative number that contained verified gene names (% accession) that were automatically parsed from the transcript identifier. 141
Table 4.2.2 Significant (Bonferroni adjusted $p \leq 0.05$) Cellular Component GO terms associated with each fraction. 136
Table 4.3.1 Number of proteins differentially expressed at 8 and 24 hr after 31 ^{o}C temperature shift and over 16hr at 31 ^{o}C and 37 ^{o}C in each of the three enriched fractions membrane, cytoplasm and nuclear. 144
Table 4.3.2 Summary of significantly (adjusted Bonferroni $p \leq 0.05$) enriched GO biological process, molecular function and cellular component terms using DAVID associated with proteins up-regulated (↑) and down-regulated (↓) 8 and 24 hr after 31 ^{o}C temperature shift. 151
Table 4.3.3 Summary of significantly (adjusted Bonferroni $p \leq 0.05$) enriched PANTHER biological process, molecular function and cellular component terms using PANTHER associated with proteins up-regulated (↑) and down-regulated (↓) 8 and 24 hr 153

after 31 ^{o}C temperature shift.

Table 4.3.4 Number of proteins differentially expressed at 8 and 24 hr after 31 ^{o}C temperature shift and in each of the three enriched fractions membrane, cytoplasm and nuclear. 158

Table 4.3.5 Summary of significantly (adjusted Bonferroni $p \leq 0.05$) enriched GO biological process and molecular function terms using DAVID that are uniquely associated with proteins up-regulated (↑) and down-regulated (↓) over 16hr at 31 ^{o}C temperature shift or 37 ^{o}C. 164

Table 4.3.6 Summary of significantly (adjusted Bonferroni $p \leq 0.05$) enriched PANTHER biological process, molecular function terms using DAVID that are uniquely associated with proteins up-regulated (↑) and down-regulated (↓) over 16hr at 31 ^{o}C temperature shift or 37 ^{o}C. 166

Table 4.4.1 Cyclon and Moesin were shown to have a significant effect on cell proliferation as well as cell perimeter and cell area. 176

Table 5.1.1 Summary of number of potential membrane expressed targets for follow up. 184

Table 5.1.2 Target candidates satisfying criteria for protein validation follow up. 185

Table 5.1.3 Summary of 5 potential membrane expressed breast cancer targets for validation. 186

Table 5.2.1 Cell lines used for validation of target expression with IGSF9 and KLRG2 expression indicated. 188

Table 5.3.1 IGSF9 immunoreactivity across a cohort of 18 available normal tissues. 197

Table 5.3.2 KLRG2 immunoreactivity across a cohort of 18 available normal tissues. 199

Table 5.4.1 Summary of numbers of HER2+ and TNBC tissues producing staining with IGSF9 and their localisation. 201

Table 5.4.2 Tissue matrix array sample numbers for normal breast and HER2+, ER+, PR+ and TNBC tumor tissue. 205

Table 5.4.3 HER2+ breast tumor tissue matrix array sample numbers with ER2+, ER-, LN+ and LN- tumor status. 205

Table 5.5.1 Summary of numbers of HER2+ and TNBC tissues producing staining with KLRG2 and their localisation. 212

Abbreviations

2D-PAGE	Two dimensional polyacrylamide gel electrophoresis
ACN	Acetonitrile
ADC	Antibody drug conjugate
ANOVA	Analysis of variance
BBCHO	Bielefeld-BOKU-CHO
BHK	Baby hamster kidney
BP	Biological process
BRCA1	Breast Cancer 1
BSA	Bovine serum albumin
CAT	Catalase
CC	Cellular component
CCDC86	Cyclon
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)
CHO	Chinese hamster ovary
CID	Collision induced dissociation
CIRP	Cold-inducible RNA-binding protein
CMV	Cytomegalovirus
CSP	Cold shock protein
DAVID	Database for annotation, visualisation and integrated discovery
DFS	Disease-free survival
DGCR8	DiGeorge Syndrome Critical Region 8
DHFR	Dihydrofolate reductase
DIGE	Differential in-gel electrophoresis
DMEM	Dulbeco's minimum essential medium
DMSO	Dimethyl sulfoxide
DTT	1,4-Dithiothreitol
ECD	Electron capture dissociation
eIF2α	eukaryotic initiation factor 2 eIF2α
ER+	Estrogen receptor positive
ESI	Electron spray ionisation
ETD	Electron transfer dissociation
EZR	Ezrin
FACS	Fluorescent activated cell sorter
FCS	Fetal calf serum
FDA	Food and drug administration
FWHM	Full width half maximum
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GCLM	Glutamate-cysteine ligase, modifier subunit
GI	Gene index
GO	Gene ontology
GOI	Gene of interest
GRP78	78 kDa glucose regulated protein
GS	Glutamine synthase
GSS	Glutathione Synthase
GSTA2	Glutathione S-transferase alpha 2
GSTA3	Glutathione S-transferase alpha 3

GSTM1	Glutathione S-transferase mu
GSTP1	Glutathione S-transferase pi
HCD	High collision dissociation
HDAC1	Histone deacetylase 1
HER2+	Human epidermal growth factor receptor 2 positive
hGH	Human growth hormone
HMBA	hexamethylene bisacetamide
hESC	Human embryonic stem cells
hMSC	Human mesenchymal stem cells
HNRPC	Heterogenous ribonuclear protein C
HPLC	High pressure liquid chromatography
HSP90	Heat shock protein 90
HSPA8	Heat shock conate 71 kDa protein
ID	Identification
IDH1	Isocitrate dehydrogenase 1 (NADP+), soluble
IGF1Rβ	Insulin like growth factor 1 receptor β
IgG	Immunoglobulin G
IMS	Industrial methylated spirits
KEGG	Kyoto encyclopaedia of genes and genomes
LC	Liquid chromatography
LMNA	Lamin A/C
LN+	Lymph node positive
LRP8	Apolipoprotein e receptor
m/z	Mass to charge ratio
mAb	Monoclonal antibody
MALDI	Matrix assisted laser desorption ionisation
MF	Molecular function
miRNA	micro RNA
mESC	Mouse embryonic stem cells
MS	Mass spectrometry
MSN	Moesin
MTX	Methotrexate
NaBu	Sodium butyrate
NCBI	National Centre for Biotechnology Information
p53	Cellular tumor antigen p53
PANTHER	Protein analysis through evolutionary relationships
PBS	Phosphate buffered saline
PCA	Principle component analysis
PCR	Polymerase chain reaction
PDIA3	Protein disulfide isomerase family A member 3
PDIA6	Protein disulfide isomerase family A member 6
PM	Pre-miRNA
PR	Progesterone receptor positive
pre-miR-7	Preliminary microRNA-7
PSME3	Proteosome activator subunit 3
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
RAD54L	RAD54-like
RBM3	RNA-binding motif protein 3
RISC	RNA induced silencing complex

RPL	Ribosomal Protein L
RPS	Ribosomal protein S
SEAP	Secreted alkaline phosphatase
SILAC	Stable isotope labelling with amino acids in cell culture
SKP2	S-phase kinase associated protein 2
SLAMF8	SLAM family member 8
T-DM1	Trastuzumab-maytaninoid emtansine
TBS	Tris buffered saline
TFA	Trifluoroacetic acid
TFA	Trifluoroacetic acid
TMA	Tissue matrix array
TNBC	Triple negative breast cancer
tPA	Tissue plasminogen ativator
TRBP	TAR RNA binding protein
TSPAN13	Tetraspanin 13
VCP	Valosine containin protein
VEGFR	Vascular endothelial growth factor receptor

Publications

Gallagher, M[1]., Meleady, P.[1] , Clarke, C., Henry, M., Sanchez, N., Barron, N. and Clynes, M. (2012). Impact of miR-7 over-expression on the proteome of Chinese hamster ovary cells. *Journal of Biotechnology*, 160(3–4), pp.251-262

Clarke, C., Barron, N., Gallagher, M., Henry, M., Meleady, P. and Clynes, M. (2012). Target Prediction Algorithms and Bioinformatics Resources for miRNA Studies. Barron, N (Ed.). *MicroRNAs as Tools in Biopharmaceutical Production*. Netherlands: Springer. pp15-27

Sanchez N, Gallagher M, Lao N, Gallagher C, Clarke C, Doolan P, Aherne S, Blanco A, Meleady P, Clynes M and Barron N. (2013) MiR-7 triggers cell cycle arrest at the G1/S transition by targeting multiple genes including Skp2 and Psme3. PLOS One.8(6):e65671

Scientific talks and presentations

European Society for Animal Cell Technology (ESACT) - 23rd biennial meeting 2013. Lille, France, 23rd-26th of June.

I presented a research poster focusing on an aspect of my work entitled "Impact of miR-7 over-expression on the proteome of Chinese hamster ovary cells" at this highly respected animal cell technology conference adhering to the theme of "Better Cells for Better Health".

The Irish Mass Spectrometry Society (IMSS) annual meeting 2013. Dublin, 1st of May.

I presented a talk entitled "Impact of miR-7 over-expression on the proteome of Chinese hamster ovary cells" at this annual meeting which encourages nationwide communication and networking in the field of mass spectrometry research and development.

Limerick Postgraduate Research Conference (LPRC) - 3rd annual meeting 2014. Limerick, 29th of May.

I presented a talk focusing on an aspect of my work entitled "Label-free mass spectrometry with sub cellular fractionation to identify key proteins involved in temperature shift phenotype of Chinese hamster ovary cells" at this cross disciplinary postgrad research conference with the theme of "Developing a culture of research".

Title: Analysis of protein expression in Chinese hamster ovary cells and breast cancer

Name: Mark Gallagher

Abstract

Genomic tools in the last few decades have made it possible to produce proteins for pharmaceutical use in mammalian cells such as Chinese Hamster Ovary cells (CHO). As such there is an ever increasing demand protein biopharmaceuticals. By identifying key proteins involved in the growth patterns of mammalian cells we hope to be able to manipulate CHO cell for biopharmaceutical production. We used two different methods to manipulate cell growth and then identified the proteins involved using quantitative label free LC-mass spectrometry. Previous studies in our lab showed up-regulation of microRNA-7 (miR-7) reduces proliferation in CHO-K1-SEAP cells but increases productivity over time. A similar phenotype is observed in temperature shifted (31 ^{o}C) CHO cells and is often used in industry for increased productivity. The mechanism of both these phenotypes in CHO are largely unknown at the protein level. Using label-free LC-MS/MS we identified catalase and stathmin as potential targets of miR-7, the potential role of glutathione metabolism up-regulation and the potential role of structural process inhibition in causing this phenotype. Using the same techniques combined with subcellular fractionation to analyse the temperature shift phenotype in CHO-K1-SEAP cells we were able to double the number of protein identifications from 960 with no fractionation to 2298 using fractionation. Two differentially regulated proteins, cyclon and lamin A/C, were identified as significantly reducing cell proliferation and cell size and may have potential as targets to induce an industrially relevant phenotype in CHO cells.

Breast cancer is the leading cause of cancer death in women, there is an urgent need to identify new molecular targets for certain aggressive breast cancer subtypes to lead to improved treatments for patients. We used bioinformatics profiling of publicly available data-sets to compare gene expression across breast cancer sub-types compared to normal breast tissue to identify a panel of differentially expressed membrane candidate targets. Candidate target expression was validated in membrane enriched breast cancer cell line extracts and extensive immunohistochemical (IHC) analysis of target expression in breast cancer subtypes, normal breast tissue and highly proliferating tissues was carried out. Two proteins, IGSF9 and KLRG2, not previously associated with breast cancer, were demonstrated to show significantly higher expression in triple negative (TNBC) and HER-2 positive breast cancers than in normal breast tissue and also have very low presence in other normal (and highly proliferating) tissues. These two protein targets may have potential to be further investigated as ADC molecular targets for these aggressive breast cancer subtypes.

CHAPTER 1

Background

1.1 General Introduction - Chinese hamster ovary and bioprocessing

With the elucidation of the structure of DNA the field of genomics emerged and revolutionised the scientific landscape. With the profiling of genes it was possible to identify key effectors in disease, predict and prevent pathologies and eventually control genetic elements to alter cellular behaviour. While differential expression of genes can be used to diagnose and predict disease more often it is the protein generated by these genes which result in presentation of pathologies. By targeting these proteins it became possible to treat cellular dysfunctions. With the advent of genetic engineering techniques in the 1980s it became possible to produce such protein therapeutics.

Among the many expression systems available Chinese hamster ovary cells emerged as one of the most suitable expression system with their high proliferation and amenability to genetic manipulation to produce human recombinant proteins as therapeutics. Today they are widely used in the biopharmaceutical industry. With an ever increasing demand on the supply of protein based therapeutics and the ever increasing cost of research and development the maximising of yield of protein based biopharmaceuticals is of great interest to the biopharmaceutical industry. While the optimisation of cell culture and bioreactor operation has been an area of optimisation for several decades there is now a move toward more molecular based techniques to maximise biopharmaceutical production from Chinese hamster ovary cells. Despite their huge industrial relevance the molecular mechanisms underlying Chinese hamster ovary cell functions are under studied.

1.2 Chinese hamster ovary cells

Over the course of several decades the biopharmaceutical industry has utilised several different cellular expression systems to produce recombinant proteins for therapeutic applications. The first recombinant therapeutic made in 1982 was insulin using *Escherichia coli* for diabetes mellitus (Berger, Lowe and Tesar 2015). Other non mammalian expression systems include yeast and insect cells. Mammalian cells however are required for protein biopharmaceuticals that require protein folding and post translational modifications (PTM) that can only be performed by mammalian cells (Jenkins, Parekh and James 1996). This is a particular advantage of mammalian expression systems over non mammalian systems to produce proteins with consistent glycosylation and prevent an immune response in the end user of the recombinant protein (Walsh and Jefferis 2006) Among the most widely used mammalian expression systems are baby hamster kidney (BHK) and human embryonic HEK-293 cells. Chinese hamster ovary (CHO) are the most popular of these expression systems.

CHO cells were first isolated in 1957 (Puck et al. 1958, Gamper, Stockand and Shapiro 2005). In the decades that followed CHO cells were found to be well suited to genetic mutagenesis and cellular function experiments in culture (Schneiderman, Dewey and Highfield 1971, Hieber, Beck and Lucke-Huhle 1981, Rajaraman and Faulkner 1984). The large amount of studies established CHO cells as ideal candidates for biopharmaceutical production having high proliferation rates, capable of growing in large scale capacities, easily manipulated in culture and also perform the necessary glycosylation PTMs and protein folding for recombinant proteins to be suitable for human use. By 1986 recombinant protein production research culminated in the first FDA approved recombinant therapeutic protein in Activase®, a recombinant form of tissue plasminogen activator (TPA) produced by Genentech (Wurm 2004). Increasing the recombinant protein yield in CHO has been a focus for industrial research in the decades that followed the approval of Activase® and the many other FDA approved recombinant proteins that followed (**Figure 1.2.1).**

In order to increase recombinant protein yield from CHO cells several culture environment based methods such as temperature shift (Butler 2005), chemical supplements and more targeted genetic engineering approaches to modify industrially desirable characteristics such as apoptosis (Lee et al. 2013) and metabolic rate (Jeon, Yu and Lee 2011) have been routinely used. Using these methods recombinant protein

yields are now in a 5-10 g/litre range compared to 50 mg/litre in the 1980s.. More recently however the publication of the CHO genome has made the prospect of even more targeted approaches possible (Xu et al. 2011). This in turn has allowed transcriptomic data to be made available for proteomic profiling (Becker et al. 2011). Using these "omics" platforms the profiling of CHO cell lines with desirable phenotypic characteristics is now more detailed than ever. Profiling of characteristics such as high productivity (Kelly et al. 2015) or proliferation (Doolan et al. 2010) now allows for the targeted screening of potential gene or protein candidates to optimise CHO cell behaviour for biopharmaceutical production. This will invariably lead to far more focused CHO cell engineering in the future compared to typical mutagenic selection methods in the past and will also improve current methods such as the utilisation of endogenous CHO promoters (Pontiller et al. 2008) that will benefit from CHO "omics" profiling.

1.2.1 Biopharmaceutical uses

Currently CHO cells account for approximately 70% of all recombinant biopharmaceutical production (Jayapal et al. 2007, Huggett and Lahteenmaki 2012). The largest specific categories of biopharmaceuticals are monoclonal antibodies (mAbs) followed by growth factors and hormones (**Figure 1.2.1**). A selection of FDA commercially approved biopharmaceuticals can be seen in **Table 1.2.1**. The complexity of mAbs compared to more simple protein molecules such as growth factors makes them expensive to produce but due to their specific binding to protein targets they are set to gain wider use in diagnostics as well as for disease treatments (Nelson, Dhimolea and Reichert 2010, Jayapal et al. 2007). This further emphasises the requirement for CHO cell optimisation for industrial applications.

Furthermore the emergence of antibody drug conjugates (ADC), which are antibodies to specific disease associated protein targets coupled to a cytotoxic agent, are set to increase the requirement for mABs in the future (Reichert 2013). This may see CHO moving away from its original role in other biopharmaceutical categories and further towards mAbs. A move toward biosimilars (Beck and Reichert 2013) also means that less difficult to produce non mAb biopharmaceuticals may become less profitable further increasing the use of CHO for mAb production.

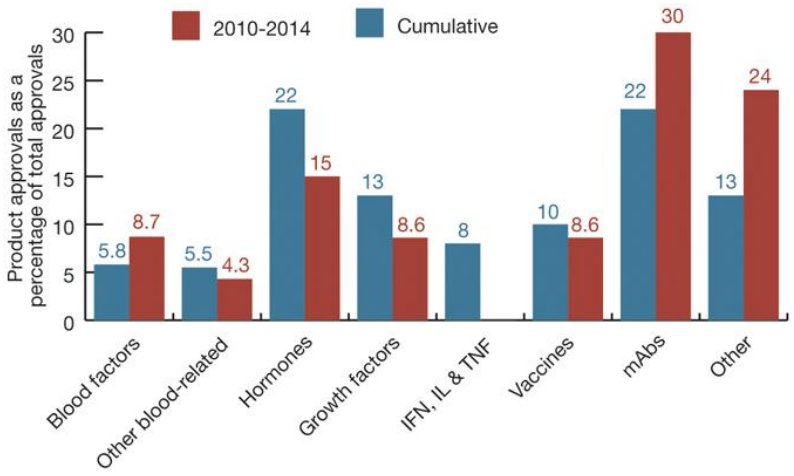


Figure 1.2.1 Top 8 biopharmaceutical product approval categories from 1982-2014 (cumulative) and 2010 to 2014.. Adapted from (Walsh 2014).

Product	Manufacturer	Used for	Protein/*mAb*	Year of FDA approval
Activase	Genetech	Acute myocardial infarction	TPA	1987
Epogen	Amgen	Anemia	EPO	1989
Pulmozyme	Genetech	Cystic fibrosis	DNase1	1993
Rituxan	Genentech	Non-Hodgkin□fs lymphoma	*CD20*	1997
Benefix	Wyeth	Hemophilia B	Factor IX	1997
Herceptin	Genentech	Breast cancer	*ErbB2*	1998
Campath-1H	Genzyme	Lymphocytic leukemia	*CD52*	2001
Zevalin	Spectrum Pharmaceuticals	Non-Hodgkin□'s lymphoma	*CD20*	2002
Humira	Abbott	Rheumatoid artheritis/Crohn's disease	*TNF-ð¢*	2002
Xolair	Genentech	Asthma	*IgE*	2003
Raptiva	Genentech	Psoriasis	*CD11a*	2003
Luveris	Serono	Infertility	FSHB	2004
Avastin	Genentech	Colon or rectum cancer	*VEGF*	2004
Vectibix	Amgen	Colorectal cancer	*EGFR*	2006
Simponi	Centocor	Rheumatoid and psoriatic arthritis	*TNF-α*	2009
Arzerra	GSK/Genmab	Chronic lymphocytic leukemia	*CD20*	2009
Elonva	Howmedica	Ovarian stimulation	*FSHB*	2010

Table 1.2.1 Selection of biopharmaceutical products produced from 1987 to 2010 in Chinese hamster ovary cells.

1.2.2 Cell factory optimisation

1.2.2.1 CHO cell productivity optimization

Increasing the productivity of recombinant CHO cell lines has been one of the key goals of the biopharmaceutical industry for many years and has largely focused on the transfection of selection methods associated with dihydrofolate reductase (DHFR), methotrexate (MTX) and glutamine synthase (GS).

DHFR is required for the catalytic production of glycine and purines from folic acid and tetrahydrofolate. CHO cells deficient in DHFR are transfected with a plasmid containing a gene of interest (GOI) which codes for the recombinant protein to be produced and a functional DHFR gene (Cacciatore, Chasin and Leonard 2010). Without glycine, hypoxyanthine and thymidine media supplementation transfected high producing cells with the GOI are selected for. High DHFR producers can also be selected using MTX which is a chemotherapeutic drug with a long standing association with acquired resistance in cancer and cancer cells via the amplification of DHFR gene (Goker et al. 1995). GS has been another exploited productivity gene. By using a GS inhibitor CHO cells containing extra copies of the glutamine synthase gene along with the GOI are selected (Cacciatore, Chasin and Leonard 2010).

More labour intensive approaches like single cell clonal selection still remain a powerful method for obtaining a high producing cell line (Jun et al. 2005). The inclusion of cell sorting methods has made cloning somewhat faster. Fluorescent-activated cell sorters (FACS) has been used to select for high producers via antibody binding to secreted product on the cell surface (Manz et al. 1995, Borth et al. 2000). Another method ultizes fluorescent MTX to bind to high DHFR producing cells which can then be selected using FACS (Yoshikawa et al. 2001).

1.2.2.2 Media formulation and waste management

Culture media can broadly be divided into serum containing and serum free formulations. Serum containing cultures are mostly advantageous in providing a complete nutrient profile for the large scale clinical use of stem cells. It has been reported that both human mesenchymal stem cells (hMSC) and human embryonic stem cells show a greater than 3 fold difference in expression in 600 genes and greater than 2

fold difference in expression of 95 genes respectively when grown in serum free media (Chase et al. 2010, Skottman et al. 2006). Use of serum free media at least with stem cell production poses problems for the consistency of cell and cell products. Conversely the presence of pathogens and batch to batch variation is a major problem inherent to using serum. For this reason in large scale bioprocessing a chemically defined serum free medium is favoured for quality control, safety and downstream processing purposes. Media formulation consists of specific vitamins, minerals, amino acids and hormones the quantities of which have differing effects on different CHO cell lines (Kim et al. 2005). The requirements of the specific cell line and it's recombinant product have to be taken into consideration when optimising media components and has in turn led to many specific findings related to many recombinant CHO cell lines.

For example a CHO glutamine synthetase producing cell line found that the addition of glutamine at 8 mM reduced the requirement of base buffers and improved antibody yield with no adverse effects (Xu et al. 2014). Conversely with an IgG producing CHO cell line it was found that low concentrations of 2 mM glutamine was optimal for improved titres of antibody production (Rajendra et al. 2011, Parampalli et al. 2007). As glutamine metabolism produces ammonia as a waste product both of these studies show that the cell lines tolerance to the growth inhibitory effects of ammonia play a large part on glutamine supplementation concentration. The large scale culture and manufacture of hMSC has in recent years opened the way for further development of biological supplementation. Humanised media containing human platelet lysate has been shown to expand hMScs in culture (Lange et al. 2007, Gruber et al. 2004). While these methods are successful in some bioprocessing regimes the use of biological material, like with serum use, still poses pathogenic risks that require screening and pathogen inactivation if they are to be viable for large scale approved bioprocessing (Castiglia et al. 2014). Synthetically derived peptides and other synthetic additives as part of a chemical defined medium do not pose these risks.

Synthetic additives also show similar variable effects and require optimisation. One of the most common and earliest used supplements is sodium butyrate (NaBu) (Kruh 1982) which arrests cell growth by the inhibition of histone deactylase 1 (HDAC1) (Davie 2003) and simultaneously enhances CMV promoters ultimately resulting in increased specific productivity (Kim et al. 1999). NaBu has recently been found to have an additive positive effect on specific productivity of CHO cells when used with other synthetic additives caffeine and hexamethylene bisacetamide (HMBA) (Fomina-Yadlin

et al. 2015) showing that even an additive with a long history of use may require further optimisation. In this study it was also shown that caffeine supplementation displayed specific productivity and gene expression similar to that of the no additive control group despite caffeine being known to have a dramatic effect on other cell types (Fredholm et al. 1999). The effect of synthetic additives can be recombinant product specific also such as with the steroid hormone dexamethasone which has been shown to reduce CHO cell death (Jing et al. 2012) and also increases recombinant glycoprotein stability through increased sialylation (Jing, Qian and Li 2010).

Waste products produced by host cell metabolism can be problematic for biopharmaceutical production. Acetate, ammonia, lactate, formate, nitrate and urea are typical waste products produced in culture leading to reduced yield of product, reduced cell growth and reduced cell viability. Each waste product is more often associated with specific expression systems. Acetate accumulation is a particular problem in bacterial fermentation. Acetate accumulation in *E. coli* fermentations has long been known to reduce recombinant protein output (Jensen and Carlsen 1990). While the generation of low acetate producers and acetate tolerant *E. coli* strains has improved yields (Eiteman and Altman 2006) it has been found that an alkaline pH shift to 7.5 reduced acetate up take by 50% and improved cell growth by 71% in an acetate high containing media (Wang et al. 2014). Waste products in culture therefore can be managed by increasing the tolerance of the host organism to the waste products not just be reducing the accumulation of waste products or their removal.

1.2.2.3 Bioprocess optimisation

Increased product yields have mostly been attributed to changes in media formulation (De Jesus and Wurm 2011). Limited batch cultures in large scale is suitable where the desired product is the cells themselves or the cells are unsuitable to be maintained for a long period of time in culture otherwise high product yield is achieved through fed batch systems which typically increase life in culture. This leads to a significant build up of waste metabolites and several modifications to the fed batch system exist such as perfusion culture which allows waste products to be removed from the culture media and maintain cell viability (Bleckwenn and Shiloach 2004). Other physical characteristics such as sheer, pH and temperature are vital in the bioprocess.

The physical apparatus varies depending on the type of culture used. These can classified into two categories with batch and fed batch. Batch cultures consist of growing the culture and harvesting the culture such that the cells are destroyed at the end of the bioprocessing. The culture is then started again and the process repeated. For stem cells and stem cell related products this type of culture is more suitable than continuous culture as stem cells differentiate over time in culture and stem cells are themselves the culture product. Cell line and cell product specific alterations to large scale bioprocessing are required, as with mouse embryonic stem cells (mESC) various strategies such as encapsulation in poly-L-lysine have been shown to reduce aggregation in large scale agitated culture and improve yield of cardiac cell markers for heart cell generation (Jing, Parikh and Tzanakakis 2010). Stirred suspension large scale bioreactor production of human embryonic stem cells (hESC) however is more challenging than mESC as hESC require some level of aggregation to prevent uncontrolled differentiation (Kehoe et al. 2010). Agitation and aggregation therefore play a critical role in the scale up and differentiation during the culture of these cells. Dissolved oxygen has also been shown to be critical in the generation of cardiac cells in both mESC and hMSC (Kehoe et al. 2010, Niebruegge et al. 2009, Bauwens et al. 2005).

The same parameters are important in fed batch cultures. Fed batch culture maintain the cells in culture for as long as applicable to produce the desired product. This can often be a far more complex process in terms of instrumentation involving the continuous monitoring and management of nutrient composition and waste metabolite build up. Bio reactors designed for this type of culture contain many inlet and outlet channels to regulate parameters other than nutrient supply (**Figure 1.2.2**). The life of the culture can be prolonged by the manipulation of many of these parameters. For example, pH control is used to achieve high biomass in *Bifidobacterium* large scale cultures. It has been reported that at a controlled pH 7 it was the concentrations of acid anions at this neutral pH that significantly affected the growth rate of various *Bifidobacterium* strains (Cui et al. 2016). This has also been demonstrated in CHO cell bioprocessing with CO_2 concentration, ammonia and lactate all producing CHO cell line and recombinant product specific effects (Zhou et al. 2010). Gas composition can also prolong culture stages as seen with hESC where increased dissolved oxygen prolonged their undifferentiated culture stage and decreased dissolved oxygen promoted differentiation toward cardiac cell types (Niebruegge et al. 2009). Temperature is another well established parameter to maintain cells at a high biomass for an extended period of time

(see **Section 1.3**) by arresting cell growth and subsequently reducing nutrient requirements, dissolved oxygen for high density cultures (Zimmer et al. 2014) and increasing apoptosis resistance (Kaufmann et al. 1999)

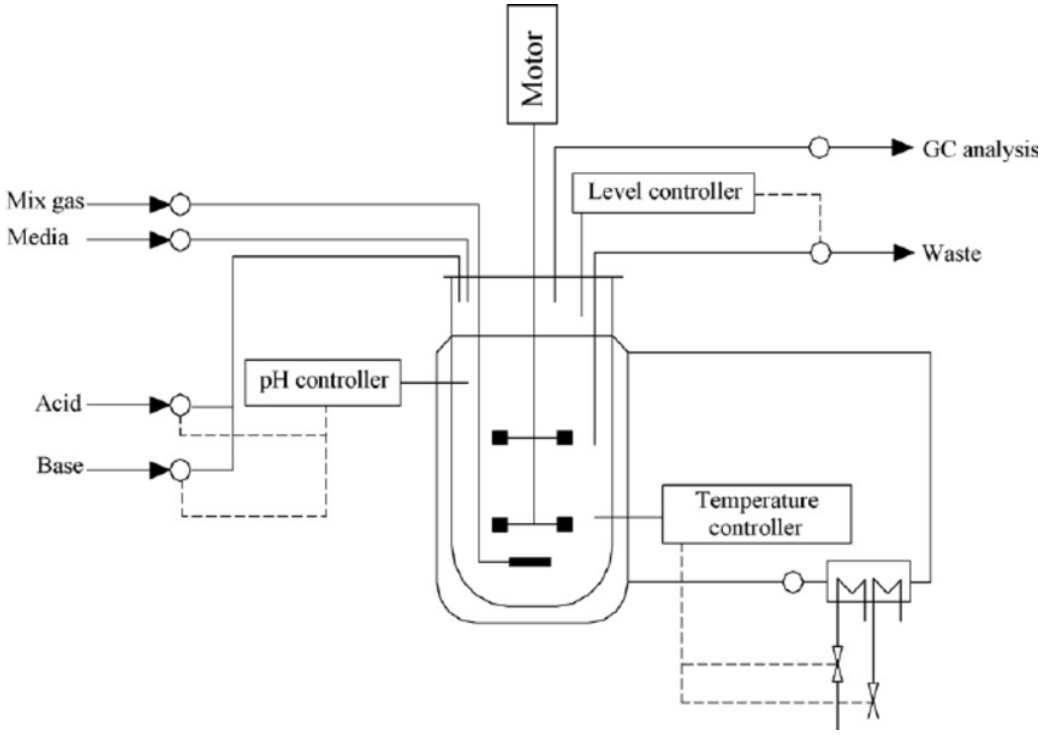


Figure 1.2.2 Bioreactor schematic. Inlets and outlets labelled above allow the culture environment to be controlled. Adapted from (Ismail et al. 2008).

A key contributor to recombinant protein loss is shear stress. Shear stress is an essential parameter in keeping suspension host cells agitated in culture media but can be problematic at high levels. With CHO cells producing recombinant tPA it was found that high shear stress causes significant cell death but also at sublytic levels it can cause recombinant tPA to be only partially glycosylated (Senger and Karim 2003). Similarly in the production of human growth hormone (hGH) it was found that sheer stress at high levels using an anti sheer stress protectant resulted in reduced specific activity of hGH (Keane, Ryan and Gray 2003). In recent years computational fluid dynamics simulations is used to optimise rotor design to deliver the appropriate shear stress to agitate cells while also producing uniform turbulence (Francis et al. 2006, Szymczak and Cieplak 2007). Shear can also occur in equipment further downstream of the bioreactor in the clarification and purification steps. Bioprocessing and its optimisation also includes upstream and downstream operations from the bioreactor (**Figure 1.2.3**).

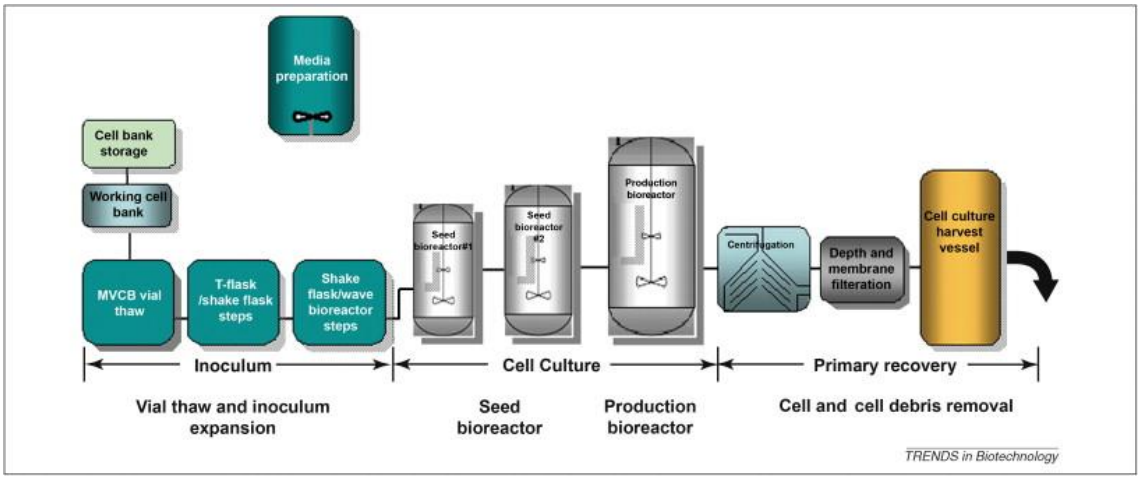


Figure 1.2.3 Bioprocess schematic including upstream and downstream of the bioreactor. Colure and product specific considerations related must be optimised upstream and downstream of the bioreactor. Some of these parameters are shared with bioreactor optimisation such as media formulation upstream and shear damage downstream. Figure adapted from (Shukla and Thömmes 2010).

Clarification is typically employed to remove cell debris and insoluble matter after cell lysis. This is achieve through the use of various pumps to deliver the lysate to the chosen filtration or centrifugation system. Sudden changes in pressure and flow rate during the transportation of the media can adversely affect the protein product leading to aggregation (Gomme et al. 2006) and product loss. In a study investigating the effects of high shear on recombinant hGH and recombinant hDNase it was found that aggregation did not occur but the presence of low molecular weight fragments of hGH suggested peptide breakage occurred (Maa and Hsu 1996).

Contrary to this other proteins studied under high shear conditions such as horse cytochrome c were found to not unfold in response to shearing (Jaspe and Hagen 2006). The negative impact of shear force on the final recombinant product is therefore reliant on a complex relationship between the agitation required, tolerance of the cells to shear force and also the tolerance of the recombinant protein product to shear.

1.3 Temperature Shift

In order to increase the yield of product produced by CHO cells over time the cells are grown at a suboptimal temperature. This is achieved typically by growing CHO cells to a high density biomass at the end of exponential phase before being shifted to 31 ^{o}C which reduces and almost arrests cell cycle progression. This high density biomass of CHO cells can then spend a longer time in culture before declining due to nutrient depletion or waste metabolite build up resulting in increased product yield. This is a very well established method in achieving increased productivity and low to severe hypothermic conditions are well documented as inducing cell survival longevity in different cell types (Schultheiss et al. 2016, Chiou et al. 2013).

Lowering temperature in culture with recombinant CHO cells, however, has long been shown to produce different effects in different producers (Barnabe and Butler 1994, Yoon, Song and Lee 2003) and has varying effects on recombinant antibodies produced (Sou et al. 2015, Kishishita et al. 2015). It is therefore important to assess the molecular effects of temperature shift on specific cell lines to better understand how it produces an industrially desirable phenotype, if any, and how cellular mechanisms may alter antibody production. The identification and subsequent manipulation of key regulators of the temperature shift phenotype will also be valuable in the targeted design of CHO cell lines suited for industry.

1.3.1 Temperature shift mechanism

The phenotype induced by temperature shift and mild hypothermia has long since been observed as reduced cell growth and metabolic activity in mammals (Van Breukelen and Martin 2002). The molecular reasons for this can also be divided into several gene and protein classes attributed with a reduction in transcription and translation, inhibition of RNA degradation, increased transcription of specific gene promoters, alternative mRNA splicing and the differential expression of cold/heat shock proteins (Sonna et al. 2002).

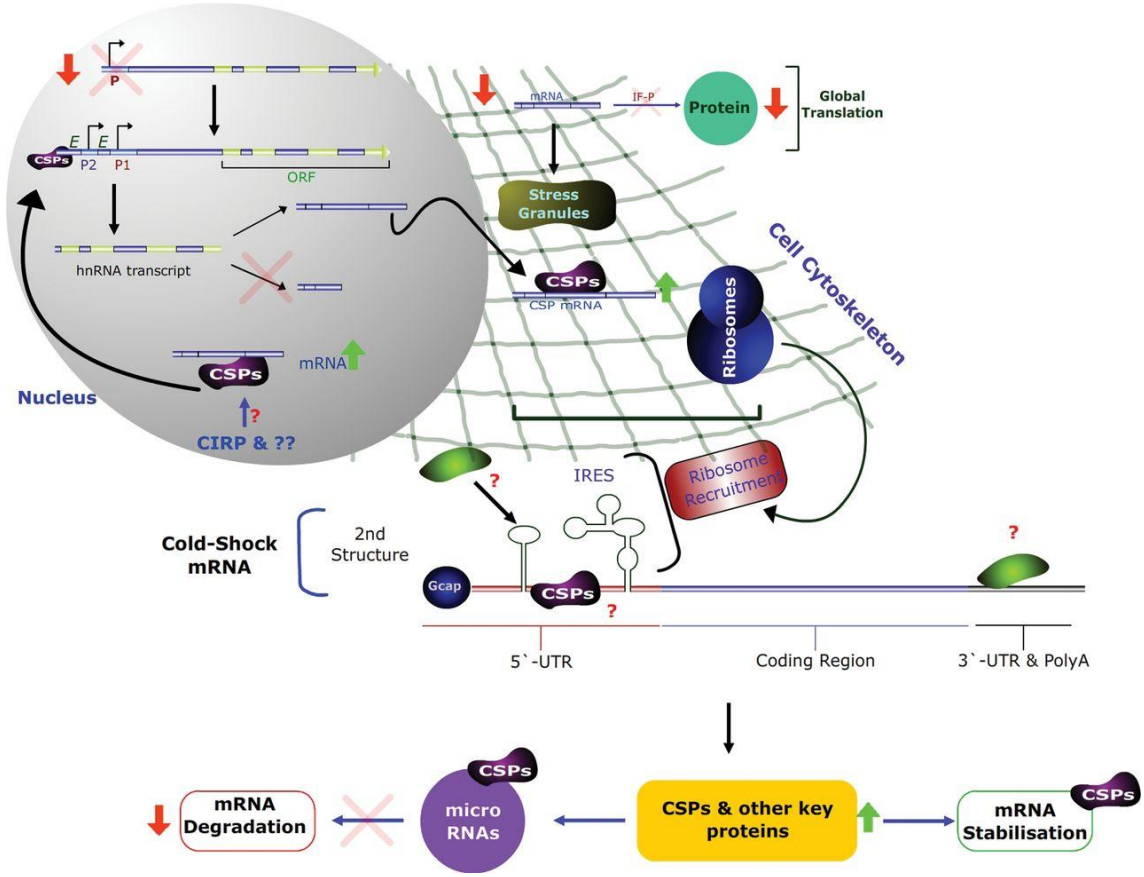


Figure 1.3.1 Temperature shift mechanism (Al-Fageeh and Smales 2006)

A comprehensive review by Al-Fageeh and Smales details what is currently understood about mild hypothermic (25-35 ^{o}C) response of cells in culture (**Figure 1.3.1**). They have proposed that a co-ordinated response occurs between the up regulation of CSPs such as cold inducible RNA binding protein (CIRP) and RNA-binding motif protein 3 (Rbm3) result in the stabilisation of mRNA directly and inhibition of microRNA (miRNA) degradation of mRNA which also leads to increased mRNA levels. Simultaneously there is decrease in mRNA translation in the cytoplasm due to phosphorylation of initiation factors. Initiation factor phosphorylation also allows specific mRNA to be processed into stress granules which contain transcripts necessary for cold shock recovery upon warming. The phosphorylation of elongation initiation factor alpha (eIF2α) specifically also mediates ribosome recruitment to internal ribosome entry segment (IRES) containing mRNA which are possibly mRNA that are required for cold shock adaptation as normal cap dependant translation is reduced during mild hypothermia. Further disruption to the cellular translational machinery is caused by microtubule disassembly associated with the cytoskeleton. The exact mechanism of action of these hypothermia induced genes and proteins still remains poorly understood (Zhu, Buhrer and Wellmann 2016).

1.3.2 Temperature shift and CHO

Despite the exact mechanism behind temperature shift not being known it has seen wide use in CHO cell culture and bioprocessing optimisation. Many studies have shown increased production yields of recombinant protein across various CHO cell lines.

rCHO product	Temperature (^{o}C)	Fold increase in rProtein yield	Source
Antibody-IL-2-fusion	30	3.4	(Shi et al. 2005)
Chimeric Fab	28	38	(Schatz et al. 2003)
Colony stimulating factor	33	2.3	(Fogolin et al. 2004)
C-terminal amidating enzyme	30	4.3	(Furukawa and Ohsuye 1998)
Erythropoietin	33	2.5	(Yoon, Song and Lee 2003)
Interferon-y	32	1.9	(Fox et al. 2004)
IgG	33	2	(Kantardjieff et al. 2010)
SEAP	30	3.4	(Kaufmann et al. 1999)
SEAP	33	8	(Nam, Ermonval and Sharfstein 2009)
Tissue Plasminogen Activator	32	1.7	(Hendrick et al. 2001)

Table 1.3.1 Increased recombinant protein yield in temperature shifted CHO cells.

The specific mechanism behind temperature shift in CHO is even less widely studied than in other cell types but it has been reported that miRNA expression (Gammell et al. 2007), promoter expression (Al-Fageeh and Smales 2013, Thaisuchat et al. 2011) and protein PTMs such as phosphorylation (Kaufmann et al. 1999) all display differential expression patterns in response to temperature shift in CHO cells. As **Table 1.3.1** shows it is also well observed that temperature shift increases recombinant production yield in CHO cells (Al-Fageeh et al. 2006). By indentifying genes and proteins that may be

responsible for this industrially desirable phenotype it would be possible to design a CHO cell line with the same properties as temperature shifted CHO cells.

The temperature shift phenotype in CHO cells has been investigated both at the transcriptional and proteomic level. CHO cells grown at 31 ^{o}C were found to differentially express 26 miRNAs (Gammell et al. 2007). Of these it was found that miR-7 up regulation was found to induce a temperature shift like phenotype (Barron et al. 2011a). A more recent study compared a non producing CHO cell line with a high producer and low producing CHO cell line under temperature shift conditions. In this study 19 miRNA were validated as producing industrially relevant phenotypes (Stiefel et al. 2016).

Transcriptional studies have also identified key candidates involved in CHO temperature shift with 237 genes transcripts having been reported as being between 1.4 and 2 fold differentially regulated in temperature shifted CHO cells (Yee, Gerdtzen and Hu 2009). These transcripts were found to be associated with glycolosis, TCA cycle, pentose phosphate pathway, lipid/cholesterol metabolism, the electron transport chain, protein synthesis and cytoskeletal elements.

Proteomic profiling has been used to identify 53 proteins differentially regulated in CHO as a result of temperature shift (Kumar et al. 2008). Among these identifications were some well known hypothermia associated proteins such as elongation initiation factors and IRES interacting proteins such as heterogeneous ribonucleoprotein C (HNRPC), the knockdown of which has been found to inhibit cell growth (Schepens et al. 2007).

Potential targets for cell line engineering derived from the temperature shift phenotype in CHO cells have therefore already emerged. Individual rate limiting targets in transcriptomic and proteomic data as well have been identified. Potential miRNA effectors of temperature shift have also been identified. These may prove more promising in determining how the mechanism of temperature shift in CHO and also in inducing a temperature shift phenotype as the large amount of transcript and protein dysregulation suggests that a complex multitude of genes and proteins are affected by temperature shift.

1.4 microRNA interference

In 1993 the first miRNA, lin-4, was discovered in *C. Elegans*. (Lee, Feinbaum and Ambros 1993). A 22 nucleotide non coding RNA it was found to negatively regulate the expression of LIN-14 protein by binding to the 3'UTR of lin-4 mRNA. Since their initial discovery the processing of miRNA is now well understood with the identification and functional role of many determined and well established prediction algorithms to determine the direct targets of miR.

The post transcriptional regulatory effect of miRNA are now known to potentially affect ~30% of all protein coding genes (Filipowicz, Bhattacharyya and Sonenberg 2008). As miRNA in mammalian systems do not require full complementarity to their mRNA targets and considering target mRNA binding is largely dependent on complementarity of the miRNA seed sequence (first 2-8 nucleotides) each miRNA can potentially recognise hundreds of targets (Chi, Hannon and Darnell 2012, Hata and Kashima 2015).

This specific but multiple targeting function of miRNA has implicated their involvement in a wide variety of physiological processes and dysfunctions such as insulin secretion and pancreas development (Poy et al. 2004, Correa-Medina et al. 2009) as well as pancreatic cancer phenotypes and diabetes (Park et al. 2009, Pandey et al. 2009). The association between miRNA and phenotype can therefore be exploited as a useful tool for disease diagnosis (Li et al. 2011), treatment (Iorio and Croce 2012) and to improve the production of biopharmaceuticals through the manipulation of specific phenotypes (Cheng et al. 2005, Barron et al. 2011b).

1.4.1 microRNA formation

The biogenesis of miRNA begins with RNA polymerase II/III transcription into a hairpin looped pri-miRNA (primary) several thousand bases in length. The Drosha RNase type III enzyme together with an RNA-binding adaptor protein DGCR8 (DiGeorge Syndrome Critical Region 8) processes the pri-miRNA into a short hairpin loop of approximately 70 nucleotides. The resulting pre-miRNA (preliminary) is then exported across the nuclear membrane into the cytoplasm by Ran-GTP activated Exportin-5 (Bohnsack, Czaplinski and Gorlich 2004). Once localised to the cytoplasm pre-miRNA is further processed by another RNase type III enzyme Dicer together with RNA- binding protein TRBP (TAR RNA binding protein) to produce a double stranded ~22 nucleotide mature miRNA. Argonaute protein family members then load the

miRNA-Dicer-TRBP complex into the RNA induced silencing complex (RISC). In its double stranded form the miRNA is still not active. Thermodynamic stability of the complex largely determines which strand of the duplex will remain in the RISC and become an active miRNA and which strand (passenger/guide strand) will be degraded (Schwarz et al. 2003, Cai et al. 2009).

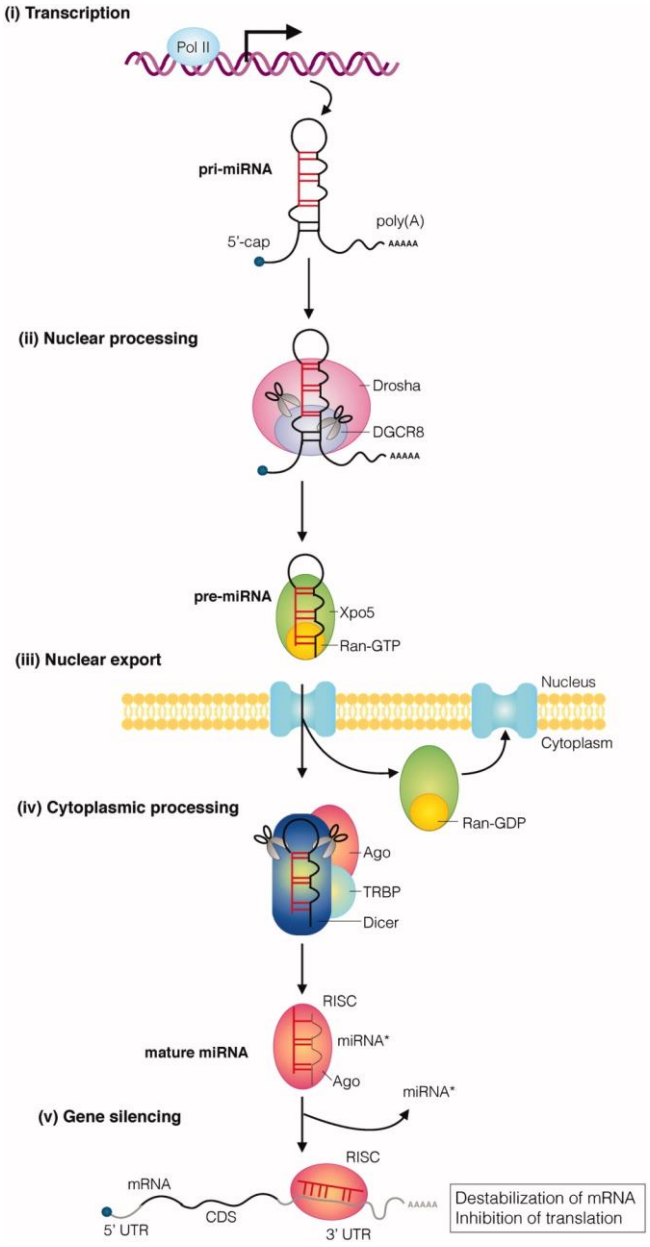


Figure 1.4.1 The biogenesis and processing of microRNA in eukaryotes. Figure sourced from (Hata and Kashima 2015).

1.4.2 microRNA utility and target prediction

The ability of miRNA to inhibit the translation of mRNA is widely recognised as a powerful tool in influencing cell behaviour. In order to understand the targets upon which miRNA act on several algorithms have been developed which predict potential targets. As mentioned previously a single miRNA can have many targets (Chi, Hannon and Darnell 2012). Target prediction allows for more guided approach in determining the most likely targets of a given miRNA. This not only allows study of the phenotype induced by a miRNA but can also identify protein candidates for knock down/out to achieve a phenotype.

The most common way a potential target is found is by searching for complementary base pairing between the 5' 6-8 nucleotide seed sequence of the miRNA and the 3'UTR of potential targets (Peterson et al. 2014, Clarke et al. 2012). Using several prediction algorithms is generally accepted as good practice (Clarke et al. 2012) as output from each can be scored differently. In **Table 1.4.1** are listed some of the most widely used miRNA target searching algorithms.

With that said wet lab validation of miRNA targets and associated targets is always necessary. Recently experimental data has been included in some prediction methods particularly high throughput sequencing technologies capable of analysing miRNA-mRNA complexes (Lu and Leslie 2016). With 28,645 miRNA sequences in the miRNA database miRBase (Kozomara and Griffiths-Jones 2014) target prediction and prediction validation will only become more prominent in the future.

Program	Species specificity	Algorithm description	Web server
miRNA target EMBL	Drosophila	Complementarity with 3′UTR	http://www.russell.embl-heidelberg.de/miRNAs/
miRNAanda	Flies, vertebrates	Complementarity with 3′UTR, thermodynamic stability, duplex species conservation	http://www.microrna.org/microrna/home.do
RNAhybrid	Any	Complementarity with 3′UTR, thermodynamic stability, binding conservation	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/
TargetBoost	Worm and fruit fly	miRNA-mRNA binding site characteristics	https://demo1.interagon.com/targetboost/
miTarget	Any	Thermodynamic stability and sequence complementarity	http://cbit.snu.ac.kr/~miTarget/
Pictar	Flies Vertebrates Worm	Perfect and partial complementary sequence with 3′UTR, Thermodynamic stability	http://pictar.mdc-berlin.de/
RNA22	Any	miRNA-mRNA binding sites characteristics, Complementarity with 3′UTR, no cross-species conservation	http://cbcsrv.watson.ibm.com/rna22.html
MicroTar	Any	Complementarity with 3′UTR and thermodynamic stability	http://tiger.dbs.nus.edu.sg/microtar/
EIMMo	Humans, mice, fish, flies, worms	miRNA binding sites conservation	http://www.mirz.unibas.ch/ElMMo3/
GenMiR++	Any	Sequence complementarity, based on expression data sets	http://www.psi.toronto.edu/genmir/
PITA	Any	Target site accessibility thermodynamic	http://genie.weizmann.ac.il/pubs/mir07/mir07_data.html
NBmiRNATar	Any	Sequence and duplex characteristics, no sequence conservation	http://wotan.wistar.upenn.edu/NBmiRTar/login.php
Sylamer	Any	Based on microarray data to identify 3′UTR sites	http://www.ebi.ac.uk/enright/sylamer/
MiRTarget2	Vertebrates	Based on microarray data to identify 3′UTR sites	http://mirdb.org/miRDB/
TargetScan TargetScan S	Vertebrates	Complementarity with 3′UTR, thermodynamic stability, duplex species conservation	http://www.targetscan.org/
DIANA-microT	Any	Complementarity with 3′UTR, thermodynamic stability, duplex species conservation, combined experimental data sets	http://diana.cslab.ece.ntua.gr/microT/

Table 1.4.1 List of online prediction tools for miRNA targets. (Barron et al. 2011b)

1.4.3 CHO and microRNA

The potential use of microRNA as a tools to manipulate CHO cell phenotypes was only recently considered (Barron et al. 2011b). In the last few years it has been demonstrated that it is possible to manipulate several industrially relevant characteristics of CHO cells using miRNA.

Several studies have used miRNA to successfully target bioprocess relevant phenotypes such as apoptosis, cell cycle and metabolism (Lee et al. 2013, Kelly et al. 2015, Sunley and Butler 2010, Kim and Lee 2007). Currently there are a number of miRNA that have been shown to have potential industrial utility in CHO cells.

miRNA (↑ or ↓)	**Effect (↑ or ↓)**	**Source**
miR-7 ↑	Proliferation ↓, Productivity ↑	(Sanchez et al. 2013)
miR-466h-5p ↓	Apoptosis ↓, Productivity ↑, Product yield ↑	(Druz et al. 2013)
miR-577 ↑	Proliferation ↑	(Strotbek et al. 2013)
miR-1287 ↑	Productivity ↑	(Strotbek et al. 2013)
miR-17 ↑	Proliferation ↑, Productivity ↑, Product yield ↑	(Jadhav et al. 2014)
miR-30 ↑	Proliferation ↑, Productivity ↑	(Fischer et al. 2014)
miR-2861 ↑	Productivity ↑	(Fischer et al. 2015)

Table 1.4.2 List of miRNA shown to produce industrially desirable phenotype changes in CHO cells. "↑" denotes increase, "↓" denotes decrease

Combining industrially relevant culture methods with miRNA screening has also provided many potential candidates for follow-up validation. Differentially expressed miRNAs were identified in nutrient depleted conditions (Druz, Betenbaugh and Shiloach 2012), temperature shift (Gammell et al. 2007, Barron et al. 2011a) and in different growth phases of batch culture (Hernandez Bort et al. 2012). The miRNA identified from these studies can then be used to understand key regulators in CHO cell phenotypes allowing for targeted CHO cell engineering.

1.4.4 microRNA-7 for CHO cell engineering

One such microRNA which has shown a phenotypic effect on CHO cell function is miR-7. As a potential target for CHO cell engineering miR-7 was found to be associated with the temperature shift phenotype in CHO (Barron et al. 2011a). An 8 fold decrease in miR-7 expression after 24 hr at 31 ^{o}C indicated that down regulation of miR-7 was associated with reduced cell density and increased productivity phenotype of temperature shifted CHO cells. Upon transient over expression of miR-7 it was found that CHO cell growth was arrested and specific productivity increased.

A subsequent study investigating the transciptomic effect of miR-7 over expression in CHO revealed potential key targets of miR-7 (Sanchez et al. 2013). In this study it was determined that proteasome activator subunit 3 (PSME3), RAD54-like (RAD54L) and S-phase kinase associated protein 2 (SKP2) were potential direct targets of miR-7. It was proposed that miR-7 binds to SKP2 causing an up-regulation of p27 protein initiating cell cycle arrest. It was also suggested that miR-7 inhibits PSME3 thereby reducing p53 protein and increasing apoptosis resistance. The result from this hypothesis is that these may induce a temperature shift like phenotype in CHO cells.

Several problems are present though in the model. In particular HDAC1 was not confirmed conclusively to be up-regulated which the model suggests as being another potential source of apoptosis resistance. Furthermore the CHO cell line used in the study was not confirmed to express p53. Further study and validation is required to determine the effects of miR-7 and identifying the regulators of the industrially relevant phenotype it produces.

1.5 Proteomics

Proteins form an integral part of cellular function. Their localisation, post translational modification, thermodynamic stability, binding partners and abundance all play a role in their ability to modulate cell function. They study of all of these states in relation to protein is known as "proteomics". Compared to large scale genetic profiling high throughput technologies for proteomic studies are relatively young. As proteomic profiling techniques mature it is becoming clearer that proteins have just as influential a role in cell behaviour as gene expression and in many cases have powerful overriding regulation on cell phenotype (Tsujimoto 2003). A comprehensive profile of protein expression within a cell can therefore provide much information about the driving forces of a given phenotype (Matejovic et al. 2016). The following section will give a brief overview of the current state of proteomic analysis.

1.5.1 Sample preparation

Sample preparation is dependent on the sample source and the method of analysis. There are general and situational considerations when dealing with tissue compared to cells and also the nature of the protein derived from those samples together with the downstream analysis to be carried out.

Clinical or in vivo derived sample tissue requires mechanical or chemical breakdown to access and solubilise protein. In circumstances where a heterogeneous tissue sample is used careful isolation of homogenous tissue samples may be required. This depends on the nature of the experiment but generally the contamination of tissue samples with blood serum is undesirable as high abundant protein such as albumin can mask the detection of lower abundant proteins (Liu et al. 2011a). The overall goal in preparing these samples is to solubilise the desired tissue.

For cultured cell sample preparation there are less concerns about sample homogeneity. What can be challenging with cell samples however is achieving high concentrations of protein for analysis. In growth rate limiting experiments it may be necessary to perform replicates to achieve enough protein. As with tissue analysis this can be aided significantly by achieving a high degree of protein solubilisation and cell disruption.

Cellular disruption and protein solubilisation is carried out for both tissue and cell culture samples. Mechanical disruption of cells often includes sonication or narrow gauge syringe disruption in a chemical solvent. Sodium dodecyl sulphate (SDS) is commonly used to both destabilise the cell membrane and solubilise protein (Xu and Keiderling 2004). Other solvents such as 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and urea are also used to solubilise protein (Fountoulakis and Takacs 2001). Specific considerations may be required for hydrophobic protein solubilisation. Membrane proteins are notoriously difficult to solubilise and as such many protocols focus on increasing their recovery (Duquesne and Sturgis 2010) as well as enriching for them (S**ection 1.5.2.2**)

Sample preparation is therefore primarily focused on obtaining as much protein or proteins of interest form a given sample. Secondary to this the downstream analysis needs to be considered as solvent exchange maybe be required to remove solvents that interfere with gel based separation methods and liquid chromatography/mass spectrometry (LC/MS) based methods. With sufficient extraction of protein from the sample however sample processing methods are easily performed. If sample material and subsequent protein concentrations are limited then compatible solvents may be more suitable from the beginning to avoid protein loss with many purification and processing steps before analysis.

1.5.2 Sample separation

The complexity of a protein sample can be vast. Differential gene splicing, RNA editing and posttranslational modifications of proteins mean that there are >1,000,000 different proteins in the human genome which contains 20,000 - 25,000 genes (**Figure 1.5.1**) (Jensen 2004). This sheer magnitude of complexity results in protein identifications being obscured by more abundant proteins. Separating proteins out based on physical and/or chemical properties allows more identifications to be achieved in a given protein sample.

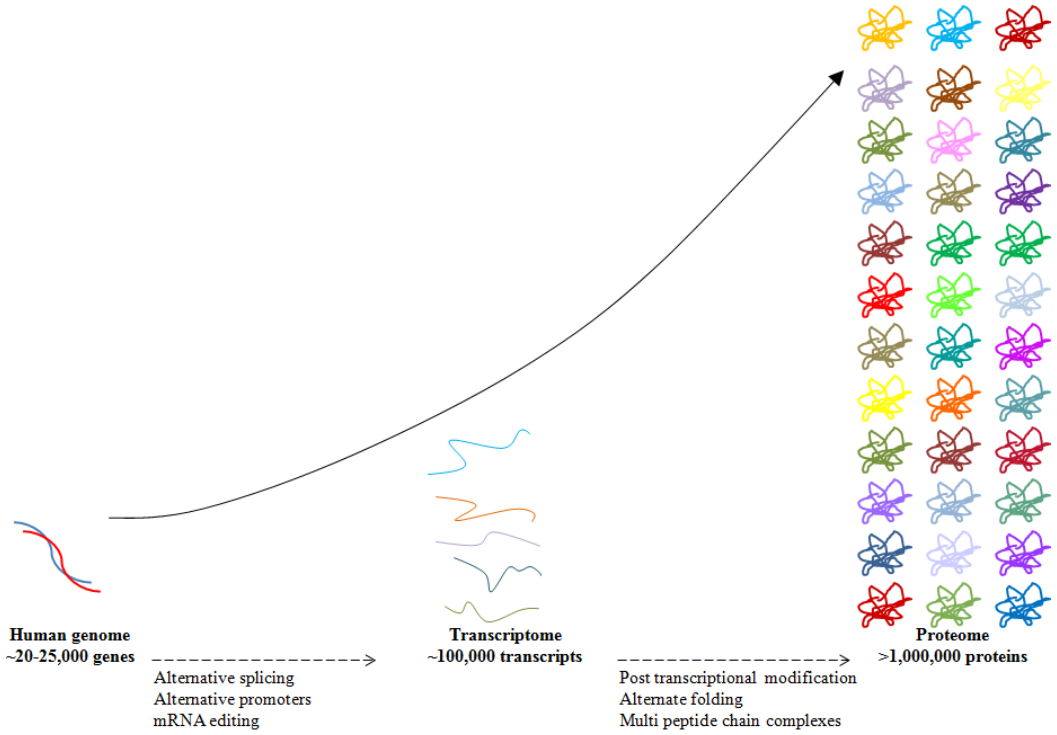


Figure 1.5.1 Proteome complexity in humans. There are an estimated >$1x10^6$ proteins produced from only ~$2.5x10^4$ protein coding genes.

1.5.2.1 Gel based separation

Sample separation was utilised early on in proteomic analysis to allow for proteins to be analysed (Shaw and Prasad 1970). Traditionally this consist of gel based separation techniques such as two dimensional Polyacrylamide Gel Electrophoresis (2D PAGE). This method consists of separation of proteins based on pH on a gel strip followed by the separation of these pH resolved proteins by mass - 2D separation. The result is thousands of resolved proteins on a gel which can be picked individually and processed for mass spectrometry identification. Spots are visualised with either fluorescent dyes or colourmetric such as coomassie or silver stain. Fluorescent labels can be used to perform differential protein analysis known as differential in-gel electrophoresis (DIGE). The various dyes required for DIGE however can be expensive as well as requiring a fluorescent scanner and special quantitative software. Coomassie and silver stain are cheaper but less sensitive alternatives (Winkler et al. 2007). Coomassie is compatible with downstream mass spectrometry analysis but silver stain is not compatible. Silver stain gels require a separate spot picking gel to be run for mass spectrometry analysis and therefore availability of sample material may require

consideration with this method. Ultimately however 2D PAGE techniques, while offering a high degree of separation, are time consuming and expensive and cheaper gel staining alternatives lack the sensitivity of 2D PAGE.

One dimensional gels that only separate by mass are also routinely used but are usually reserved for proteomic technique that do not require high levels of protein separation. Western blot analysis and immuno-precipitation are two immunological based techniques using specific antibodies to bind to a specific protein. Separation on a 1D gel therefore is used to determine a resolved molecular weight relative to a molecular weight standard in order to confirm the specificity of the antibody binding on the gel. A high degree of separation may not be required if the protein of interest has a known molecular weight. A one dimensional gel in this case can be used to isolate proteins within a given weight, cutting out the gel in the correct area of the gel and preparing it for mass spectrometry analysis. Several studies even use this technique to maximise identifications by cutting up a protein lane into multiple bands, analysing each band and combining the resulting data (Petushkova et al. 2015). This has been met with relative success but one dimensional approaches are still more labour intensive and less high throughput than gel-free methods.

1.5.2.2 Fractionation and enrichment

The heterogonous properties of proteins are a characteristic that make them hard to analyse in a high throughput fashion like genes and RNA. This feature of proteins however can be exploited to achieve separation of protein sub populations. Characteristics of proteins are often spread across a variable range such as highly hydrophobic all the way up to highly hydrophilic (Sengupta and Kundu 2012). The word "fraction" usually refers more to the experimental parameters (e.g. Hydrophobic fraction) when in fact the proteins in the fraction have been preferentially enriched (e.g. hydrophobic protein enriched sample). This is an important distinction to clarify as many studies use this term interchangeably (Ly and Wasinger 2008). Some of the most common attributes exploited for fractionation or enrichment are protein mass, protein charge, PTM, differential solubilisation and cellular localisation.

Fractionation by mass can be performed by size exclusion chromatography. These techniques utilise a porous agarose bead matrix within a vertical column. Protein sample

is added to the column with the highest mass proteins eluting at the bottom first. The smallest proteins are the last to elute from the column taking a more convoluted path through the pores in the matrix. Pore size can be used to preferentially separate proteins based on size (Irvine 2001).

Proteins can also be separated based on charge. Similar to size exclusion a charged resin can be used to bind proteins with strong cation exchange resins binding positively charged proteins and strong anion exchange binding negatively charged proteins (Nakatani et al. 2012). Furthermore the proteins bound to the resin can be eluted off the column with increasing salt concentrations achieving greater separation between weak and strongly charged bound proteins (Schmidt, Hafner and Frech 2014).

Charge state can also be exploited in relation to separation of specific post translationally modified proteins. Phoshopeptides carry a positive charge and can therefore be preferentially enriched using cation exchange methods. Specifically titanium dioxide is one of the main resins used for phosphopeptide enrichment (Ruprecht et al. 2015).

Another characteristic often exploited is differential solubilisation. Most solubilisation buffers used in protein sample preparations are aqueous based. These solvents do not facilitate the solubilisation of hydrophobic proteins such as membrane proteins (Devraj et al. 2009). To solubilise these proteins organic non aqueous buffers can be used to solubilise hydrophobic proteins. Membrane proteins in particular don't represent a large proportion of the proteome in the cell which makes these techniques even more important for enrichment. Hydrophobic proteins represent a major class of proteins often associated with cell membranes and often mediate important cell signalling and therapeutic targets (Cho and Stahelin 2005).

Cellular membranes, cytoplasmic and nuclear structures and organelles represent large structures in terms of proteins. As these structures are made of proteins it is possible to isolate them based on their cellular localisation. These methods usually involve centrifugation combined with some of the principles already mentioned such as differential solubility and differential mass of organelles. Gentle buffers are used to first disrupt the cell structure. Low speed centrifugation is then used to isolate the heaviest organelles such as the nucleus, ribosomes, endoplasmic reticulum, mitochondria, membrane (Graham 2015). With heavier organelles removed the retrieval of membrane proteins can be achieved by using differential solubility by using a non aqueous buffer

to enrich for hydrophobic proteins with a final centrifugation step to remove any remaining cell debris. These methods in particular are better described as enrichment as contamination with proteins from other cell structures commonly occurs (Murray, Barrett and Van Eyk 2009).

1.5.2.3 Liquid Chromatography

The separation methods discussed in **Section 1.5.2.2** are often performed in relatively large volumes compared to liquid chromatography (LC). LC can encompass the principles of many of these methods including size exclusion, differential solubility, charge and PTMs but using smaller µl volumes of sample. This can prove particularly advantageous with limited sample material. The physical principles are the same as benchtop chromatography but on a smaller scale with sample passing through a resin (stationary phase) in a liquid (mobile phase).

The nature of the stationary and mobile phases is dependent on the desired separation. Varying the flow rate of mobile phase and the organic solvent composition of the mobile phase over time is often used to manipulate the elution of peptide fragments. The degree of separation therefore can be finely tuned and optimised. Furthermore LC can be used in conjunction with mass spectrometry (MS) achieving in line separation before analysis. Multiple different columns can also be used to achieve even greater separation (Shen et al. 2001).

1.5.3 Mass spectrometry

Mass spectrometry (MS) has become the method of choice for most protein identification and quantification analysis. MS allows for a large number of proteins to be identified in a given sample based on the behaviour of their peptide fragments in a magnetic field. It is compatible with gel based, liquid based and LC based protein separation techniques. The following will outline the steps required to prepare protein samples for MS and analyse the resulting data.

1.5.3.1 Mass spectrometry sample preparation

Sample preparation for MS proteomic analysis encompasses two approaches - "top down" and "bottom up". Top down approaches involve the analysis of whole proteins in their native structural form (Scheffler 2014). Bottom up approaches involve enzymatically digesting the proteins into peptide fragments and then analysing the fragments with MS. The data associated with each peptide fragment can then used to identify the protein and it's abundance. While top down approaches require less sample preparation it is the bottom up approaches which have shown to produce greater number of identifications when dealing with complex protein samples such as cell lysates and tissue (Resing and Ahn 2005).

Protein samples for bottom up proteomics are typically denatured, reduced, alkylated and then finally enzymatically digested. Each step is performed to allow the protease enzyme access to the protein polypeptide chain and facilitate cleavage. Denaturation at 95 ^{o}C and reduction in the presence of dithiothreitol breaks weak hydrogen bonds and strong disulfide bonds respectively. Alkylation, by iodoacetamide, inhibits the activity of cysteine peptidase preventing disulphide bonds to reoccur in a random manner (Boja and Fales 2001). The last step involves the denatured protein being digested with protease enzymes. Trypsin and Lys-C are some of the most commonly used enzymes with trypsin cleaving proteins at lysine and argenine residues (except after a proline) generating peptides 7-20 peptides in length (Siepen et al. 2007) while Lys-C cleaves at the C-terminal side of lysine residues (Gershon 2014). Enzymes that cleave at different residues can also be used in tandem to increase the number of cleavages and produce smaller peptide fragments and ultimately increase the amount of a identified sequence in a given protein (Gauci et al. 2009). The cleavage patterns are predicable and therefore can be used to determine the identity of the protein it belongs to hence the term "bottom up" analysis.

1.5.3.2 Nano-liquid chromatography coupled to mass spectrometry

Protein identification using MS is largely dependent on separation methods. As outlined in **Section 1.5.2** there are various separation methods that can be used to process protein samples before MS preparation. As also mentioned the proteome is estimated to contain >1x10^{6} different proteins. Peptide digestion results in multiple fragments often up to 20

peptides in length and with hundreds of peptides in contained in some proteins the peptide fragmentation produces an even more complex sample by numbers than the protein source sample. Separation of peptides is therefore essential in allowing the MS analyser time to detect as many peptides as possible. High pressure liquid chromatography (HPLC) coupled with a reverse phase column is the most commonly used form of separation coupled with the MS instrumentation. Reverse phase columns achieve separation by exploiting peptide hydrophilic/hydrophobic (polarity) interactions with an inert C18 stationary phase and a charged aqueous/organic solvent mobile phase. By altering the composition of the mobile phase peptides can be retained and eluted over time. Solvent ratio gradients over long periods of time have been shown to increase peptide separation and hence increase the number of protein identifications achieved (Ma et al. 2011). A low flow rate of 200-400 nL/min is also desirable to achieve greater separation (Luo et al. 2006).

The peptides that elute from the HPLC are then submitted to the MS. In order to be suitable for analysis the peptides need to be ionised. The most common ionisation methods are matrix assisted laser desorption ionisation (MALDI) and electron spray ionisation (ESI). MALDI involves mounting peptides onto a plate before they are irradiated by a laser. The resulting ablated gas ionises the peptides which then can be accelerated into the MS instrument. ESI can be easily applied to the terminal end of the HPLC column where an applied voltage disperses the peptides and solvent into an ionised aerosol that is drawn into the vacuum of the MS instrument (Ho et al. 2003).

Mass analysers within the MS instrument operate on various principles all involving the migration of peptides based on their mass and charge through a magnetic field. Time of flight, ion trap, quadrupole and orbitrap mass analysers are all based on this principle. The clearest way to describe this principle is in a quadrupole mass analyser, which contains four rods. Two rods on the same plane (horizontal or vertical) have a direct current positive voltage while the other two rods have a negative direct current voltage applied (**Figure 1.5.2**). On top of this direct current an alternating current is also applied to all four rods. This means that all four rods switch between positive and negative voltages with two on the same plane having a stronger positive charge and the two rods on the opposite plane having a stronger negative charge.

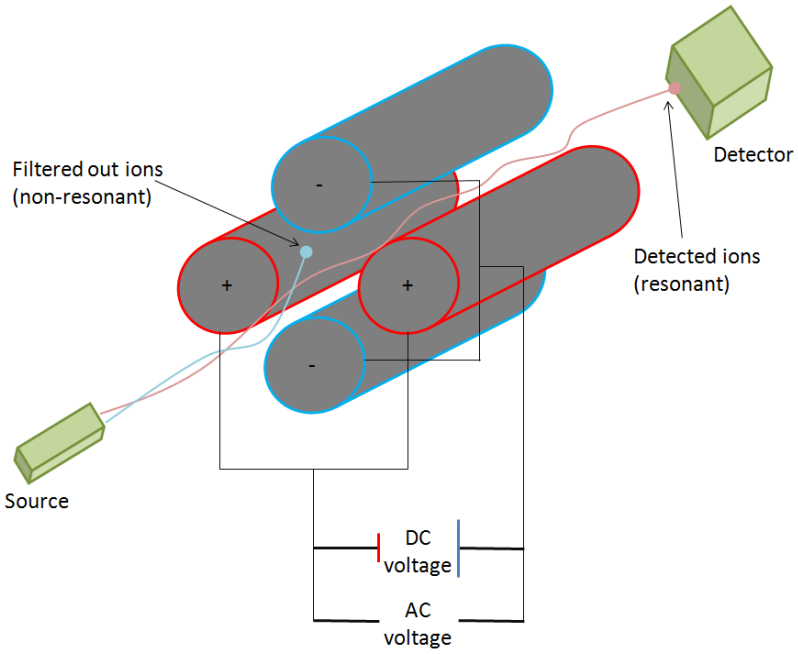


Figure 1.5.2 Quadrupole mass analyser schematic. A direct and alternating current is applied to four rods. The oscillation of the generated magnetic field allows ionised peptides of a specific mass to charge (m/z) ratio to pass through to the detector (resonant ions) while those with an m/z ratio too high or too low collide with the rods or are expelled from the system (non-resonant ions).

This oscillation perturbs the migration of the charged peptides and impedes their migration through the mass analyser to the detector. The overall mass of the peptides also dictates how they migrate through the magnetic field of the mass analyser. In the case of the quadrupole instrumentation the heavier the mass of the peptide the less influenced it will be by weak electromagnetic forces. Conversely strong magnetic forces mean that heavy peptide fragments have a tendency to crash into the rods and cannot readjust their path through the mass analyser as quickly as lighter fragments with the aid of the magnetic repulsion from the voltage switching.

The end result is that the mass analyser acts like a filter for peptides of specific mass (m) and charge (z) states, Anything below a certain mass/charge (m/z) ratio ends up crashing into one set of rods and anything above a certain m/z ratio crashes into the other set of rods. This leaves only a small m/z ratio window that peptides are allowed through the analyser to the detector. By varying the voltage over time this window can be shifted to higher or lower m/z ratios allowing the detector greater time to analyse each m/z ratio subset of peptides effectively amplifying detection sensitivity (Douglas and Konenkov 2014). A linear ion trap mass analyser is a modification of the

quadrupole which allows ions to be trapped and focused tightly before being released through an opening to a detector. This ability to filter ions means that quadrupole mass analysers are now often used in conjunction with other mass analysers and instead of being released to a detector are coupled with other mass analysers.

The orbitrap mass analyser in recent years has been coupled with the linear ion trap to produce a "hybrid" system, "hybrid" referring to the coupling of two different mass analysers in the same instrument. The orbitrap mass analyser can accurately measure mass as low as 2 ppm and has a much higher resolution at 100,000 FWHM (full width at half maximum) compared to quadrupole ion trap instruments at 50 ppm mass accuracy and 10,000 FWHM (Krauss, Singer and Hollender 2010). Orbitrap mass analysers work on a similar principle to quadrupole instruments in that a direct current is applied to a central electrode which is surrounded by a vacuum and enclosed in the excitation electrode. The application of voltage across the external electrode allows for the excitation and expulsion of the ions from the orbitrap towards the detector.

In a linear-trap-quadrupole/orbitrap (LTQ-Orbitrap) hybrid such as the Thermo LTQ-Orbitrap XL^{TM} the linear trap serves to trap and focus peptide ions and selectively deliver focused packets of ions to the orbitrap. Ion fragmentation occurs at a faster rate within the linear ion trap than in the orbitrap while the orbitrap has much higher resolving power than the linear ion. High fragmentation and resolution are essential in achieving a high degree of protein sequence coverage and increasing the number of protein IDs.

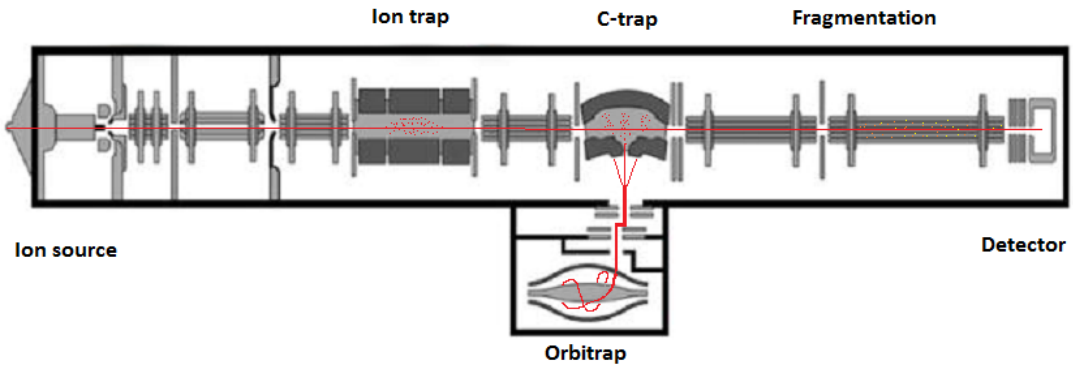


Figure 1.5.3 Thermo Hybrid LTQ-Orbitrap XL^{TM} schematic. Peptide ions enter from the ion source, are filter based on m/z ratio in the ion trap and c-trap, analysed at high resolution in the orbitrap and further fragmented before detection for peptide fragmentation data. Figure adapted from (Rumachik et al. 2012).

This first round of MS can also be followed by further MS analysis. This is referred to as tandem mass spectrometry (MS/MS) (**Figure 1.5.3)**. The peptides of a specific m/z ratio from the first MS are submitted to the second MS where they are fragmented further. Fragmentation is typically achieved using one of several methods such as electron transfer dissociation (ETD), collision induced dissociation (CID), high collision dissociation (HCD) or electron capture dissociation (ECD). CID is one of the most common methods and results in fragmentation of the peptide backbone by an inert gas resulting in different N terminal and C-terminal fragments. Instruments such as the triple quadrupole mass spectrometer perform this type of analysis with an initial m/z sorting quadrupole MS followed by a collision chamber for fragmentation and then a final quadrupole MS. The resulting MS/MS data provides further information on the MS ions, strengthing the quantification and providing more specific IDs.

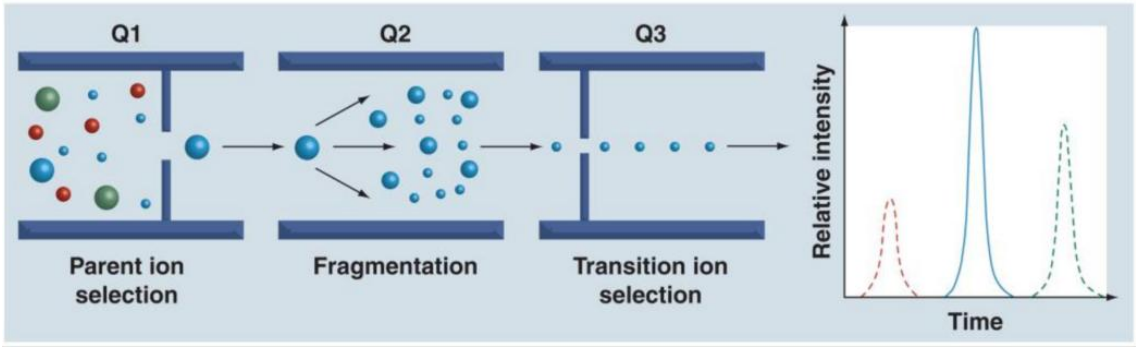


Figure 1.5.4 Work flow of tandem mass spectrometry (MS/MS) analysis (Boja et al. 2010). Q1 involves MS separation of peptide ions, Q2 contains the fragmentation chamber where peptides are further fragmented and Q3 contains the MS that provides the final mass to charge ratio data on the fragmented peptide ions**.**

1.5.3.3 Mass spectrometry data output

At its most fundamental level MS/MS fragmentation allows the mass of each peptide in a peptide ion to be determined. As the peptide ion is fragmented it loses mass and this loss in mass can be used to determine the mass of each peptide (**Table 1.5.1**). Charge state of the peptide ion needs to also be considered as this shifts the mass of the peptide from the first MS m/z ratio. The MS/MS data however for each peptide will produce the same peptide fragmentation peak pattern for the same isotopes. The distance between peaks chromatographic peaks relating to the different masses of fragmented peptides can then be used to determine each amino acid in sequence (**Figure 1.5.4**).

Amino Acid	Letter	Average Mass (Da)
Glycine	G	57.0519
Alanine	A	71.0788
Serine	S	87.00782
Proline	P	97.1167
Valine	V	99.1326
Threonine	T	101.1051
Cycteine	C	103.1388
Isoleucine	I	113.1594
Leucine	L	113.1594
Aspargine	N	114.1038
Aspartic acid	D	115.0886
Glutamine	Q	128.1307
Lysine	K	128.1741
Glutamic acid	E	129.1155
Methionine	M	131.1926
Histidine	H	137.1411
Phenylalanine	F	147.1766
Arginine	R	156.1875
Tyrosine	Y	163.176
Tryptophan	W	186.2132

Table 1.5.1 Amino acids and their average mass in daltons (Da) as determined by MS/MS analysis.

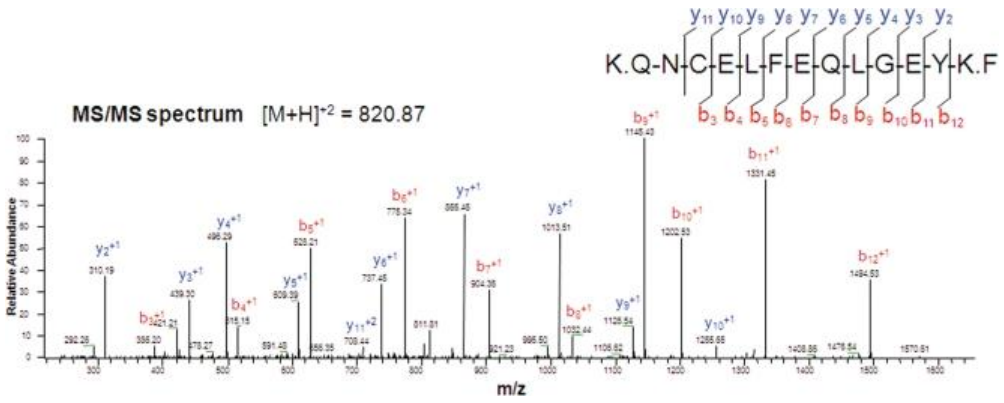


Figure 1.5.5 MS/MS spectrum of peptide fragments for Albumin (Ahn et al. 2008). The peptide sequence QNCELFEQLGEYK is fragmented from the N-terminal end producing "b" ions and from the C-terminal producing "y" ions. Distance between "b" ions or "y" ions which are used to determine the amino acids in the peptide sequence. The "b" ions provide an N to C terminus sequence while the "y" ions provide a C to N terminus sequence. Combining both provides a more confident peptide sequence to determine the protein identification. Unfragmented peptide m/z ratio is also used to obtain the identification with the MS/MS fragmentation data above.

All of this data is collected in a proprietary mass spectrometry computer file. Thermo Scientific software for example outputs a RAW file where as Ab Sciex produce WIFF files. This means that proprietary software is usually required from the same vendor to access the data for further analysis. The overall file sizes can also vary depending on the complexity of the samples.

1.5.3.4 Bioinformatics processing

At the end of the LC-MS/MS analysis there is several pieces of data that are obtained for each protein. From the bottom up there is m/z ratio of fragmented peptide ions, m/z ratio of peptide ions, the abundance of both peptide fragment and peptide ions and finally the retention time associated with the peptides as they emerge from the HPLC. For the purposes of identifying proteins the m/z information is required. This is all contained within the resulting file output from the MS software.

The data from these files can then be used to search spectral libraries and assign protein identifications. Search algorithms such as MASCOT and SEQUEST are commonly used. These algorithms are used to determine if peptide fragmentation patterns and sequences match that of proteins contained in an annotated spectral library (Sadygov,

Cociorva and Yates 2004). By inputting the enzyme used to facilitate protein cleavage the peptides can be predicted for every protein in the protein database. These theoretical fragments can then be searched against the fragment data obtained from the experiment. The experimental conditions have to be accounted for when processing the peptide data.

Trypsin, for example, only cleaves at lysine and arginine residues except after a proline. Missed cleavages can also occur due to various amino acid sequence combinations after or before the lysine-arginine cleavage site (Schechter and Berger 1967, Monigatti and Berndt 2005, Yen et al. 2006). These missed cleavage events can be accounted for in MASCOT and SEQUEST allowing more of the experimental peptide data to be compared to the peptide database. Other modifications to in the peptides from the digestion and separation can also be accounted for such as the fixed carbamidomethylation of cysteine due to iodoacetamide in digestion procedure and the variable oxidation of methionine. For example the oxidation of methionine can be detected by a neutral loss of 64 kDa in the MS/MS data attributed to CID fragmentation removing the methanesulfenic acid (CH_3SOH) group of oxidised methionine (Griffiths and Cooney 2002). These experimental parameters can be accounted for in MASCOT and SEQUEST with the users input but the way in which each algorithm scores the matching between theoretical and experimental peptides differs. MASCOT calculates the probability that each peptide ion mass or MS/MS fragment ion mass from the experimental data is a match with the calculated peptide masses in the protein database (Perkins et al. 1999). SEQUEST compares fragment ions to the MS/MS spectrum and performs a cross correlation analysis against the top 500 scoring peptides (Sadygov 2015). Scoring algorithms are however reliant on input for statistical cut offs from the user, therefore the level of stringency in scoring is user dependant.

1.5.3.5 CHO database

While automated database searching using search algorithms is high throughput it is dependent on the quality of the protein database used to search for theoretical peptides. Up until 2011 there was no genomic or proteomic database available for CHO with studies relying on sequence homology between Chinese hamster and well annotated mouse and rat databases (Doolan et al. 2010, Pontiller et al. 2008, Carlage et al. 2009).

With the release of the CHO genome by Xu et al (Xu et al. 2011) this changed allowing the generation of a CHO specific database. Subsequent additions to this including a cDNA library produced by Meleady et al and a transcriptomic study by Becker et al (Becker et al. 2011, Meleady et al. 2012a) have resulted in a reliable and tested database of identifications that can be used for proteomics. Access to the CHO genome presents numerous possibilities in the way of CHO cell engineering and understanding CHO phenotypes through profiling (Kildegaard et al. 2013).

1.5.3.6 Quantitative label-free LC-MS/MS

Previous sections (**Section 1.5.1.2** and **1.5.1.3**) have discussed how proteins are identified, however the quantification of protein involves further steps. Protein labelling methods have been used over many years producing a great deal of information on cellular physiology. The most widely used labelling methods include iTRAQ, iCAT and SILAC. Labelling strategies are capable of producing very low inter sample variation (Piehowski et al. 2013) but are subject to many limitations such as labelling inefficiency, poor dynamic range and a limit to the number of comparisons that can be made (Chandramouli and Qian 2009).

With the increased capabilities of software analysis and LC-MS/MS separation it is now possible to use a label-free approach (Patel et al. 2009). This method uses spectral counting or ion peak intensity measurements (area under peak curve) as well as the retention time from the LC separation (**Figure 1.5.6**). Spectral counting is calculated from the linear relationship between MS spectra and peptide abundance. Peak ion intensity uses the chromatographic peak intensity of peptide precursors to calculate protein abundance (Neilson et al. 2011, Wong, Sullivan and Cagney 2008).

Label-free methods are more susceptible to sample variation as there is no label used and variation in calculated abundance can be as a result of poor quantification in sample preparations. Label-free data analysis therefore relies heavily on computational data normalisation to minimise sample variation effects leading to inaccurate differential identifications (Listgarten and Emili 2005). As label-free analysis simultaneously identifies and quantifies proteins the balance between MS and MS/MS mode must also be considered. A separate analysis excluding (Hodge et al. 2013) or including (Jaffe et al. 2008) specific m/z ratio peptides for MS/MS spectra analysis can be conducted if

proteins of a specific m/z are being looked for by the user i.e. biomarker discovery, increase sequence coverage. This can serve to improve quantitation or identify low abundant proteins. Overall though these limitations mean that confirmation of differential expression at the protein level or RNA level is routinely performed to validate label-free data.

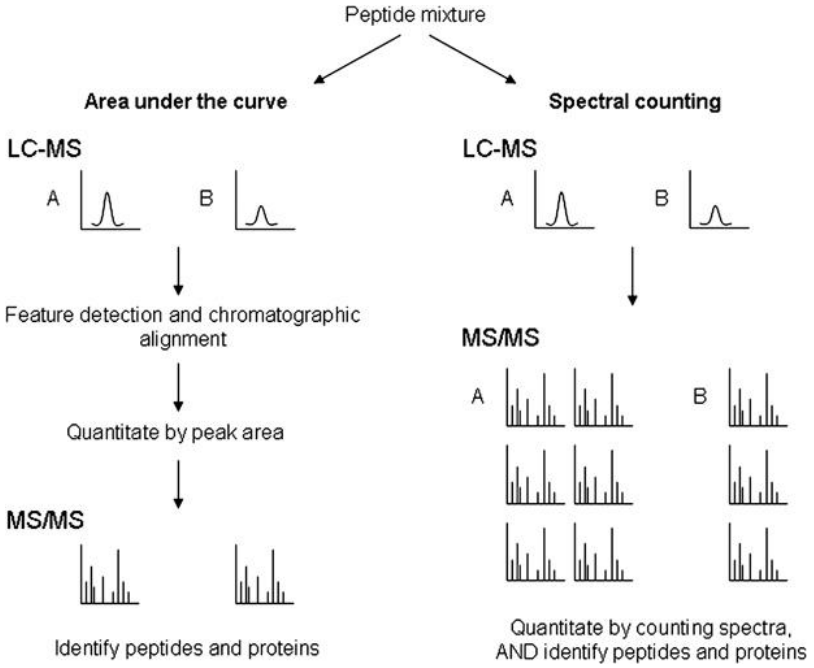


Figure 1.5.6 Schematic of label-free LC-MS/MS peptide and proteins quantitation. The peak represents a m/z ratio and retention time matched between two samples, A and B. Sample A has greater abundance than B for this particular peptide feature (left) or using the spectral count method Sample A has more MS/MS spectra associated with the particular peptide feature than B (right). Figure adapted from (Neilson et al. 2011)

While comparatively new when compared to labelled quantification methods the label-free LC-MS/MS method has now gained considerable popularity mostly due to the large experimental comparisons possible and the reduced sample preparation steps. Studies incorporating label free analysis range from cell studies (Ahn et al. 2008, Van Dyk et al. 2003, Chong et al. 2012) all the way up to the analysis of tissues (Christin, Bischoff and Horvatovich 2011, Maltman et al. 2011, Katz et al. 2010, Piovesana et al. 2016) with associated optimised sample preparations. The limitations, as mentioned in the previous paragraph, require that experimental design considerations are crucial in generating useful and statistically accurate data.

1.6 General introduction - Novel breast cancer proteins

Breast cancer is the most commonly diagnosed cancer in women accounting for 14% of all female cancer deaths worldwide (Jemal et al. 2011). Risk factors include, but are not limited to, age, ethnicity, obesity and lifestyle. The incidence of breast cancer is more common in women over 50 and those with a strong family history of breast cancer. In the last twenty years mortality rates related to breast cancer have decreased sharply due mostly to the implementation of screening and early detection. Administering treatments early on has proven to reduce mortality rates as well and also increase 5 year and 10 year disease free survival (DFS).

Overtime these treatments have become more targeted based on profiling of breast cancer subtypes. The classic subtypes of human epidermal growth factor receptor 2 over expression (HER2+), oestrogen receptor positive (ER+) and progesterone receptor positive (PR+) have been exploited since the early 1990s and remain strong prognostic markers combined with histological grading, node involvement and tumour size (Simpson et al. 2005). While these breast cancers now have several targeted therapies and hormone based therapeutic options available there are still considerable unmet needs associated with triple negative breast cancers (TNBC) which lack expression of HER2, ER and PR and chemoresistant HER2+ disease.

In recent years however both HER2+ and TNBC breast cancers have been further classified into subcategories such as Luminal A and Luminal B subtypes in HER2- and HER2+ breast cancer respectively and Basal like 1 (BL1) Basal like 2 (BL2) subtypes in TNBC (Lehmann, Pietenpol and Tan 2015). Further to this high Ki-67 expression in Luminal B breast cancer has been associated with increased auxiliary lymph node involvement compared to Luminal A subtype (Inic et al. 2014). Another tumour subtype classified as claudin-low expressing are a less common subtype within TNBC breast cancers although not all TNBC breast cancers are claudin-low (Prat et al. 2010). Specific clinical outcomes are also associated with the expression of Cytokeratin 5/6 (CK5/6) (Nielsen et al. 2004), Epithelial growth factor receptor (EGFR) and Vimentin in Basal-like breast cancer (Livasy et al. 2006). Molecular profiling has therefore emerged as an effective tool to unravel the complexity of this disease. With an ever increasing number of breast cancer subtypes there is also a growing need to identify new therapeutic targets.

These relatively recent developments in profiling produced by the fields of genomics, transcriptomics, proteomics and metabolomics have become immense. The omics sciences have since become extremely useful in the field of breast cancer and have given rise to a growing field of computational biology surrounding breast cancer target prediction and discovery (Wirapati et al. 2008, Volinia et al. 2006). With highly successful targeted therapies already available for subtypes such as HER2+ breast cancer there is now a push to identify and exploit more breast cancer specific targets (Vanneman and Dranoff 2012). Acquired drug resistance in HER2+ tumors also means that these subtypes too require further target specific treatments (Vrbic et al. 2013). Identifying a single agent such HER2 that has a dramatic effect on cancer phenotypes is uncommon therefore antibody drug conjugates (ADCs) have been proposed as method to exploit cancer specific antigens that lack therapeutic activity (Sassoon and Blanc 2013). This involves using an antibody against a cancer associated membrane protein for the targeted delivery of drugs to cancer cells. Profiling data will become invaluable in allowing researchers to mine large profiling data sets associated with the many varying breast cancer subtypes and establish subtype specific novel ADC targets.

1.7 Breast cancer subtypes

Breast cancer classification is achieved through several methods from grading according to immuno-histology observations and staging according to the local or metastatic development of the cancer. This is often referred to as the TNM method, tumor size (T), involvement of nearby lymph nodes (N) and presence or absence of metastatic tumors (M) and has been an established and standardised method for decades (Printz 2010). Molecular sub-typing has however also proven to hold great predictive and prognostic value (Carlson et al. 2009) with the use of multi-gene expression arrays and biomarkers (Rakha, Reis-Filho and Ellis 2010, Buyse et al. 2006) The following section will outline some of these subtypes and their clinical relevance.

1.7.1 ER positive

Estrogen receptor (ER) is one of the earliest molecular clinical markers being used since the mid 1970s. It is widely known as a key indicator for early reoccurrence and responsiveness to endocrine treatment (Rakha, Reis-Filho and Ellis 2010). The presence

of ER in breast cancer has now become a key component of much larger gene expression profiles defining the molecular profile of breast cancer (Dai, Chen and Bai 2014). ER over-expressing tumors have been reported to comprise 75% of breast cancers with a higher percentage of 80% associated with patients over 50 (Anderson et al. 2002). Despite its high prevalence ER+ tumors are associated with less aggressive pathologies and have a far better outcome from surgery (Dunnwald, Rossing and Li 2007) than ER- tumors (Putti et al. 2005). Furthermore ER+ tumors have been proven to be responsive to hormone therapies with 50% of ER+ responding to eostrogen inhibitors (Tamoxifen for early breast cancer: An overview of the randomised trials. early breast cancer trialists' collaborative group. 1998) and only a small documented number of ER- responding to hormone therapy (Dowsett et al. 2006). This variability clearly points to other key regulators within the ER subtype. ER status on its own while powerful in predicting breast cancer has only limited use in stratifying patient survival due to high rates of long term survival. It has also been reported that ER tumors differ greatly at the transcriptional and gene level (Farmer et al. 2005, Natrajan et al. 2010). With the application of molecular profiling ER+ breast cancers are often classified as belonging to the luminal subtype**.**

1.7.2 PR positive

Progesterone receptor (PR) is another endocrine activated protein like ER. It has been reported that up to 75% of breast cancers are PR+ (Colomer et al. 2005). Its clinical relevance has been questioned in the past with ER status being a stronger predicting factor in hormone treatment response rates (Olivotto et al. 2004). It has been shown however that up to 10% of all PR+ tumors are ER- (Rakha, Reis-Filho and Ellis 2010) making the vast majority of PR+ tumors also ER+. In contrast to this approximately 40% of ER+ tumors are PR- (Rakha et al. 2007b). These ER+PR- tumors in turn are less responsive to hormone therapy than ER+PR+ tumors (Arpino et al. 2005). Lack of PR expression in ER+ breast tumors may therefore be linked to hormone therapy resistance (Bardou et al. 2003). PR and ER expression have become highly connected and as such ER and PR are often reported together in their respective combinations. The double positive group ER+PR+ accounts for 55 to 65% of breast tumors (Dunnwald, Rossing and Li 2007) with approximately 80% of these being responsive to hormone treatment (Dowsett et al. 2006). Double negative PR-ER- breast cancers on the other hand account

for approximately 25% of tumors and are associated with higher reoccurrence, no response to hormone therapy and poorer overall survival (Bardou et al. 2003). Single positive tumors such as PR+ER- or PR-ER+ groups respond less well to hormone treatment than double positive tumors and show increased expression of other proliferative markers such as epidermal growth factor (EGFR) and HER2 (Bardou et al. 2003). The complex relationship between ER and PR have more recently been given further levels of sub classification. Categorising PR and ER in terms of percentage has shown clinical significance. Over-expressers defined as ER>50%, PR>50% have been shown as highly responsive to hormone therapy, ER/PR<10%, PR/ER>50% are less responsive to hormone therapy and finally ER<10%, PR<10% tumors show no benefit from hormone therapy (Goldhirsch et al. 2007). Both PR and ER then, taken together, represent powerful prognostic markers for patient treatment. Still, poor outcome combinations such as PR-ER- still require further treatment options and classification.

1.7.3 HER2 positive

Human epithelial growth receptor 2 (HER2) has been identified as clinically significant in breast cancer since the late 1980s when it was observed to be 2 to 20 fold amplified in 30% of breast tumors (Slamon et al. 1987). HER2 tumors are now known to account for approximately 50% of invasive ductal breast cancers which are ER-PR- (Dandachi, Dietze and Hauser-Kronberger 2002, Quenel et al. 1995) and therefore do not respond well to ER and PR associated hormone therapies. There are however a number of other solutions such as antibody mediated inhibition of HER2 (Piccart-Gebhart et al. 2005), tyrosine kinase inhibitors (Wang 2014) and aromatase inhibitors (Rasmussen et al. 2008) which are effective against HER2 over-expressing tumors as well as or in conjunction with chemotherapy treatments (Pritchard et al. 2008). While there are a number of treatment options available for this molecular subtype, specific combinations of HER2, PR, ER have more favourable clinical results than others with specific treatments. ER+, PR+, HER2+ for example show very little benefit from single hormone therapy and often benefit more from an anti-HER2 therapy (Cuzick et al. 2011). Combinations with the worst outcome have been shown as ER-, PR-, HER2+ and ER-, PR-, HER2- (Lehmann et al. 2011). Taken together then ER+, PR+, HER2- have tumors have the best outcome while ER-, PR-, HER2- has the poorest outcome with the fewest targeted treatment options.

1.7.4 TNBC and Basal-like

Triple negative breast cancers (TNBC) by definition do not express ER, PR and HER and are resistant to hormone therapies involving these molecular receptors. In reality however these tumors can lowly express these molecular markers depending on the methodology used (Dowsett et al. 2005, Regitnig et al. 2002, Stendahl et al. 2006). This lead to a TNBC false positive or negative diagnosis (Regitnig et al. 2002, Stendahl et al. 2006). While this subtype is attributed to only approximately 15% of all breast cancers its poor prognosis disproportionately makes up a large portion of meta-studies on breast cancer deaths (Boyle 2012, Harris et al. 2006). As with other breast cancer subtypes there are several reported subtypes within TNBC.

While "Basal-like" is sometimes used interchangeably with "TNBC" it has been reported that 77% of Basal-like breast cancers, as defined by phenotype and gene expression signature, are TNBC (Bertucci et al. 2008). Basal-like encompasses a expanding number of immunhistochemistry based markers from the cytokeratin family and epithelial growth factor receptor (Rakha et al. 2009, Cheang et al. 2008) but the most widely accepted definition is ER-, PR-, HER2- with positive expression of EGFR and cytokine 5 and 6. It has even been reported that cytokine expression profiling alone can identify carcinomas that display typical basal morphology (Rakha et al. 2007a, Fulford et al. 2006). Other markers of the basal subtype include several proteins involved in adhesion and metastasis such as vimentin and P-cadherin (Pan et al. 2010).

The basal subtype, while included in the TNBC subtype, have a distinctly worse prognosis than ER-, PR-, HER2- TNBC (Perou 2011) with a poorer response to neoadjuvent chemotherapy (Fan et al. 2006). It is therefore important to have robust clinical markers to distinguish between these closely related subtypes. The BRCA1 mutation has been suggested as one potential marker to make this distinction as this mutation has been repeatedly associated with basal type keratins (Turner et al. 2007). To further complicate basal identification a small percentage (1% to 18%) of these tumors exhibit ER+, PR+ activity (Rakha, Reis-Filho and Ellis 2010). Similarly HER2+ has been observed in a small number of basal marker defined tumors which show resistance to trastuzumab anti-HER2 antibody treatment (Harris et al. 2007). Taking all of this into consideration TNBC and its associated subtypes are often hard to identify. Identifying the specific type of TNBC breast cancer is crucial for individualised treatment, treatments of which are very limited with only modest advances in targeted

therapies compared to hormone receptor positive breast cancers (Crown, O'Shaughnessy and Gullo 2012).

Effective treatment choices are currently limited to chemotherapy (anthracycline) which has proven to be quite effective and surgery. Anti-angiogenesis, such as anti-VEGF, therapy has also shown to be successful in triple negative groups in some data sets as well as in other subtypes. VEGFR inhibition is also emerging as successful. Suggested BRCA1 driven therapy. BRCA1"ness" has also recently been proposed as a method to stratify TNBC into further subcategories for PARP inhibitor therapies (Severson et al. 2015). To date there are no approved targeted therapies for TNBC

1.8 Breast cancer treatment

1.8.1 Conventional therapy

The first line of cancer treatment is usually surgical removal but this depends on a number of factors such as multicentricity and tumor size where larger tumors may first be subjected to chemotherapy (neoadjuvant) in order to save breast tissue (Carey et al. 2007, Liedtke et al. 2008). For small tumors however surgery has proven the most successful and is usually combined with post operative chemotherapy, radiation or targeted drug therapy (Ruiterkamp and Ernst 2011).

1.8.2 Chemotherapy

Chemotherapeutic agents consist of a variety of chemical compounds that cause cytotoxicity. Ideally these have the effect of shrinking tumors, inhibiting tumor growth and ultimately destruction of tumor tissue. These agents can also be used as an adjuvant (after first presentation) therapy to prevent reoccurrence and metastasis. It is particularly useful in treating metastatic tumors as chemotherapeutic agents are delivered intravenously or orally. This however also means that chemotherapeutic agents can have a great have cytotoxic effects on healthy cells leading to the characteristic symptoms of hair loss, nausea and vomiting as well as many other side effects (Tao, Visvanathan and Wolff 2015).

Chemotherapeutics can be classified into several groups depending on their mechanism of action. The platinum group of which cisplatin is a commonly used agent cause DNA damage by the addition of a platinum adduct (Dasari and Tchounwou 2014). The alkylation agents of which cyclophosphamide is a member damage cellular DNA by addition of alkyl groups (McCarroll et al. 2008). Another mechanism is through the inhibition of pyramiding and purine production is by the anti-metabolites group of which methotrexate is a member (Tian and Cronstein 2007). DNA intercolators such as those of the anthracycline group function inhibit replication and DNA repair (Szulawska and Czyz 2006).

1.8.3 Targeted therapy

As mentioned in **Section 1.7** various subtypes of breast cancer are successfully treated using targeted therapies. These therapies target clinically relevant proteins associated with the specific phenotype of a breast cancer subtype inhibiting growth and tumor progression. Most of these targets consist of hormone receptors meaning TNBC breast cancer subtypes do not respond to these treatments.

HER2 positive breast cancer treatment shows some of the greatest examples of sucessful targeted therapies. These therapies act on several members of the epithelial growth factor receptor family including EGFR, HER2, HER3 and HER4. The first of these therapeutics developed was trastuzumab which is a humanised mAb. The antibody binds to the extracellular domain of HER2 receptor which inhibits downstream pathways such as PI3K-AKT-mTOR resulting in inhibition of proliferation (Slamon et al. 2001). While using trastuzumab on its own has proven successful it has also been shown that using it in conjunction with chemotherapy significantly reduces relapse (Piccart-Gebhart et al. 2005). Another highly effective targeted HER2 treatment called lapatinib (a small molecule reversible inhibitor) which targets both HER2 and EGFR (Burris et al. 2005). Lapatinib prevents the activation of pro cancer pathways such as Erk/MAPK (extracellular-signal-regulated kinase/mitogen-activated protein kinase). Afatinib works in a similar manner but has been shown to be more potent than lapatinib (Khelwatty et al. 2011).

While all these therapies are useful they only relate to a sub categories of hormone receptor positive tumors. Tumor resistance associated with trastuzumab treatment has

necessitated its conjugation with a cytotoxic agent emtansine 1 to produce Trastuzumab-maytansinoid emtansine (T-DM1) an ADC (Lu et al. 2013). This approach of conjugating an antibody that targets a cancer specific antigen to a cytotoxic agent is now emerging as a viable targeted therapy strategy known as antibody drug conjugates (ADC) (**Figure 1.8.1**). While the targeted therapies mentioned related to antigens that have a significant role in cell function an ADC target antigen does not necessarily need to have such a function. The requirements for a useful ADC target the antigen must be present on the outside of the cell, ideally be cancer specific to minimise any toxicity to healthy tissue and lastly the antigen will internalise bound antibody thereby incorporating the antibody-drug conjugate into the cell.

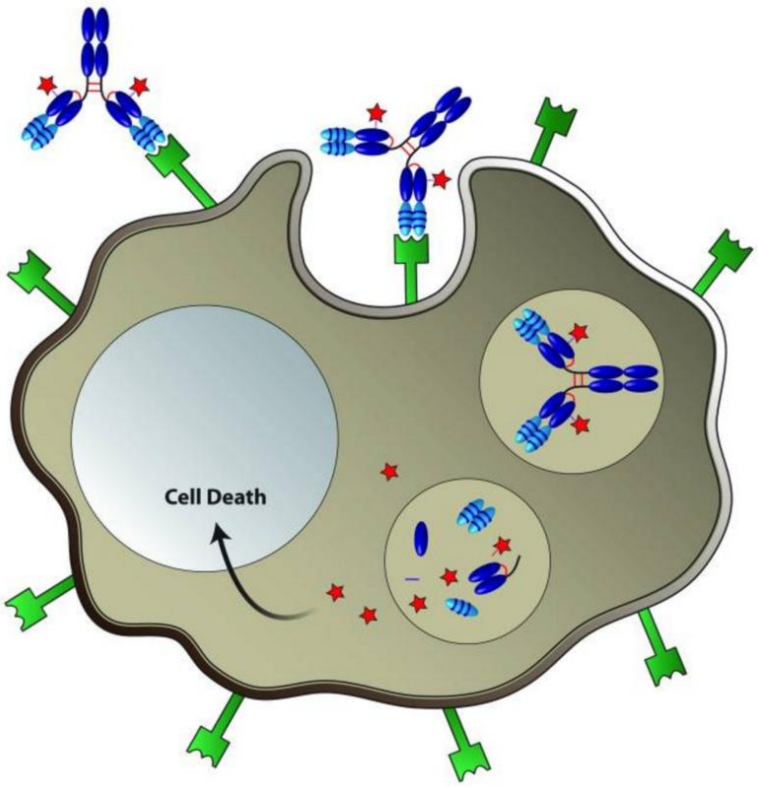


Figure 1.8.1 Delivery of cytotoxic drugs conjugated to a monoclonal antibody (ADC). Binding of cell surface antigen (green) causes internalisation of ADC. Upon internalisation the linker (light blue) is cleaved from ADC releasing the cytotoxic agent (red) (Panowski et al. 2014)

Identifying novel target antigens will be important in developing these new targeted methods.

As well as T-DM1 there are several other targeted therapies available based on monoclonal antibody specificity to cancer specific membrane antigens (**Table 1.8.1**).

ADC (market name)	**Lead**	**Application**	**Target**	**Payload**
T-DM1 (Kadcyla)	Roche	Breast cancer	HER2	Maytansinoid DM1
Brentuximab vedotin (Adcertis)	Seattle Genetics	Hodgkin lymphoma	CD30	Monomethyl auristatin E
Gemtuzumab ozogamicin (Mylotarg)*	Pfizer	Acute myeloid leukemia	CD33	Calicheamicin cytotoxin

Table 1.8.1 ADCs approved for therapeutic use as of January 2016. The use of ADCs is relatively novel with far more ADCs currently awaiting further clinical trial results (see **Table 1.8.2**). *Mylotarg was approved but has since been removed from market by Pfizer after it was shown to have no therapeutic benefit (Petersdorf et al. 2013).

There are also many more currently in the clinical trial pipeline. As of 2013 there were 30 ADCs of varying applications, targets and cytotoxic payload (**Table 1.8.2**). Many of the cytotoxic agents used in these new ADCs are P-glycoprotein substrates such as monomethyl auristatin E (MMAE) and calicheamicin. MMAE and other auristatins inhibit tubulin assembly causing cell cycle arrest at G2/M phase (Sapra and Shor 2013). Calicheamicin agents causes double-strand DNA breaks in cancer cells by binding to DNA's minor groove and can be useful for cancer types with low proliferation rates as cytotoxicity is independent of cell cycle progression (Sissi, Moro and Crothers 2015). The maytansinoid class tubulin inhibitors emtasine (DM1) and ravtansine (DM4) are also used in some clinical trial ADCs.

As with many targeted therapies acquired drug resistance can develop through reduced expression of the target antigen which has been observed with T-DM1 treatment (Barok, Joensuu and Isola 2014). There is also evidence that activation of the P13K/AKT, MEK/ERK and JAK/STAT pathways increases ADC resistance and combination therapies may still be important with ADC treatment (Shefet-Carasso and Benhar 2015).

ADC	Lead	Application	Target	Payload	Phase
Inotuzumab ozogamicin (CMC-544)	Pfizer	Aggressive non-Hodgkin's lymphoma; acute lymphoblastic leukaemia	CD22	Calicheamicin	III
RG-7596	Genentech	DLBCL and follicular non-Hodgkin's lymphoma	CD79b	MMAE	II
Pinatuzumab vedotin (RG-7593)	Genentech	DLBCL and follicular non-Hodgkin's lymphoma	CD22	MMAE	II
Glembatumumab vedotin	Celldex	Breast cancer	GPNMB	MMAE	II
SAR-3419	Sanofi	DLBCL; acute lymphoblastic leukaemia	CD19	DM4	II
Lorvotuzumab mertansine (IMGN-901)	ImmunoGen	Small-cell lung cancer	CD56	DM1	II
BT-062	BioTest	Multiple myeloma	CD138	DM4	II
PSMA-ADC	Progenics	Prostate cancer	PSMA	MMAE	II
ABT-414	AbbVie	Glioblastoma; non-small-cell lung cancer; solid tumour	EGFR	N/D	I/II
Milatuzumab doxorubicin	Immunomedics	Chronic lymphocytic leukaemia; multiple myeloma; non-Hodgkin's lymphoma	CD74	Doxorubicin	I/II
IMMU-132	Immunomedics	Solid tumour	TACSTD2 (TROP2/ EGP1)	Irinotecan metabolite	I
Labetuzumab-SN-38	Immunomedics	Cancer; colorectal cancer	CEA (CD66e)	Irinotecan metabolite	I
IMGN-853	ImmunoGen	Ovarian tumour; solid tumour	Folate receptor 1	DM4	I
IMGN-529	ImmunoGen	B cell lymphoma; chronic lymphocytic leukaemia; non-Hodgkin's lymphoma	CD37	DM1	I
RG-7458	Genentech	Ovarian tumour	Mucin 16	MMAE	I
RG-7636	Genentech	Melanoma	Endothelin receptor ETB	MMAE	I
RG-7450	Genentech	Prostate cancer	STEAP1	MMAE	I
RG-7600	Genentech	Ovarian tumour; pancreatic tumour	N/D	N/D	I
RG-7598	Genentech	Multiple myeloma	N/D	N/D	I
RG-7599	Genentech	Non-small-cell lung cancer; ovarian tumour	N/D	N/D	I
SGN-CD19A	Seattle Genetics	Acute lymphoblastic leukaemia, aggressive non-Hodgkin's lymphoma	CD19	MMAE	I
Vorsetuzumab mafodotin	Seattle Genetics	Non-Hodgkin's lymphoma; renal cell carcinoma	CD70	MMAF	I
ASG-5ME	Agensys	Pancreatic tumour; stomach tumour	SLC44A4 (AGS-5)	MMAE	I
ASG-22ME	Agensys	Solid tumour	Nectin 4	MMAE	I
AGS-16M8F	Agensys	Renal cell carcinoma	AGS-16	MMAF	I
MLN-0264	Millennium	Gastrointestinal tumour; solid tumour	Guanylyl cyclase C	MMAE	I
SAR-566658	Sanofi	Solid tumour	Mucin 1	DM4	I
AMG-172	Amgen	Cancer; renal cell carcinoma	CD70	N/D	I
AMG-595	Amgen	Glioma	EGFRvIII	DM1	I
BAY-94-9343	Bayer	Mesothelioma	Mesothelin	DM4	I

Table 1.8.2 ADCs for solid tumors in clinical trial phases I, II and III. Some ADC targets were not disclosed as denoted by N/D. Most payload agents are maytaninoid (DM1, DM4), auristatin (MMAE, MMAF) or caliceamicin class molecules. Data adapted from (Mullard 2013).

1.9 Aims of thesis

The overall aim of the Chinese hamster ovary (CHO) proteomic profiling is to identify proteins involved in CHO cell phenotypes and potentially identify those which could be used as targets for CHO cell engineering in industry. The following are the specific aims for each CHO study.

- **To investigate the effect of microRNA-7 on the CHO proteome**

Previous work in our laboratory revealed several microRNAs, e.g. miR-7, to be differentially regulated in temperature shifted CHO cells. Growing cells at a lower temperature (temperature shift) produces a phenotype that prolongs growth in culture, reduces proliferation and overall increases productivity over time.

Transient miR-7 over expression in cells grown at 37 ^{o}C was found to produce a temperature shift phenotype (Barron et al. 2011a). As microRNA regulate protein translation we proposed using quantitative label-free LC-MS/MS analysis to identify proteins differentially regulated in response to miR-7 over expression. The resulting data would allow us to investigate predicted miR-7 direct targets from bioinformatics tools and to determine the specific effect of miR-7 using pathway analysis. Also since microRNAs negatively regulate protein translation it would allow us to identify potential knock down targets in the down-regulated proteins to potentially induce a temperature shift phenotype.

- **To investigate the deeper proteome of temperature shifted CHO cells**

In the last number of years in our laboratory we have developed a quantitative label-free LC-MS/MS platform for proteomic research. The principal method used previous to this was 2 dimensional gel electrophoresis (2D-DIGE) (Kumar et al. 2008). Quantitative label-free LC-MS/MS methods have a higher throughput than methods such as 2D DIGE. One of the key disadvantages of quantitative label-free LC-MS/MS however is the volume of peptide data that the mass analyser has to process compared to labelled methods. Highly abundant peptides can mask the identification (ID) of less abundant peptides and ultimately reduces the number of and quality of the resulting protein IDs.

We proposed using simple benchtop cellular component enrichment kits to reduce sample complexity. This should lead to a larger number of IDs overall between enriched fractions compared to the unfractionated sample. It is also expected to produce a number of IDs that are unique to the enriched fractions compared to the unprocessed sample. The functional effect of these proteins on CHO cell phenotype was investigated with the aim to replicate the characteristics of temperature shift.

The breast cancer study was conducted with the overall aim of identifying novel membrane expressed protein targets showing high differential tumor/normaltissue expression and also to confirm these targets suitability as potential antibody drug conjugate (ADC) candidates. The following were the specific aims of the novel breast cancer target study presented in this thesis.

- **To identify potential membrane proteins differentially expressed in breast cancer versus normal breast tissues.**

(A) Using bioinformatics profiling of publicly available, differentially expressed, transcriptomic data on various breast cancer subtypes to identify potential membrane proteins highly expressed in breast cancer sub-types with respect to normal breast tissue; (i) triple negative breast cancer (TNBC), (ii) oestrogen receptor positive (ER+), (iii) lymph node positive (LN+) and (iv) human epidermal growth factor positive (HER2+) sub types.

(B) Using uniprot database information to predict potential membrane localised proteins. We wished to identify proteins that were, ideally, novel breast cancer candidate targets, i.e. protein targets which previously had not been documented to have a functional role in breast cancer and furthermore could be demonstrated to be present in the membrane. The approach was to validate their expression across a large panel of breast cancer cell lines (whole cell and membrane enriched extracts) representing the above subtypes.

- **To investigate the expression of these potential novel membrane protein targets in breast cancer sub-types, and normal breast tissues.**

With new breast cancer subtypes and profiles being generated with better profiling and computational techniques as well as resistance to certain treatments there is still a constant search for potential drugable targets to widen treatment options.

The overall aim was to validate the expression of potential targets that could be used for antibody drug conjugate treatment (ADC) as targeted delivery method for cytotoxic drugs. Such a target will ideally be present at high levels in the membrane of breast cancer cells so as to be accessible to the ADC and have low to negligible expression in other healthy tissues. The work presented in this thesis aimed to investigate a panel of membrane associated targets in specific breast cancer sub-types to address if these candidate proteins may have the potential to be further investigated as ADC molecular targets for breast cancer.

CHAPTER 2

Materials and Methods

2.1 Cell culture techniques

2.1.1 Preparation of culture media

Glassware used for cell culture was soaked in a 2% RBS-25 (Chemical Products R. Borghgraef S.A.) for 1 hr and washed in an industrial dishwasher with Neodisher detergent and rinsed twice with UHP. Ultrapure water (UHP) was purified with a reverse osmosis system (Millipore MillI-RO 10 Plus, Elgastat UHP) to a standard of 12-18 MΩ/cm resistance. Autoclave sterilisation was carried out at 121 ^{o}C for 20 min at 15 bar (Thermolabile solutions were filtered through 0.22 μm sterile filters (Millipore, Millex-GV SLGV025BS)).

2.1.2 Suspension culture

Cell culture was carried out in a class II down flow re-circulating laminar flow cabinet (Nuaire Biological Cabinet). Laminar flow cabinets are swabbed with 70% industrial methylated spirits (IMS) before and after use. A 15 min clearing step was observed in between working with different cell lines. Operation of laminar flow cabinets was done under strict aseptic techniques. Cell culture cabinets and incubators were also cleaned with industrial detergents (Virkon) and IMS on a weekly rota cleaning schedule

Suspension cultures were grown in either sterile disposable 250 ml vented cap flasks (Corning, cat. 431144) with a 30-50 ml working media volume or a disposable 50 ml vented cap spin-tube (Sartorius, DF-050MB-SSH) with a 2-5 ml working media volume.

Suspension adapted CHO-K1-SEAP cell cultures were maintained at 37 ^{o}C in an atmosphere with 5% CO_2 with 80% humidity in an ISF1-X (Climo-Shaker) Kuhner incubator at 170 rpm. Cells were passaged every 3 days, spun down at 170 x g (1000 rpm) for 5 min and the pellet resuspended in 10 ml fresh media for a new flask. Suspension CHO-K1-SEAP were grown in serum-free medium in CHO-S-SFM II (Gibco, 12052114), a complete, serum free, low protein (<100 μg/ml). Media was supplemented with geneticin (Sigma-Aldrich, A7120) selection agent at 1000μg/ml. Cells were counted as in **Section 2.2.1** and seeded at a concentration of 2x10^{5}cells/ml.

2.1.3 Adherent culture

A total of 10 adherent breast cancer cell lines were used. As listed in **Table 2.1.1** cells were grown in RPMI 1640 (Gibco, 52400025), Dulbeco's minimum essential medium (DMEM) (Gibco, 52400-025) or high glucose media DMEM GlutaMAX™ (Gibco, 61965-026) as required. Media was supplemented with 10% (v/v) foetal calf serum (FCS) (PAA, GE Healthcare BioScience Corp), 2% (v/v) L-Glutamine (L-Glut) (Gibco, 11140-0350) and/or 0.5% (v/v) insulin (Thermo, 41400045) as required and detailed in **Table 2.1.1.** Cells were grown in vented 75 cm^2 (Costar, 3276) or vented 175 cm^2 flasks (Costar, 431466) as required at 37 ^oC in an atmosphere with 5% CO_2. Specifically MDA-MB-157 required 175 cm^2 flasks for cell lysate preparations due to low protein concentrations in this cell line.

As with suspension culture in **Section 2.1.2** cell culture was carried out in a class II down flow re-circulating laminar flow cabinet (Nuaire Biological Cabinet). Laminar flow cabinets are swabbed with 70% industrial methylated spirits (IMS) before and after use. Only one cell line at a time was manipulated inside the laminar flow cabinet with a 15 min clearing step observed in between working with different cell lines to avoid cross contamination.

Cell line	Basal media	Supplementation	Source
BT20	DMEM	10% FCS	NICB
BT474	RPMI 1640	10% FCS, 2% L-Glut	NICB
HS578T	DMEM	10% FCS, 2% L-Glut, 0.5% insulin	NICB
MCF7	DMEM	10% FCS, 2% L-Glut	NICB
MDA-MB-157	DMEM GlutaMAX™	10% FCS	NICB
MDA-MB-231	RPMI 1640	10% FCS	NICB
MDA-MB-361	RPMI 1640	10% FCS	NICB
MDA-MB-468	RPMI 1640	10% FCS	NICB
SKBR3	RPMI 1640	10% FCS	NICB
T47D	RPMI 1640	10% FCS, 2% L-Glut	NICB

Table 2.1.1 Breast cancer cell lines used in Chapter 6. All were grown in vented 75 cm^2 or vented 175 cm^2 flasks at 37 ^oC in a 5% CO_2 atmosphere.

Once media exhaustion was observed the flask was rinsed with 2 ml PBS solution to remove spent medium. Depending on the size of the flask, 2-5 ml of trypsin solution (0.25% (v/v) of trypsin (Gibco, 043-05090) and 0.01% (v/v) of EDTA (Sigma, E9884) solution in PBS (Oxoid, BRI4a)) was then added. Cells were incubated at 37 ^{o}C for approximately 2-5 min or until all of the cells detached from the inside surface of the flask. To deactivate trypsin an equal volume of complete media (containing FCS) was added to the flask.

For membrane protein isolation preparations of adherent cells we used Cell Dissociation Buffer (Gibco,13151-014) instead of trypsin due to the likely hood of membrane protein cleavage by trypsin (See **Section 3**). Detachment was monitored by microscope.

Cells were then spun at 170 x g for 5 min in a sterile universal container (Sterilin, 128a). The supernatant was removed and the cell pellet was resuspended in fresh medium. A cell count was performed as described in **Section 2.2.1** with an aliquot taken to seed a new flask. Waste media and cells were sent for autoclave inactivation.

2.1.4 Cryopreservation of cells

Cells for cryopreservation were harvested in mid-log phase of growth and were counted as described in **Section 2.2**. Cell pellets were resuspended in a suitable volume of serum and an equal volume of an ice cold filter sterilized (0.22 μm) solution of 10% (v/v) DMSO (Sigm-Aldrich, D5879) in serum was added dropwise while mixing the cell suspension. 1 ml of cell suspension was aliquoted into the cryovials (Greiner, 122278) and stored on ice during immediate transport to the -20 ^{o}C freezer for 1 hr. Cryovials were then placed in a -80 ^{o}C freezer for four hr or overnight followed by transfer to liquid nitrogen tank for long term storage at -196 ^{o}C.

2.1.5 Thawing cells

Upon removal from liquid nitrogen the cryovial was thawed in warm water. Following observed thaw, pre-warmed media (37 °C) was added to thaw the pellet fully. The cell suspension was centrifuged at 170 x g for 5 min, the supernatant removed and the resulting pellet resuspended in warmed fresh media. To remove any remaining DMSO the suspension cells were grown in 5 ml suspension spin tubes for 24 hr and the media

replaced. Similarly adherent cells were grown for 24 hr in 75 cm^2 vented culture flask and media replaced.

2.1.6 Mycoplasma testing

Routine screening for Mycoplasma was carried out every 4 months on cell lines using the Fluorescent Hoechst stain method by Mr. Michael Henry. Mycoplasma-negative NRK (Normal rat kidney fibroblast) cells were used as indicator cells for this analysis. NRK cells were incubated with a sample of supernatant from the cell lines being tested for the presence of mycoplasma and then stained.

2.2 Cell counting and viability

2.2.1 Trypan blue

Trypan blue (Gibco, 525) is a dye exclusion technique that penetrates and stains dead cells blue, excluding live cells with intact membranes. This was the most routine counting method used for cell culture experiments. Equal amount of cells and trypan blue were mixed. 10 µl of this mixture was transferred to a haemocytometer (Neubauer) and covered with a coverslip. Live cells and dead cells were counted in four corner grids **(Figure 2.2.1**) in the haemocytometer and the average calculated.

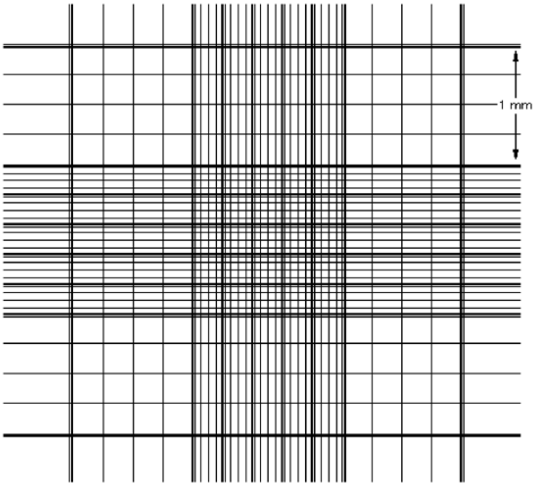


Figure 2.2.1 Haemocytometer grid. Four corner grids are counted and the average viable (non dyed) and dead (dyed) cell number is determined.

The cells/ml concentration was calculated by multiplying by 10^4 (the volume of the grid) and the dilution factor used when mixing with trypan. Percentage viable cells were calculated by counting the number of non dyed live cells and dyed dead cells to get the total cell number. Viable cells percentage is then expressed as ("dead cell number"/"live cell number") x 100).

2.2.2 Cedex automated cell counter

The Cedex Automated Cell Counter (Roche Innovatis AG) is an automated cell counting system based on the Trypan Blue exclusion method. As well as cell counts and viability it also determines cell size. It was employed in the cell density, cell viability and cell size measurements in **Section 4.4.**

Each Cedex XS Smart Slide contains 8 chambers that can hold 10 µl. Cells were diluted with Trypan Blue below 1 x 10^7 cells/ml in order to obtain an accurate cell count. Four chambers on each slide can be read after insertion into the slide carrier of the instrument. To read the 4 chambers at the other side the slide is removed from the slider and positioned in the opposite direction. The slide can be pushed into the instrument into four positions to read each chamber.

To avoid cell stress 4 samples were prepared at a time. Initially 5 µl of cells was mixed with 5 µl of Trypan blue in an eppendorf tube. Following Cedex XS readings dilutions were performed as necessary to achieve a countable cell number.

2.2.3 Flow cytometry

Guava Viacount® reagent (Merck-Millipore) uses two DNA dyes to count cell through the EasyCyte flow cytometry system. The membrane permeable-dye (LDS-751) stains all nucleated cells and is detected by photomultipler tube 2. The membrane impermeant-dye (propidium iodide) stains only damaged cells distinguishing viable, apoptotic, and dead cells and is detected by photomultiplier 1. Furthermore viable cell fluorescence is accompanied by a forward light scattering measurement. If the forward light scattering is appropriately large then the event is counted as a live cell but if it is too small it is counted as cell debris.

Cell counting was performed in a round bottom 96 well plate format (Costar, 10308005) which was loaded into the sample tray. No less than 100 µl should be used per well to allow the sample syringe to acquire. Cell count must also be >10 cells/µl and <500 cells/µl to obtain an accurate reading.

Guava Viacount® reagent was allowed to reach room temperature in the dark before aliquoting the required amount into a sterilin (100 µl per sample). Cells in serum free media were then diluted to 100 µl to achieve a >10 cells/µl and <500 cells/µl concentration of cells. The 100 µl volume of diluted cells were then added to the 96 well round bottom plate followed by the room temperature Guava Viacount®. Cells and dye were left to complex for 10 min before reading on the Guava flow cytometer. WorkEdit™ Software was used to acquire the data with care taken to enter the correct cell dilution factor. EasyFit software analysis processes the data and reports viable cells/ dead/apoptotic cells and cellular debris.

This counting method is capable of counting very low cell concentrations and was employed in **Section 3.1** as this experiment resulted in cell cycle arrest and low seed concentrations of cells (1 x 10^5 cells/ml). These would have been impossible to count using the Trypan Blue methods in **Section 2.2.1** and **Section 2.2.2** and have sample remaining for proteomic analysis.

2.3 Molecular techniques

2.3.1 Transfection with miRNA/siRNA

CHO-K1 SEAP cells seeded at 1 x 10^5 cells/ml with a viability at minimum 90%, were transiently transfected with a total concentration of 50 nM double stranded miRNA mimic molecules (GenePharma #M-01-D), non specific double stranded controls (GenePharma, #M-03-D, double stranded microRNA mimic negative control, GenePharma) or siRNAs (Custom design, Integrated DNA Technologies). Both miRNA and siRNA were purchased in a lyophilised form and reconstituted using nuclease-free water (Ambion®, AM9932) in a laminar flow cabinet. miRNAs and siRNAs were made up to a final stock concentration of 50 µM. Reconstituted miRNA/siRNA were vortexed for 1 min, centrifuged briefly and stored at -20 ^{o}C. Sequences for miRNA can been seen in **Table 2.3.1** and siRNA in **Table 2.3.2** below.

Transfections were carried out using SiPORT® NeoFx™ transfection reagent (Ambion, AM4510), a lipid based transfection reagent. The following describes a typical transfection using a final concentration of 50 nM miRNA/siRNA. In all subsequent volumes 10% extra is included to account for potential pipetting errors. To start 2.2 µl of 50 µM siRNA/miRNA stock was added to 110 µl warmed (37 ^{o}C) serum free CHO-S-SFM II media. NeoFX™ was allowed to reach room temperature before adding 2.2 µl to a separate pre warmed (37 ^{o}C) 110 µl volume of CHO-S-SFM II. To complex the miRNA/siRNA with the NeoFX™ both 112.2 µl mixtures are added together. The resulting 224.4 µl siRNA/miRNA/NeoFX™/Media complex was allowed to incubate for 10 min at room temperature. Inoculation of 1.8 ml of cells at 1.1 x 10^5 cells/ml (2 x 10^5 cells) with 200 µl of the 224.4 µl siRNA/miRNA/NeoFX™/Media complex results in 2 x 10^5 cells/ml transfected either an miRNA/siRNA concentration of 50 nm. Care was taken to add the 200 µl complex drop wise to the cells with constant swirling agitation to minimise transfection cell stress. All transfections were performed in triplicate and for relevant controls. Volumes above were multiplied accordingly depending on the number of samples required.

hsa-mir-7 mimic (functional strand)	$^{5'}$uggaagacuagugauuuuguugu$^{3'}$
Double stranded non specific control	$^{5'}$uucuccgaacgugucacgutt$^{3'}$ $^{5'}$acgugagacguucggagaatt$^{3'}$

Table 2.3.1 Sequence of miRNA mimic and control used in Chapter 3. Shown are the functional strand of the double stranded pre-miR-7 mimic (Genepharma, #M-01-D) and sequence of non specific double stranded control (Genepharma, #M-03-D).

Target	Sequence
Negative Control (NC5 - IDT premade nonspecific sequence)	CAUAUUGCGCGUAUGUCGCGUUAG CUAACGCGACUAUACGCGCAAUAUGGU
Cyclon (Oligo 1)	AGAGACGUCAUCCAAAUCUCUUCCC GGGAAGAGAUUUGGAUGACGUCUCUCG
Cyclon(Oligo 2)	CACCGUGUCUGAAACAGGAAGCAGG CCUGCUUCCUGUUUCAGACACGGUGAU
Cyclon (Oligo3)	CUGGACUCAGAGGUGGUACACGCTA UAGCGUGUACCACCUCUGAGUCCAGUU
Ezrin (Oligo 1)	CUUUUUGAUCAGGUAGUAAAGACTA UAGUCUUUACUACCUGAUCAAAAAGCU
Ezrin (Oligo 2)	CGCUAUGUUGGAAUACCUGAAGATT AAUCUUCAGGUAUUCCAACAUAGCGCU
Ezrin (Oligo 3)	GGACUUAAUAUUUAUGAGAAAGATG CAUCUUUCUCAUAAAUAUUAAGUCCAA
Moesin (Oligo 1)	GAAUGAGCGUGUGCAGAAGCAUCTT AAGAUGCUUCUGCACACGCUCAUUCUU
Moesin (Oligo 2)	GCAGAUUGAAGAGCAGACUAAGAAG CUUCUUAGUCUGCUCUUCAAUCUGCUU
Moesin (Oligo 3)	AGCGUCAAGAAGCUGAAGAAGCCAA UUGGCUUCUUCAGCUUCUUGACGCUCU
Lamin A (Oligo 1)	GACUUGGUGUGGAAGGCACAGAACA UGUUCUGUGCCUUCCACACCAAGUCAG
Lamin A (Oligo 2)	AGGCUAAGAAGCAACUUCAGGAUGA UCAUCCUGAAGUUGCUUCUUAGCCUCA
Lamin A (Oligo 3)	GAACUGGACUUCCAGAAGAACAUCT AGAUGUUCUUCUGGAAGUCCAGUUCCU

Table 2.3.2 Sequences of custom CHO double stranded siRNA. A pre made negative control was supplied (IDT, NC5). For each target Cyclon, Ezrin, Moesin and Lamin A there were 3 alternative oligomers designed based on the CHO sequence submitted.

2.3.2 RNA extraction

Tri Reagent® is a mixture of guanidine thiocyanate and phenol in a monophase solution that separates DNA, RNA and proteins into 3 phases: an aqueous phase (RNA), the interphase (DNA), and an organic phase (proteins). 1ml is capable of lysing 5-10 x 10^6 cells in accordance with manufacturer's instructions.

To achieve this 1 x 10^7 cells were counted (see **Section 2.2.1**), centrifuged and the supernatant removed. Lysis was preformed with the addition of 1ml of Tri Reagent®. Cells were incubated at room temperature for 5 min. Phase separation was achieved with the addition of 200 µl chloroform (Sigma-Aldrich, C2432). The sample was then vortexed for 15 sec and allowed to stand for 10 min. Subsequent centrifugation at 13,709 x g (12,000 rpm) for 15 min at 4 ^oC separated the solution into 3 visibly distinct phases - red organic layer containing proteins, intermediate phase containing DNA and the upper aqueous phase containing RNA. The aqueous phase is removed and transferred to a fresh tube. A 500 µl volume of isopropanol was added the aqueous phase. The mixture is allowed to stand for 10 min, centrifuged for 10 min at 13,709 x g at 4 ^oC and the supernatant was removed. Precipitated RNA may be observed at the side of the tube. RNA was then washed with 1 ml 75% ethanol. Sample was mixed thoroughly by vortex before being centrifuged at 5,355 x g (7,500 rpm) for 10 min at 4 ^oC. The ethanol was then poured off and the RNA pellet allowed to air dry for 10 min. Before becoming totally dry the RNA pellet was resuspended in nuclease-free water. The resulting isolated RNA was resuspended in appropriate volume (20 µl) in nuclease-free water.

2.3.3 RNA quantitation

A NanoDrop 2000 (Thermo scientific), a spectrophotometer for nucleic acid and protein quantitation, was used for DNA/RNA quantitation. The instrument arm was cleaned with a lint free wipe and UHP before sample analysis. An RNA/DNA sample solution of 1.5 µl was applied to the pedestal of the NanoDrop. The upper arm was lowered until a column of liquid was observed to from the DNA/RNA solution and the upper arm making contact. Light absorbance at 260 nm was measured and sample purity determined 260/280 nm with a reading of ~1.8-2.2 being indicative of pure RNA/DNA.

2.3.4 RNA to cDNA synthesis

Total RNA was extracted as in **Section 2.3.2** and isolation confirmed by RNA quantitation as in **Section 2.3.3.** Reverse transcription was performed to convert RNA into complementary DNA (cDNA). RNA samples were reverse transcribed into cDNA using Taqman microRNA Reverse Transcription kit (Applied Biosystems, 4366596). 10 µl of master mix, containing 10x reverse transcription buffer, 25x dNTP Mix (100 mM), 10x Random primers, 1 µl of Multiscribe® Reverse Transcriptase (50 U/µl) and 1 µl of RNase Inhibitor, was added to 2 µg of RNA in 5 µl volume into a 500 µl eppendorf tube.

Component	**Master Mix single reaction (µl)**
100mM dNTPs (with dTTP)	0.15
MultiScribe® Reverse Transcriptase, 50 U/µl	1.00
10x Reverse Transcriptase, 50 U/µl	1.5
5x Reverse Transcriptase Primer	3
RNase Inhibitor, 20 U/µl	0.19
Nuclease -free water	4.16
Total Volume	**10**

Table 2.3.3 Reverse transcription master mix volumes for a single reaction.

The final volume of 15 µl (10 µl Master Mix and 5 µl RNA) was then submitted to the thermocycler to undergo amplification. Step one was run at 25 °C for 10 min, step 2 at 37 °C for 120 min, step 3 at 85 °C for 5 min and temperature was held at 4 °C after the end of the reverse transcription cycle.

2.3.5 Real-time PCR

PCR reaction for miRNA was performed with TaqMan® Small RNA Assay 20x (Applied Biosystems, 4440418) and TaqMan® Universal PCR Master Mix 2x (Applied Biosystems, 4324018). A single reaction was prepared with the following volumes (**Table 2.3.4**).

Component	Single reaction (µl)
TaqMan® Small RNA assay (20x)	1.00
Reverse transcription cDNA product (1/15 dilution minimum)	1.33
TaqMan® Universal Master Mix 2x	10.00
Nuclease-free water	7.67
Total Volume	**20**

Table 2.3.4 Real-time PCR single reaction mix. Included is the reverse transcription product from **Section 2.3.4**

The reaction was mixed by pipetting and applied to a well of the reaction plate with hsa-miR-7 primer (Applied Biosystems, 4427975). The plates were run in the PCR thermal cycler (HT7600) at 95 °C for 10 min in the first step for enzyme activation then 95 °C for 15 sec in the second step and 40 cycles at 95 °C for 15 sec for the denaturation step and 60 °C for 60 sec for the annealing step. **Figure 2.3.1** shows how transcription occurs in the PCR reaction and **Figure 2.3.2** shows how amplification is reported.

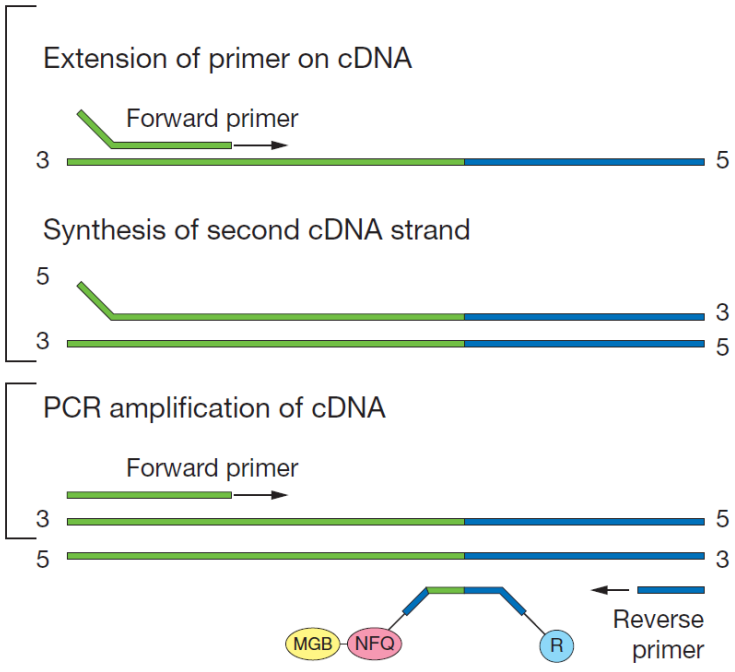


Figure 2.3.1 Real-time PCR using miRNA-specific primer. CT values are used to quantify miRNAs. MGB: Minor Groove Binding. NFQ: Non-Fluorescent Quencer. R: Reporter dye (FAM™).

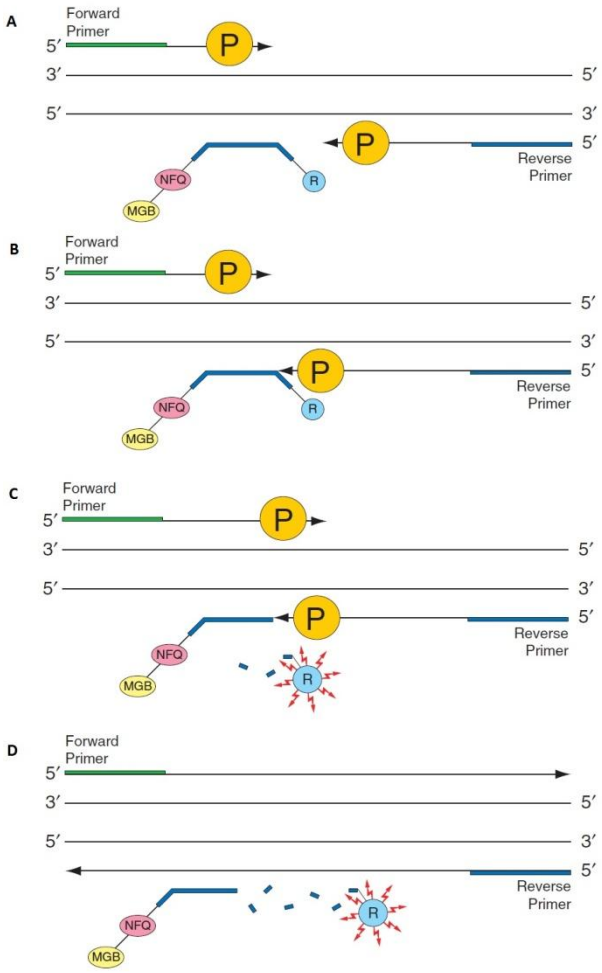


Figure 2.3.2 Diagram of 5' nuclease reporter process. Polymerisation (A), strand displacement (B), cleavage (C) and the final completion of polymerization (D) cleaves the reporter (R) from the quencher (NFQ).

2.4 Proteomic Techniques

2.4.1 Cell lysate preparation

Cells were pelleted by centrifugation at 170 x g for 5 min and the media was removed. Cells were washed with ice cold 50 mM HEPES. The cell suspension was then transferred to an eppendorf and centrifuged at 14000 x g for one minute in a 4 ^{o}C pre-cooled micro-centrifuge. HEPES was then decanted, the eppendorf containing the cell pellet was submersed in liquid nitrogen for 30 sec before being stored at -80 ^{o}C until required. All procedures from this point forward were performed on ice. Cells were lysed using an appropriate volume of lysis buffer (7 M Urea, 2 M Thiourea, 4% CHAPS, 30 mM Tris, pH 8.5) and incubated on ice for 20 min with occasional

vortexing. Halt protease inhibitor (PI-78415 Thermo Scientific), Halt phosphatase inhibitor (78428 Thermo Scientific) and Nuclease mix (80-6501-42 GE Healthcare) were also added to the lysis buffer to form a final concentration of 1x for each inhibitor. Cells were lysed by pipetting up and down and by vortexing. Sample lysates were centrifuged at 19,000 x g for 15 min at 4 ^{o}C. Supernatant containing extracted protein was transferred to a fresh chilled eppendorf tube and stored in aliquots at -80 ˚C if they were not going to be used immediately.

2.4.2 Nuclear/Cytoplasmic enrichment

The nuclear/cytoplasmic extraction kit works on the principle of swelling the cell membrane with a hypotonic buffer to increase its fragility. Addition of detergent causes cytoplasmic proteins to leak out into the supernatant which can then be collected and stored. The remaining nuclei are then lysed and solubilised in a detergent free lysis buffer.

Nuclear/cytoplasmic enrichment was carried out as per the manufacturer's instructions (Active Motif, 40410). The kit includes 10x PBS, Protease Inhibitor Cocktail, 10x Hypotonic Buffer, Phosphatase Inhibitor, 10 mM DTT, Lysis Buffer AM1 and Detergent which were used in the nuclear/cytoplasmic protein enrichment. **Table 2.4.1** details volumes required for solutions mentioned in the protocol. The following protocol was used to extract nuclear and cytoplasmic enriched fractions from 2 x 10^7 CHO-K1-SEAP suspension cells. Note all reagents are kept on ice.

Cells were counted using the trypan blue exclusion method (**Section 2.2.1**) and a volume of cells was taken representing 2 x 10^7 cells. Cells were centrifuged at 170 x g for 5 min. Meanwhile 1.6 ml of 10x PBS was diluted with 13.6 ml UHP and 0.8 ml of Phosphatase Inhibitor was added making a final volume of 16 ml 1x PBS/Phosphatase Inhibitor solution. The supernatant was removed from the centrifuged cells and the cell pellet washed in 8 ml of 1x PBS/Phosphatase Inhibitor. Cells were centrifuged once more as above and washed in the final 8 ml volume of 1x PBS/Phosphatase Inhibitor. The cell pellet was then resuspended in 1 ml of 1x Hypotonic buffer and incubated on ice for 15 min. After this 50 µl of Detergent was added and the sample was vortexed at high speed for 10 sec. Samples were then added to eppendorf tubes and centrifuged in a 4 ^{o}C pre-cooled microcentrifuge at 14,000 x g for 30 sec. The resulting supernatant was removed and stored. This is the cytoplasmic enriched fraction.

For the nuclear fraction the pellet was resuspended in 100 µl complete lysis buffer. Complete lysis buffer contains 10 µl of 10 mM DTT, 89 µl Lysis Buffer AM1 and 1 µl Protease Inhibitor Cocktail. The sample was incubated shaking and on ice for 30 min. After incubation the sample was vortexed at high speed for 30 sec and centrifuged in a 4 ^{o}C pre-cooled microcentrifuge at 14,000 x g for 10 min. The final supernatant was removed and stored as the nuclear enriched fraction.

Reagent	**Kit Component**	**Volume required for 2 x 10^7 cells**
1x PBS/Phosphatase Inhibitor	10x PBS	1.6 ml
	UHP	13.6 ml
	Phosphatase Inhibitors	0.8 ml
	Total Required	**16 ml**
1x Hypotonic Buffer	10x Hypotonic Buffer	0.1 ml
	Distilled Water	0.9 ml
	Total Required	1.0 ml
Complete Lysis Buffer	10 mM DTT	10.0 µl
	Lysis Buffer AM1	89.0 µl
	Protease Inhibitor Cocktail	1.0 µl
	Total Required	100.0 µl

Table 2.4.1 Volumes required for nuclear/cytoplasmic extraction of 2 x 10^7 cells.

2.4.3 Membrane enrichment

Membrane enrichment was carried out using CALBIOCHEM® ProteoExtract™ Native Membrane Protein Extraction Kit (Merck Millipore, 444-810KIT). The extraction buffers within the kit are proprietary. The mechanism however has been described as being gentle enough to retrieve membrane proteins in their native conformation. Membrane proteins are first pelleted and then subsequently solubilised resulting in a membrane protein rich supernatant (**Figure 2.4.1**). It has been shown that this kit when compared to 5 other membrane protein extraction kits produced the highest enrichment of membrane associated proteins and at the second highest concentration (Bunger, Roblick and Habermann 2009).

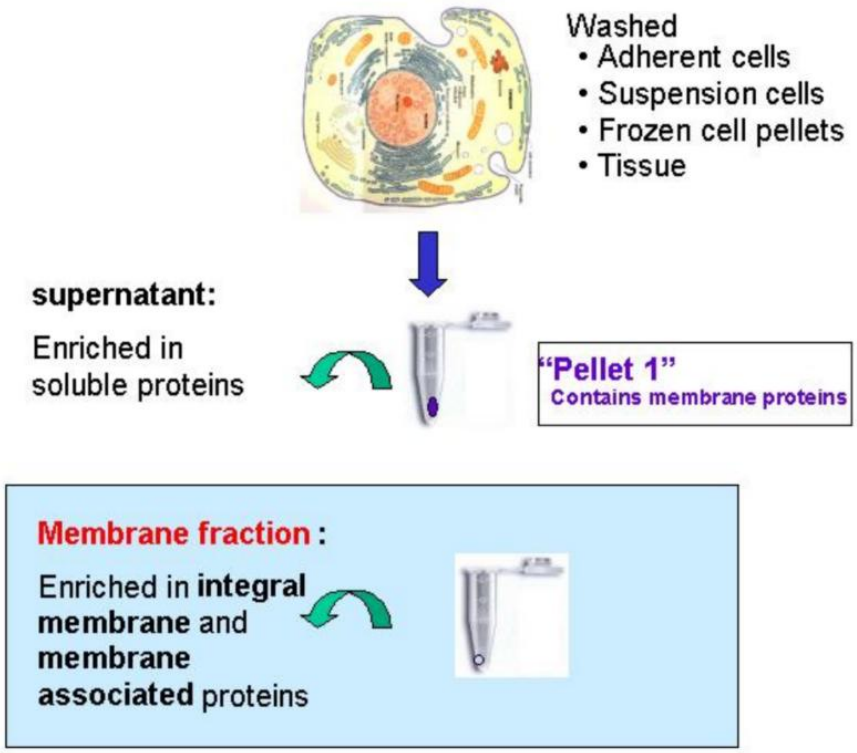


Figure 2.4.1 Schematic of the membrane enrichment methodology. Lysis and differential centrifugation steps carried out using CALBIOCHEM® ProteoExtract™ Native Membrane Protein Extraction Kit (Merk Millipore, 444-810KIT) result in a membrane protein enriched sample.

Membrane enrichment was carried out as per manufacturer's recommendations (Merck Millipore, 444-810KIT) for suspension cells and with some modifications for adherent cells. The components of the kit were as follows: Wash Buffer, Extraction Buffer I, Extraction Buffer II and Protease Inhibitor Cocktail. The following protocol outlines the steps taken to obtain a 500 µl membrane protein enriched fraction from ~90% confluent 75 cm^2 adherent culture flask or 8 x 10^6 suspension cells. All reagents were kept on ice except Protease Inhibitor Cocktail which was stored in DMSO and was allowed to come to room temperature to thaw.

For adherent cells conditioned media was removed and 5 ml of ice cold PBS was used to wash the cells. Cells were washed two more times in PBS. Extraction Buffer I was prepared by taking 2 ml and adding 10 µl of Protease Inhibitor Cocktail. 2 ml of Extraction Buffer I/Protease Inhibitor Cocktail was added to the washed cells in the flask. The cells were then scraped into the buffer using a sterile disposable cell scraper.

Suspension cells were counted using the trypan blue exclusion method (**Section 2.2.1**) and a volume of cells was used to obtain a cell pellet containing 6 x 10^8 cells. Cells were washed three times in ice cold PBS with centrifugation at 1000 x g for 5 min in-between each wash. The last PBS wash was decanted and cells are resuspended in 2 ml of Extraction Buffer I/Protease Inhibitor Cocktail (2 ml of Extraction Buffer I and 10 µl of Protease Inhibitor Cocktail).

From here both adherent and suspension cells are treated the same. 2 ml samples were divided into two 1 ml volumes in eppendorf tubes and incubated at 4 ^{o}C for 10 min. Samples were then centrifuged in a 4 ^{o}C pre-cooled microcentrifuge at 16,000 x g for 5 min. Following this the supernatant was discarded. Sample pellets were carefully resuspended in 500 µl Extraction Buffer II/Protease Inhibitor Cocktail (500 µl of Extraction Buffer II and 2.5 µl of Protease Cocktail Inhibitor). Samples were incubated and agitated at 4 ^{o}C for 30 min. To remove any remaining debris the samples were then centrifuged in a 4 ^{o}C pre-cooled microcentrifuge at 16,000 x g for 15 min. The supernatant containing solubilised membrane proteins was removed and stored.

2.4.4 Protein quantitation

Protein was quantified using the Bio-Rad Quick Start™ Bradford Dye Reagent (Bio-Rad, 500-0205) as follows. The required volume of Bio-Rad Quick Start™ Bradford Dye Reagent (250 µl/sample) was aliquoted into a sterilin tube and allowed to reach room temperature in the dark. A 2 mg/ml bovine serum albumin (BSA) stock solution (Sigma, A9543) was prepared in lysis buffer. A protein standard curve (0, 0.125, 0.25, 0.5, 0.75 and 1.0 mg/ml) was prepared from the BSA stock with dilutions made in lysis buffer. An aliquot of the protein samples were diluted as appropriate with lysis buffer to fit the linear range detection limits of the assay (125 - 1000 µg/µl). 5 µl of standards and samples were added to a minimum of three separate wells in a 96-well plate. 250 µl of the Bio-Rad dye solution was added to each well. After 5 min incubation, the absorbance was measured at 595 nm and the concentration of the protein samples was determined from the plot of the absorbance at 595 nm versus concentration of the protein standard.

2.4.5 Brilliant Blue G Colloidal Coomassie staining

After gel electrophoresis gels containing separated proteins were placed in a fixation solution which contained 7% glacial acetic acid in 40% (v/v) methanol. To prepare a 1X working solution of Brilliant Blue G colloidal coomassie (Sigma, B2025) 800 ml UHP was added to the stock bottle. Following the fixing step of 30 min, a dye solution containing 4 parts of 1x working colloidal coomassie solution and 1 part methanol was made, mixed by vortexing for 30 sec. Fixing solution was removed from the gels and then the dye solution was placed on top of the gels. The gels were left to stain for 2 hr. Following this the dye solution was poured off and the gels were destained to remove background dye. Destain solution containing 10% acetic acid in 25% methanol was poured over the gels on a shaker for 60 sec. The gels were then rinsed with 25% methanol for 30 sec and then destained with 25% methanol for 24 hr.

2.5 LC-MS/MS protein identification

2.5.1 Protein sample cleanup

Sample clean up was performed using ReadyPrep™ 2-D Cleanup Kit (Bio-Rad, 1632130). The end product was a washed and precipitated protein sample which was resuspended in a buffer compatible for mass spectrometry analysis. The clean up kit can process a maximum of 500 µg of protein in one sample preparation and a maximum sample volume of 100 µl as this method is performed in 1.5 ml eppendorf tubes. The following protocol details the steps taken to process whole lysate samples (**Section 2.4.1**) and the subcellular enriched samples (**Section 2.4.2** and **2.4.3**) . Note all steps are carried out on ice unless stated otherwise. Also Wash Reagent 2 must be pre chilled to -20 ^{o}C.

A 100 µl volume of protein sample was clean up which was previously quantified using the Bradford method (**Section 2.4.4**). For low volume samples (Nuclear enrichment fractions) volume was made up to 100 µl with mass spec grade water. To 100 µl of protein sample was added 300 µl of Precipitating Agent 1. Samples were mixed thoroughly for 1 min by vortexing. Samples were then incubated on ice for 15 min. Following this 300 µl of Precipitating Agent 2 was added to each sample and mixed for 1 min by vortexing. The samples were then centrifuged in a 4 ^{o}C micro-centrifuge at 12,000 x g for 5 min. The orientation of the tube in the centrifuge was noted as to repeat the same orientation for further spins and minimise sample loss. The supernatant was

carefully pipetted off and discarded. The tube was centrifuged for an additional 15 sec to remove any remaining precipitating agent which was then discarded.

A 40 µl aliquot of Wash Reagent 1 was applied over the pellet and the tubes were returned to the microcentrifuge in the same orientation and spun at 12,000 x g for 5 min. The wash reagent was then pipetted off and discarded. 25 µl of mass spec grade water was added to each sample and the samples were then mixed by vortexing. After this 5 µl of Wash 2 Additive and 1 ml of pre chilled -20 ^{o}C Wash Reagent 2 was added to each sample. They were then incubated overnight at -20 ^{o}C.

Following overnight incubation samples were centrifuged at 12,000 x g for 5 min in a 4 ^{o}C micro-centrifuge, the supernatant was removed and discarded with additional spin for 15 sec to remove and discard any remaining supernatant. The resulting pellet was air dried at room temperature for a maximum of 5 min until its edges appear translucent. Pellets were then resuspended in an appropriate volume of MS suitable solubilisation buffer (See **Section 2.5.2**) and quantified (**Section 2.4.1**) for further applications.

2.5.2 In solution digest for mass spectrometry

As outlined in **Section 2.5.1** a 100 µl of volume of whole cell lysate or cell enrichment fraction was cleaned to remove any agents that could interfere with mass spectrometry analysis. The resulting protein pellets were resuspended in label-free lysis buffer (6 M Urea, 2 M Thiourea, 10 mM Tris, pH 8), quantified using Bio-Rad Quick Start™ Bradford Dye Reagent (**Section 2.4.4**) and volume of cleaned up protein representing 5 µg of protein was taken for digestion. All volumes were brought up to 15 µl with label-free lysis buffer and 15 µl of 8 M urea was added to further solubilise samples.

A 20 µl volume of 0.2% ProteaseMax (Promega V2071) solution (1 part 1% ProteaseMax to 4 parts 50 mM ammonium bicarbonate) was added to each sample. ProteaseMax is a surfactant which increases protein solubilisation thereby increasing protein digestion. Samples were mixed by vortexing for 3 min. 58.5 µl of 50 mM ammonium bicarbonate was added to each sample.

Protein samples were denatured by the addition of 1 µl 0.5 M DTT (5mM DTT final concentration required). Samples were incubated for 20 min at 60 ^{o}C. After this 2.7 µl of 0.55 M iododacetamide was added to each sample (15 mM final concentration required) and incubated in the dark at room temperature for 15 min.

For the digestion step 1 µl of 1% ProteaseMax was added to each sample along with 2.5 µl of 0.1 µg/µl trypsin (Promega, V5111). A 1:20 ratio of Trypsin: Protein was required. Samples were then incubated for 6 hr at 37 ^{o}C. Samples were then quickly spun in a microcentrifuge for 10 sec to collect any condensation. Digestion reaction was then stopped with the addition of 5 µl trifluoroacetic acid (TFA) for a final required concentration of 0.5% TFA.

This method was used in **Chapter 4** for the preparation of whole cell lysates and subcellularly enriched protein fractions. An alteration to this method was used for **Chapter 3** and the analysis of the effects of miR-7 on the CHO cell proteome. Instead of a trypsin and ProteaseMax method a double digestion was used with trypsin and Lys-C. Briefly the alterations in this protocol were as follows.

10 µg of protein sample was resuspended in 50 µL of 50 mM ammonium bicarbonate. Reduction was performed by adding 1 µL of 100 mM DTT and incubated at 60 °C for 30 min. Samples were then alkylated by adding 5 µL of 0.3 mM Iodoacetamide and then incubated in the dark at room temperature for 30 min.

Digestion was carried out with sequence grade Lys-C (Promega, V1071) at a ratio of 1:20 Lys-C:Protein at 37 °C for 4 hr, followed by a second digestion with sequence grade Trypsin (Promega, V5111) at a ratio of 1:25 Trypsin: Protein. Incubation was carried out overnight at 37 °C. After 24 hr the digestion reaction was terminated by the addition of TFA to a final required concentration of 0.5% .

2.5.3 C18 peptide cleanup

Using C18 resin based centrifuge columns (Thermo, 89870) the peptides from enzymatic digestion (**Section 2.5.2**) were purified for MS analysis. A number of buffers required preparation involving methanol, TFA and acetonitrile (ACN). These were as follows: Activation Solution - 50% Methanol, Equilibrium Buffer - 0.5% TFA in 5% ACN, Sample Buffer - 2% TFA in 20% ACN, Wash Solution - 0.5% TFA in 5% ACN, Elution Buffer - 70% ACN.

Columns were prepared by tapping until the resin was collected at the bottom end and then placed into a collection tube. The top and bottom cap of the C18 column was removed and 200 µl of Activation Solution was added. Samples were then centrifuged at 1,500 x g for 1 min and flow through discarded. This step was repeated once more.

Following this 200 µl of Equilibration Buffer was added and the samples centrifuged at 1,500 x g for 1 min with flow through discarded. This step was repeated once more. Peptide sample was then added to the column. C18 columns were placed into fresh collection tubes and centrifuged at 1,500 x g for 1 min. The flow through was then added to the column again and centrifuged at 1,500 x g for 1 min. The flow through was stored in case of insufficient peptide binding and the collection tubes for the C18 columns replaced. 200 µl Wash Buffer was added to the column and tubes were again centrifuged at 1,500 x g for 1 min with the flow through discarded. The wash step was repeated once more. Finally the collection tube was replaced and 20 µl of Elution Buffer was added to the C18 column, centrifuged at 1,500 x g for 1 min and repeated once more resulting in 40 µl of purified peptides. Samples were then dried using a vacuum centrifuge.

2.5.4 LC/MS analysis

Samples were resolubilised in a volume of LC-MS grade water with 0.1% TFA and 2% ACN such that 5 µl contained 1 µg of peptides i.e. 50 µl suspension for 10 µg of peptides (digested protein). 12.5 µl was transferred to glass vials and the remaining reconstituted sample stored at -80 ^{o}C. Mass spectrometry was performed by Mr. Michael Henry. Nano LC–MS/MS analysis was carried out using an Ultimate 3000 nano LC system (Dionex) coupled to a hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific). Nano electrospray tips used were SilicaTip™ Standard Coating Tubing OD/ID 360/20 µm Tip ID 10 µm Length 5cm, (NewObjective, FS360-20-10-CE-5).

A 2.5 µl injection volume of digested sample were picked up using an Ultimate 3000 nano LC system (Dionex) auto sampler using direct injection pickup onto a 20 microlitre injection loop. The sample was loaded onto a C18 trap column (C18 PepMap, 300 µm ID × 5 mm, 5 µm particle size, 100 Å pore size; Dionex) and desalted for 10 min using a 25 µl/min flow rate in 0.1% TFA containing 2% acetonitrile. The trap column was switched online with the analytical column (PepMap C18, 75 µm ID × 250 mm, 3 µm particle and 100 Å pore size; (Dionex)) using a column oven at 35 °C and peptides were eluted with the following binary gradients.

Peptides were eluted in a binary gradient of Mobile Phase Buffer A and Mobile phase buffer B: 0–25% solvent B in 120 min and 25–50% solvent B in a further 60 min, where

solvent A consisted of 2% acetonitrile (ACN) and 0.1% formic acid in water and solvent B consisted of 80% ACN and 0.08% formic acid in water. Column flow rate was set to 350 nl/min. Data were acquired with Xcalibur software, version 2.0.7 (Thermo Fisher Scientific).

The Hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific) was run in data-dependent mode and externally calibrated. Survey MS scans were acquired in the Orbitrap in the 300–2000 m/z range with the resolution set to a value of 60,000 at m/z 400. Up to seven of the most intense ions (1+, 2+ and 3+) per scan were CID fragmented in the linear ion trap. All tandem mass spectra were collected using a normalised collision energy of 35%.

2.5.5 Mass spectrometry generated protein identifications

Mass spectrometry data is generated as .RAW files from LTQ XL instruments. Data was analysed using the search algorithms TurboSequest (Thermo Fisher Scientific) and MASCOT (v2.3.01, Matrix Science, London, UK) through Proteome Discover 1.4 (Thermo Fisher Scientific) against Bielefeld-BOKU-CHO database (BBCHO) (Meleady et al. 2012a) and CHO-K1 genomic data as published by Xu et al. (Xu et al. 2011) for all LC-MS analysis of samples originating from CHO cultures. For samples searched against other species, Swissprot-uniprot Human, Mouse or Rat databases (fasta file Jan 2014) were used. All data base searches were performed through Proteome Discoverer 1.4 software. This method was used to obtain qualitative identifications and identifications of proteins after quantitative differential analysis (**Section 2.5.6**).

The following search parameters were used: Allowed two missed cleavages, fixed modification of cysteine (carbamidomethyl-cysteine), variable modifications of methionine (oxidised), The Hybrid linear ion trap/Orbitrap mass spectrometer used a peptide tolerance of 20 ppm and the MS/MS tolerance set at 0.5 Da.

Statistical criteria used for peptide filtering were as follows: A MASCOT criteria of 95% confidence interval threshold ($p<0.05$), with a minimum MASCOT score of ≥ 30 while the following TurboSequest filters were applied: for charge state 1, XCorr > 1.9; for charge state 2, XCorr > 2.2; for charge state 3, XCorr > 3.75 and a delta CN of 0.1.

2.5.6 Label-free LC-MS data analysis

Quantitative analysis was performed using Progenesis QI for proteomics software form Non-Linear Dynamics by importing the Raw MS data file into the software package. The software graphically represents retention time and m/z ratio in spot map format. Automatic alignment was performed which aligns all the samples with the sample that contains most peptide features. This allows the corresponding peptide identifications to be carried across samples in the experiment regardless of slight differences in retention time. Manual alignment correction was performed to improve overall alignment which is reported as a percentage.

Upon satisfactory alignment of all sample runs (>80% for each sample) peptide features were filtered based on an Anova p-value of less than 0.05 between experimental groups. A principal component analysis was generated from this filter feature list to ensure separation between sample groups. The MS/MS data from the filtered peptide feature list was exported in and searched using MASCOT or SEQUEST through Proteome Discoverer 1.4 (**Section 2.5.5**). Once identifications were assigned to each peptide they were imported back into Progenesis. Identification conflicts were inspected and resolved manually by looking at the peptide sequence and strength of overall protein identification (higher ion scores were selected over lower ion scores). Conflicts occur when one peptide is incorrectly assigned to two or more protein identifications. A fold change cut off in differential expression of 1.2 fold between sample groups was used with an Anova p-value less than 0.05.

2.6 Western Blot analysis

2.6.1 Gel electrophoresis

Proteins for analysis by Western blotting were separated using SDS-polyacrylamide gel electrophoresis (SDS-PAGE) with precast 4-12% Bis-Tris gels (Life Technologies, NP0322BOX (12 well) or NP0321BOX (10 well)). Protein was quantified using the Bradford method (**Section 2.4.4)** and 5-20 µg of protein was diluted in 2x Laemmli loading buffer (Sigma, S3401-1VL). A 2 µl volume of Pageruler™ Plus prestained protein ladder (Pierce, 26619) was loaded alongside samples. The gels were run at

constant voltage (200 V) until the Bromophenol blue dye contained in the 2x laemmli loading buffer reached the end of the gel. This was typically achieved after 55 min at which time sufficient separation resolution of the molecular weight markers was achieved.

2.6.2 Western blot transfer

After electrophoresis had been completed, the gel was removed from the precast casing, rinsed in distilled water and was equilibrated in 10 ml of 1x tris-glycine transfer buffer (Bio-Rad, 161-0734) containing 20% methanol for 15 min. Approximately 50 ml of prepared transfer buffer was required for one Western Blot. Two sets of five sheets of Whatman 3 mm filter paper (Whatman, 1001824) were each soaked in 10 ml freshly prepared transfer buffer. PVDF membrane (GE Healthcare, 10600021) was cut to filter paper size, activated for 5 sec with 100% methanol, rinsed with distilled water and equilibrated in 10 ml of prepared transfer buffer for 10 min.

Following this the transfer unit was set up. One set of soaked filter papers were placed on the cathode plate of a semi-dry blotting apparatus (Bio-Rad), followed by the PVDF, then the gel and then the final set of soaked filter paper.. Air pockets were then removed from between the sandwiched layers with gentle rolling using a disposable 10 ml pipette taking care not to disturb the layers.. All proteins were transferred from the gel to the membrane at a current of 340 mA at 15 V for 23 min. Lower molecular weight proteins such as Histone H3 and Histone H4 (~15 kDa) were transferred for 18 min to avoid transfer off of the PVDF .

The membranes were then blocked for 1 hr using 5% blocking grade blocker (Bio-Rad 1706404) in pre-prepared TBS-T (1X tris-buffered saline (TBS) (Sigma, t5912) containing 0.05% Tween 20 (Sigma, P5927)). The membranes were washed with TBS-T prior to the addition of the primary antibody, prepared in 5% blocker in TBS-T at recommended dilutions. Membranes were then incubated overnight at 4 °C on a shaker ensuring membrane was adequately covered with primary antibody solution. After incubation the membranes were rinsed 3 times with TBS-T for a total of 30 minutes. Relevant secondary antibody (1/2000 dilution of anti-mouse (Dako, P0260) or anti-rabbit (Dako, P0448) or anti-rat (Dako, P0450) or anti-goat (Santa Cruz Biotechnology, Sc2098) IgG peroxidase conjugate in 5% blocking grade blocker/TBS-T) was added for 1 hr at room temperature. Membranes were wash again in TBS-T three times to ensure unbound secondary was washed off. The following primary antibodies were used.

Protein target CHO	Symbol	Supplier	Product #	Host	Dilution
Histone H3	HISTH3	Cell Signalling	4499	Rabbit	1/2000
Acetyl Histone H3	HISTH3	Millipore	17-245	Rabbit	1/2000
Histone H4	HISTH4	Cell Signalling	13919	Rabbit	1/2000
Acetyl Histone H4	HISTH4	Cell Signalling	2594	Rabbit	1/2000
Protein disulfide isomerase A6	PDIA6	Abcam	ab11432	Rabbit	1/1000
78 kDa glucose-regulated protein	GRP78	Sigma	4501452	Rabbit	1/750
Heat shock protein family (Hsp70) member 8	HSPA8	Abcam	ab19136	Rat	1/1000
14-3-3 protein epsilon	YWHAE	Abcam	ab40117	Rabbit	1/2000
Glyceraldehyde 3-phosphate dehydrogenase	GAPDH	Abcam	ab8245	Mouse	1/10000
Catalase	CAT	Abcam	ab1877	Rabbit	1/2000
Stathmin	STMN1	Abcam	ab52630	Rabbit	1/50000
Ezrin	EZR	Invitrogen	357300	Mouse	1/2000
Moesin	MSN	Thermo	38/87	Mouse	1/200
Cyclon	CYC	SantaCruz	sc-82568	Goat	1/1000
Lamin A/C	LMNA	Abcam	ab26300	Rabbit	1/1000
Protein target breast cancer					
Immunoglobulin Superfamily Member 9	IGSF9	Atlas	HPA037753	Rabbit	1/50
Killer cell lectin like receptor 2	KLRG2	Abcam	ab121563	Rabbit	1/200

Table 2.6.1 List of primary antibodies used for Western blot analysis. Antibodies in the "Protein targets used for CHO" list were used in **Chapter 3** and **4** while IGSF9 and KLRG2 antibodies were used on breast cancer samples in **Chapter 5.**

2.6.3 Enhanced chemiluminescent detection

To develop the immunoblots an enhanced chemiluminescence (ECL) kit (GE Life Sciences, RPN2106) was used, which allowed for the detection of bound peroxidase-conjugated secondary antibody. After the final wash in TBS-T the blots were placed on a transparent plastic sheet and 3 ml of a freshly prepared 1:1 (v/v) mixture of ECL reagent A and B was applied ensuring full coverage the membrane. Chemiluminescence reaction was allowed to develop for 5 min in the dark. Following this Another clear plastic sheet was applied over the blot and excess ECL reagent mixture was completely

removed as well as any bubbles that formed. The membrane was then exposed to autoradiographic film (GE Life Sciences, 95017-681) for various times (from 10 sec to 30 min depending on the intensity of the signal). The exposed autoradiographic film was developed for 3 min in developer solution (Kodak, LX24, diluted 1:5 in water) or until clear band development in the predicted area could be seen. The film was then washed in water for 15 sec and transferred to a fixative solution (Kodak, FX-40, diluted 1:5 in water) for 5-10 min. The film was washed with water for 5-10 min and allowed to dry at room temperature.

2.7 Bioinformatics and statistical analysis

2.7.1 Pathway analysis

Three separate pathway analysis tools were used in the analysis of qualitative lists and quantitative differential lists in **Chapter 3** and **Chapter 4**. The following will explain how this was performed for each tool.

In preparation for pathway analysis the protein identifications obtained from the 2 CHO specific databases (BBCHO and NCBI see **Section 2.5.5**) was changed to corresponding human gene symbols. Protein identifications were in a format unsuitable for pathway analysis tools to recognise BBCHO identifiers were in a transcript isotig labelled format while the NCBI database produced ChiTaRS (Chimeric Transcripts and RNA-seq database) identifiers. Labels within each database were not consistent which severely limited the amount of automated parsing of labels. Each protein identified had its peptide sequences BLAST searched through UniProt (http://www.uniprot.org/blast/) to confirm its correct protein name and associated mouse gene name. Annotated proteins were added to a reference list for increasingly automated searching by matching the unique CHOisotig number of each identification in the BBCHO database or the NCBI accession number of the NCBI database with that of already annotated proteins.

When gene names were assigned to the relevant list the pathway analysis tool DAVID (https://david.ncifcrf.gov/home.jsp) was used. Gene names were pasted into the submission window, identifier selection was "OFFICIAL_GENE_SYMBOL", list type selection was "Gene List" and then the list was submitted. Next the "Mus musculus" species background was selected. The drop down menu for "Gene_Ontology" was selected. From here three categories of enrichment groupings were obtained in

molecular function (GOTERM_MF_FAT), biological process (GOTERM_BP_FAT) and cellular component (GOTERM_CC_FAT). It the reported chart view for the data a dropdown menu for options was selected and "Bonferroni" was selected followed by "Rerun using options". The resulting data was copied into Excel and sorted according to the Bonferroni significance value $p<0.05$. DAVID also allowed KEGG pathway analysis by selecting "Pathways" and selecting "KEGG_PATHWAY". Again Bonferroni significance values were used. By selecting the "Term" in the data table (e.g. Glutathione metabolism, Ribosome, etc.) a graphical representation of the protein pathway was obtained.

For PANTHER analysis (pantherdb.org) the gene list was pasted into the submission box, "Mus musculus" organism was selected and "Statistical overrepresentation test" was chosen with default settings. Again the three categories of annotation were chosen molecular function (PANTHER GO-Slim Molecular Function), biological process (PANTHER GO-Slim Biological Process) and cellular component (PANTHER GO-Slim Cellular Component), Bonferroni significance of $p<0.05$ was used to sort enriched terms in Excel.

2.7.2 Statistical significance

Statistical significance was calculated where indicated throughout the results section using an unpaired t-test. This assesses the difference between population means (type 2 t-test in Excel).

High variance between runs in the miR-7 PCR experiment meant that there was not enough data to perform such a stringent test on this data.

2.7.3 Bioinformatic dataset analysis

Gene lists of potentially novel membrane expressed proteins associated with breast cancer were generated by Dr. Stephen Madden and Dr. Padraig Doolan, NICB. The following procedure was applied to produce this data.

Publicly available datasets containing both normal tissue and breast cancer tissue together with gene profiling were obtained from Sircoulomb et al. (Sircoulomb et al.

2010), Pau et al. (Pau Ni et al. 2010) and Tollet-Egnell et al. (Tollet-Egnell et al. 2001) as well as this an in house data set of breast cancer vs normal tissue was also used resulting in the comparisons and differential gene expression numbers in **Table 2.7.1.**

	In-house breast cancer dataset		Publically available breast cancer datasets	
	Genes 1.3 fold UP	Genes 1.3 fold DOWN	Genes 1.3 fold UP	Genes 1.3 fold DOWN
Norm vs ER+	8103	8313	122	144
Norm vs HER2+	6870	7333	1718	3060
Norm vs LN+	8364	8127	314	169
Norm vs. Triple Neg.	4236	3368	2248	4918

Table 2.7.1 Number of genes ≥1.3 fold increased in ER+, HER2+, LN+ or triple negative breast cancer compared to normal breast tissue. These were generated using an in house transcript profiling dataset and 3 publicly available transcript profiling datasets.

Predicted protein localisation was then assessed using a uniprot subcellular location batch search. As an over-expressed target was required only the UP gene lists were assessed further for commercial antibody availability, high fold change, no literature association with breast cancer functional studies and strong association with the cell membrane.

2.8 Immunohistochemistry

All immunohistochemical (IHC) staining was performed using the DAKO Autostainer (DAKO, S3800) (table 2-4). Deparaffinisation and antigen retrieval was performed using Epitope Retrieval 3-in-1 Solution (pH 6) (DAKO, S1699) or the Epitope Retrieval 3-in-1 Solution (pH 9) (DAKO, S2375) and the PT Link system (DAKO, PT101).

For epitope retrieval, slides were heated to 97 °C for 20 min and then cooled to 65 °C. The slides were then immersed in wash buffer (DAKO, S3006). On the Autostainer slides were blocked for 10 min with 200 µL HRP Block (DAKO, S2023). Cells were washed with 1x wash buffer and 200 µL of antibody (KLR2 or IGSF9 antibody) added to the slides for 27 min. Slides were washed with 1x wash buffer and then incubated with 200 µL Real EndVision (DAKO, K4065) for 30 min. All slides were

counterstained with haematoxylin (DAKO) for 5 min, and rinsed with deionised water, followed by wash buffer.

Slides were then dehydrated in graded alcohols (2 x 3 min each in 70% IMS, 90% IMS and 100% IMS), and cleared in xylene (2 x 5 min), and finally mounted with coverslips using DPX mountant (Sigma, 44581).

2.8.1 Scoring guidelines

Immunoreactivity was measured based on the staining intensity as outlined by DAKO "Guidlines for Scoring HercepTest^{TM} - Breast" which used criteria outlined by Wolff et al (Wolff et al. 2007). These guidelines were established using HER2 immunoreactivity as a scoring model in patients with invasive breast cancer. Immunoreactivity was designated as absent (0), weak (1), moderate (2) or strong (3) (**Figure 2.8.1**). The guideline IHC images also show clear examples of membrane staining as HER2 is expressed in the cell membrane.

Staining intensity was scored for individual tissue sections and tissue matrix array (TMA) cores. The TMA cores were also scored for percentage coverage of staining and the data was graphed in an x, y, z scatter plot.

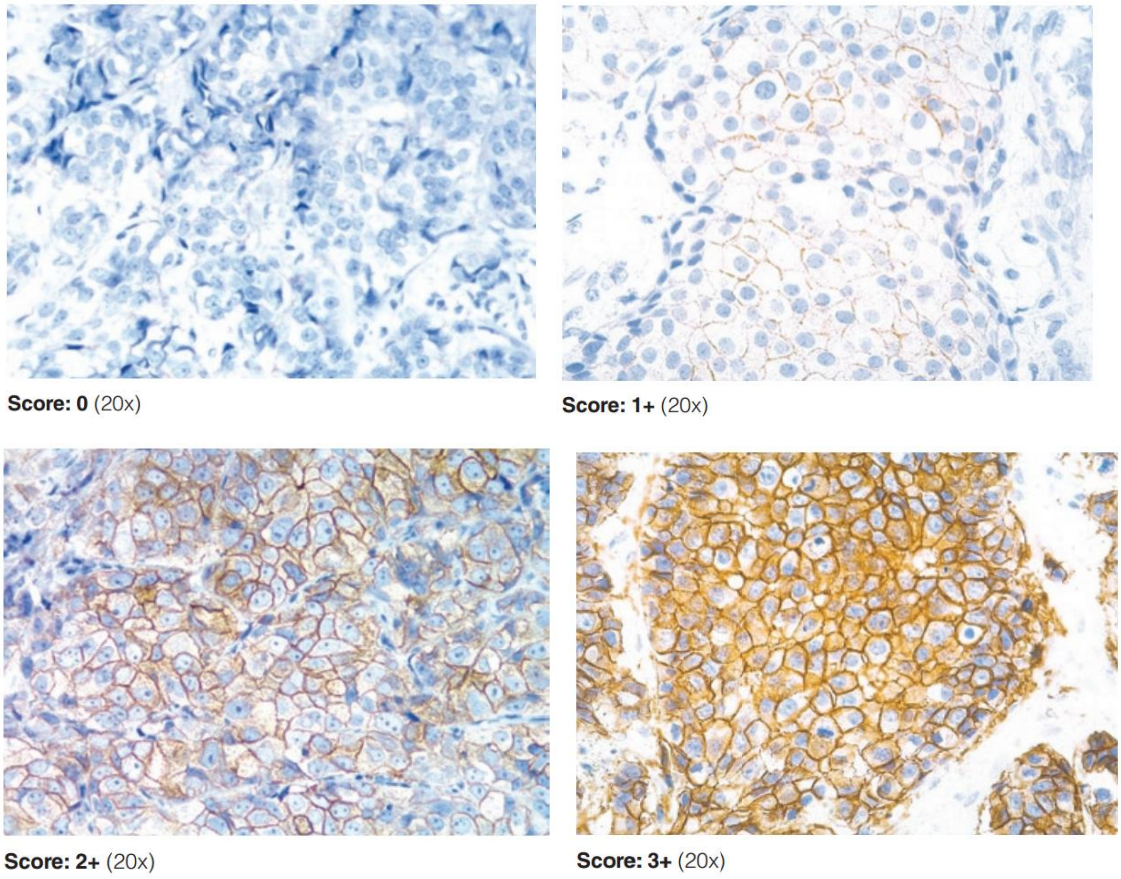


Figure 2.8.1 Immunoreactivity guidelines used for staining intensity. Strength of staining was marked based on the examples above as provided by DAKO. Absent, weak, moderate and strong immunoreactivity was denoted by 0, 1, 2 and 3 respectively.

CHAPTER 3

Effect of miR-7 on CHO cells

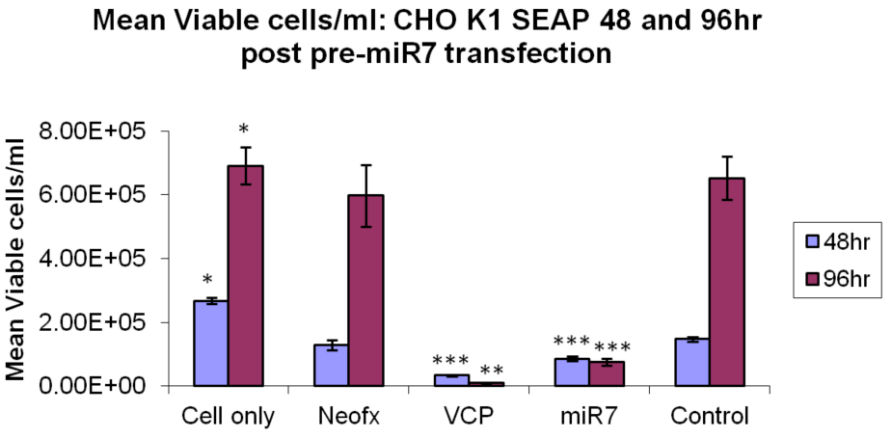
3.1 Impact of miR-7 on CHO cells

Micro RNA (miR) are a potential tool for the manipulation of Chinese hamster ovary (CHO) cells in industry. They do not impose any translational burden on cellular processes and can be used to manipulate a wide variety of phenotypes. In previous work conducted in our lab miR-7 was found to dramatically reduce cell proliferation without affecting viability in CHO-K1 SEAP cells ultimately leading to increased productivity over time (Barron et al. 2011a). This is a phenotype that holds industrial relevance, therefore miR-7 could be used as a potential target for cell line engineering or by targeting the proteins that it affects. The objective of this study is to identify these potential targets of miR-7 and what molecular mechanisms miR-7 affects.

A model producer CHO cell line which produces secreted alkaline phosphatase (CHO-K1 SEAP) were previously adapted to suspension culture in our lab and optimised for transfection with miR-7 ((Barron et al. 2011a)). The resulting growth conditions were similar to those used in industry with cells being grown in serum free CHO-S-SFM II (Gibco) media, 37 ^{o}C and 5% CO_2. Suspensions were maintained at 170 rpm for cultures ranging from 2 - 50 ml.

To investigate the effect of miR-7 on the phenotype of CHO-K1 SEAP cells we transfected $2x10^5$ cells in a 2ml volume (1 x 10^5 cells/ml) with a double stranded pre-miR-7 (Genepharma). The effect on cell growth (**Figure 3.1.1**) was confirmed using an automated cell counter (Guava Technologies). As cell growth is reduced it led to a number of challenges including the accurate counting of cells and low protein yields in the pre-miR-7 transfected samples. For sufficient protein yield for both nano LC-MS/MS and Western blot analysis the transfections were performed in technical duplicate and then pooled for lysates. To further validate the transfection we used RTPCR to confirm miR-7 up regulation following transfection with pre-miR-7 (**Figure 3.1.2**). These were the conditions that were used going forward for further mass spectrometry and Western blot protein profiling.

A



B

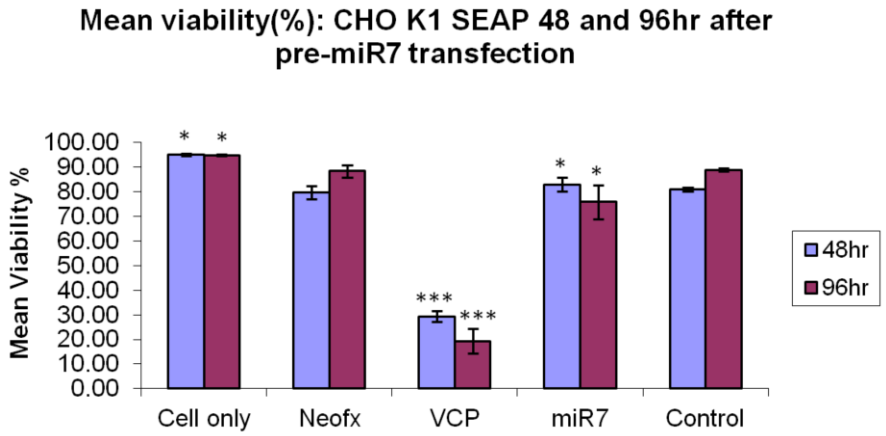


Figure 3.1.1 Impact on growth and viability in CHO K1 SEAP cells at 48 and 96 hr after transfection with pre-miR-7. Exogenous pre-miR-7 is incorporated into the RISC complex and processed into mature miR-7. The resulting phenotype in CHO-K1 SEAP cells is that of reduced cell proliferation (A) with no effect on viability (B). Error bars represent the standard deviation from three biological replicates. Neofx was the transfection reagent used and VCP (valosin containing protein) knockdown was used as a positive control as it was shown in our lab previously to reduce cell viability (Doolan et al. 2010). Scramble miR control is represented on the far right (Control). * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$ when compared to Control. All samples groups n=3.

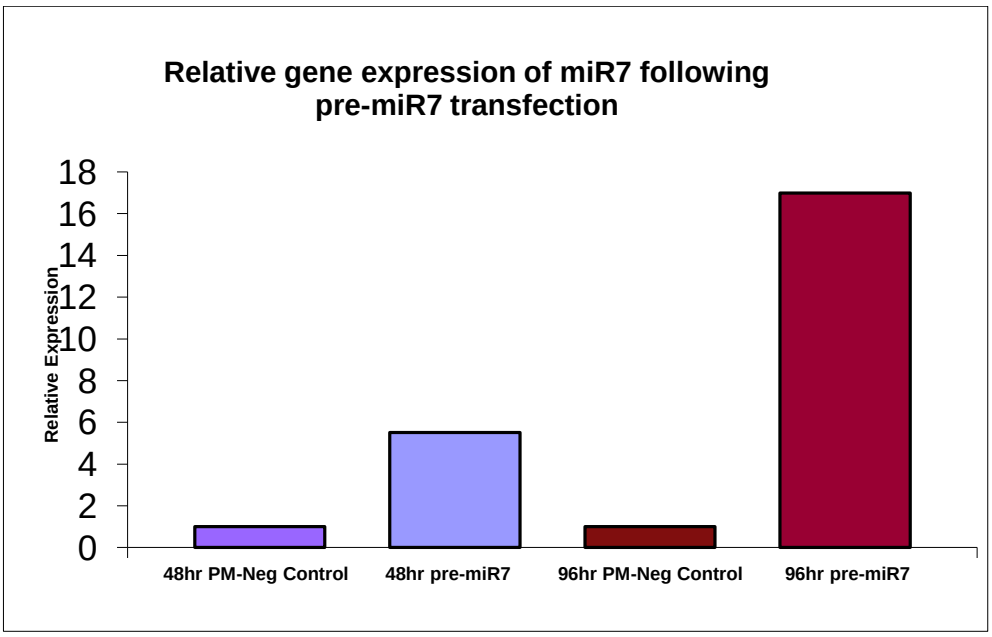


Figure 3.1.2 Relative expression of miR-7 at 48 and 96 hr after transfection with pre-miR-7 compared to endogenous miR-7 in pre-miR negative control (PM-Neg). It was confirmed with RT-PCR that exogenous pre-miR-7 increased miR-7 in CHO cells. As the transfection is transient it was particularly important to confirm if miR-7 was increased at the later time point of 96 hr. One replicate is shown (n=1).

3.2 Quantitative proteomic profiling of pre-miR-7 transfected CHO-K1 SEAP

To identify what proteins were involved in the phenotype observed from up-regulation of miR-7 we used a quantitative label-free LC-MS/MS proteomics approach. This method is capable of simultaneously identifying and quantifying hundreds or thousands of proteins as opposed to older methods such as labelled methods such 2D DIGE which are not as high throughput and more labour intensive. As unlabelled peptide fragments are submitted directly to the LC-MS/MS the resulting MS/MS data is mapped to protein databases allowing for protein identification and quantification. This proved particularly challenging with CHO as database annotations required further processing (See **Section 2.5.6** for further details)

The resulting IDs were used for further analysis using with miR-7 target prediction software and pathway tools such as DAVID to determine the effect of miR-7 on the cellular processes both of which required official UniProt accession numbers or gene names. Many of these identifications therefore required manual annotation. This in itself required a significant amount of time in particular for the CHO database. The results from the miR-7 target prediction, manual protein ID annotation and pathway analysis can be seen in detail below in **Section 3.4**.

3.2.1 Label-free analysis

Cells were grown in biological triplicate and lysed 48 and 96 hr after transfection with pre-miR-7 or pre-miR-7 scramble negative control (PMNeg). Cell lysates were taken and processed for MS analysis by using a Ready Prep 2-D clean up kit (Biorad) to remove incompatible solvents for enzymatic digestion and a double digestion strategy using Trypsin (Promega) and Lys-C (Promega). This respectively extracts the proteins from the cells, removes incompatible buffers from the lysate, cleaves the proteins at specific amino acids into peptides ready to be extracted and concentrated. Cleaning and concentration of peptides was performed with C-18 resin based micro centrifuge columns, lyophilised and rehydrated for submission to the LC-MS for analysis.

Peptides were separated over a 3 hr reverse phase gradient. Peptide features with +1, +2 and +3 charge state and ANOVA <0.05 were then searched using Mascot with UniProt-Swissprot human, mouse and rat databases and CHO specific database searched with Mascot and Sequest. Principal component analysis (PCA) was used to assess the

separation between protein IDs selected with these cut offs (**Figure 3.2.1**).Distinct separation between groups showed that differentially expressed IDs were selected.

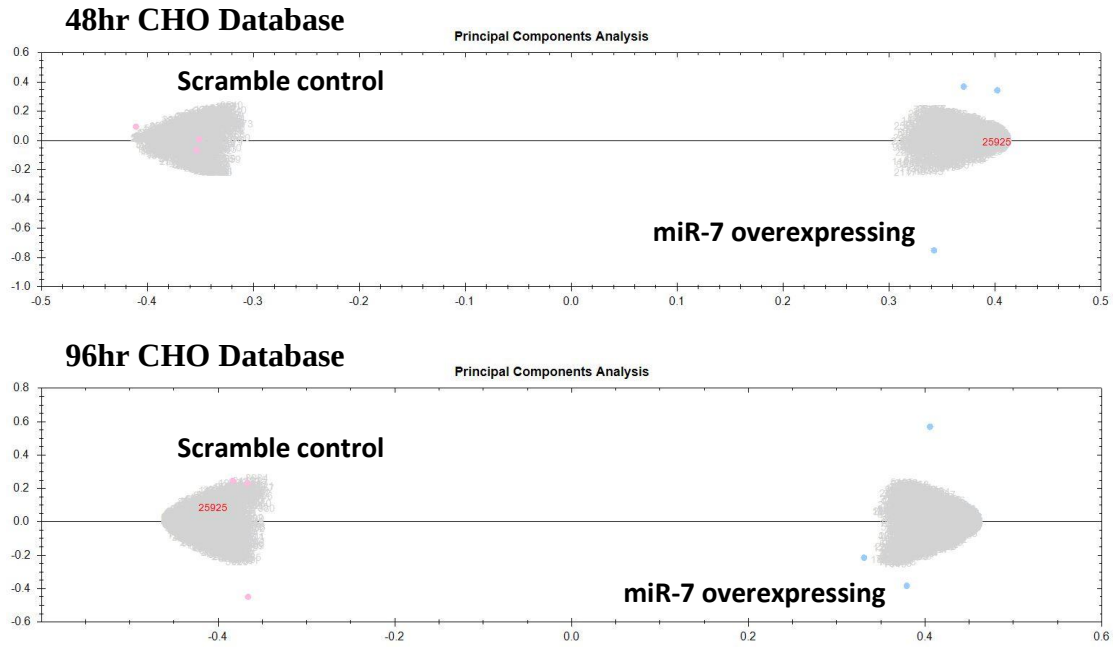


Figure 3.2.1 Principal component analysis (PCA) of differentially regulated proteins with statistical cut off of ANOVA <0.05 and features with +1, +2 and +3 charge state using the CHO database. Large separation shows the criteria used has resulted in differentially expressed proteins identified between the scramble control group on the left and the pre-miR-7 up-regulated group on the right (n=3).

During the time of the initial published study there was no CHO specific protein database available so a multi-species approach was taken() (**Table 3.2.2 and 3.2.3**). More recently the data was reanalysed with Mascot and Sequest search algorithms with a CHO specific database (Meleady et al. 2012a) developed in partnership with Professor Nicole Borth at BOKU in Vienna and using an NCBI non redundant CHO sequence database (**Table 3.2.4 and 3.2.5**). Statistically significant differentially expressed protein identifications were considered as having an ANOVA score ≤0.05 and a fold change >1.2 between experimental groups. A relatively low fold change cut off was chosen based on the previous work in our lab showing the effect of miR on protein dysregulation (Muniyappa et al. 2009). The only criteria that was used differently in both was the minimum peptide number. For the mammalian IDs one peptide hits were excluded but for the CHO database 1 hit peptide IDs were included resulting in 217 and

265 differentially regulated proteins respectively (**Table 3.2.1**). A comparison between the resulting identifications can be found in **Section 3.2.2**

		↓ regulated		↑ regulated		
Database	**Includes 1 peptide IDs**	**48 hr**	**96 hr**	**48 hr**	**96 hr**	**TOTAL**
Mammalian	**No**	48	73	29	67	217
Chinese hamster ovary	**Yes**	63	98	38	66	265

Table 3.2.1 Overview of the number of down-regulated (↓) and up-regulated (↑) proteins 48 and 96 hr after miR-7 over-expression compared to scramble negative control in CHO-K1-SEAP cells using a mammalian database search and a custom Chinese hamster ovary database search. Relying on sequence homology between species the 1 peptide IDs are often ignored however with reduced sequence coverage in the CHO database and the presence of a high number of CHO 1 peptide identifications being present with greater than one peptide in the mammalian list (see **Section 3.2.3** for comparison) these CHO 1 peptide IDs were included.

			PM-7 v PM-Neg 48 h			PM-7 v PM-Neg 96 h		
Accession [a]	Gene name [b]	Protein description	Peptide	Fold change [c]	ANOVA	Peptide	Fold change [c]	ANOVA
Q3T0F4	RPS10	40S ribosomal protein S10				3	1.25	0.05
Q3T0V4	RPS11	40S ribosomal protein S11				2	3.8	1.94×10^{-3}
P25398	RPS12	40S ribosomal protein S12	2	1.32	0.03			
Q9WVH0	RPS13	40S ribosomal protein S13				3	17.21	3.16×10^{-3}
P62265	RPS14	40S ribosomal protein S14	2	1.68	0.04	2	1.71	0.01
Q5R938	RPS15A	40S ribosomal protein S15a				3	13.21	0.01
Q3T0X6	RPS16	40S ribosomal protein S16	2	5.09	7.10×10^{-3}	2	5.43	1.04×10^{-3}
P63274	RPS17	40S ribosomal protein S17				2	2.67	0.05
Q3T0R1	RPS18	40S ribosomal protein S18	3	2.97	5.12×10^{-3}	5	5.69	2.03×10^{-3}
Q5R8M9	RPS19	40S ribosomal protein S19	3	1.54	0.02	6	1.27	4.13×10^{-4}
P25444	RPS2	40S ribosomal protein S2				3	5.81	0.03
Q3ZBH8	RPS20	40S ribosomal protein S20				3	1.83	6.09×10^{-4}
Q3T199	RPS23	40S ribosomal protein S23				3	3.21	6.50×10^{-3}
Q6Q311	RPS25	40S ribosomal protein S25				2	4.22	0.04
Q0Z8U2	RPS3	40S ribosomal protein S3				4	1.49	0.03
P49242	RPS3A	40S ribosomal protein S3a				5	2.25	0.04
P47961	RPS4	40S ribosomal protein S4				2	4.12	5.74×10^{-3}
P38982	RPSA	40S ribosomal protein SA	7	1.7	0.02	7	1.4	0.05
P05388	RPLP0	60S acidic ribosomal protein P0				5	1.46	0.05
Q5R931	RPL10	60S ribosomal protein L10				2	2.93	0.05
Q3T087	RPL11	60S ribosomal protein L11				4	1.39	0.02
Q6QMZ7	RPL12	60S ribosomal protein L12				2	1.69	0.02
Q5EAD6	RPL15	60S ribosomal protein L15				2	2.55	0.04
P35980	RPL18	60S ribosomal protein L18				3	5.25	0.02
Q3T0W9	RPL19	60S ribosomal protein L19				2	6.36	0.02
Q4R5I3	RPL22	60S ribosomal protein L22	3	1.74	0.01	3	1.73	2.74×10^{-3}
P21531	RPL3	60S ribosomal protein L3				3	1.79	0.04
Q6QMZ4	RPL6	60S ribosomal protein L6				2	11.97	0.01
Q2TBQ5	RPL7A	60S ribosomal protein L7a				3	3.47	6.46×10^{-3}
Q4R596	AHCY	Adenosylhomocysteinase	3	2.89	3.53×10^{-3}			
Q8SQH5	SLC25A5	ADP/ATP translocase 2	4	2	0.04	4	4.83	1.92×10^{-3}
Q5R874	DHX9	ATP-dependent RNA helicase A				2	2.32	0.01
Q9CWJ9	ATIC	Bifunctional purine biosynthesis protein PURH	2	2.55	4.52×10^{-3}	2	1.84	1.55×10^{-3}
P24270	CAT	Catalase	4	4.07	4.34×10^{-3}	4	3.14	8.43×10^{-3}
Q5R6X7	CBX3	Chromobox protein homolog 3	3	2.05	0.04			
P62629	EEF1A1	Elongation factor 1-alpha 1	11	1.43	0.04	13	1.26	3.81×10^{-3}
Q9D8N0	EEF1G	Elongation factor 1-gamma	4	1.37	0.03	6	1.26	0.01
Q3SZC0	ERH	Enhancer of rudimentary homolog				2	3.5	0.02
Q3SZ54	EIF4A1	Eukaryotic initiation factor 4A-I				13	1.37	8.01×10^{-3}
Q99PF5	KHSRP	Far upstream element-binding protein 2	2	4.87	5.07×10^{-3}			
Q14444	CAPRIN1	Caprin-1	3	2.41	4.45×10^{-3}			
Q3T054	RAN	GTP-binding nuclear protein Ran				6	1.3	0.01
Q4R7Y4	GNB2L1	Guanine nucleotide-binding protein subunit beta 2-like 1	4	1.57	0.03	4	1.48	0.01
Q9N1U2	HSPA6	Heat shock 70 kDa protein 6				3	1.2	0.01
P19378	HSPA8	Heat shock cognate 71 kDa protein	18	1.47	4.32×10^{-3}			
P46633	HSP90AA1	Heat shock protein HSP 90-alpha	8	1.32	0.024			
Q4R4T5	HSP90AB1	Heat shock protein HSP 90-beta	8	1.64	7.11×10^{-3}			
Q9Z2X1	HNRNPF	Heterogeneous nuclear ribonucleoprotein F	5	1.48	6.21×10^{-3}			
Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U	4	1.73	0.01	9	1.66	1.94×10^{-3}
P02253		Histone H1.1				6	4.48	3.90×10^{-3}
P0C0S8		Histone H2A type 1				3	2.78	1.27×10^{-3}
Q96QV6	HIST1H2AA	Histone H2A type 1-A				3	2.85	1.53×10^{-3}
P0C169		Histone H2A type 1-C				3	2.78	1.48×10^{-3}
P0C0S4	H2AFZ	Histone H2A.Z				4	3.06	6.79×10^{-4}
Q2M2T1	HIST1H2BK	Histone H2B type 1-K				4	2.76	4.34×10^{-4}
Q32L48	HIST1H2BN	Histone H2B type 1-N	3	2.09	0.04	3	2.81	4.59×10^{-4}

P00494	HPRT1	Hypoxanthine-guanine phosphoribosyltransferase				8	1.6	0.04
Q14974	KPNB1	Importin beta-1 subunit				8	1.44	0.04
P12269	IMPDH2	Inosine-5?-monophosphate dehydrogenase 2				8	5.2	4.76×10^{-5}
Q5R1W9	LDHA	L-lactate dehydrogenase A chain	5	2.01	0.05	8	2.06	7.69×10^{-3}
Q09666	AHNAK	Neuroblast differentiation-associated protein AHNAK	2	2.96	8.07×10^{-3}			
Q63525	NUDC	Nuclear migration protein nudC	3	1.59	4.33×10^{-3}	3	1.43	0.01
Q28618	YBX1	Nuclease sensitive element-binding protein 1	3	1.36	7.84×10^{-3}	3	1.25	0.04
P08199	NUCL	Nucleolin	12	1.76	6.51×10^{-3}	16	1.3	3.74×10^{-3}
P06748	NPM1	Nucleophosmin	6	1.82	5.65×10^{-3}			
P28656	NAP1L1	Nucleosome assembly protein 1-like 1	2	1.92	0.05			
Q8NC51	SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein				7	1.44	4.36×10^{-3}
Q5E9A3	PCBP1	Poly(rC)-binding protein 1				2	2.31	4.92×10^{-3}
Q9EPH8	PABPC1	Polyadenylate-binding protein 1	7	1.45	0.03	9	1.34	0.01
Q2NL22	EIF4A3	Eukaryotic initiation factor 4A-III				2	1.36	5.85×10^{-4}
P57761	PCNA	Proliferating cell nuclear antigen	2	1.88	3.04×10^{-3}	6	1.41	1.40×10^{-3}
Q9UQ80	PA2G4	Proliferation-associated protein 2G4	4	1.8	0.02	7	1.5	3.80×10^{-3}
Q9JIF0	PRMT1	Protein arginine N-methyltransferase 1	6	1.39	0.02			
Q9EQU5	SET	Protein SET (Phosphatase 2A inhibitor I2PP2A)	5	1.3	7.21×10^{-3}			
P26350	PTMA	Prothymosin alpha				3	36.14	6.96×10^{-3}
P14618	PKM2	Pyruvate kinase isozymes M1/M2	3	1.4	0.01			
P43487	RANBP1	Ran-specific GTPase-activating protein	2	1.68	1.58×10^{-3}			
Q2HJ58	PRPS1	Ribose-phosphate pyrophosphokinase I				2	2.57	0.03
Q5RE47	BAT1	Spliceosome RNA helicase BAT1				2	1.37	0.05
Q3YLA6	SRSF1	Splicing factor, arginine/serine-rich 1				4	2.01	5.45×10^{-3}
Q3MHR5	SRSF2	Splicing factor, arginine/serine-rich 2				3	2.89	4.67×10^{-4}
Q8VIJ6	SFPQ	Splicing factor, proline- and glutamine-rich	3	3.29	0.03			
Q3T0C7	STMN1	Stathmin				2	6.88	0.04
P80318	CCT3	T-complex protein 1 subunit gamma	3	1.63	3.25×10^{-3}			
Q86V81	THOC4	THO complex subunit 4	2	2.03	0.02	2	1.61	0.01
P37802	TAGLN2	Transgelin-2	6	1.35	0.02			
Q2XVP4	TUBA1B	Tubulin alpha-1B chain	13	1.67	7.24×10^{-3}	14	1.3	0.02
Q3MHM5	TUBB2 C	Tubulin beta-2 C chain	12	1.91	7.95×10^{-3}	14	1.21	0.05
Q3ZBU7	TUBB4	Tubulin beta-4 chain	9	1.85	0.02	9	1.29	7.41×10^{-3}
Q922F4	TUBB6	Tubulin beta-6 chain	7	1.68	0.04			
P69893	TUBB5	Tubulin beta-5 chain	2	1.96	4.91×10^{-3}	16	1.27	0.03

[a] Uniprot accession number from MASCOT search of UniProtKB-SwissProt, taxonomy Mammalia.

[b] Official recommended gene name taken from UniProtKB-SwissProt (in some cases there is no official gene name available so it has not been included).

[c] Fold change showing decreased protein expression in PM-7 transfected cells compared to PM-Neg transfected cells at 48 and at 96

Further information on MASCOT scores, numbers of peptides matched and molecular weight for each protein can be found in Supplementary Table 1A.

Table 3.2.2 List of proteins derived from human, mouse and rat UniProt-SwissProt databases with decreased expression following transient tansfection of miR-7. CHO cells transiently over expressing miR-7 were compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 and >1 peptide between experimental groups being deemed significant.

			PM-7 v PM-Neg 48 h			PM-7 v PM-Neg 96 h		
Accession [a]	Gene name [b]	Protein description	Peptide	Fold change [c]	ANOVA	Peptide	Fold change [c]	ANOVA
Q4R572	YWHAB	14-3-3 protein beta/alpha				8	1.3	0.03
P62258	YWHAE	14-3-3 protein epsilon (14-3-3E)				15	1.44	7.37×10^{-4}
Q5RC20	YWHAG	14-3-3 protein gamma				5	1.41	0.02
Q3SZI4	YWHAQ	14-3-3 protein theta				6	1.32	0.02
Q5R651	YWHAZ	14-3-3 protein zeta/delta				6	1.32	0.04
P17980	PSMC3	26S protease regulatory subunit 6A	3	1.35	0.04			
P07823	HSPA5	78 kDa glucose-regulated protein	18	1.58	0.02	20	2.06	2.48×10^{-4}
P60712	ACTB	Actin, cytoplasmic 1	4	1.42	0.05			
Q4R4I6	CAP1	Adenylyl cyclase-associated protein 1				3	2.2	0.02
P14550	AKR1A1	Alcohol dehydrogenase [NADP+]				3	1.72	0.05
O60218	AKR1B10	Aldo-keto reductase family 1 member B10				2	2.55	2.74×10^{-5}
P16116	AKR1B1	Aldose reductase	4	1.49	6.39×10^{-3}	4	1.93	7.93×10^{-4}
O08782	AKR1B8	Aldose reductase-related protein 2	5	1.73	0.05	7	2.3	2.02×10^{-3}
Q9XSJ4	ENO1	Alpha enolase	2	1.34	0.02			
P12763	AHSG	Alpha-2-HS-glycoprotein				3	7.01	1.12×10^{-3}
P07150	ANXA1	Annexin A1	5	1.66	0.03	5	1.49	8.26×10^{-3}
P07356	ANXA2	Annexin A2				10	1.72	2.89×10^{-3}
Q4R4H7	ANXA5	Annexin A5	3	2.01	6.95×10^{-3}	6	3.41	8.44×10^{-4}
Q03265	ATP5A1	ATP synthase subunit alpha, mitochondrial				4	1.43	2.56×10^{-3}
P56480	ATP5B	ATP synthase subunit beta, mitochondrial				14	1.37	0.02
Q5RAD2	CALM	Calmodulin				4	2.2	0.04
Q8K3H7	CALR	Calreticulin				4	1.74	3.18×10^{-3}
Q5R957	CLIC4	Chloride intracellular channel protein 4				2	5.46	2.56×10^{-3}
Q68FD5	CLTC	Clathrin heavy chain 1				6	1.77	8.83×10^{-3}
Q9D1A2	CNDP2	Cytosolic nonspecific dipeptidase				2	2.19	7.61×10^{-3}
P00639	DNASE1	Deoxyribonuclease-1				5	8.15	1.83×10^{-3}
P08113	HSP90B1	Endoplasmin				21	1.49	8.15×10^{-3}
P15311	EZR	Ezrin				6	1.77	1.95×10^{-3}
P29389	FTH1	Ferritin heavy chain				6	2.51	4.5×10^{-3}
O46638	FKBP3	Peptidyl-prolyl cis-trans isomerase FKBP3	2	1.48	5.99×10^{-5}	2	1.42	0.01
Q923D2	BLVRB	Flavin reductase				4	2.06	1.64×10^{-3}
P05064	ALDOA	Fructose-bisphosphate aldolase A	7	1.26	0.05	8	1.45	6.33×10^{-3}
P48538	LGALS1	Galectin-1	3	1.94	6.75×10^{-3}	3	1.64	2.28×10^{-3}
P30116		Glutathione S-transferase				3	1.54	7.62×10^{-3}
P08263	GSTA1	Glutathione S-transferase A1				2	5.12	1.29×10^{-3}
P30115	GSTA3	Glutathione S-transferase A3	3	3.14	0.02	4	4.94	1.12×10^{-3}
P04905	GSTM1	Glutathione S-transferase Mu 1				3	1.57	0.03
P08010	GSTM2	Glutathione S-transferase Mu 2				4	1.79	7.00×10^{-3}
O35660	GSTM6	Glutathione S-transferase Mu 6				4	1.56	0.02
P46424	GSTP1	Glutathione S-transferase P	7	2.02	6.85×10^{-7}	8	2.93	3.59×10^{-4}
Q00285		Glutathione S-transferase Y1	5	2.16	0.03	5	1.71	9.47×10^{-3}
P08009	GSTM3	Glutathione S-transferase Yb-3				4	1.62	0.02
P46413	GSS	Glutathione synthetase				2	3.16	0.03
P15991	HSPB1	Heat-shock protein beta-1	4	1.53	0.01	7	1.5	1.93×10^{-3}
Q9Z2K8	IDH1	Isocitrate dehydrogenase [NADP] cytoplasmic				2	2.63	0.04
P48678	LMNA	Prelamin-A/C				18	1.91	5.87×10^{-4}
P49129	LAMP1	Lysosome-associated membrane glycoprotein 1	2	2.25	8.14×10^{-3}	2	3.23	6.45×10^{-4}
P24452	CAPG	Macrophage capping protein	2	2.28	9.77×10^{-5}	2	1.62	0.02
Q2HJ49	MSN	Moesin				10	1.92	4.04×10^{-5}
Q9JKY1	PRDX1	Peroxiredoxin-1				11	2.21	7.43×10^{-4}
Q9BGI2	PRDX4	Peroxiredoxin-4				3	2.02	7.34×10^{-3}
Q5RFB8	PGAM1	Phosphoglycerate mutase 1				4	1.29	4.38×10^{-3}
P11598	PDIA3	Protein disulfide-isomerase A3	12	1.89	1.98×10^{-3}	13	2.4	1.75×10^{-4}

Q5R6T1	PDIA6	Protein disulfide-isomerase A6	2	2.15	1.16×10^{-3}	5	2.88	6.37×10^{-3}
P04785	P4HB	Protein disulfide-isomerase	14	1.46	4.48×10^{-3}	15	2.19	4.03×10^{-6}
P05964	S100A6	Protein S100-A6				2	1.4	0.01
P11980	PKM2	Pyruvate kinase isozymes M1/M2				10	1.28	0.02
P50399	GDI2	Rab GDP dissociation inhibitor beta				2	1.51	0.01
Q5R9L3	G3BP2	Ras GTPase-activating protein-binding protein 2	2	44.11	0.01			
Q8BH97	RCN3	Reticulocalbin-3				3	1.7	0.02
P02787	TF	Serotransferrin				8	4.99	1.56×10^{-3}
P19324	SERPH	Serpin H1	7	1.64	0.02			
Q9BYN0	SRXN1	Sulfiredoxin-1	2	5.28	9.02×10^{-4}	2	21.9	1.47×10^{-4}
P08228	SOD1	Superoxide dismutase [Cu-Zn]				2	2.78	2.91×10^{-3}
Q62465	VAT1	Synaptic vesicle membrane protein VAT-1 homolog				2	3.96	1.19×10^{-3}
Q2IBA3	TES	Testin				2	3.61	8.29×10^{-4}
P10639	TXN	Thioredoxin	2	1.78	4.54×10^{-4}			
Q9JMH6	TXNRD1	Thioredoxin reductase 1, cytoplasmic	5	2.43	0.01	6	2.84	2.80×10^{-3}
P37802	TAGLN2	Transgelin-2				6	1.43	2.47×10^{-4}
Q01853	VCP	Transitional endoplasmic reticulum ATPase	7	2.01	4.86 × 10-4	12	2.08	1.75×10^{-3}
P40142	TKT	Transketolase	10	1.57	0.03	11	2.16	6.93×10^{-4}
P48500	TPI1	Triosephosphate isomerase	3	1.26	0.04	3	1.65	3.43×10^{-3}
P02544	VIM	Vimentin				24	1.32	0.04
Q2KJH4	WDR1	WD repeat protein 1	4	3.81	0.02			

[a]

Uniprot accession number from MASCOT search of UniProtKB-SwissProt, taxonomy Mammalia.

[b]

Official recommended gene name taken from UniProtKB-SwissProt (in some cases there is no official gene name available so it has not been included).

[c]

Fold change showing increased protein expression in PM-7 transfected cells compared to PM-Neg transfected cells at 48 and at 96 h.

Further information on MASCOT scores, numbers of peptides matched and molecular weight for each protein can be found in Supplementary Table IB.

Table 3.2.3 List of proteins derived from human, mouse and rat UniProt-SwissProt databases with increased expression following transient tansfection of miR-7. CHO cells transiently over expressing miR-7 were compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 and >1 peptide between experimental groups being deemed significant.

		PM-7 v PM-neg 48hr			PM-7 v PM-neg 96hr		
Gene name	Description	Peptide	Anova (p)	Fold Change [b]	Peptide	Anova (p)	Fold Change [b]
YWHAB	14-3-3 protein beta/alpha	1	3.64E-02	2.26			
YWHAH	14-3-3 protein eta	1	1.54E-02	317.64			
PSMC1	26S protease regulatory subunit 4				1	3.93E-02	1.70
RPS10	40S ribosomal protein S10				1	1.95E-04	1.63
RPS11	40S ribosomal protein S11				1	3.80E-03	3.96
RPS12	40S ribosomal protein S12	1	7.17E-03	1.48			
RPS13	40S ribosomal protein S13				1	3.30E-03	27.99
RPS14	40S ribosomal protein S14				1	1.36E-03	2.42
RPS15A	40S ribosomal protein S15a	1	9.22E-04	5097.92	2	1.12E-03	24.21
RPS16	40S ribosomal protein S16	1	5.33E-03	3.90	1	1.35E-02	6.22
RPS17	40S ribosomal protein S17				1	4.51E-02	3.89
RPS18	40S ribosomal protein S18	2	9.91E-03	2.62	2	8.41E-03	6.41
RPS19	40S ribosomal protein S19				1	3.74E-02	1.42
RPS2	40S ribosomal protein S2	1	4.76E-02	4.96	1	4.17E-03	4.69
RPS20	40S ribosomal protein S20				1	6.88E-04	2.26
RPS21[a]	40S ribosomal protein S21	2	1.31E-02	2.10			
RPS23	40S ribosomal protein S23	2	4.93E-02	46.65	1	2.23E-02	2.74
RPS28	40S ribosomal protein S28	1	4.55E-02	1.65			
RS3A	40S ribosomal protein S3a	1	6.57E-03	Infinity	2	1.97E-02	2.40
RPS7	40S ribosomal protein S7				1	1.33E-02	2.63
RPSA	40S ribosomal protein SA	4	5.92E-03	2.01	4	4.03E-03	1.51
CH60	60 kDa heat shock protein, mitochondrial				1	4.57E-02	1.88
RLA0	60S acidic ribosomal protein P0				2	3.14E-03	1.58
RPL10	60S ribosomal protein L10				1	4.47E-02	2.46
RPL10A	60S ribosomal protein L10a				1	4.33E-04	9.03
RPL11[a]	60S ribosomal protein L11				1	2.75E-02	1.34
RPL13	60S ribosomal protein L13				2	4.86E-03	4.08
RPL13A	60S ribosomal protein L13a				1	1.12E-02	4.98
RPL14	60S ribosomal protein L14				2	2.93E-03	6.50
RPL15	60S ribosomal protein L15				1	7.15E-04	Infinity
RPL18	60S ribosomal protein L18				1	1.75E-02	5.44
RPL19	60S ribosomal protein L19	1	4.71E-02	3.53	1	2.24E-02	6.89
RPL21	60S ribosomal protein L21				2	1.36E-02	2.65
RPL22	60S ribosomal protein L22	2	6.86E-03	1.76	2	9.83E-03	1.83
RPL24	60S ribosomal protein L24				1	4.06E-02	30.06
RPL26L1	60S ribosomal protein L26-like 1	1	2.96E-03	Infinity	3	7.52E-03	9.33
RPS27A	60S ribosomal protein L27a	1	3.82E-02	52.06			
RPL35	60S ribosomal protein L35				1	1.92E-03	18.76
RPL38	60S ribosomal protein L38	1	4.45E-02	1.55	1	5.86E-03	1.31
RPL6	60S ribosomal protein L6				1	3.42E-02	16.78
RPL7	60S ribosomal protein L7				1	3.85E-02	12.15
RPL7A	60S ribosomal protein L7a				1	7.74E-03	2.49
RPL9	60S ribosomal protein L9	1	3.08E-02	82.76			
ACTB	ACTB Actin, cytoplasmic 2				1	8.47E-03	1.81
SUB1	Activated RNA polymerase II transcriptional coactivator p15	1	3.33E-04	Infinity			
AHCY	Adenosylhomocysteinase	2	1.01E-03	2.60			
SLC25A5	ADP/ATP translocase 2	1	5.66E-03	11.44	1	6.07E-03	3.52
ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	1	2.29E-02	101.90			
DDX3Y	ATP-dependent RNA helicase	1	4.63E-03	2.43			
EIF4A1	ATP-dependent RNA helicase eIF4A	1	4.82E-02	1.23	2	1.52E-02	1.54
DDX39B	BAT1 HLA-B associated transcript 1	1	1.19E-04	Infinity			
CAT	Catalase	1	3.39E-03	4.15	3	1.16E-02	3.11
CS	Citrate synthase				1	1.18E-02	2.30
CDA	Cytidine deaminase				1	1.68E-03	1.56
DDX17	DDX5 DEAD (Asp-Glu-Ala-Asp) box polypeptide 17				1	2.58E-03	1.40
TOP2	DNA topoisomerase II				1	2.60E-07	Infinity

TADBP	DNA-binding protein 43				1	1.36E-02	2.00
EF1G	Elongation factor 1-gamma				1	8.42E-04	1.30
EEF2[a]	Elongation factor EF-2	2	6.24E-03	1.59	3	9.55E-03	2.55
EEF1A1	Eukaryotic translation elongation factor 1 alpha 1	1	4.54E-02	1.50	3	1.03E-04	1.45
EIF3I	Eukaryotic translation initiation factor 3 subunit I	1	2.14E-02	3.59			
EIF4A3	Eukaryotic translation initiation factor 4A3				1	4.89E-02	1.45
CSE1L	Exportin-2				1	1.17E-02	10.03
FUBP2	Far upstream element-binding protein 2	1	4.76E-02	3.04			
FKBP4	FK506 binding protein 4, 59kDa	1	6.00E-03	1.70	1	9.54E-03	1.42
GSR	Glutathione reductase (NADPH)	1	1.36E-03	Infinity			
RAN[a]	GTP-binding nuclear protein Ran	3	2.83E-03	1.62			
GNB2L1[a]	Guanine nucleotide-binding protein subunit beta-2 1				2	1.47E-02	1.64
HSPA8	Heat shock 70kDa protein 1/8	4	8.77E-04	1.73			
HSP90A[a]	Heat shock protein HSP 90-alpha	3	2.05E-03	1.46			
HSP90AB1	Heat shock protein HSP 90-beta	2	2.23E-03	1.72	1	9.86E-03	2.06
HNRNPH2	Heterogeneous nuclear ribonucleoprotein A/B	1	4.60E-03	1.51	1	7.38E-03	3.99
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1				2	2.43E-02	1.72
HNRNPCL1	Heterogeneous nuclear ribonucleoprotein C-like 1	1	3.82E-02	1.50			
HNRNPF	Heterogeneous nuclear ribonucleoprotein F	1	3.68E-02	1.41	1	4.46E-03	1.63
HNRPK	Heterogeneous nuclear ribonucleoprotein K				1	1.62E-02	1.67
HNRNPU	Heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)				2	3.35E-02	1.96
RIR2	High mobility group protein B1				1	2.56E-02	1.27
HIST1H1C	Histone H1.2	1	1.42E-02	1.91	4	1.02E-04	8.02
HIST1H2AG[a]	Histone H2A type 1				6	4.28E-05	4.85
H2AFV	Histone H2A.V				3	3.16E-04	3.86
H2AZ[a]	Histone H2A.Z				1	1.64E-03	3.12
HIST1H2BM	Histone H2B type 1-M				1	3.58E-02	17.66
HIST3H3[a]	Histone H3.1t				3	9.65E-05	11.28
H33	Histone H3.3	1	3.91E-02	4.36			
H3F3A	Histone H3.3				2	9.67E-05	11.28
HISTH4A	Histone H4				1	1.35E-03	65.80
N/A[a]	Hypothetical protein LOC100766349				1	8.56E-03	1.68
HPRT1	Hypoxanthine phosphoribosyltransferase 1				1	3.75E-02	1.62
IMPDH1	IMP dehydrogenase	2	9.57E-03	2.87	1	2.77E-04	4.08
IMB1	Importin subunit beta-1				2	8.94E-04	2.98
IMDH2[a]	Inosine-5'-monophosphate dehydrogenase 2	2	9.57E-03	2.87			
KRT10[a]	Keratin, type I cytoskeletal 10				1	7.59E-04	1.84
LDHA	Lactate dehydrogenase A	2	2.26E-02	2.08	2	4.13E-03	2.01
MIF	Macrophage migration inhibitory factor	1	1.24E-06	Infinity			
MAP6	Microtubule-associated protein 6				1	7.40E-03	3.91
MAPRE1	Microtubule-associated protein RP/EB family member 1				1	6.07E-04	Infinity
HTPG	Molecular chaperone HtpG	3	4.44E-03	1.49			
MYL9	Myosin regulatory light polypeptide 9	1	2.89E-03	Infinity			
AHNAK	Neuroblast differentiation-associated protein AHNAK	2	6.06E-03	3.02			
NUDC	Nuclear migration protein nudC	1	7.27E-03	1.98	1	2.20E-04	1.48
NUCL	Nucleolin	5	1.03E-03	1.81	4	4.78E-03	1.38
NPM1	Nucleophosmin				1	3.77E-03	2.91
NAP1L4	Nucleosome assembly protein 1-like 1				1	2.03E-02	1.55
ATIC	Phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP cyclohydrolase				2	8.41E-03	2.64
PCBP1	Poly(rC)-binding protein 1				1	3.31E-02	1.99
PABP1	Polyadenylate-binding protein 1	2	7.25E-04	1.52	3	1.17E-02	1.50
PA2G4	Proliferation-associated protein 2G4	1	8.27E-03	1.87	1	2.61E-02	1.77

PSME3	Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma)	1	1.21E-02	11.31	1	4.19E-04	32.13
PRMT1	Protein arginine methyltransferase 1	2	4.61E-02	1.70			
RANBP1	Ran-specific GTPase-activating protein	2	5.13E-04	1.72	1	1.04E-02	1.94
RBMX	RNA binding motif protein, X-linked	1	1.20E-03	1.69			
STRAP	Serine-threonine kinase receptor-associated protein	1	3.48E-02	5.92	1	1.27E-02	2.61
PTMA	Similar to prothymosin, alpha (gene sequence 28)				4	3.19E-04	95.19
SNRP70[a]	Small nuclear ribonucleoprotein polypeptide F, isoform CRA_b				1	1.54E-02	1.59
SNRPD3	Small nuclear ribonucleoprotein Sm D3				1	2.42E-02	1.34
SUMO2	Small ubiquitin-related modifier 2				1	4.80E-02	1.35
SRSF1	Splicing factor, arginine/serine-rich 1				2	2.05E-03	2.02
SRSF7	Splicing factor, arginine/serine-rich 7	1	3.21E-02	9.22			
STMN1	Stathmin				1	1.18E-04	22.20
CCT3	T-complex protein 1 subunit gamma				1	1.89E-03	3.95
CCT6A	T-complex protein 1 subunit zeta	1	1.08E-03	Infinity	1	3.23E-02	7.35
BTF3	Transcription factor BTF3				1	2.11E-02	1.53
RHOA	Transforming protein RhoA				1	1.12E-03	2.35
TPM4	Tropomyosin alpha-4 chain	1	6.04E-03	2.34			
TUBA1C	Tubulin alpha-1C chain	2	1.77E-04	1.91			
TUBB	Tubulin beta	4	2.01E-02	2.01			
TUBB2C[a]	Tubulin beta-2C chain	4	2.01E-02	2.01			
TUBB4	Tubulin beta-4 chain				5	3.37E-03	1.90
FAU	Ubiquitin-like protein FUBI				1	4.48E-02	5.98

[a]

Protein identification obtained from NCBI nonredudant database which was not present in BBCHO or had a higher confidence score compared to BBCHO database

[b]

Fold change showing increased protein expression in PM-7 transfected cells compared to PM-Neg transfected cells at 48 and at 96 h.

Table 3.2.4 List of proteins derived from CHO database with decreased expression following transient tansfection of miR-7. CHO cells transiently over expressing miR-7 were compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 and >1 peptide between experimental groups being deemed significant.

		PM-7 v PM-neg 48hr			PM-7 v PM-neg 96hr		
Gene name	Description	Peptide	Anova (p)	Fold Change [b]	Peptide	Anova (p)	Fold Change [b]
YWHAB	14-3-3 protein beta/alpha				1	4.98E-02	1.29
YWAHE	14-3-3 protein epsilon				2	5.84E-03	1.62
YWHAH	14-3-3 protein eta				1	3.46E-02	25650.97
YWHAG	14-3-3 protein gamma	1	2.68E-03	1.77	1	4.91E-03	2.01
YWHAZ	14-3-3 protein zeta/delta				1	1.69E-02	1.31
PSMD4	26S proteasome non-ATPase regulatory subunit 4				1	2.50E-02	1.74
MRPL12	39S ribosomal protein L12, mitochondrial				1	4.58E-04	Infinity
GRP78	78 kDa glucose-regulated protein	5	1.05E-02	2.17	8	6.40E-04	2.64
CAP1	Adenylyl cyclase-associated protein 1				2	3.74E-03	2.21
AKR1A1	Aldehyde reductase				1	8.03E-05	1.28
AKR1B8	Aldo-keto reductase family 1, member B8	1	7.56E-03	1.78	1	6.72E-05	2.56
FBAA	Aldolase A, fructose-bisphosphate				1	4.40E-02	1.21
ACTN1	Alpha-actinin-1	1	3.86E-02	3.75			
ANXA1	Annexin A1				2	5.69E-03	1.58
ANXA2	Annexin A2				2	1.19E-02	1.81
ANXA5	Annexin A5	1	2.37E-02	1.51	3	7.67E-04	3.02
BLVRB	Biliverdin reductase B (flavin reductase (NADPH))	1	2.89E-02	2.38	4	5.98E-04	1.99
CALR	Calreticulin	1	3.92E-02	7.54	2	1.54E-02	1.80
CES1	Carboxylesterase 1 (monocyte/macrophage serine esterase 1)				1	1.84E-07	Infinity
CLTC	Clathrin heavy chain 1				1	2.45E-02	2.19
CLU	Clusterin	2	2.07E-04	16.60	3	1.75E-03	23.74
ENPL	Endoplasmin	1	3.51E-02	1.71	4	8.01E-04	1.68
ERP57[a]	ERP57 protein				9	3.66E-04	2.45
ESD	Esterase D/formylglutathione hydrolase	1	2.56E-02	1.69	2	2.46E-02	1.84
EIF5A	Eukaryotic translation initiation factor 5A-1				1	4.06E-02	110.86
FTH1	Ferritin, heavy polypeptide 1				1	1.01E-02	1.96
ATP5B	F-type H+-transporting ATPase subunit beta				3	2.49E-02	1.65
LEG3	Galectin-3				1	1.50E-02	3.53
GAA	Glucosidase, alpha	1	5.13E-03	6.37	1	5.05E-06	24.18
GCLM	Glutamate-cysteine ligase, modifier subunit	1	1.22E-02	1.62	1	2.55E-02	2.37
GSTA3	Glutathione S-transferase alpha 3				1	2.56E-03	7.91
GSTM1	Glutathione S-transferase mu 1	2	3.91E-02	1.93	1	2.25E-03	1.90
GSTP1	Glutathione S-transferase P	5	3.24E-04	2.83	5	4.89E-04	3.60
GSTP2[a]	Glutathione S-transferase P 2				3	1.18E-03	173.37
GSS[a]	Glutathione synthetase	1	1.88E-02	2.25			
G3P[a]	Glyceraldehyde-3-phosphate dehydrogenase	1	1.22E-02	Infinity			
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase pseudogene	1	1.22E-02	Infinity			
HSPB1	Heat shock protein beta-1				1	1.11E-02	1.53
ROA1	Heterogeneous nuclear ribonucleoprotein A1	1	2.23E-02	132.87			
HNRNPA1[a]	Heterogeneous nuclear ribonucleoprotein A1 isoform a	1	2.23E-02	132.87			
HNRDL	Heterogeneous nuclear ribonucleoprotein D-like				1	1.36E-02	1.40
ROA2	Heterogeneous nuclear ribonucleoproteins A2/B1	1	3.37E-03	18.72			
CDC37	Hsp90 co-chaperone Cdc37	1	4.31E-02	21.57			
IDH1	Isocitrate dehydrogenase 1 (NADP+), soluble				1	3.76E-04	1.87
KRT77[a]	Keratin, type II cytoskeletal 1b	1	4.66E-02	1.94			
MFGE8	Lactadherin				1	8.95E-05	10.83
LMNA	Lamin-A/C				6	2.18E-03	1.94
LGMN	LGMN Legumain				1	5.61E-03	60.91
LAMP1	Lysosome-associated membrane glycoprotein 1				2	1.07E-03	3.20
LAMP2	Lysosome-associated membrane glycoprotein 2				1	9.61E-03	10.42
MIF	Macrophage migration inhibitory factor				1	3.92E-03	1.28
MOES	Moesin				3	8.32E-04	2.42
NME1	NME1 non-metastatic cells 1, protein (NM23A) expressed in	1	4.30E-02	61.52			
PRDX1	Peroxiredoxin 1				3	4.53E-03	2.14
PGAM1	Phosphoglycerate mutase	1	2.17E-02	88.19	1	7.48E-04	1.40
P4HB	Prolyl 4-hydroxylase, beta polypeptide	2	1.80E-04	1.57	9	3.71E-04	2.45

PDIA6	Protein disulfide isomerase family A, member 6				4	2.70E-03	2.29
S100A13	Protein S100-A13	1	2.32E-02	2.51	1	4.38E-02	2.85
PKLR	Pyruvate kinase				2	1.32E-02	1.25
GDI2	Rab GDP dissociation inhibitor beta				1	7.74E-03	1.40
G3BP2	Ras GTPase-activating protein-binding protein 2	1	2.90E-02	38.04			
GDI1	Rho GDP-dissociation inhibitor 1				1	1.29E-04	1.71
RPL37A	Ribosomal protein L37a	1	1.38E-02	158.89			
SERPH	Serpin H1	1	1.55E-02	1.79			
SOD2	SOD2 superoxide dismutase 2, mitochondrial				1	2.09E-02	5.26
SRSF3	Splicing factor, arginine/serine-rich 3				1	4.45E-04	40989.80
STIP1	Stress-induced-phosphoprotein 1	1	7.38E-04	Infinity			
SRXN1	Sulfiredoxin	1	1.18E-03	4.59	1	2.44E-04	18.07
SOD1	Superoxide dismutase 1, soluble				2	3.15E-03	2.78
TXN	Thioredoxin	1	2.03E-02	4.40			
TXNRD1	Thioredoxin reductase 1	2	1.60E-04	2.15	4	5.56E-03	2.59
TYMS	Thymidylate synthase	1	4.56E-04	Infinity			
TPI1	TPI1 Triosephosphate isomerase 1				2	2.90E-03	1.69
TAGLN2	Transgelin-2				3	4.74E-05	1.46
VCP	Transitional endoplasmic reticulum ATPase	1	1.23E-02	1.71	4	1.69E-04	1.90
TKTL2	Transketolase-like 2	3	6.16E-03	1.76	5	1.50E-03	2.35
WARS	Tryptophanyl-tRNA synthetase				1	2.88E-03	2.34
UMA1	Ubiquitin-like modifier activating enzyme 1	3	1.98E-04	1.66			
UAP1L1	UDP-N-acteylglucosamine pyrophosphorylase 1-like 1				1	4.53E-02	2.18
UBE2N	Uubiquitin-conjugating enzyme E2 N				1	4.48E-02	1.31
WDR1	WD repeat-containing protein 1				1	2.51E-02	16.26

[a]

Protein identification obtained from NCBI nonredudant database which was not present in BBCHO or had a higher confidence score compared to BBCHO database

[b]

Fold change showing increased protein expression in PM-7 transfected cells compared to PM-Neg transfected cells at 48 and at 96 h.

Table 3.2.5 List of proteins derived from CHO database with increased expression following transient tansfection of miR-7. CHO cells transiently over expressing miR-7 were compared to scramble negative control transfected cells with ANOVA ≤0.05 and fold change >1.2 and >1 peptide between experimental groups being deemed significant.

3.2.2 Mammalian vs CHO databases

Having completed the differential analysis with both the BBCHO combined with NCBI Chinese hamster databases and also the multi species human, mouse and rat database (mammalian) the two could be compared. Standard statistical cut offs were used for both searches with ANOVA ≤0.05 and fold change >1.2. The mammalian database search resulted in 217 differentially expressed proteins, excluding 1 hit peptide IDs, while the CHO database search resulted in 265 differentially expressed proteins, including 1 hit peptides. Due to the large number of 1 peptide hits from the CHO database search (66%) we decided to include them and investigate overlap with the mammalian list.

This resulted in four comparisons between 48 hr up-regulated, 48 hr down-regulated, 96 hr up-regulated and 96 hr down-regulated lists. As the 1 peptide IDs were included for the BBCHO and not the mammalian list we were more interested in the overlaps rather than comparing total numbers. There were 12, 21, 33 and 40 overlapped identifications found between the BBCHO and mammalian databases (see **Chapter 3**) in the 48 hr up, 48 hr down, 96 hr up and 96 hr down lists respectively (**Figure 3.2.2**). In all cases there were more peptides identified in each of the overlapping IDs with the mammalian list than in the CHO database list (**Figure 3.2.3 - 3.2.6**). This strengthens the 1 hit identifications found in the CHO database with the presence of multiple peptides in the mammalian search.

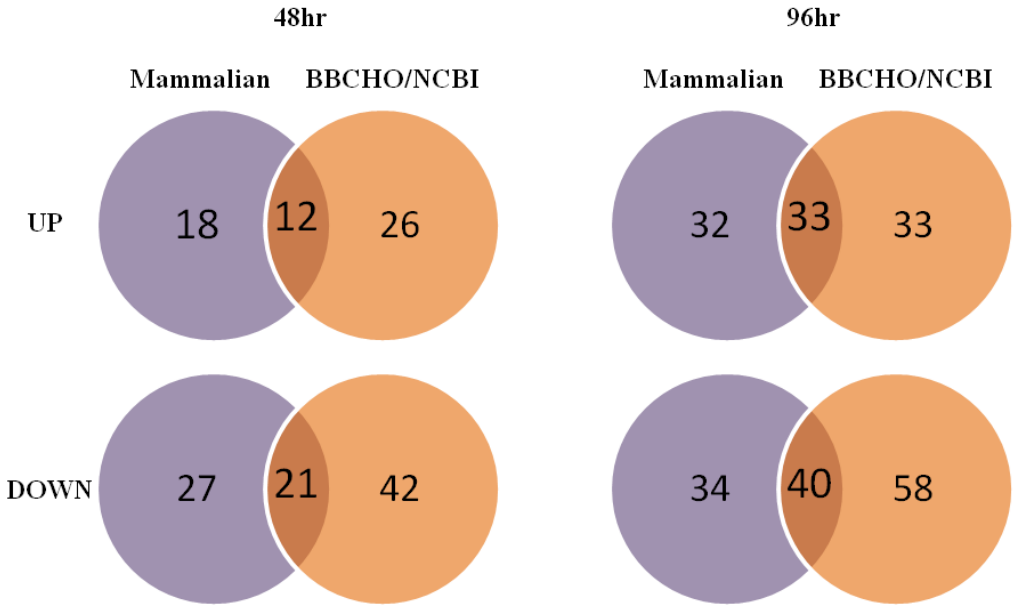


Figure 3.2.2 Overlap between gene name identifications associated with differentially expressed proteins in CHO-K1_SEAP cells over expressing miR-7 IDs were derived from a mammalian (human, mouse and rat) and Chinese hamster (BBCHO and NCBI non redundant) databases.

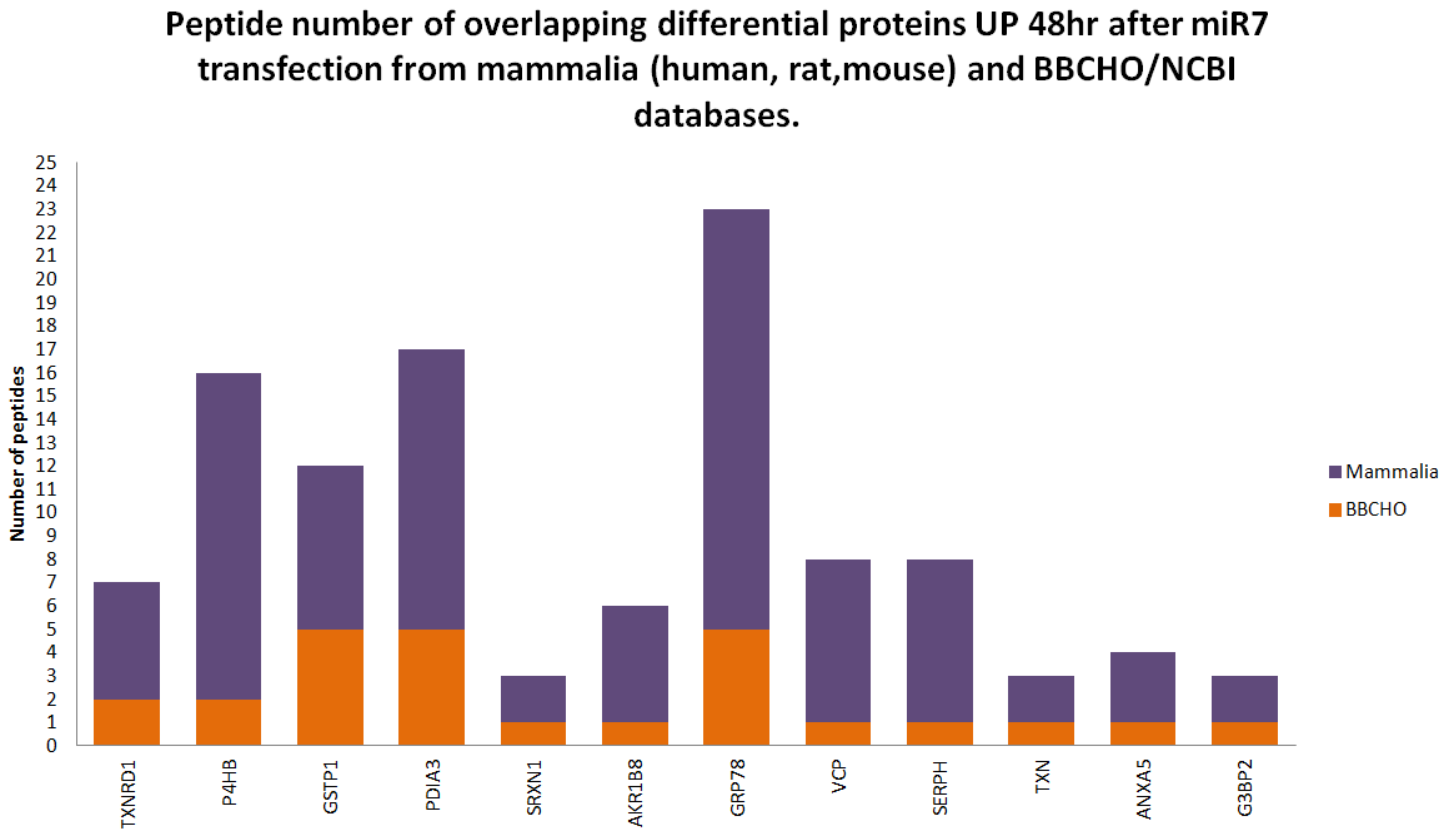


Figure 3.2.3 Graph of number of peptides found for each protein at 48 hr up-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database. Corresponding gene names shown from left to right are TXNRD1, P4HB, GSTP1, PDIA3, SRXN1, AKR1B8, GRP78, VCP, SERPH, TXN, ANXA5, G3BP2 all of which can be seen in detail on **Table 3.2.3** and **3.2.5**. In all cases more peptides were found in the corresponding multi-species IDs but this also confirmed presence of many hamster IDs within the mammalian database search.

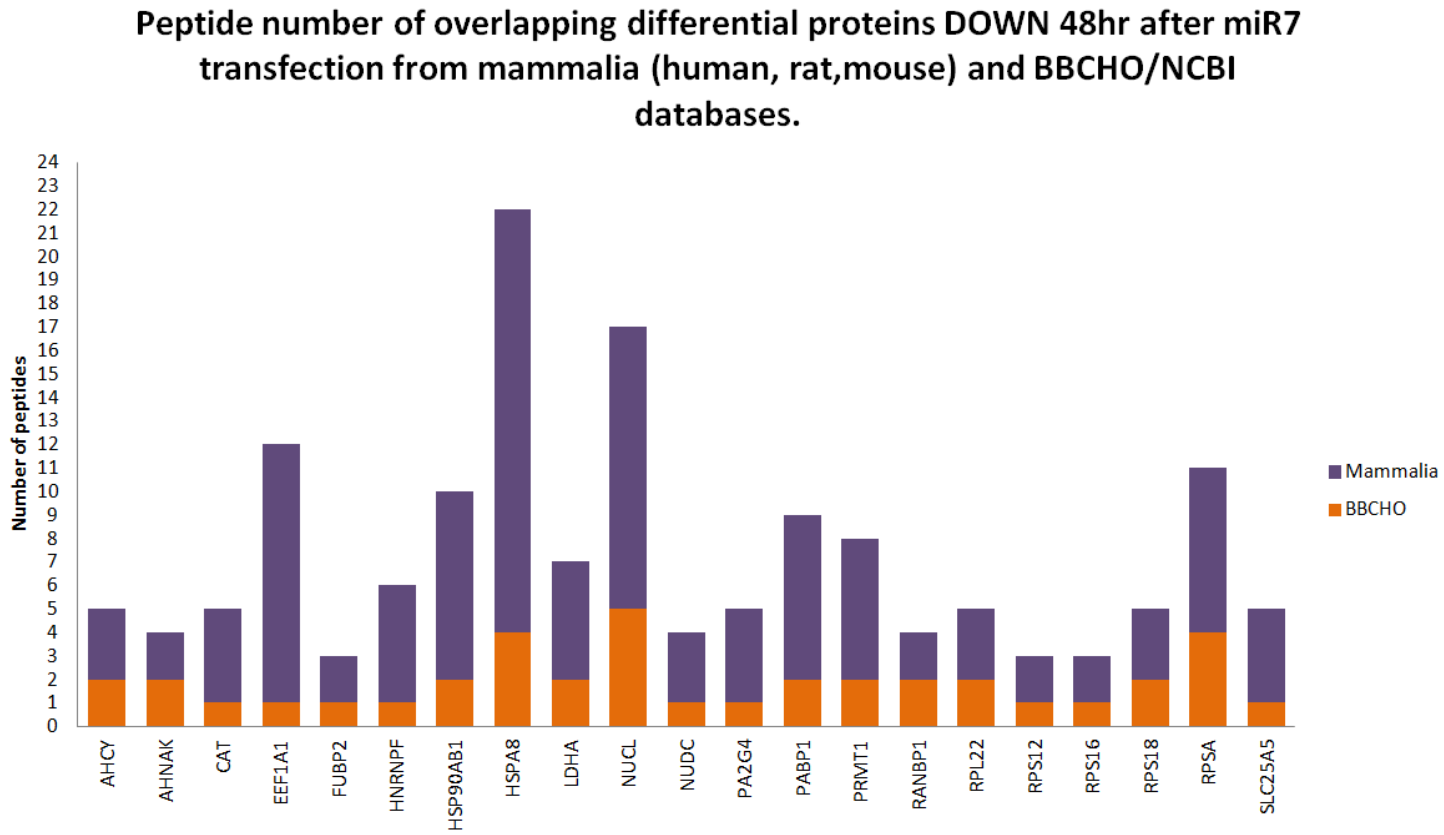


Figure 3.2.4 Graph of number of peptides found for each protein at 48 hr down-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database. Corresponding gene names shown from left to right AHCY, AHNAK, CAT, EEF1A1, FUBP2, HNRNPF, HSP90AB1, HSPA8, LDHA, NUCL, NUDC, PA2G4, PABP1, PRMT1, RANBP1,PRL22,RPS12, RPS16, RPS18, RPSA, SLC25A5 all of which can be seen in detail on **Table 3.2.2** and **3.2.4** . In all cases more peptides were found in the corresponding multi-species IDs but this also confirmed presence of many hamster IDs within the mammalian database search.

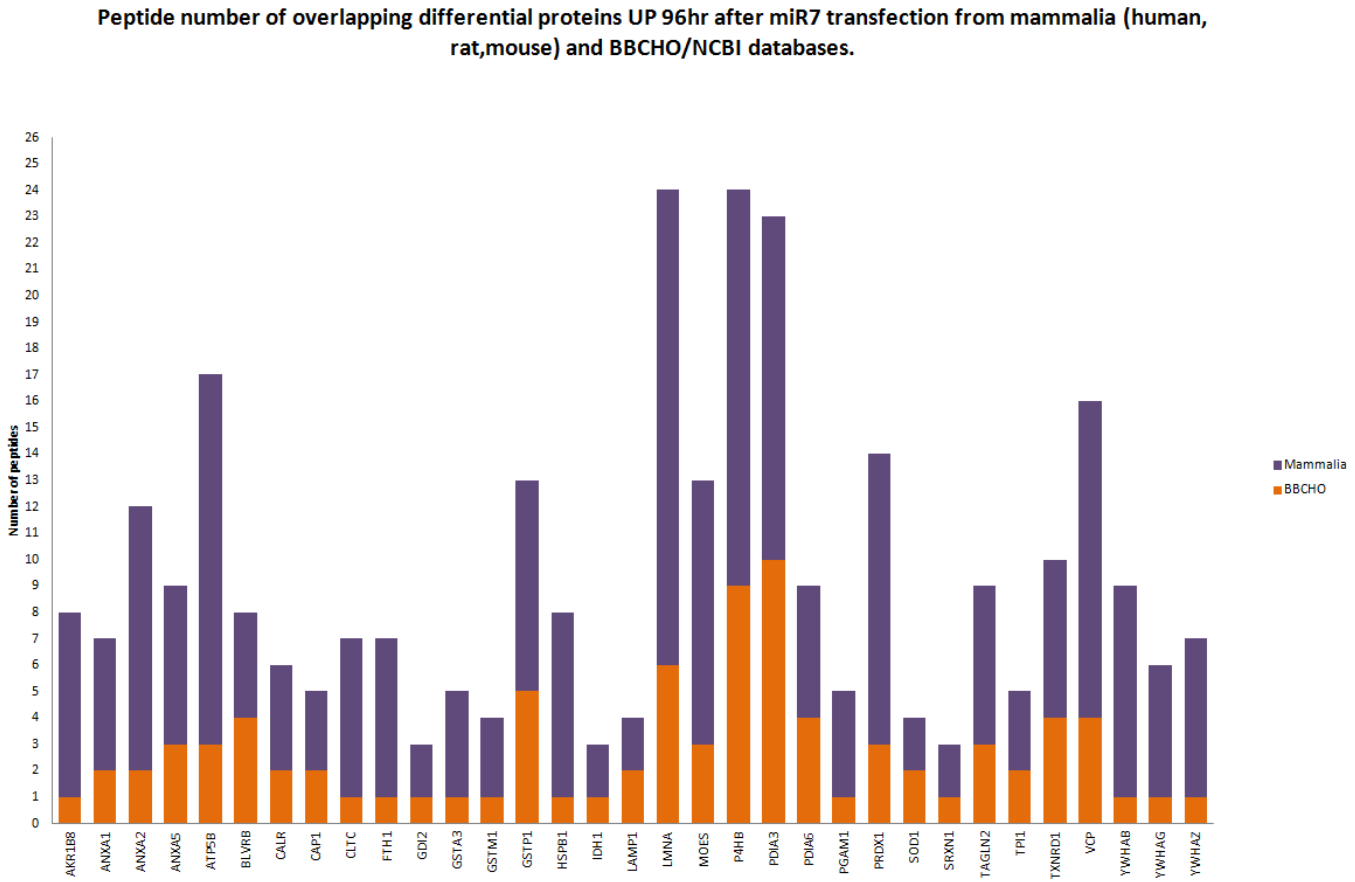


Figure 3.2.5 Graph of number of peptides found for each protein at 96 hr up-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database. Corresponding gene names from left to right AKR1B8, ANXA1, ANXA2, ANAXA5, ATP5B, BLVRB, CALR, CAP1, CLTC, FTH1, GDI2, GSTA3, GSTM1, GSTP1, HSPB1, IDH1, LAMP1, LMNA, MOES, P4HB, PDIA3, PDIA6, PGAM1, PRDX1, SOD1SRXN1, TAGLN1, TPI1, TXNRD1, VCP, YWHAB, YWHAG, YWHAZ all of which can be seen in **Table 3.2.3** and **3.2.5**. In all cases more peptides were found in the corresponding multi-species IDs but this also confirmed presence of many hamster IDs within the mammalian database search.

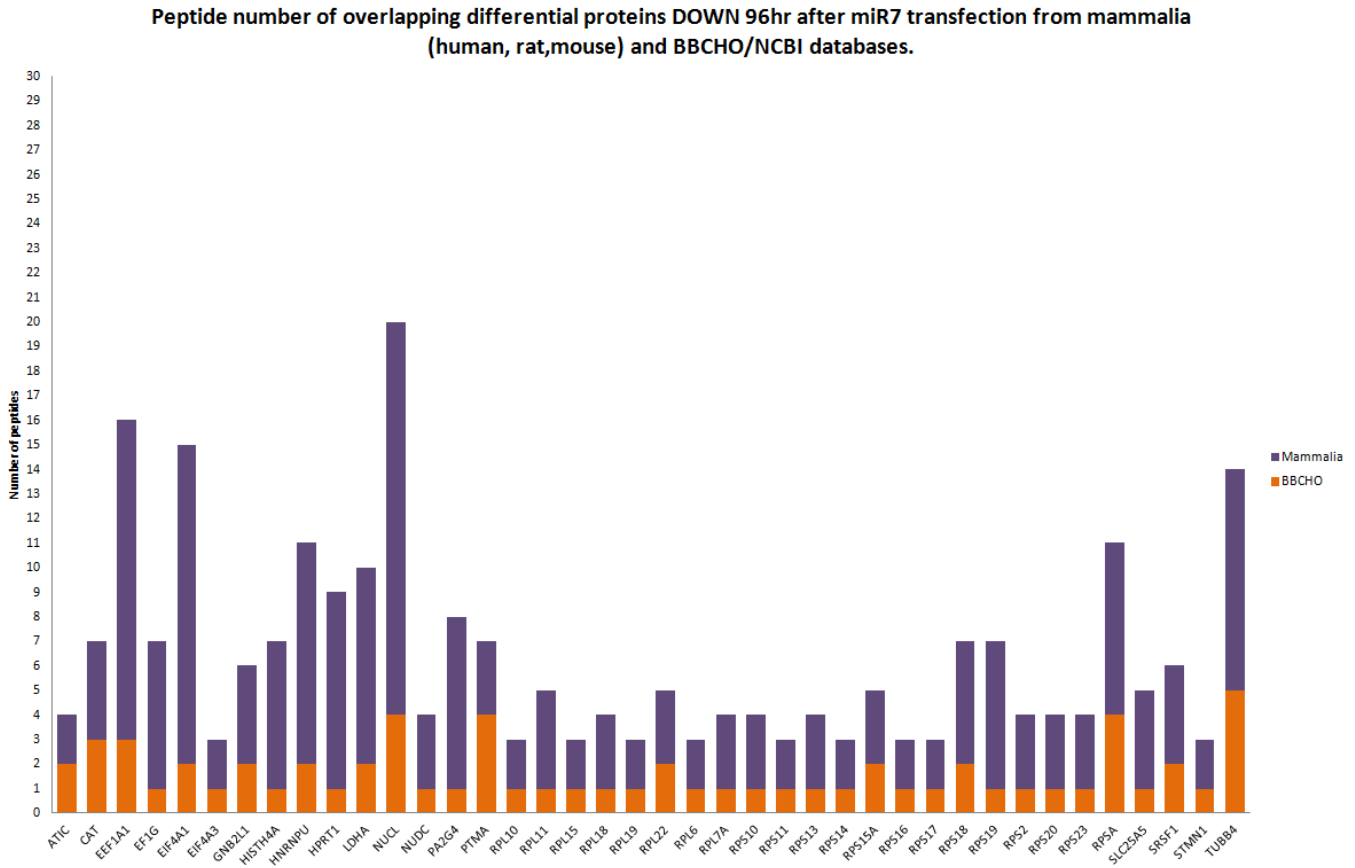


Figure 3.2.6 Graph of number of peptides found for each protein at 96 hr down-regulated between the mammalian (purple) and BBCHO/NCBI hamster (orange) database. Corresponding gene names from left to right ATIC, CAT, EEF1A1, EF1G, EIF4A1, GNB2L1, HISTH4A, HNRNPU, HPRT1, LDHE, NUCL, NUDC, PA2G4, PTMA, RPL10, RPL11, RPL15, RPL18, RPL19, RPL22, RPL6, RPL7A, RPS10, RPS11, RPS13, RPS14, RPS15A, RPS16, RPS17, RPS18, RPS19, RPS2, RPS20, RPS23, RPSA, SLC25A5, SRSF1, STMN1TUBB4 all of which can be seen in detail on **Table 3.2.2** and **3.2.4**. In all cases more peptides were found in the corresponding multi-species IDs but this also confirmed presence of many hamster IDs within the mammalian database.

3.2.3 Western blot identification validation

Having completed the differential analysis with both the BBCHO/NCBI Chinese hamster databases and the multi species databases (human, mouse, rat) a number of proteins were chosen for validation by Western blot analysis. Often commercially available antibodies raised against proteins of interest derived from a different species to Chinese hamster will not react with CHO samples. The targets therefore were chosen first from the multi-species list as these were likely highly conserved sequences that commercial antibodies would possibly have immunoreactivity to and then followed up if that the same protein was reported dysregulated in the CHO database output.

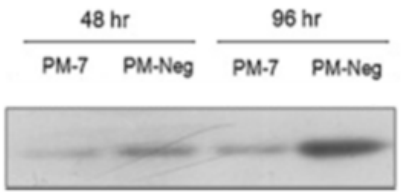
We chose 4 proteins that were reported to be >2 fold differentially regulated with ANOVA $p<0.05$ according to the mammalian list and 2 protein that was <2 fold dysregulated in the miR-7 transfected group compared to the scramble negative control group (**Table 3.2.6**).

	Mammalian fold change		**BBCHO/NCBI fold change**	
Target	**48 hr**	**96 hr**	**48 hr**	**96 hr**
Histone H3	3.83↓	9.13↓	4.36↓	11.28↓
Histone H4	3.62↓	8.07↓	-	65.08↓
PDIA6	2.15↑	2.88↑	-	2.29↑
GRP78	1.58↑	2.06↑	2.17↑	2.64↑
HSPA8	18↓	-	1.73↓	-
14-3-3 epsilon	-	1.44↑	-	1.62↑

Table 3.2.6 Label free LC-MS fold change of proteins chosen for Western blot validation. Targets were chosen initially from the mammalian list to maximise commercial antibody immunoreactivity and then follow up with their presence in the CHO list (BBCHO/NCBI). Down regulation is denoted by "↓" and up regulation is denoted by "↑".

All the proteins chosen for Western blot validation displayed the same differential expression pattern observed in the label free fold change data for miR-7 transient up regulation using the CHO database. The most pronounced differential expression was seen with Histone H3 and H4 (**Figure 3.2.7 and 3.2.8**). The other 4 proteins PDIA6, GRP78, HSPA8 and 14-3-3 epsilon were less prominently dysregulated according to Western blot results (**Figure 3.2.9 - 3.2.12**) but did still follow the same expression trend observed in the label free data. Equal loading was also observed with all samples (**Figure 3.2.13**).

Western Blot



Quantitative label free protein expression 96 hr

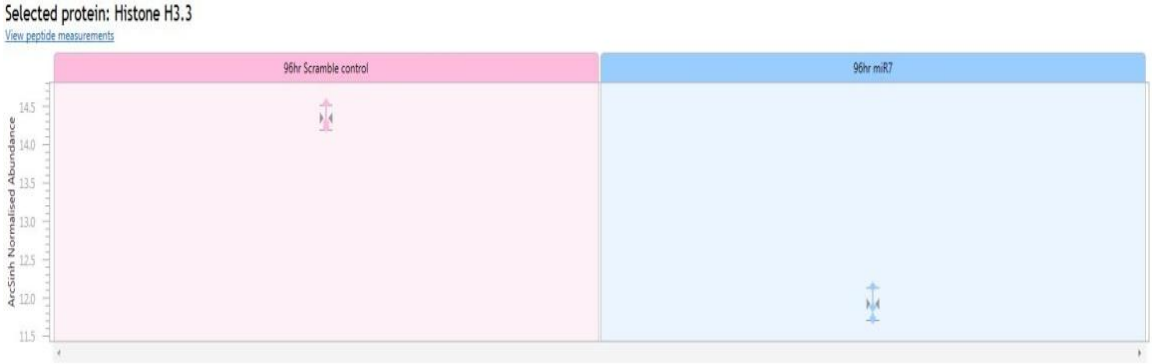
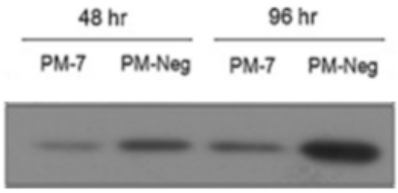


Figure 3.2.7 Western blot analysis showing down regulation of Histone H3 in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg). Quantitative label-free LC-MS/MS data using CHO database also shows reduced expression at 96 hr after transfection in the miR-7 up-regulated group (right) compared to the negative control (left) (n=3). As seen in **Table 3.2.4** Histone H3 is down-regulated 4.36 fold at 48 hr ($p=3.91 \times 10^{-2}$).

Western Blot



Quantitative label free protein expression 96 hr

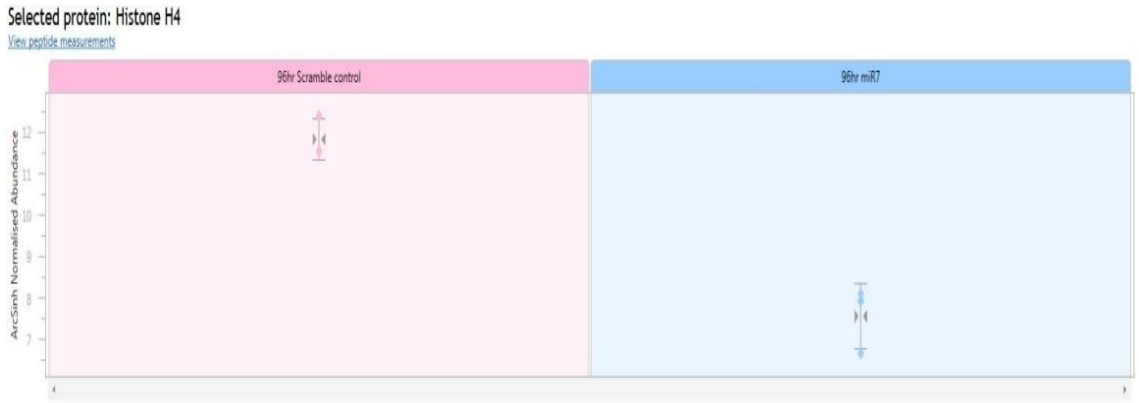
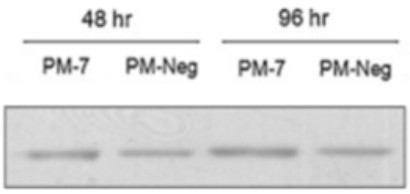


Figure 3.2.8 Western blot analysis showing down regulation of Histone H4 in CHO- SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg). Quantitative label-free LC-MS/MS data using CHO database also shows reduced expression at 96 hr after transfection in the miR-7 up-regulated group (right) compared to the negative control (left) (n=3). As seen in **Table 3.2.4** Histone H3 is down regulated 65.8 fold at 48 hr ($p=1.35 \times 10^{-3}$).

Western Blot



Quantitative label free protein expression 96 hr

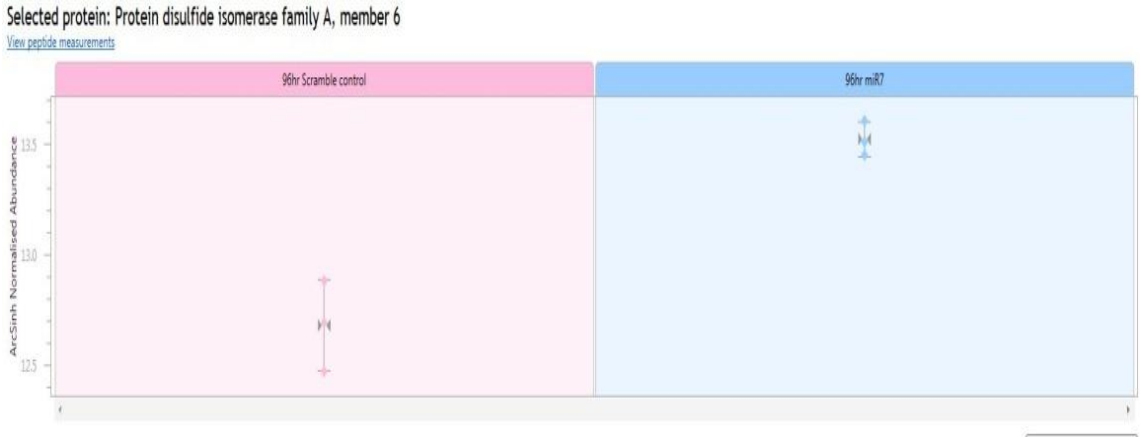
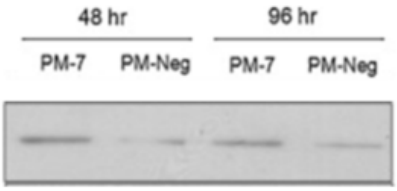


Figure 3.2.9 Western blot showing up regulation of PDIA6 in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg). Quantitative label-free LC-MS data using CHO database also shows increased expression at 96 hr after transfection in the miR-7 up-regulated group (right) compared to the negative control (left) (n=3). From **Table 3.2.5** PDIA6 is up-regulated 2.29 fold at 96 hr ($p=2.7 \times 10^{-3}$).

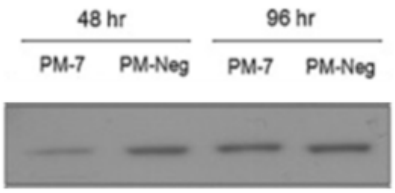
Western Blot



Quantitative label free protein expression 96 hr

Figure 3.2.10 Western blot showing up regulation of GRP78 in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg). Quantitative label free LC-MS data using CHO database also shows increased expression at 96 hr after transfection in the miR-7 up-regulated group (right) compared to the negative control (left) (n=3). From **Table 3.2.5** GRP78 is up-regulated 2.64 fold at 96 hr ($p=6.4 \times 10^{-4}$).

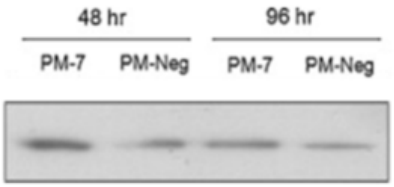
Western Blot



Quantitative label free protein expression 48 hr

Figure 3.2.11 Western blot analysis showing down regulation of HSPA8 at 48 hr in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg). Quantitative label-free LC-MS data using CHO database also shows reduced expression at 48 hr after transfection in the miR-7 up-regulated group (right) compared to the negative control (left) (n=3). As seen in **Table 3.2.4** HSPA8 is down-regulated 1.73 fold at 48 hr ($p=8.77 \times 10^{-4}$).

Western Blot



Quantitative label free protein expression 48 hr

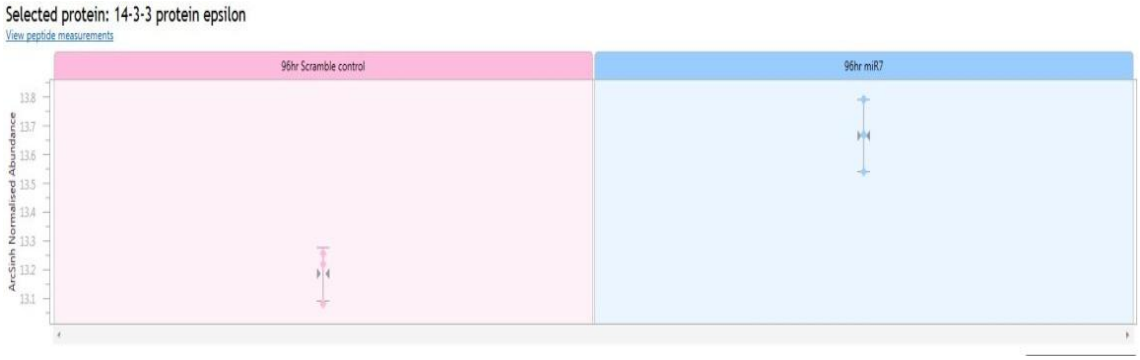


Figure 3.2.12 Western blot analysis (left) showing up regulation of 14-3-3 epsilon in CHO-SEAP cells over expressing miR-7 (PM-7) compared to a transfected negative control (PM-Neg). Quantitative label-free LC-MS data using CHO database also shows increased expression at 48 hr after transfection in the miR-7 up-regulated group (right) compared to the negative control (left) (n=3). As seen in **Table 3.2.5** 14-3-3 epsilon is up-regulated 317 fold at 48 hr ($p=1.54 \times 10^{-2}$)

Western Blot GAPDH

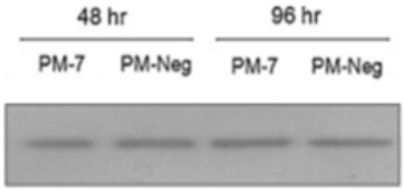


Figure 3.2.13 Western blot analysis for GAPDH showing equal loading for **Figure 3.2.7 to 3.2.12**.

3.3 Predicted miR-7 target analysis using miRWalk

The targets for translational inhibition by miR-7 in CHO cells were investigated using the quantitative label-free LC-MS data in **Table 3.2.2**. Specifically we focused on down-regulated proteins as miRNA inhibits translation of its target mRNA with one miRNA potentially having several mRNA targets while up regulation is often a by-product of miRNA mediated negative regulation. An online tool miRWALK (http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/index.html) was used to determine if any of the down-regulated proteins in miR-7 over expressing cells were significant direct targets of miR-7. These direct complementary targets of miR-7 may have a domino effect resulting in the observed proteomic profile of miR-7 up regulation. The miRWALK bioinformatics tool combines the output of several miRNA target prediction algorithms (Dweep et al. 2011). We searched RNA22, miRanda, miRDB, TargetScan and RNAhybrid across mouse, rat and human species (**Section 1.4.2**). The most probable direct targets of miR-7 were those that appeared across as many of the 5 algorithms used which is generally regarded as best practice and also across as many of the three species that were searched (**Table 3.3.2**). Following confirmation that these proteins were present in the differentially expressed lists two were chosen for validation, catalase and stathmin.

According to the label-free LC-MS data in for the mammalian database (**Table 3.2.2**) and for the CHO database (**Table 3.2.4**) catalase was down-regulated at 48 and 96 hr while stathmin was at 96 hr in miR-7 over expressing cells (**Table 3.3.1**).

	Mammalian fold change		**BBCHO/NCBI fold change**	
Target	**48 hr**	**96 hr**	**48 hr**	**96 hr**
Catalase	4.07↓	4.34↓	4.15↓	3.11↓
Stathmin	-	6.88↓	-	22.2↓

Table 3.3.1 Catalase and stathmin down regulation (↓) in miR-7 over expressing CHO-K1-SEAP cells according to label free LC-MS data from mammalian and CHO (BBCHO/NCBI) databases. (**Table 3.3.2**).

The down regulation of stathmin and catalase was further validated by Western blot (**Figure 3.3.1 and 3.3.2**). Interestingly these predicted targets were identified as differentially expressed from one peptide in the CHO database while the multi-species

search yielded 4 peptides for catalase and 2 peptides for stathmin. Going forward with this information and observing the large amount of overlap between the mammalian multi-species list and the CHO database (**Figure 3.2.2**) the CHO database was used as follow up pathway analysis and for the CHO temperature shift analysis (**Chapter 4**).

Gene name	MicroRNA	StemLoop ID	miRanda	miRDB	miRWalk	RNA22	RNAhybrid	TargetScan	Overlap
Mouse									
Cat	mmu-miR-7b	mmu-mir-7b	•	•	•		•	•	5
Cat	mmu-miR-7a	mmu-mir-7a-2	•	•	•			•	4
Stmn1	mmu-miR-7a	mmu-mir-7a-2	•		•			•	3
Serbp1	mmu-miR-7b	mmu-mir-7b	•		•	•			3
Sfrs1	mmu-miR-7a	mmu-mir-7a-2			•			•	2
Caprin1	mmu-miR-7a	mmu-mir-7a-2	•		•				2
Cct3	mmu-miR-7b	mmu-mir-7b	•				•		2
Sfrs1	mmu-miR-7b	mmu-mir-7b			•			•	2
Pa2g4	mmu-miR-7b	mmu-mir-7b	•					•	2
Caprin1	mmu-miR-7b	mmu-mir-7b	•		•				2
Stmn1	mmu-miR-7b	mmu-mir-7b			•			•	2
Impdh2	mmu-miR-7b	mmu-mir-7b			•		•		2
Rat									
Cat	rno-miR-7a	rno-mir-7a-2		•	•		•	•	4
Stmn1	rno-miR-7a	rno-mir-7a-2	•		•		•	•	4
Cat	rno-miR-7b	rno-mir-7b		•	•		•	•	4
Stmn1	rno-miR-7b	rno-mir-7b			•		•	•	3
Tubb5	rno-miR-7a	rno-mir-7a-2			•		•		2
Ran	rno-miR-7a	rno-mir-7a-2			•		•		2
Nap1l1	rno-miR-7a	rno-mir-7a-2					•	•	2
Tagln2	rno-miR-7a	rno-mir-7a-2					•	•	2
Eef1a1	rno-miR-7a	rno-mir-7a-2			•		•		2
Hspa90aa1	rno-miR-7a	rno-mir-7a-2			•		•		2
Cct3	rno-miR-7a	rno-mir-7a-2	•				•		2
Kpnb1	rno-miR-7a	rno-mir-7a-2			•		•		2
Tubb5	rno-miR-7b	rno-mir-7b			•		•		2
Ran	rno-miR-7b	rno-mir-7b			•		•		2
Nap1l1	rno-miR-7b	rno-mir-7b					•	•	2
Tagln2	rno-miR-7b	rno-mir-7b					•	•	2
Eef1a1	rno-miR-7b	rno-mir-7b			•		•		2
Hspa90aa1	rno-miR-7b	rno-mir-7b			•		•		2
Cct3	rno-miR-7b	rno-mir-7b	•				•		2
Kpnb1	rno-miR-7b	rno-mir-7b			•		•		2
Human									
PA2G4	hsa-miR-7	hsa-mir-7-3	•		•			•	3
RPL15	hsa-miR-7	hsa-mir-7-3			•			•	2
SFRS1	hsa-miR-7	hsa-mir-7-3			•			•	2
AHNAK	hsa-miR-7	hsa-mir-7-3	•					•	2
CAPRIN1	hsa-miR-7	hsa-mir-7-3			•	•			2

Table 3.3.2 Predicted direct targets of mir-7 using miRWALK across mouse, rat and human species. In total 5 search algorithms were used to search the down-regulated label-free LC-MS mammalian protein data associated with miR-7 over expression (**Table 4.2.2**). Catalase and stathmin were chosen for follow up based on the

high number of hits across multiple databases in both rat and mouse species as well as appearing down-regulated in label-free LC-MS for both the mammalian and CHO database lists.

Catalase

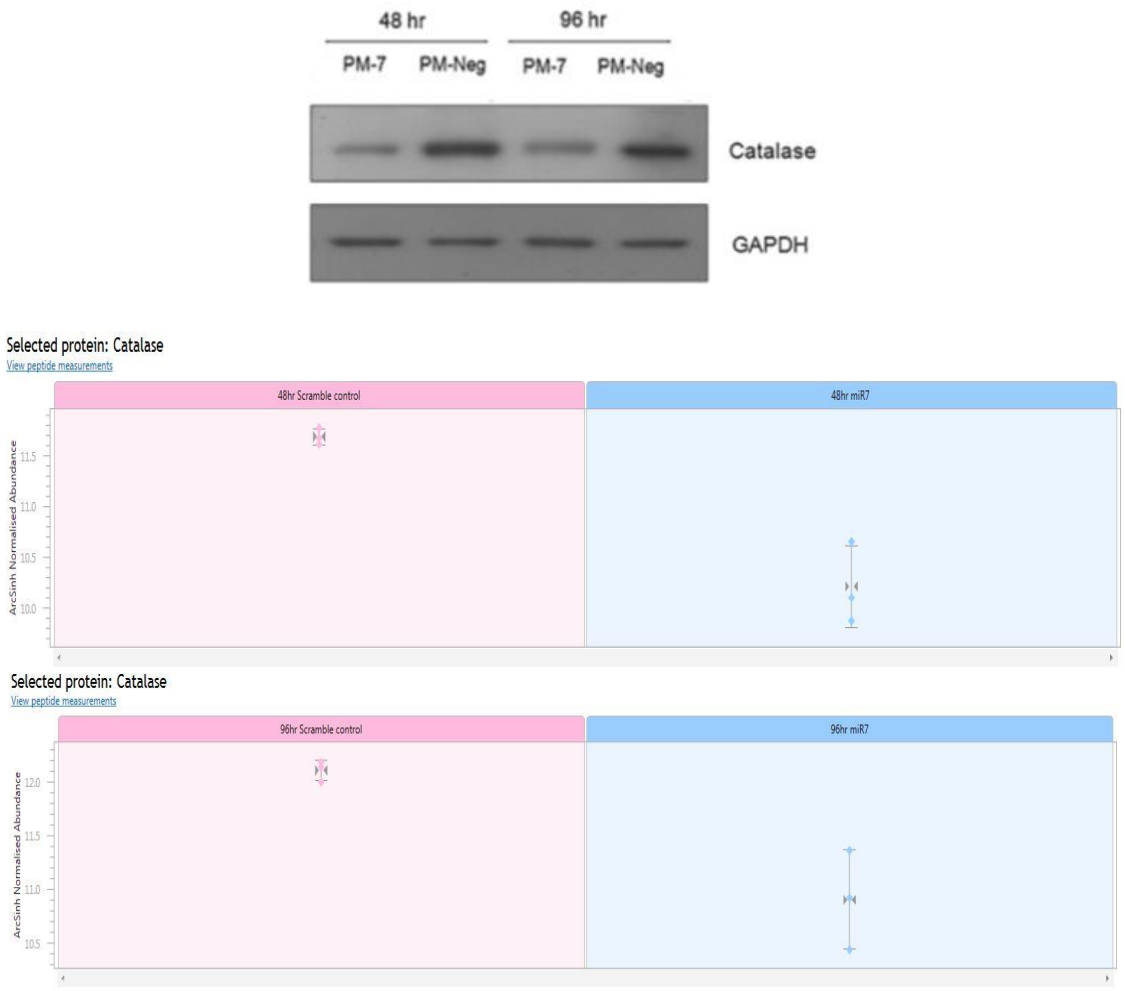


Figure 3.3.1 Normalised abundance of peptides associated with predicted direct target of miR-7 catalase from the quantitative label-free LC-MS/MS data using the CHO database (**Table 4.2.4)**. Each line represents a peptide while each point represents a sample. Expression of catalase is reduced on the right in pre-miR-7 transfected cells compared to scramble control cells on the left. Catalase expression was reduced 4.15 fold ($p=3.39 \times 10^{-3}$) at 48 hr and catalase expression was reduced 3.11 fold ($p=1.16 \times 10^{-2}$) at 96 hr (n=3).

Stathmin

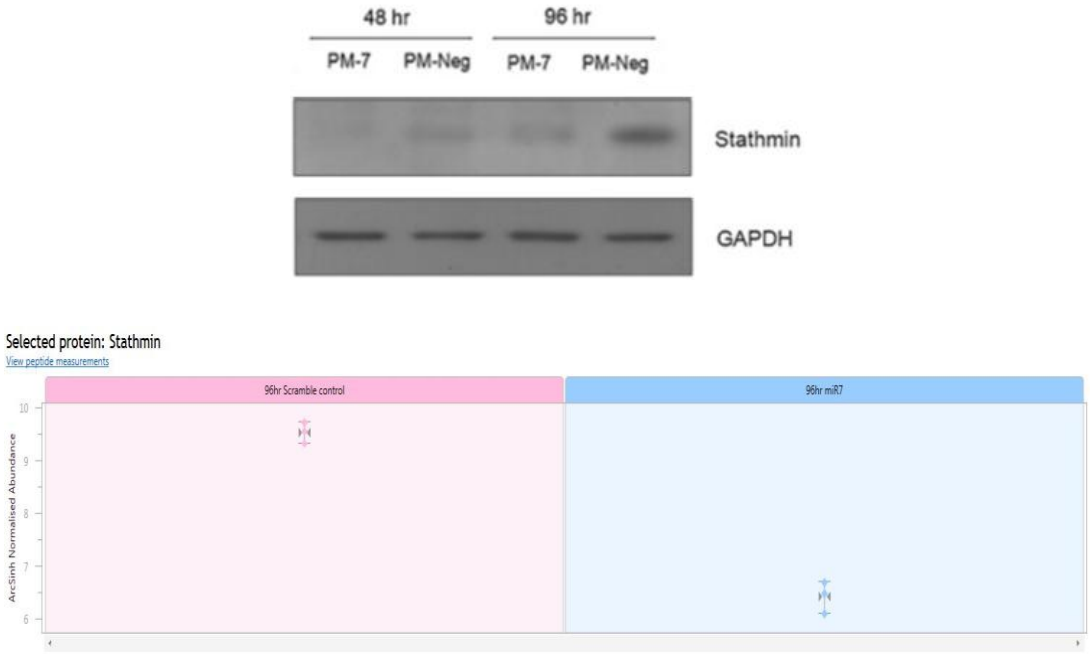


Figure 3.3.2 Normalised abundance of peptides associated with predicted direct target of miR-7 stathmin from the quantitative label-free LC-MS/MS data using the CHO database (**Table 4.2.4**). Expression of stathmin is reduced (on the right) in pre-miR-7 transfected cells compared to scramble control cells on the left. stathmin expression was reduced 22.2 fold ($p=1.18 \times 10^{-4}$) at 96 hr with no significant change at 48 hr (n=3).

3.4 Pathway analysis

In proteomics and other "omics" fields large amounts of identification data is produced. In an effort to contextualise these large lists of identifications, pathway analysis tools are used. Generally these comprise of evidence-based and literature mining associated links between genes and/or proteins and their relationships to each other. In order to use these tools to search identification data, the data must be in a usable form for the pathway tool. As the protein identifiers from the Chinese hamster proteomic databases are not in a form that is usable for pathway analysis the annotation had to be conducted manually (**Section 3.4.1**). Once the identifications were converted to corresponding gene names they were submitted to pathway analysis tools. Since different pathway tools have different hierarchical ways of assigning these relationships and determining what relationships are significant, three different pathway analysis tools were used

- DAVID - **D**atabase for **A**nnotation, **V**isualisation and **I**ntegrated **D**iscovery) (**Section 3.4.2**)
- PANTHER - **P**rotein **An**alysis **Th**rough **E**volutionary **R**elationships (**Section 3.4.3**)
- KEGG (**K**yoto **E**ncyclopaedia of **G**enes and **G**enomes (**Section 3.4.4**)

The output from each pathway tool is different but each is complementary to the other giving a greater insight to the processes, functions, components and pathways affected by increased miR-7 expression in CHO cells.

3.4.1 Protein annotation

In order to use the protein identifications for pathway analysis and predict functional effects of the differentially expressed proteins, the protein identifications must be in a suitable format. For lists such as the human, mouse and rat multi-species analysis (**Table 3.2.2 and 3.2.3**) these databases have well annotated protein accession numbers associated with each identification making them immediately ready for pathway analysis tools. The BBCHO and NCBI Chinese hamster ovary protein identifications do not have the same identifiers.

The BBCHO database being derived from transcriptomic data has "isotig" number identifiers for each protein identification as well as other information as part of the text string that is the protein identification, while the NCBI non-redundant CHO list consists of a protein name and unofficial gene name derived from ChiTaRS database (**Figure 3.4.1**). In order to use these lists the identifications had to be converted to official gene names or uniprot accession numbers (see **Section 1.5.1.4**). This required significant manual searching and required each peptide identification to be searched using BLAST to validate the correct protein identification. This was an extremely time-consuming process spanning 4 differential lists for the miR-7 study (**Table 3.2.4 and 3.2.5**).

BBCHO

Accession
CHOisotig08417v2 RecName: Full=Chromodomain-helicase-DNA-binding protein 4; Short=CHD-4; AltName: Full=ATP-dependent helicase CHD4; AltName: Full=Mi-2 autoantigen 218 kDa protein; AltName: Full=Mi2-beta

NCBI

Accession	Description
CHOisotig09352v1 product="Phosphoserine aminotransferase" homologue_to="gi;@	20140196 sp|Q99K85.1|SERC_MOUSE" gene_name="SERC"

Annotation summary

Unique gene count	Unique name count
1135	746

Figure 3.4.1 Examples of typical text strings from the BBCHO in house CHO database (top) and the NCBI non redundant CHO database (bottom) requiring annotation. A combination of text mining, manual parsing and blast search of peptide sequences through UniProt resulted in 1135 gene names and 746 protein assigned to CHO IDs.

In summary, for the BBCHO protein identification text string, the name was isolated from between "RecName: Full=" and "; Short=". For the NCBI text string the name was isolated from between "product=" and "homologue to=". Multiple names occurred in the same text string therefore the peptide sequences associated with each protein was then blast searched using UniProt, the full name confirmed and the gene name added. Frequent manual parsing was also required for text strings that followed different patterns. A master reference list was compiled for the miR-7 study and the temperature shift study in **Chapter 4** resulting in 1135 gene names and 746 protein names being manually assigned across the differential lists using the above method.

3.4.2 DAVID analysis

The database for annotation, visualisation and integrated discovery (DAVID)(https://david.ncifcrf.gov/home.jsp) is a computational tool combining several publicly available gene databases to assess the functional and biological patterns associated with a list of submitted identifications (Huang da, Sherman and Lempicki 2009a, Huang da, Sherman and Lempicki 2009b). For this analysis we looked at molecular function (MF), biological process (BP) and cellular component (CC) which uses the gene ontology (GO) database through DAVID. On its own, GO is an encyclopaedic database of gene functions and their relationships. These three contexts of MF, BP and CC are defined by GO (Ashburner et al. 2000). DAVID further categorises this information into 5 levels based on order of complexity with level 1 being general fundamental categories up to level 5 which encompasses immune response terminology. To understand the fundamentals of our observed phenotype we used the FAT category which includes the lower 3 levels of organisation in DAVID.

A process was deemed to be significantly enriched when adjusted Bonferroni p-value was ≤ 0.05. The analysis was divided into four lists based on 48 and 96 hr time points and increased or decreased protein abundance in response to transient transfection mediated over expression of miR-7. As an initial comparison the BP enriched from the multi-species human, mouse, rat list at 48 hr (**Figure 3.4.1A**) and 96 hr (**Figure 3.4.2A**) were compared to the enriched BP from corresponding CHO database lists at 48 hr (**Figure 3.4.1B**) and 96 hr (**Figure 3.4.2B**) after miR-7 up-regulation. The following observations were made on this comparison:

- At 48 hr, cell redox homeostasis was associated with up-regulated proteins in both the multi species and Chinese hamster lists with additional processes in the CHO list related to homeostasis and apoptosis.

- At 48 hr, down-regulated proteins in the multi-species list were associated with translational elongation, macromolecule, chromatin and nucleosome assembly with the CHO list showing translation and translational elongation enrichment.

- At 96 hr, up-regulated proteins in the multi-species were associated with cellular homeostasis and regulation of apoptosis processes which was also seen in the CHO list enriched BP output.

- At 96 hr down-regulated proteins in the multi-species list were related to BP terms translation, RNA processing and assembly of the nucleosome, chromatin and macromolecule complex, similar output was also observed for the corresponding CHO list.

Due to the high similarity between the output for BP between the two different database lists it was decided that the CHO database list would be used for all subsequent pathway analysis.

Analysing MF (**Table 3.4.3**) with the differentially expressed proteins derived from the CHO database in response to miR-7 over expression we observed the following:

- At 48 hr, up-regulated proteins were not significantly associated with any MF.
- At 48 hr, down-regulated proteins were associated with structural constituent of the ribosome, structural molecule activity and RNA binding.
- At 96 hr, protein up regulation was associated with intramolecular oxidoreductase activity, antioxidant activity and glutathione transferase activity.
- At 96 hr, down regulation of proteins was related to structural constituent of the ribosome, structural molecule activity and RNA binding.

The greater degree of enrichment of functions associated with down regulation potentially highlights the negative regulation of miR-7 on cellular activities. This is most evident at 48 hr, but interestingly it is in up regulation and glutathione transferase activity which we see more references to in KEGG pathway analysis output.

Finally with CC (**Table 3.4.**4) analysis on the differentially expressed proteins derived from the CHO database in response to miR-7 over expression we observed the following:

- At 48 hr, up regulation of proteins were associated with the cytosol.
- At 48 hr, down-regulated proteins were associated with a larger number of components including the cytosol, ribosome, intracellular non-membrane-bounded organelle, non-membrane bounded organelle, melanosome and pigment granule.

- At 96 hr, up-regulated proteins were associated with the melanosome, cytosol, soluble fraction and vesicle.

- At 96 hr, the down-regulated proteins were associated with the ribosome, cytosol, intracellular non-membrane bound organelle.

Again, this emphasis on negative regulatory effect of miRNA with a large number of CC terms enriched in the down-regulated proteins. By 96 hr the enriched CC terms are more balanced in number between up and down-regulated proteins.

To take BP, MF and CC into consideration together the following conclusions can be made regarding up and down regulation of proteins in response to miR-7 up regulation in CHO-K1-SEAP cells

- Up regulation of proteins are related to homeostasis processes related to antioxidant activity at 48 hr which are localised in the cytosol. This carries over to 96 hr with added up-regulation of anti-apoptotic and apoptotic regulation processes that are functionally linked to oxidoreductase activity and glutathione transferase activity and associated with the melanosomal, cytosolic and vesicle components. This suggests that early 48 hr stage oxidative stress homeostasis is activated with antioxidant proteins such as thioredoxin, thioredoxin reductase and superoxide dismutase being enriched in this pathway. The maintaining of oxidative stress response proteins may in turn be related to cell survival with the addition of anti-apoptotic up regulation at 96 hr. In that regard homeostasis process up regulation may be related to regulatory and inhibitory elements of homeostasis rather than an up regulation of homeostasis activity.

- Down-regulated proteins at 48 hr are involved in translation processes, functionally linked with RNA binding and structural molecule activity and most significantly localised to the ribonucleo complex, cytosol, ribosome and melanosome. Similarly at 96 hr there was down regulation of translational, chromatin and ribosomal process proteins which were functionally related to structural constituent of the ribosome, structural molecule activity and RNA binding with localisation associated with the ribosome and cytosolic components.

A

Biological process	Count	p-value	Adjusted p-value
Up-regulated at 48hr			
GO:0045454 ~ cell redox homeostasis	5	4.2×10^{-06}	1.4×10^{-03}
Down-regulated at 48hr			
GO:0006414 ~ translational elongation	9	4.6×10^{-10}	1.8×10^{-07}
GO:0034622 ~ cellular macromolecular complex assembly	11	2.0×10^{-08}	8.1×10^{-06}
GO:0034621 ~ cellular macromolecular complex subunit	11	6.0×10^{-08}	2.4×10^{-05}
GO:0006333 ~ chromatin assembly or disassembly	7	1.8×10^{-06}	7.1×10^{-04}
GO:0065003 ~ macromolecular complex assembly	12	2.3×10^{-06}	9.3×10^{-04}
GO:0042274 ~ ribosomal small subunit biogenesis	4	3.8×10^{-06}	1.5×10^{-03}
GO:0043933 ~ macromolecular complex subunit organisation	12	4.4×10^{-06}	1.7×10^{-03}
GO:0006334 ~ nucleosome assembly	6	4.5×10^{-06}	1.8×10^{-03}
GO:0006412 ~ translation	9	4.5×10^{-06}	1.8×10^{-03}
GO:0031497 ~ chromatin assembly	6	5.4×10^{-06}	2.1×10^{-03}
GO:0065004 ~ protein–DNA complex assembly	6	6.7×10^{-06}	2.6×10^{-03}
GO:0034728 ~ nucleosome organisation	6	7.5×10^{-06}	2.9×10^{-03}
GO:0006323 ~ DNA packaging	6	2.2×10^{-05}	9.1×10^{-03}

B

Biological Process	Count	p-value	Adjusted p-value
Up regulated at 48hr			
GO:0019725~cellular homeostasis	7	2.99E-05	1.39E-02
GO:0042981~regulation of apoptosis	8	6.84E-05	3.16E-02
GO:0043067~regulation of programmed cell death	8	7.29E-05	3.36E-02
GO:0010941~regulation of cell death	8	7.46E-05	3.44E-02
GO:0045454~cell redox homeostasis	4	8.84E-05	4.06E-02
GO:0051235~maintenance of location	4	9.27E-05	4.26E-02
Down regulated at 48hr			
GO:0006414~translational elongation	16	1.77E-21	9.79E-19
GO:0006412~translation	18	2.37E-16	1.23E-13

Table 3.4.1 GO biological process (BP) analysis through DAVID of up or down-regulated proteins 48 hr after transient over expression of miR-7 in CHO-K1-SEAP cells. The multi-species (A) and CHO databases (B) were both used and enriched BP compared in both. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

A

Biological process	Count	p-value	Adjusted p-value
Up regulated at 96hr			
GO:0019725 ~ cellular homeostasis	13	5.1×10^{-07}	4.3×10^{-04}
GO:0042981 ~ regulation of apoptosis	15	5.3×10^{-06}	4.5×10^{-03}
GO:0043067 ~ regulation of programmed cell death	15	6.0×10^{-06}	5.1×10^{-03}
GO:0010941 ~ regulation of cell death	15	6.2×10^{-06}	5.3×10^{-03}
GO:0045454 ~ cell redox homeostasis	6	7.0×10^{-06}	6.0×10^{-03}
GO:0051235 ~ maintenance of location	6	7.6×10^{-06}	6.4×10^{-03}
GO:0006916 ~ anti-apoptosis	8	2.6×10^{-05}	2.2×10^{-02}
Down-regulated at 96hr			
GO:0006414 ~ translational elongation	30	9.1×10^{-46}	4.0×10^{-43}
GO:0006412 ~ translation	31	2.9×10^{-31}	1.3×10^{-28}
GO:0006334 ~ nucleosome assembly	9	6.7×10^{-09}	2.9×10^{-06}
GO:0031497 ~ chromatin assembly	9	8.9×10^{-09}	3.9×10^{-06}
GO:0065004 ~ protein-DNA complex assembly	9	1.2×10^{-08}	5.6×10^{-06}
GO:0034621 ~ cellular macromolecular complex subunit	14	1.4×10^{-08}	6.2×10^{-06}
GO:0034728 ~ nucleosome organisation	9	1.5×10^{-08}	6.7×10^{-06}
GO:0034622 ~ cellular macromolecular complex assembly	13	3.6×10^{-08}	1.6×10^{-05}
GO:0006323 ~ DNA packaging	9	9.3×10^{-08}	4.1×10^{-05}
GO:0006333 ~ chromatin assembly or disassembly	9	1.7×10^{-07}	7.8×10^{-05}
GO:0043933 ~ macromolecular complex subunit organisation	16	1.2×10^{-06}	5.3×10^{-04}
GO:0006396 ~ RNA processing	14	1.9×10^{-06}	8.4×10^{-04}
GO:0042254 ~ ribosome biogenesis	8	2.0×10^{-06}	9.0×10^{-04}
GO:0022613 ~ ribonucleoprotein complex biogenesis	9	2.5×10^{-06}	1.1×10^{-03}
GO:0065003 ~ macromolecular complex assembly	15	3.0×10^{-06}	1.3×10^{-03}
GO:0042274 ~ ribosomal small subunit biogenesis	4	1.7×10^{-05}	7.8×10^{-03}
GO:0006364 ~ rRNA processing	6	8.4×10^{-05}	3.6×10^{-02}
GO:0000398 ~ nuclear mRNA splicing, via spliceosome	7	9.8×10^{-05}	4.2×10^{-02}
GO:0000375 ~ RNA splicing, via transesterification reactions	7	9.8×10^{-05}	4.2×10^{-02}
GO:0000377 ~ RNA splicing, via transesterification reactions	7	9.8×10^{-05}	4.2×10^{-02}
GO:0016072 ~ rRNA metabolic process	6	1.0×10^{-04}	4.4×10^{-02}

B

Biological Process	Count	p-value	Adjusted p-value
Up regulated at 96hr			
GO:0042981~regulation of apoptosis	16	6.18E-08	6.12E-05
GO:0043067~regulation of programmed cell death	16	7.04E-08	6.98E-05
GO:0010941~regulation of cell death	16	7.39E-08	7.32E-05
GO:0043066~negative regulation of apoptosis	11	3.58E-07	3.55E-04
GO:0043069~negative regulation of programmed cell death	11	4.07E-07	4.03E-04
GO:0060548~negative regulation of cell death	11	4.18E-07	4.14E-04
GO:0019725~cellular homeostasis	12	5.14E-07	5.09E-04
GO:0006916~anti-apoptosis	8	7.44E-06	7.34E-03
GO:0006979~response to oxidative stress	7	2.33E-05	2.29E-02
GO:0006886~intracellular protein transport	9	4.54E-05	4.40E-02
GO:0042592~homeostatic process	12	4.90E-05	4.74E-02
Down regulated at 96hr			
GO:0006414~translational elongation	33	8.38E-51	4.96E-48
GO:0006412~translation	34	1.92E-34	1.14E-31
GO:0042254~ribosome biogenesis	12	3.86E-11	2.29E-08
GO:0022613~ribonucleoprotein complex biogenesis	13	1.64E-10	9.68E-08
GO:0042274~ribosomal small subunit biogenesis	6	1.67E-09	9.89E-07
GO:0006396~RNA processing	17	2.07E-08	1.22E-05
GO:0006364~rRNA processing	9	2.82E-08	1.67E-05
GO:0016072~rRNA metabolic process	9	3.94E-08	2.33E-05
GO:0034621~cellular macromolecular complex subunit organi	13	3.51E-07	2.08E-04
GO:0034622~cellular macromolecular complex assembly	12	8.58E-07	5.08E-04
GO:0043933~macromolecular complex subunit organization	16	3.84E-06	2.27E-03
GO:0034470~ncRNA processing	9	6.45E-06	3.81E-03
GO:0065003~macromolecular complex assembly	15	8.86E-06	5.23E-03
GO:0042273~ribosomal large subunit biogenesis	4	1.69E-05	9.95E-03
GO:0034660~ncRNA metabolic process	9	2.90E-05	1.70E-02
GO:0006334~nucleosome assembly	6	8.26E-05	4.77E-02

Table 3.4.2 GO biological process (BP) analysis through DAVID of up or down-regulated proteins 96 hr after transient over expression of miR-7 in CHO-K1-SEAP cells. The multi-species (A) and CHO databases (B) were both used and enriched BP compared in both. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

A

Molecular Function	Count	p-value	Adjusted p-value
Up regulated at 48hr			
NONE			
Down regulated at 48hr			
GO:0003735~structural constituent of ribosome	14	1.72E-14	2.93E-12
GO:0005198~structural molecule activity	18	2.87E-11	4.88E-09
GO:0003723~RNA binding	15	1.37E-07	2.33E-05

B

Molecular Function	Count	p-value	Adjusted p-value
Up regulated at 96hr			
GO:0016862~intramolecular oxidoreductase activity, intercon	4	8.45E-06	1.96E-03
GO:0016860~intramolecular oxidoreductase activity	5	1.91E-05	4.43E-03
GO:0016209~antioxidant activity	5	4.49E-05	1.04E-02
GO:0004364~glutathione transferase activity	4	7.79E-05	1.79E-02
Down regulated at 96hr			
GO:0003735~structural constituent of ribosome	30	6.07E-37	9.54E-35
GO:0005198~structural molecule activity	32	2.07E-22	3.25E-20
GO:0003723~RNA binding	32	8.13E-21	1.28E-18

Table 3.4.3 GO molecular function (MF) analysis through DAVID of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

A

Cellular Component	Count	p-value	Adjusted p-value
Up regulated at 48hr			
GO:0005829~cytosol	10	1.56E-05	0.001136
Down regulated at 48hr			
GO:0030529~ribonucleoprotein complex	21	1.78E-16	2.55E-14
GO:0005829~cytosol	26	1.74E-13	2.00E-11
GO:0022626~cytosolic ribosome	11	2.00E-13	2.30E-11
GO:0005840~ribosome	14	3.74E-13	4.31E-11
GO:0033279~ribosomal subunit	12	6.51E-13	7.49E-11
GO:0044445~cytosolic part	12	4.36E-12	5.01E-10
GO:0022627~cytosolic small ribosomal subunit	8	8.27E-11	9.51E-09
GO:0015935~small ribosomal subunit	8	2.31E-09	2.65E-07
GO:0043232~intracellular non-membrane-bounded organelle	28	1.27E-08	1.46E-06
GO:0043228~non-membrane-bounded organelle	28	1.27E-08	1.46E-06
GO:0042470~melanosome	5	2.87E-04	0.032434
GO:0048770~pigment granule	5	2.87E-04	0.032434

B

Cellular Component	Count	p-value	Adjusted p-value
Up regulated at 96hr			
GO:0042470~melanosome	8	1.64E-08	2.96E-06
GO:0048770~pigment granule	8	1.64E-08	2.96E-06
GO:0000267~cell fraction	16	1.18E-06	2.13E-04
GO:0005829~cytosol	17	2.99E-06	5.37E-04
GO:0031410~cytoplasmic vesicle	12	6.21E-06	0.001118
GO:0005625~soluble fraction	9	8.13E-06	0.001462
GO:0031982~vesicle	12	9.33E-06	0.001678
GO:0016023~cytoplasmic membrane-bounded vesicle	11	1.06E-05	0.001914
GO:0031988~membrane-bounded vesicle	11	1.41E-05	0.002534
GO:0005788~endoplasmic reticulum lumen	5	1.46E-04	0.025912
Down regulated at 96hr			
GO:0022626~cytosolic ribosome	26	9.56E-39	1.25E-36
GO:0030529~ribonucleoprotein complex	41	6.41E-38	8.40E-36
GO:0005840~ribosome	32	4.33E-37	5.68E-35
GO:0033279~ribosomal subunit	28	7.20E-37	9.43E-35
GO:0044445~cytosolic part	27	6.92E-33	9.07E-31
GO:0005829~cytosol	46	1.37E-27	1.80E-25
GO:0043232~intracellular non-membrane-bounded organelle	52	3.04E-21	3.98E-19
GO:0043228~non-membrane-bounded organelle	52	3.04E-21	3.98E-19
GO:0022627~cytosolic small ribosomal subunit	14	1.30E-20	1.71E-18
GO:0015935~small ribosomal subunit	15	1.65E-19	2.16E-17
GO:0015934~large ribosomal subunit	14	2.44E-17	3.19E-15
GO:0022625~cytosolic large ribosomal subunit	11	5.25E-15	6.84E-13

Table 3.4.4 GO cellular component (CC) analysis through DAVID of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

3.4.3 PANTHER analysis

The second online pathway tool used was PANTHER (Protein Analysis Through Evolutionary Relationships)(http://pantherdb.org/). Much like DAVID (**Section 3.4.2**) this tool also uses Gene Ontology (GO) terms that have been associated with genes and proteins for their functions and relationships to each other based on biological processes (BP), molecular function (MF) and cellular component (CC). PANTHER classifies function related to evolutionary family trees through statistical analysis before annotating each node through manual review into three types of grouping associated with UniProt for "subfamily membership" and GO terms for "protein class" and "gene function" (Thomas et al. 2003, Mi et al. 2005). These have been continually revised and defined since the conception of PANTHER over the last 17 years and is constantly revised and currently on version 10.0 since April 2015.

Several tools are available through PANTHER including an over-representation tool and enrichment analysis tool for gene list analysis with both analyses being processed differently. For our differential lists which are proteins that are positively or negatively enriched for with respect to a given comparison the overrepresentation analysis was used. This analysis compares a differential protein list to a reference gene set. As human gene names were chosen for the corresponding CHO protein IDs a reference list of 20,814 human genes were used to assess overrepresentation within each differential list.

The first analysis focused on searching BP using PANTHER on the up and down-regulated proteins 48 hr **(Table 3.4.5A)** and 96 hr **(Table 3.4.5B)** after transient over expression of miR-7 in CHO-K1-SEAP cells. The following key findings were observed:

- At 48 and 96 hr PANTHER shows up-regulated proteins from the CHO database lists to be associated with protein folding, protein metabolic process and metabolic process.

- At 48 and 96 hr PANTHER shows down-regulated proteins from the CHO database lists to be associated with translation, protein metabolic process, metabolic process, mRNA splicing, cellular component biogenesis and nuclear transport are down-regulated at 48 and 96 hr

Of note is that metabolic processes are both up and down-regulated suggesting a non linear relationship between the proteins associated with that process. This may point

toward a level of sub categorisation of processes that PANTHER is pointing to but unable to formally identify.

Enriched MF in the up and down-regulated proteins 48 hr **(Table 3.4.6A)** and 96 hr **(Table 3.4.6B)** after transient over expression of miR-7 in CHO-K1-SEAP cells shows a similar trend to BP with many of the same functions being affected at 48 hr as at 96 hr. The key observations were as follows:

- At 48 and 96 hr up-regulated proteins a linked to oxidoreductase activity, catalytic activity and translocase with glucosidase activity specifically associated with 48 hr and protein disulfide isomerase activity specifically associated with 96 hr.

- At 48 and 96 hr down-regulated proteins are linked to terms related to structural constituent of the ribosome, structural molecule activity, nucleic acid binding, RNA binding, translational initiation, elongation and translational regulator activity.

Using CC analysis it was possible to see what components the BP and MF were associated with. As with the other analyses CC enrichment was assessed with the up and down-regulated proteins 48 hr **(Table 3.4.7A)** and 96 hr **(Table 3.4.7B)** after transient over expression of miR-7 in CHO-K1-SEAP cells. The key observations:

- At 48 and 96 hr up-regulated proteins were not significantly overrepresented in any CC terms.

- At 48 and 96 hr down-regulated proteins were associated with the ribonucleoprotein complex, ribosome, macromolecular complex and cytosol with additional CC of the macromolecular complex, organelle, intracellular and cell part being overrepresented at 48 hr.

Taking BP, MF and CC obtained through PANTHER into account as a whole there is a clear effect on metabolic, ribosomal and nuclear processes specifically related to down regulation of translational functions with an up regulation of enzymatic functions. This activity is localised to a wider range of cellular components at 48 hr than at 96 hr. Overall BP and MF was down-regulated more than up-regulated in response to miR-7 over expression at 48 and 96 hr which was a trend also seen with DAVID analysis in the previous section.

A

Biological Process Up regulated at 48hr	Count	Adjusted p-value
Protein folding	7	6.17E-06
Protein metabolic process	16	6.69E-04
Metabolic process	26	1.49E-02
Monosaccharide metabolic process	4	1.55E-02
Down regulated at 48hr		
Translation	19	3.33E-16
Protein metabolic process	33	2.11E-12
Primary metabolic process	46	1.01E-09
Metabolic process	49	5.71E-09
mRNA splicing, via spliceosome	9	1.59E-04
Cellular component biogenesis	6	1.69E-04
Regulation of translation	6	4.45E-04
MRNA processing	9	6.55E-04
Nuclear transport	5	4.68E-03
Nucleobase-containing compound metabolic process	22	2.12E-02
Purine nucleobase metabolic process	4	4.13E-02

B

Biological Process Up regulated at 96hr	Count	Adjusted p-value
Protein folding	7	2.66E-04
Protein metabolic process	21	2.82E-03
Metabolic process	39	2.08E-02
Down regulated at 96hr		
Translation	35	1.36E-33
Protein metabolic process	52	1.81E-20
Primary metabolic process	75	8.36E-19
Metabolic process	79	4.20E-17
MRNA splicing, via spliceosome	11	1.14E-04
rRNA metabolic process	7	2.92E-04
Nucleobase-containing compound metabolic process	34	3.42E-04
mRNA processing	11	6.05E-04
Cellular component biogenesis	6	2.22E-03
Nuclear transport	6	3.09E-03
RNA splicing	8	4.89E-03
RNA splicing, via transesterification reactions	8	4.89E-03
Regulation of translation	6	5.69E-03
Organelle organization	8	1.46E-02
Cellular component organization or biogenesis	15	1.95E-02

Table 3.4.5 PANTHER biological process (BP) enrichment of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

A

Molecular Function Up regulated at 48hr	Count	Adjusted p-value
Oxidoreductase activity	9	3.67E-04
Catalytic activity	23	4.49E-03
Transketolase activity	2	7.03E-03
Glucosidase activity	2	4.02E-02
Down regulated at 48hr		
Structural constituent of ribosome	19	3.40E-20
Structural molecule activity	30	1.77E-16
Nucleic acid binding	37	1.42E-10
RNA binding	18	3.79E-10
Translation initiation factor activity	9	1.15E-08
Translation regulator activity	9	1.08E-07
Translation factor activity, nucleic acid binding	9	1.49E-07
Binding	44	2.97E-07
mRNA binding	8	1.56E-03
Translation elongation factor activity	4	2.51E-03
Poly(A) RNA binding	4	3.30E-02
RNA helicase activity	4	3.75E-02

B

Molecular Function Up regulated at 96hr	Count	Adjusted p-value
Oxidoreductase activity	13	3.46E-05
Protein disulfide isomerase activity	3	4.04E-03
Transketolase activity	2	2.30E-02
Catalytic activity	33	4.16E-02
Down regulated at 96hr		
Structural constituent of ribosome	39	8.93E-48
Structural molecule activity	48	1.15E-27
Nucleic acid binding	67	3.69E-25
Binding	75	1.66E-16
RNA binding	23	2.72E-11
Translation regulator activity	10	3.21E-07
Translation factor activity, nucleic acid binding	10	4.55E-07
Translation initiation factor activity	9	5.79E-07
Translation elongation factor activity	5	5.89E-04
MRNA binding	10	8.40E-04
Poly(A) RNA binding	6	9.81E-04
RNA helicase activity	5	1.66E-02

Table 3.4.6 PANTHER molecular function (MF) enrichment of up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

A

Cellular Component	Count	Adjusted p-value
Up regulated at 48hr		
NONE		
Down regulated at 48hr		
Ribonucleoprotein complex	13	3.97E-14
Macromolecular complex	17	1.38E-09
Ribosome	5	1.96E-08
Organelle	13	1.62E-03
Cytosol	3	3.73E-03
Cytoplasm	7	4.90E-03
Intracellular	14	9.80E-03
Cell part	14	2.81E-02

B

Cellular Component	Count	Adjusted p-value
Up regulated at 96hr		
NONE		
Down regulated at 48hr		
Ribonucleoprotein complex	15	2.32E-14
Ribosome	5	1.76E-07
Macromolecular complex	18	2.43E-07
Cytosol	3	1.34E-02

Table 3.4.7 PANTHER cellular component (CC) enrichment in up or down-regulated proteins from the CHO protein database 48 hr (A) and 96 hr (B) after transient over expression of miR-7 in CHO-K1-SEAP cells. The number of proteins from the submitted list associated with each GO term is represented by the count value. Enrichment was deemed significant with a Bonferroni adjusted p-value ≤0.05.

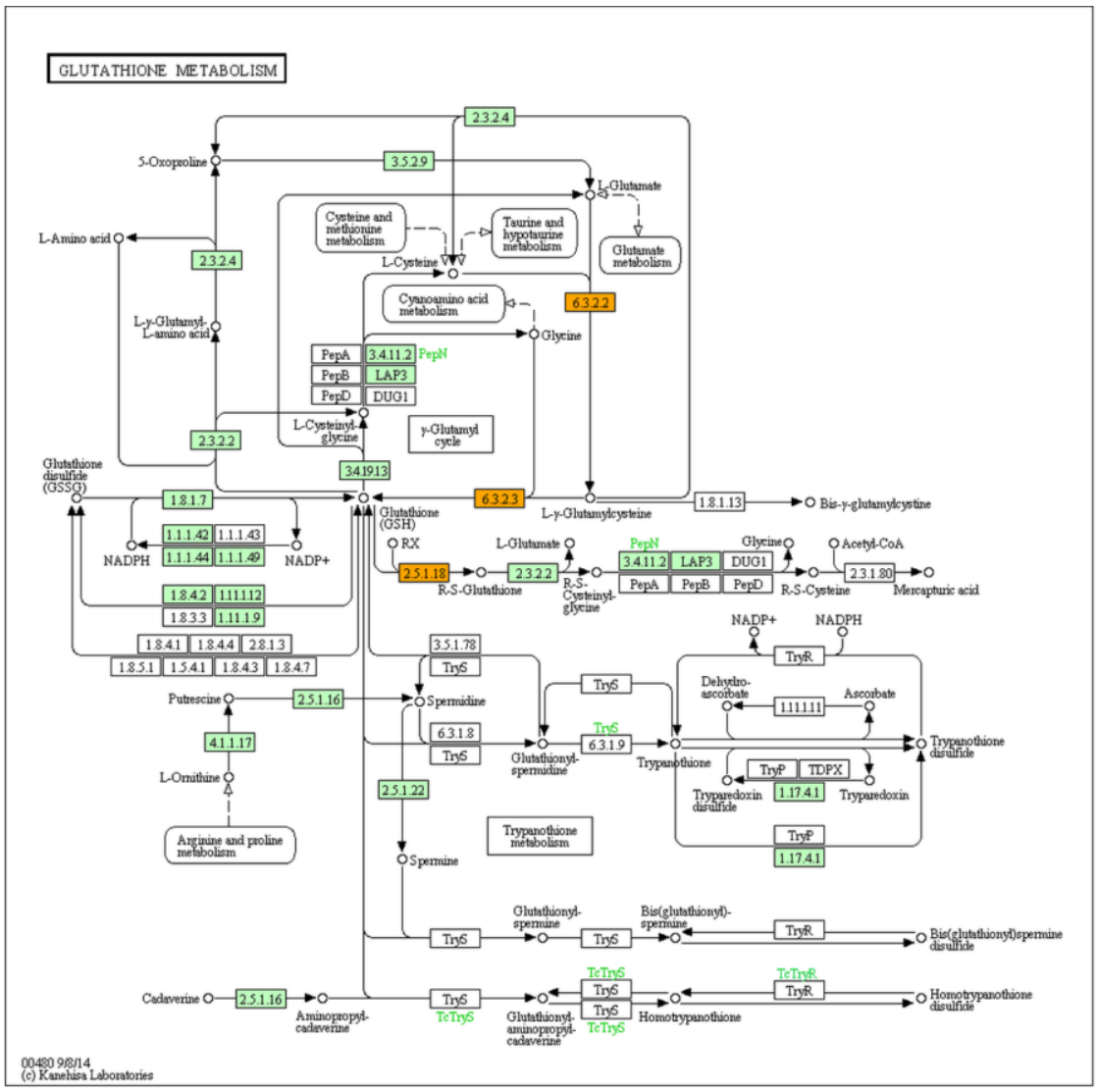
3.4.4 KEGG analysis

KEGG (Kyoto Encyclopaedia of Genes and Genomes)(http://www.genome.jp/kegg/) is a database of 472 pathways totalling 372,728 genes which are searched against a submitted list of identifiers. A number of search functions are available including pathological associations and ligands (Kanehisa et al. 2002). The basic pathway search was used for the purposes of this experiment. The lists were searched against the KEGG pathway human database through DAVID and then the corresponding significant pathway chart was obtained from KEGG.

Using KEGG with the differentially expressed proteins derived from the CHO database in response to miR-7 over expression the following was observed ;

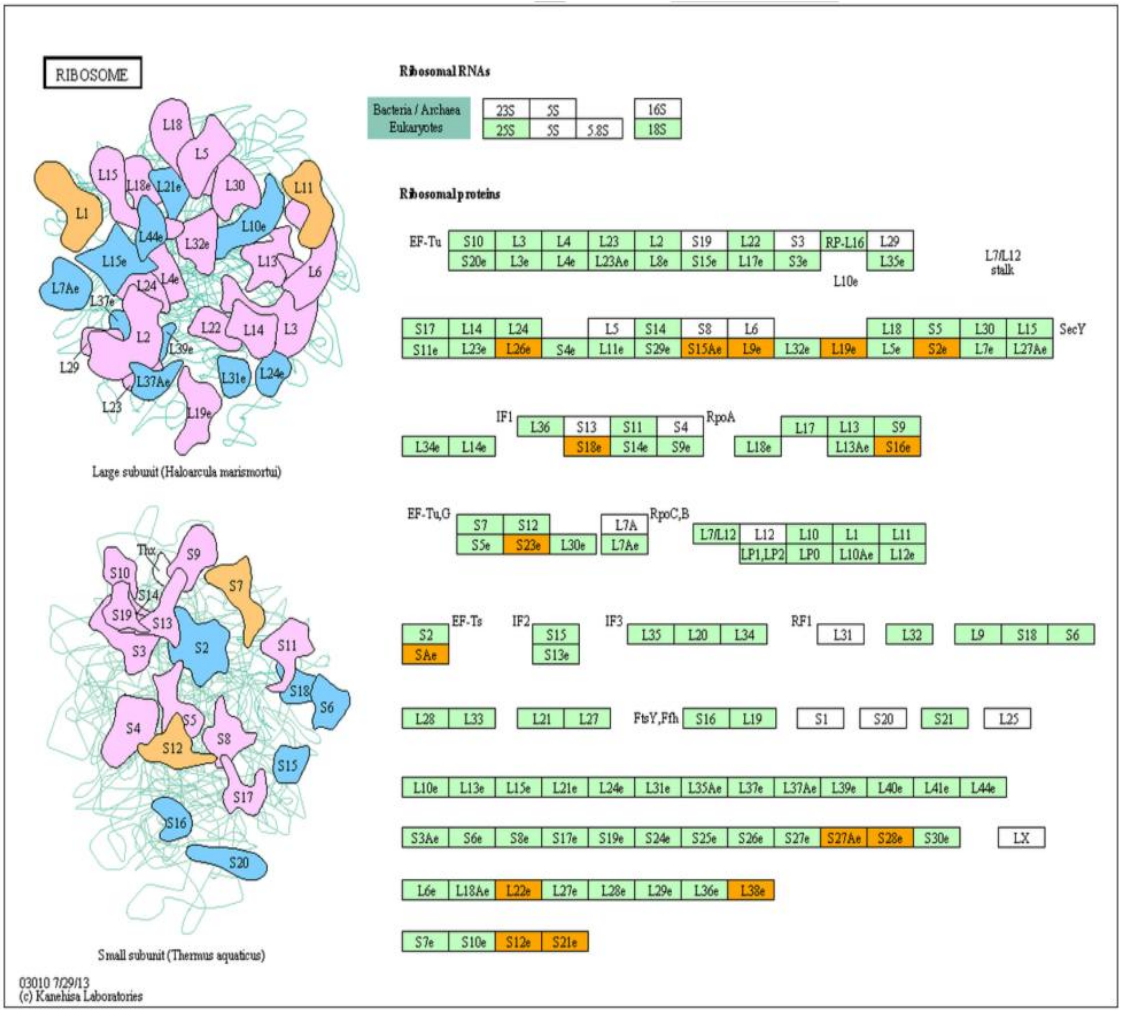
- At 48 and 96 hr up-regulated proteins were significantly associated with Glutathione metabolism (**Figure 3.4.2 and 3.4.4**). Specifically these proteins were Glutamate-cysteine ligase, modifier subunit (GCLM), Glutathione Synthase (GSS), Glutathione S-transferase mu (GSTM1), Glutathione S-transferase pi (GSTP1) at 48 hr. The same proteins were also unregulated at 96 hr in Glutathione metabolism with the exception of GSS and the addition of Isocitrate dehydrogenase 1 (NADP+), soluble (IDH1), Glutathione S-transferase alpha 3 (GSTA3). These proteins are indicated in the pathway as being involved in enzymatic reduction (RX) of glutathione which is important in detoxifying functions in the cell.

- At 48 and 96 hr down-regulated proteins were associated with the ribosome pathway (**Figure 4.4.3 and 4.4.5**). There are a large number of ribosomal proteins down-regulated in the ribosomal pathway, 18 at 48 and 33 at 96 hr respectively. Due to the highly co-operative nature of ribosomal proteins as well as some ribosomal proteins involved only in the formation of the ribosome itself it may be difficult to determine the role of this pathway. Generally however with so many proteins in this pathway down-regulated and the reduced growth phenotype observed in the miR-7 transfected cells it may indicate a large scale down regulation of protein synthesis reducing proliferation. This again would be consistent with the findings in DAVID and PANTHER related down regulation of proteins involved in ribosomal structural subunit formation, RNA synthesis, translation and down regulation related to proteins localised to the ribosome.



48hr UP			
Term	**Count**	**Genes**	**Adjusted p-value**
Glutathione metabolism	4	GSTM1, GSS, GCLM, GSTP1	3.51E-02

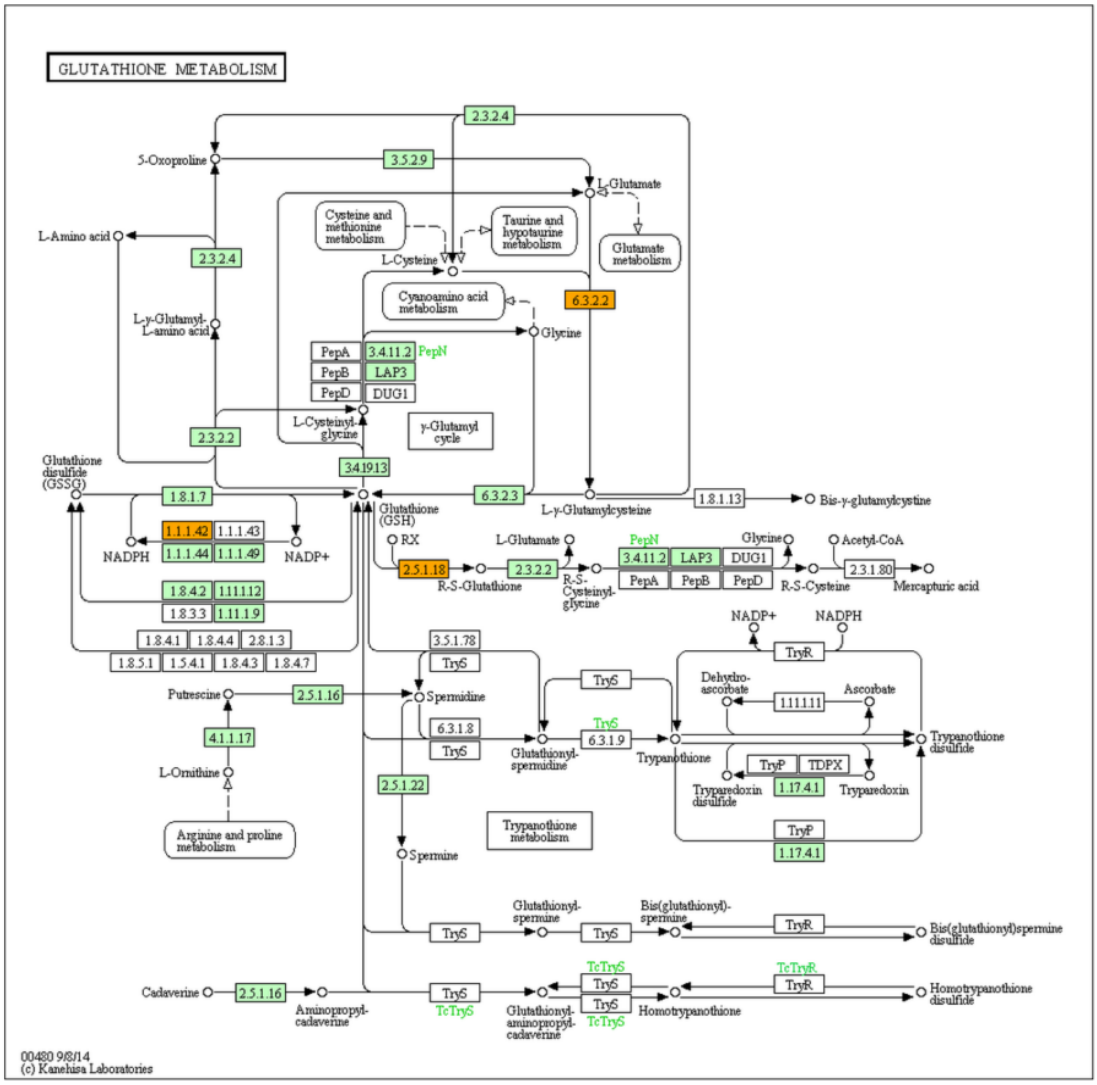
Figure 3.4.2 Diagram representing protein up-regulation in the Glutathione Metabolism pathway using DAVID and KEGG analysis. Four proteins (highlighted in orange above with two proteins represented by "2.5.1.18") were up-regulated in CHO cells at 48 hr after transfection with miR-7. The KEGG diagram denotes Glutamate-cysteine ligase, modifier subunit (GCLM) and Glutathione Synthase (GSS) as 6.3.2.2 and 6.3.2.3 respectively. Glutathione S-transferase mu (GSTM1) and Glutathione S-transferase pi (GSTP1) are denoted by 2.5.1.18 where a number of enzymes reduce (RX) glutathione in performing important detoxifying functions in the cell. Contrary to DAVID and PANTHER this analysis shows the exact locations of proteins in enriched pathways.



48hr DOWN

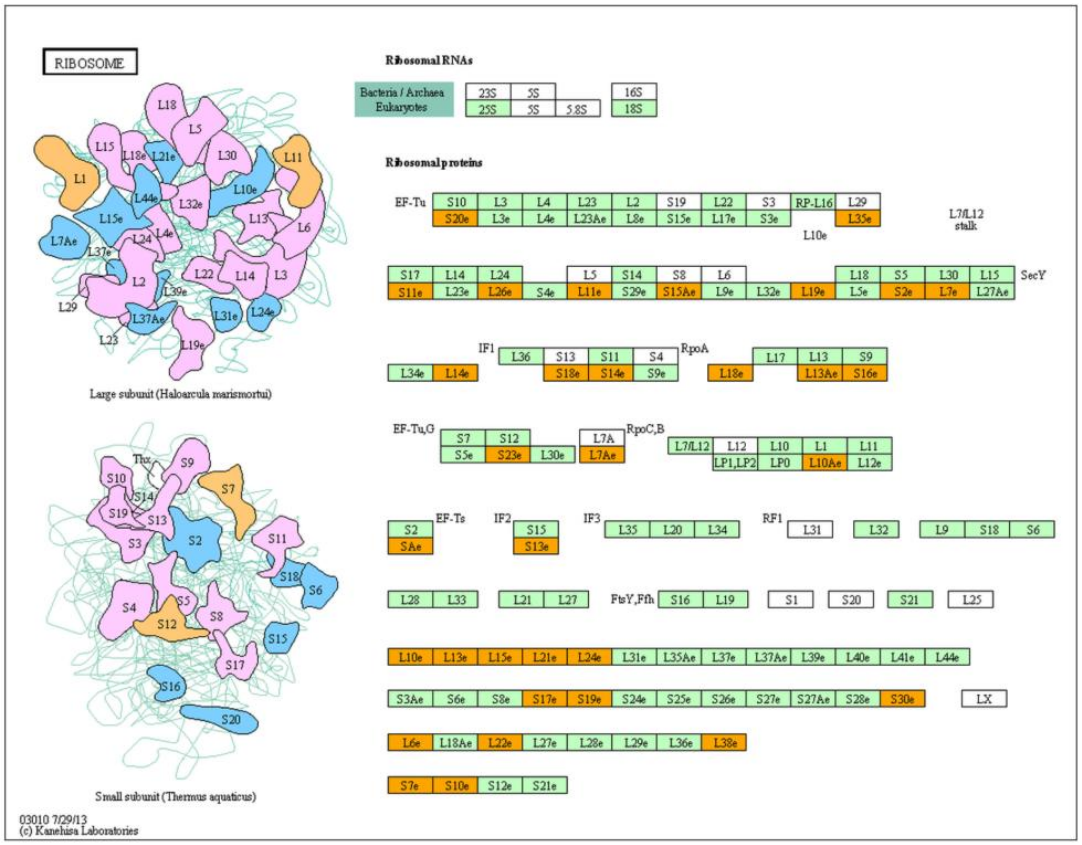
Term	Count	Genes	Adjusted p-value
Ribosome	18	RPSA, RPL19, RPL27A, RPL26, RPS15A, RPL38, RPS2, RPS18, HNRNPH2, RPS28, RPS16, RPL22, RPL9, RPS12, RPL26L1, RPS21, RPS27A, RPS23	8.10E-19

Figure 3.4.3 Diagram representing protein down-regulation in the ribosome pathway using DAVID and KEGG analysis. From the differential list of down-regulated proteins at 48 hr in miR-7 over expressing CHO cells 18 ribosomal members (highlighted in orange above, some contain more than one ID) were found to enrich for the ribosomal KEGG pathway (Bonferonni $p<0.05$). The above KEGG diagram represents the large ribosomal (50S) and small (30S) ribosomal subunit.



96hr UP			
Term	**Count**	**Genes**	**Adjusted p-value**
Glutathione metabolism	6	GSTM1, GSTA2, GSTA3, IDH1, GCLM, GSTP1	1.86E-03

Figure 3.4.4 Diagram representing protein up-regulation in the Glutathione Metabolism pathway using DAVID and KEGG analysis. Six proteins (highlighted in orange above with "2.5.1.18" representing 3 proteins) were up-regulated in CHO cells at 96 hr after transfection with miR-7. The KEGG diagram denotes Glutamate-cysteine ligase, modifier subunit (GCLM) as 6.3.2, Isocitrate dehydrogenase 1 (NADP+), soluble (IDH1) as 1.1.1.42 and the remaining four proteins Glutathione S-transferase alpha 3 (GSTA3), Glutathione S-transferase alpha 2 (GSTA2), Glutathione S-transferase, mu 1 (GSTM1) and Glutathione S-transferase P (GSTP1) are denoted by 2.5.1.



96hr DOWN

Term	Count	Genes	Adjusted p-value
Ribosome	33	RPL38, RPS2, RPL7, RPL6, RPL26L1, RPL10, FAU, RPL11, RPL7A, RPL10A, RPS20, RPS23, RPSA, RPL26, RPL24, RPS7, RPS18, HNRNPH2, RPS19, RPS16, RPL13A, RPL22, RPS17, RPS13, RPS10, RPS11	1.57E-41

Figure 3.4.5 Diagram representing protein down-regulation in the ribosome pathway using DAVID and KEGG analysis. All 33 members of the ribosomal members down-regulated in CHO cells 96 hr after miR-7 transfection are labelled (highlighted in orange above) in the KEGG diagram representing the large ribosomal (50S) and small (30S) ribosomal subunit.

3.5 Proteomic and microarray overlap of targets associated with miR-7

By comparing both proteomic and microarray differential expression data from miR-7 over expressing CHO cells additional potential targets of miR-7 can be found. The proteomic down-regulated differential data in **Table 3.2.2** was overlapped with down-regulated microarray data generated by Dr. Noelia Sanchez (Sanchez et al. 2013). Part of her work involved gene profiling of miR-7 transfected CHO cells.. The proteomic data profiles differential protein abundance at 48 and 96 hr while the differential transcript data from the microarray relates to a 72 hr time point after transfection with miR-7. In both the microarray and proteomic data differential expression was considered as being >1.2 fold increased or decreased when compared to a miRNA scramble control with significance of $p \leq 0.05$. Such a low fold change in protein levels was previously observed in our lab as being associated with miR29a regulation of invasion and proliferation (Muniyappa et al. 2009). The gene names from the proteomic data and the microarray data were then overlapped. It was found that 6 gene names overlapped and all of these had the same up or down differential regulation at the transcript and protein level (**Table 3.5.1**). Further to this catalase (CAT), a predicted direct target of miR-7, was found to be differentially regulated in the microarray and proteomic data.

		Fold change and signaicance (p=-value) across 3 datasets					
	Gene name	72hr Microarray		48hr Label Free		96hr Label Free	
Down in pre-miR7	PRPS1	2.15	(p=8.25E-04)	-		2.57	(p=3.0E-02)
	CAT	1.96	(p=2.07E-02)	4.07	(p=4.34E-03)	3.14	(p=8.43E-03)
	AHCY	1.61	(p=2.86E-02)	2.89	(p=3.53E-03)	-	
	ATIC	1.36	(p=1.61E-04	2.55	(p=4.52E-03)	1.84	(p=1.55E-03)
Up in pre-miR7	FTH1	1.67	(p=4.5E-03)	-		2.51	(p=4.5E-03)
	CLIC4	1.39	(p=1.43E-02)	-		5.46	(p=2.56E-03)

Table 3.5.1 Gene names associated with both transcriptional and proteomic differential regulation in response to miR-7 over expression compared to a miR Scramble control. The direction of differential regulation is the same for all overlapping targets suggesting direct association between transcriptional and translational regulation with these proteins. The presence of catalase (CAT) in the list also strengthens the overlap and the association of catalase with miR-7 as it was predicted as one of the direct targets of miR-7.

Among the differentially regulated transcripts in the microarray was Histone deacetylase 1 (HDAC1) which was down-regulated in miR-7 over expressing CHO cells compared to negative transfected control (**Figure 3.5.1A**). We had previously

observed through Western blot analysis that acetylation of Histones was reduced in miR-7 over expressing CHO cells (**Figure 3.5.1B**). As a result we decided to investigate if HDAC1 was down-regulated in miR-7 over expressing CHO cells. Our results show down-regulation of HDAC1 in miR-7 transiently over expressing CHO-K1-SEAP with Coomassie loading stain control (**Figure 3.5.1C**).

A

	Transcriptional downregulation in PM-7 transfected cells		
	Fold change	p-value	adj. p-value
HDAC1	-2.53	1.56E-07	3.08E-03

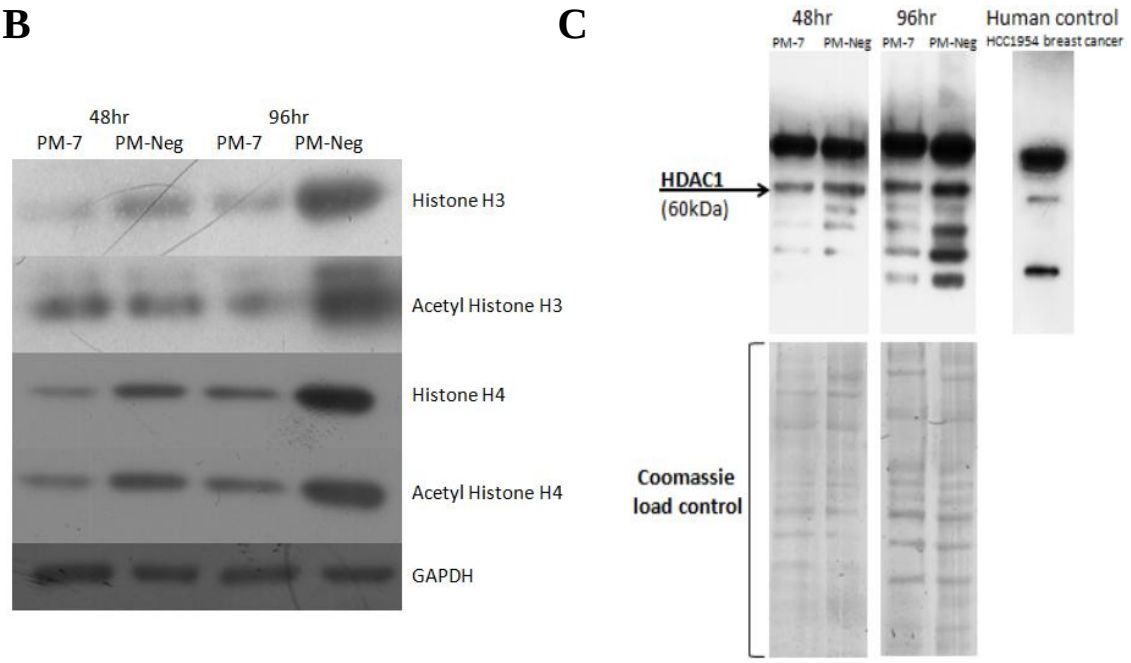


Figure 3.5.1 Transcriptional expression of HDAC1 (A) and Western blot analysis of Histone/Acetyl Histone (B) and of HDAC1 expression (C). Transcriptional down-regulation of Histone Deacetylase 1 (HDAC1) from microarray (A) showing a 2.53 fold decrease 72 hr after transient over expression of miR-7 together with Western blot for total Histone H3, Acetyl Histone H3, total Histone H4 and Acetyl Histone H4 (B) showing acetylation of Histone H3 is reduced at 96 hr and Histone H4 is reduced at 48 and 96 hr in premiR-7 transfected CHO cells. Western confirms down-regulation of HDAC1at 48 and 96 hr after miR-7 over expression (C).

From the Western blot analysis and acetylated histones it would be expected that HDAC1 would be increased in transfected cells. Therefore deacetylation of Histones H3 and H4 may be occurring without the interaction of HDAC1 or alternatively the histone acetylation is reduced in miR-7 over expressing cells and the reduction in acetylation is passive in that histones aren't being acetylated rather than increased deacetylation occurring.

Results summary

In order to determine the effect of miR-7 on the CHO cell proteome we conducted a quantitative label-free LC-MS/MS analysis on CHO-K1-SEAP cells 48 and 96 hr after transient over expression of miR-7 and compared them to a negative scramble miR control. Protein identifications were assigned using a species homology approach, searching peptide sequences against a human, mouse and rat database, as a CHO specific one was not available. Using miRWALK software we identified a number of potential miR-7 direct targets with the two strongest candidates, stathmin and catalase, being validated by Western blot. This work was published (Meleady et al. 2012b) and has since included further validative Western blot analysis and pathway analysis using DAVID, PANTHER and KEGG.

Following this we gained access to two CHO specific protein databases (See **Chapter 1**) and the analysis was conducted again. After the manual correction and annotation of the identifiers from the resulting differential protein lists was completed (combined with annotation from **Chapter 4**) a reference list of 1135 gene name annotated CHO protein identifications were generated. This allowed a larger pathway analysis investigation using 3 pathway analysis tools, DAVID, PANTHER and KEGG. The combination of all three confirmed that translational inhibition, anti-apoptotic activity and heat shock were all responsible for the miR-7 over expression phenotype. Novel findings suggested that glutathione metabolism was also involved in this phenotype.

Overlapping proteomic identifications with transcriptomic from a previous study in our lab (Sanchez et al. 2013) re-established catalase as a target of miR-7. It was also observed that Histone deacetylase 1 (HDAC1) was transcriptionally down-regulated in miR-7 over expressing CHO cells and simultaneously acetylation of Histones H3 and H4 was reduced at the protein level as seen by Western blot in miR-7 over-expressing CHO cells. HDAC1 was then confirmed to be down-regulated at the protein level by Western blot suggesting a more complex mechanism involving miR-7 and Acetylation in CHO cells.

CHAPTER 4

Effect of temperature shift on CHO-K1-SEAP cells using subcellular fractionation

4.1 Effect of temperature shift and deep proteome target discovery

Temperature shift is a technique used to increase the life span and productivity of cells in culture but also decreases cell proliferation. The technique involves lowering culture temperature from 37 ^{o}C to 31 ^{o}C and maintaining cells at this lower temperature. The molecular basis for this phenotype is still poorly understood in CHO cells but is of great interest for industrial applications. Previous work in our lab has shown differential micro RNA expression (Gammell et al. 2007) and differential protein expression linked to the temperature shift phenotype (Kumar et al. 2008). In this study a large scale global proteomic analysis was used to investigate the molecular mechanisms involved in temperature shift. CHO K1-SEAP cells were grown at 37 ^{o}C for 72 hr before one of two sets was transferred to 31 ^{o}C. Cells were cultured for an additional 8 and 24 hr before being pelleted and processed for analysis.

To obtain more detailed information on the processes involved in temperature shift each sample was fractionated into membrane, cytoplasmic and nuclear fractions. This reduces sample complexity and allows for more protein identifications. It also allows the analysis of various pathways associated with each fraction to be analysed and determine the localisation of activity. Additionally it was possible to identify proteins differentially regulated over time between the 8 and 24 hr time points as well as between 37 and 31 ^{o}C. Therefore a temperature shift comparison at 8 and 24 hr was possible and also a time course comparison of proteins differentially regulated between 8 and 24 hr at 37 and 31 ^{o}C was possible. These multiple comparisons allowed the identification of proteins that were differentially regulated in cells grown at 31 ^{o}C and also over time in cells grown at 31 ^{o}C that aren't differentially regulated over time at 37 ^{o}C.

Most of the proteins differentially regulated were found to be unique to each fractionation and so can be considered that of "deep proteome identifications" (see **Section 1.5.2**). Four of these proteins Ezrin, Moesin, Lamin A/C and Cyclon were chosen for siRNA knockdown in an attempt to induce a temperature shift-like phenotype. Of these Cyclon and Moesin were found to have a significant effect on cell proliferation as well as average cell perimeter and average cell area.

4.2 Fractionation validation

We used simple benchtop fractionation kits to enrich for proteins localised to the membrane, cytoplasm and nucleus in CHO-K1 SEAP cells. For membrane enrichment we used ProteoExtract® Native Membrane Protein Extraction Kit (Calbiochem) and for cytoplasmic and nuclear enriched fractions we used a Nuclear Extract Kit (Active Motif).

These kits were used in other fractionation studies in our laboratory using human cancer cell lines and were demonstrated to be user friendly, high throughput and result in reducing sample complexity. By reducing sample complexity for mass spectrometry analysis, more IDs as well as more sequence coverage for low abundant protein IDs could be achieved. As these kits had not been previously applied to CHO cells enrichment was validated by Western blot, protein identification overlapping and annotated enrichment analysis before a larger scale temperature shift experiment.

4.2.1 Western blot validation of enrichment

CHO-K1-SEAP cells were grown for 72 hr at 37 ^{o}C and an aliquot was taken for membrane enrichment and for nuclear/cytoplasmic enrichment. This resulted in a membrane, cytoplasmic and nuclear enriched sample for Western blot as well as the unfractionated "whole" sample. The four samples were then probed for proteins that are primarily associated with each subcellular fraction according to COMPARTMENTS, an open source tool which uses text mining and protein curation information for visualisation of protein localisation (Binder et al. 2014).

Expression of PDIA3, which is known to be associated with the endoplasmic reticulum and known to translocate to the nucleus (Wu et al. 2010b)(), was used to assess nuclear enrichment. HSP90 was used to assess cytoplasmic enrichment and is mostly associated with cytoplasm localisation (Langer, Rosmus and Fasold 2003). IGF1R is well established in the literature to promote cell-cell adhesion and cellular cross talk in cancer (Guvakova and Surmacz 1997, Hellawell et al. 2002) therefore IGF1Rβ was used to assess membrane enrichment. All three proteins showed a marked increase in abundance in their associated enriched fraction (**Figure 4.2.1**). PDIA3 and IGF1Rβ in particular showed higher abundance in nuclear and membrane fractions respectively compared to corresponding un-fractionated whole samples. This demonstrated that

enrichment had occurred across all three fractions compared to the un-fractionated sample. It also showed that each fraction was enriched for proteins that are heavily associated with the corresponding subcellular localisation of each enrichment (e.g. membrane fraction contains membrane proteins). Following this enriched samples were analysed using LC-MS/MS to assess the enrichment using pathway analysis on the qualitative IDs.

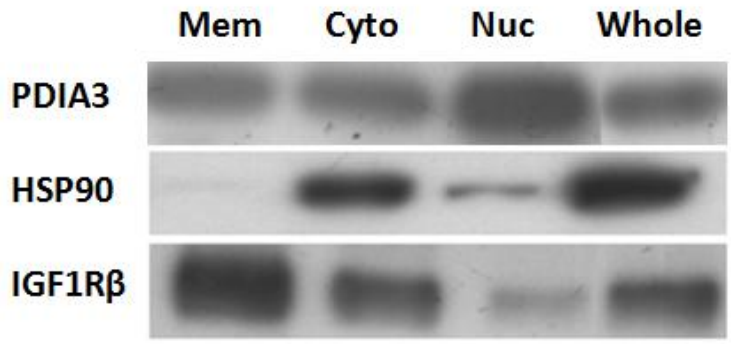


Figure 4.2.1 Validation of membrane, cytoplasmic and nuclear enrichment in CHO-K1 SEAP cells at 72 hr using commercially available benchtop fraction kits. The Western blot analysis shows that PDIA3, HSP90 and IGF1Rβ are enriched in nuclear, cytoplasmic and membrane fractions respectively. All three proteins are known to be localised in these corresponding fractions demonstrating that enrichment of the samples was achieved with these three proteins.

4.2.2 Mass spectrometry validation of enrichment

CHO-K1 SEAP cells were grown at 37 ^{o}C for 72 hr and processed for whole cell lysates and for membrane, cytoplasmic and nuclear enriched fractions. These samples were then submitted for LC-MS/MS analysis.

As well as generating the necessary quantitative data for differential expression information the LC-MS also generates qualitative identifications. For the purposes of this experiment these identifications were used to further investigate the extent of enrichment of each fraction. This was achieved by identifying the number of overlapping identifications between fractions and also assessing enrichment through pathway analysis tools.

While GI numbers are unique to each protein and overlap comparison is possible using them, they are not a usable format for pathway analysis. As mentioned in **Section 3.4.1** the manual annotation is extremely time-consuming and as such was not justified for this qualitative analysis which would require the annotation of 2197 unique GI taking into account overlaps between fractions (**Figure 4.2.2**) with 4546 including overlapping IDs. Of these 2197 unique GI numbered proteins 757 (33%) had gene names automatically assigned (**Table 4.2.1**). These proteins with gene name identifiers were used for GO analysis through DAVID to show the localisation of the proteins associated with each list (**Table 4.2.2**) and the top 10 significant terms ($p \leq 0.05$) were graphically represented in pie charts for each fraction (**Figure 4.2.3**).

4.2.2.1 Protein identification overlap

The qualitative IDs were searched against the BBCHO protein identification database. This resulted in 960 proteins identified from the un-fractionated lysate, 953 from the nuclear, 1295 from the cytoplasmic and from the 1335 membrane enrichment based on gene index numbers (GI) (**Table 4.2.1**). The number of proteins that were identified that contained gene names (No. Accession) were also counted as these would be used later for enrichment analysis which requires these identifiers.

As all the protein IDs had GI numbers these were then overlapped from each fraction to determine the number of unique IDs associated with each fraction. The overlap showed that there were 241 proteins that were unique to nuclear fraction, 309 proteins unique to the cytoplasmic fraction and 412 proteins unique to the membrane fraction (**Figure 4.2.2**). These proteins could not be identified without using the enrichment. Conversely there were 137 proteins uniquely associated with the un-fractionated protein lysate.

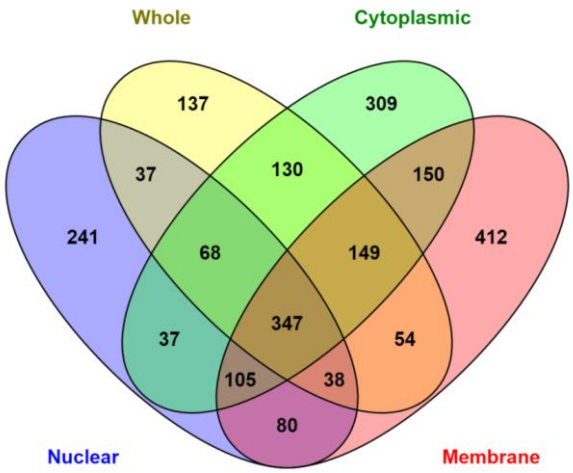


Figure 4.2.2 Venn diagram showing the overlap of protein IDs using gene index (GI) identifiers from the BBCHO database using CHO-K1-SEAP cells grown for 72 hr at 37 ^{o}C. The sum total of IDs between whole un-fractionated, nuclear, cytoplasmic and nuclear samples was 4546, while removing repeat IDs resulted in 2298 unique IDs which is evident from the high number of overlaps between fractions. Larger number of proteins unique to each enriched fraction (412, 309 and 241 in membrane, cytoplasmic and nuclear respectively) compared to those unique to the un-fractionated (whole) lysate (137) shows that the sample complexity of the un-fractionated sample has been reduced and enrichment has occurred. If there were no IDS unique to each fraction and all could be accounted for in the un-fractionated sample then enrichment would not have occurred.

4.2.2.2 Annotated enrichment analysis

The BBCHO database generated a total of 2298 unique protein IDs from combining the nuclear, cytoplasmic and membrane enriched fractions along with the un-fractionated lysate. These IDs from the BBCHO database are derived from transcript IDs which are not suitable for use with pathway analysis tools. As detailed in **Section 3.4** these IDs require manual annotation involving a BLAST search of the peptides for each ID, acquiring the exact protein ID and assigning a protein accession number or gene name for pathway analysis. With a total of 2298 unique identifications according to the assigned gene index (GI) number from transcriptomic information, this annotation would take unjustifiably long for a qualitative assessment of enrichment. Therefore the IDs that had gene names that could be automatically parsed from the transcript label were verified and included in the pathway analysis.

Of the unique pool of 2298 identifications 33% had correct verified gene names with 48% containing correct gene names in the pool when including repeat IDs (**Table 4.2.1**). The relative percentage for each fraction was different, however, with 42%, 37%, 51% and 46% of un-fractionated, nuclear, cytoplasmic and membrane IDs containing correct gene names respectively.

BBCHO Qualitative ID Summary			
Fraction	**No. GI**	**No. Accession**	**% Accession**
Whole	960	404	42.08%
Nuclear	953	356	37.36%
Cytoplasmic	1295	658	50.81%
Membrane	1335	613	45.92%
Total Pool	**4546**	**2197**	**48.33%**
Unique Pool	**2298**	**757**	**32.94%**

Table 4.2.1 Summary of the number of qualitative IDs obtained from the BBCHO database and the relative number that contained verified gene names (% accession) that were automatically parsed from the transcript identifier. The samples included the un-fractionated lysate, nuclear, cytoplasmic and membrane enriched fractions from CHO-K1 SEAP cells after 72 hr at 37 ^{o}C. All protein IDs contained a gene index (GI) with a smaller number indicated that had a clearly assigned gene name (Accession). Sum totals of IDs across the four fractions was calculated as "Total Pool". "Unique Pool" was the total after removing replicate IDs based on the GI number in the "Total Pool". The percentage values on the far right denote the number of GI numbers that were converted to gene names.

These gene names for each ID list were analysed with GO analysis through DAVID as in **Section 3.4.2** with a cut-off Bonferonni p<0.05 using the cellular component (CC) category to assess enrichment. We found 11 localisation terms are significantly enriched in the un-fractionated lysate (whole) compared to 28, 22 and 17 in the membrane, cytoplasmic and nuclear fractions respectively as well as bias for specific terms that would have been expected for each fraction (**Table 4.2.2**). This shows more CC are enriched and that sample complexity has been reduced. Visualising the top ten most significant of these terms it was found that each fraction has a unique enrichment profile which is unlike the un-fractionated sample, showing again the high degree of enrichment (**Figure 4.2.3**).

A

Whole

Cellular component	Count	*p*-value	Adjusted *p*-value
Ribonucleoprotein Complex	30	3.43E-16	7.63E-14
Ribosome	17	4.46E-11	1.02E-08
Intracellular Non-Membrane-Bounded Organelle	47	8.89E-10	2.04E-07
Non-Membrane-Bounded Organelle	47	8.89E-10	2.04E-07
Proteasome Complex	9	7.79E-08	1.78E-05
Cytosol	18	2.07E-05	4.73E-03
Cytosolic Large Ribosomal Subunit	4	3.01E-05	6.86E-03
Large Ribosomal Subunit	6	3.46E-05	7.88E-03
Melanosome	7	1.70E-04	3.82E-02
Pigment Granule	7	1.70E-04	3.82E-02
Cell Cortex	8	2.18E-04	4.87E-02

B

Membrane

Cellular Component	Count	*p*-value	Adjusted *p*-value
Mitochondrial Part	57	3.77E-26	1.19E-23
Mitochondrion	85	9.92E-24	3.12E-21
Mitochondrial Inner Membrane	40	1.64E-21	5.17E-19
Organelle Inner Membrane	40	1.12E-20	3.52E-18
Mitochondrial Envelope	43	1.05E-19	3.31E-17
Mitochondrial Membrane	41	5.64E-19	1.78E-16
Ribonucleoprotein Complex	45	1.38E-18	4.33E-16
Organelle Envelope	46	9.07E-17	3.50E-14
Envelope	46	1.04E-16	3.50E-14
Organelle Membrane	54	4.28E-15	1.36E-12
Mitochondrial Matrix	23	9.97E-13	3.14E-10
Mitochondrial Lumen	23	9.97E-13	3.14E-10
Membrane-Enclosed Lumen	59	3.04E-11	9.59E-09
Organelle Lumen	56	2.43E-10	7.67E-08
Intracellular Organelle Lumen	55	6.58E-10	2.07E-07
Ribosome	18	2.50E-07	7.88E-05
Spliceosome	14	9.98E-07	3.14E-04
Ribosomal Subunit	10	5.49E-06	1.73E-03
Melanosome	11	6.43E-06	2.02E-03
Pigment Granule	11	6.43E-06	2.02E-03
Endoplasmic Reticulum Part	17	1.39E-05	4.38E-03
Proteasome Complex	9	1.69E-05	5.31E-03
Proton-Transporting Atp Synthase Complex	6	2.68E-05	8.40E-03
Respiratory Chain	9	3.95E-05	1.24E-02
Endoplasmic Reticulum	35	4.76E-05	1.49E-02
Endoplasmic Reticulum Lumen	9	8.32E-05	2.59E-02
Large Ribosomal Subunit	7	1.06E-04	3.27E-02
Proton-Transporting Two-Sector Atpase Comple	7	1.62E-04	4.96E-02

C

Cytoplasm

Cellular Component	Count	*p*-value	Adjusted *p*-value
Cytosol	42	4.88E-16	1.29E-13
Ribonucleoprotein Complex	36	5.64E-14	1.64E-11
Proteasome Complex	15	5.20E-13	1.51E-10
Mitochondrion	60	1.10E-12	3.20E-10
Organelle Envelope	32	1.99E-09	5.78E-07
Envelope	32	2.18E-09	6.32E-07
Organelle Inner Membrane	24	2.81E-09	8.16E-07
Mitochondrial Inner Membrane	23	5.47E-09	1.59E-06
Ribosome	18	2.53E-08	7.33E-06
Mitochondrial Part	29	6.07E-08	1.76E-05
Mitochondrial Membrane	24	6.21E-08	1.80E-05
Melanosome	12	1.74E-07	5.05E-05
Pigment Granule	12	1.74E-07	5.05E-05
Mitochondrial Envelope	24	1.88E-07	5.45E-05
Organelle Membrane	34	1.94E-06	5.62E-04
Cytosolic Part	9	1.01E-05	2.91E-03
Proteasome Core Complex	6	1.65E-05	4.76E-03
Vesicle	23	6.60E-05	1.90E-02
Intracellular Non-Membrane-Bounded Organelle	55	6.97E-05	2.00E-02
Non-Membrane-Bounded Organelle	55	6.97E-05	2.00E-02
Ribosomal Subunit	8	1.18E-04	3.37E-02
Cytoplasmic Vesicle	22	1.38E-04	3.93E-02

D

Nuclear

Cellular Component	Count	*p*-value	Adjusted *p*-value
Ribonucleoprotein Complex	33	3.26E-22	7.27E-20
Intracellular Non-Membrane-Bounded Organelle	50	7.79E-16	1.73E-13
Non-Membrane-Bounded Organelle	50	7.79E-16	1.73E-13
Ribosome	17	1.61E-12	3.58E-10
Spliceosome	11	4.28E-08	9.54E-06
Ribosomal Subunit	8	7.81E-07	1.74E-04
Membrane-Enclosed Lumen	26	2.63E-06	5.85E-04
Cell Cortex	9	6.27E-06	1.40E-03
Organelle Lumen	24	1.68E-05	3.73E-03
Cytoskeleton	23	4.36E-05	9.69E-03
Intracellular Organelle Lumen	23	5.06E-05	1.12E-02
Cytoskeletal Part	18	9.67E-05	2.13E-02
Mitochondrial Inner Membrane	11	1.06E-04	2.33E-02
Mitochondrial Membrane	12	1.40E-04	3.07E-02
Nuclear Lumen	19	1.51E-04	3.32E-02
Organelle Inner Membrane	11	1.63E-04	3.56E-02
Actin Cytoskeleton	9	2.03E-04	4.42E-02

Table 4.2.2 Significant (Bonferroni adjusted p≤0.05) Cellular Component GO terms associated with each fraction. The unfractionated lysate (A), membrane (B), cytoplasmic (C) and nuclear (D) fractions all show bias for terms that would be expected to be associated with each fraction such as mitochondrial, envelope and organelle membrane for the membrane fraction, cytosol, mitochondrion and ribosome for the cytoplasmic fraction and ribonucleoprotein and spliceosome for the nuclear fraction as well as a larger number of enriched terms in the enriched samples.

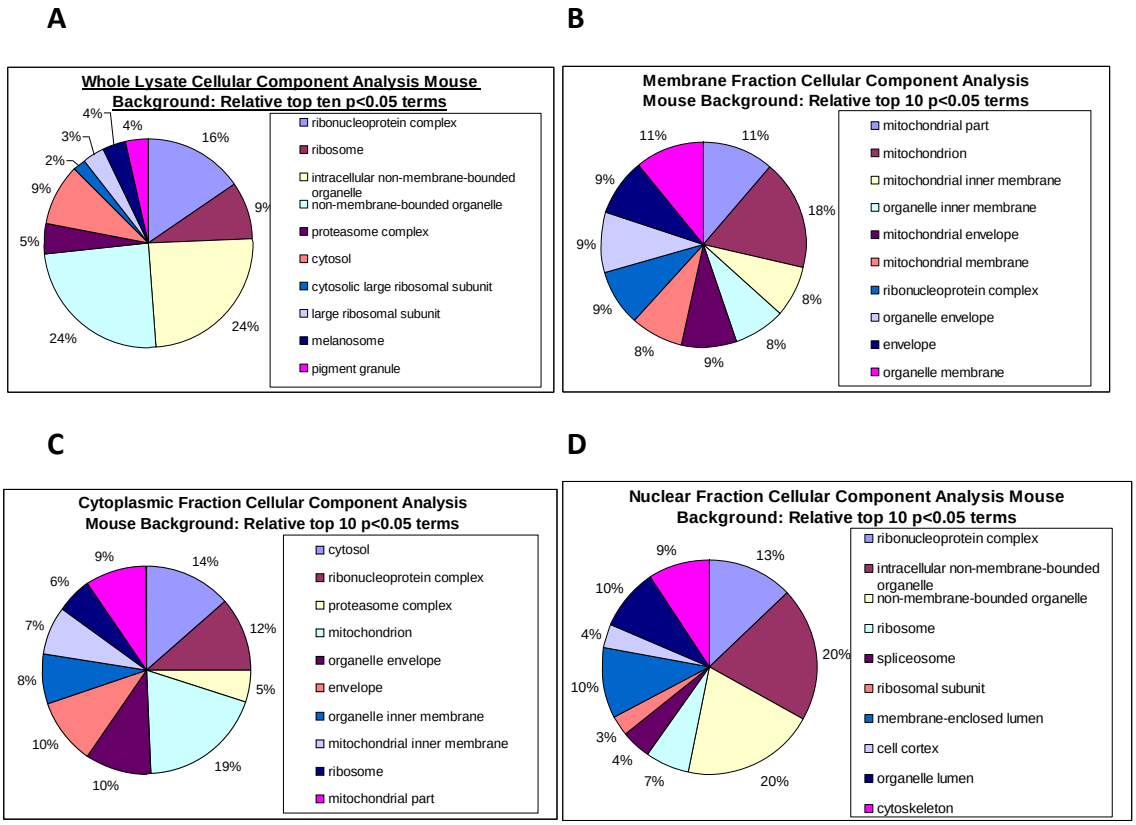


Figure 4.2.2 Representation of the top ten significant (Bonferroni adjusted p≤0.05) Cellular Component GO terms associated with each fraction from Table 4.2.3. This shows that each enriched samples (B, C, D) show a bias for specific sub-cellular components compared to the unfractionated lysate (A) and confirms enrichment in the fractionated samples. Specifically, it shows more clearly the differing profiles of each fraction and the unfractionated sample compared to each other.

4.3 Quantitative label-free LC-MS/MS analysis

Having confirmed enrichment through Western blot analysis (**Figure 4.2.1**) and pathway analysis which indicated a bias towards the expected cellular components in each enriched fraction (**Table 4.2.2** and **Figure 4.2.3**) a quantitative label free experiment was conducted on CHO-K1-SEAP cells subjected to temperature shift. Combined with enrichment, the label free analysis would allow for a detailed insight into the subcellular processes involved in temperature shift and identify potential functional targets to induce a similar phenotype.

Cells were grown at 37 ^{o}C for 72 hr before half of the culture flasks were transferred to 31 ^{o}C for temperature shift. This resulted in 2 identical sets of flasks grown at 37 ^{o}C and 31 ^{o}C. Samples were taken at 8 and 24 hr time point after temperature shift which were subsequently fractionated with three biological replicates of the experiment conducted. Two large analyses were possible by comparing each temperature **at** a given time point (**Section 4.4.2**) and also comparing differentially expressed proteins **over** time (**Section 4.4.4**). The resulting comparisons led to multiple lists of up and down-regulated proteins at different time points and for each fraction with different numbers of proteins differentially regulated with each list (**Table 4.3.1**). These were the 24 final lists (**Figure 4.3.1**) but from using the NCBI and BBCHO database there were 48 lists for annotation before these were overlapped (**Section 3.4.1**) and used for subsequent pathway analysis (**Section 4.3.3 and 4.3.5).**

Temp	Fraction	↑ at 8hr	↑ at 24hr	↓ at 8hr	↓ at 24hr	↑ over 16hr	↓ over 16hr
31 ^{o}C	Membrane	22	49	36	12	82	55
	Cytoplasm	11	46	16	66	80	130
	Nuclear	6	16	68	69	52	27
37 ^{o}C	Membrane					59	65
	Cytoplasm					78	98
	Nuclear					57	62

Table 4.3.1 Number of proteins differentially expressed at 8 and 24 hr after 31 ^{o}C temperature shift and over 16hr at 31 ^{o}C and 37 ^{o}C in each of the three enriched fractions membrane, cytoplasm and nuclear. Up-regulated proteins are denoted by "↑" and down-regulated proteins by "↓"(**Appendix i., ii., ix., x.** for full tables).

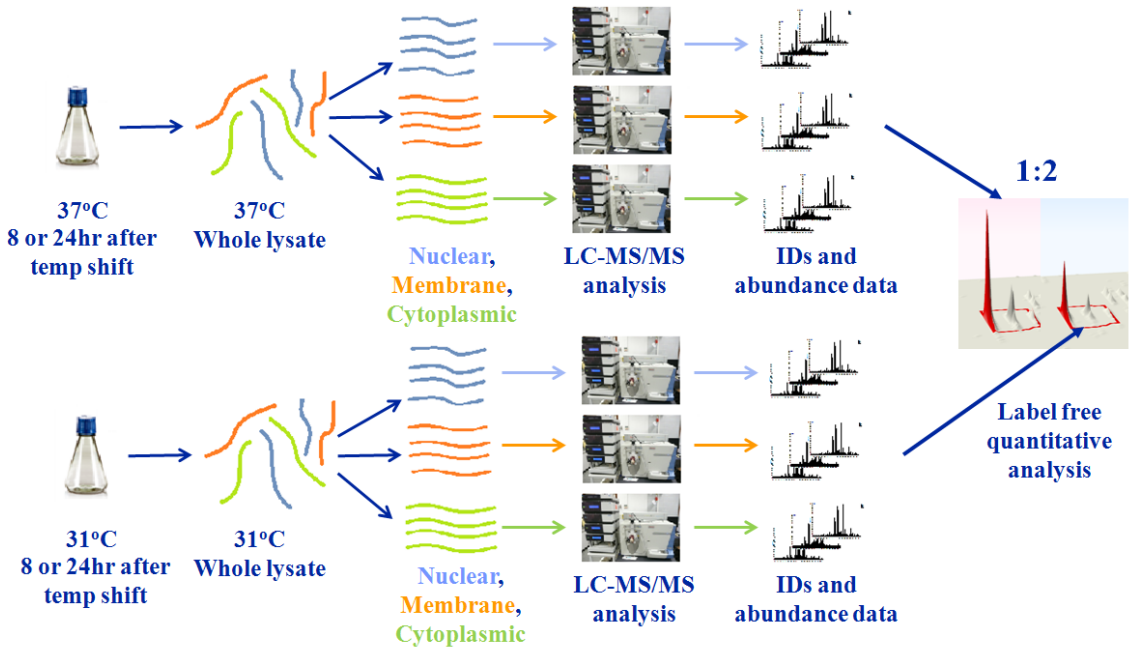


Figure 4.3.1 Overview of the experimental workflow of temperature shift study with subcellular fractionation. With the above number of variables (temperature, time and fractions) there are 24 differential comparisons possible; up-regulation in 31 ^{o}C temperature shift at 8 hr in each fraction (3 lists), down-regulation in 31 ^{o}C temperature shift at 8 hr in each fraction (3 lists), up-regulation in 31 ^{o}C temperature shift at 8 hr in each fraction (3 lists), down-regulation in 31 ^{o}C temperature shift at 8 hr in each fraction (3 lists), up regulation between 8 and 24 hr at 37 ^{o}C in three fractions (3 lists), down regulation between 8 and 24 hr at 31 ^{o}C in three fractions (3 lists), down regulation between 8 and 24 hr at 31 ^{o}C in three fractions (3 lists).

4.3.1 Database comparison BBCHO vs NCBI

Two databases were used in order to obtain the CHO peptide identifications. BBCHO database is an in house database derived from CHO transcript data in collaboration with Professor Nicole Borth of BOKU University, Vienna. The NCBI database is derived from non-redundant Chinese hamster identifications submitted to NCBI. Both databases were searched using the same criteria and the output was overlapped to assess any large difference in the output between the two. Specifically the time course analysis of the fractionation label free experiment was used as the larger differential lists allowed for an assessment of a larger number of protein IDs compared to the smaller differential lists from the miR-7 study in **Chapter 3**. It also allowed for a comparison between the multiple sample preparations of fraction, experimental variables of different growth temperatures and the differential identifications over time in culture.

The comparisons show that most of the differential identifications between the databases overlap with slightly more identifications unique to the BBCHO than the NCBI database (**Figure 4.3.2**). As differential lists are in and of themselves small the relatively small number of identifications that were unique to each database justifies this output merging method for more thorough searching using our CHO specific databases.

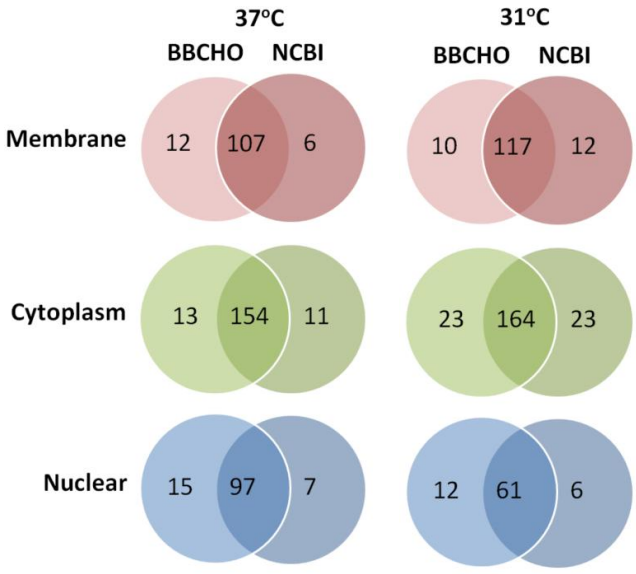


Figure 4.3.2 Venn diagram showing the overlap in CHO IDs between the up and down-regulated IDs of the BBCHO and non redundant NCBI databases using the 16hr time course of CHO-K1-SEAP cells grown at 37 ^{o}C and 31 ^{o}C. The large numbers of proteins and multiple fractions show the clear overlap between the two with a slight bias for unique identifications in the BBCHO database.

4.3.2 Differentially expressed/abundant proteins at 31 ^{o}C temperature shift

The first comparison within the quantitative label-free experiment was between 37 ^{o}C and 31 ^{o}C at 8 and 24 hr temperature shift. As previously mentioned in **Section 3.3** the 1 peptide hit differential identifications were included. Proteins were considered differentially expressed with a fold change ≥1.2 fold and a significance of $p \leq 0.05$ between experimental groups. In total there were 12 individual lists. These were split between the 8 and 24 hr time point (6 each) with up-regulated at 31 ^{o}C and down-regulated at 31 ^{o}C lists for each of the three fractions membrane, cytoplasmic and nuclear (**Table 4.3.1**).

To better interpret the numerous lists the total number of identifications (**Figure 4.3.2**) and the overlap between fractions (**Figure 4.3.3** and **4.3.4**) were graphed. The observed trend is that of a steady increase in protein up regulation in each fraction from 8 to 24 hr with an overall larger number of proteins down-regulated in response to temperature shift. The response in down regulation is less linear in nature with a sharp down regulation of proteins associated with the nuclear enriched fraction observed at 8 hr which is reduced at 24 hr. At 24 hr the majority of down-regulated proteins shifts to the cytoplasmic enriched fraction. This shows that the initial response to temperature shift (8 hr) could be associated with the down regulation of nuclear proteins which may be a precursor for cytoplasmic protein down regulation with a steady increase in proteins associated with each fraction over time. Conversely there cytoplasmic proteins are up regulated which may indicate the reorganisation of cytoplasmic proteins by the later time point (24 hr)

The overlap between the differential proteins in each fraction was compared to further assess the enrichment and to determine if the trend observed was represented by unique IDs in each fraction. Comparing the fractions in up and down-regulated proteins (**Figure 4.3.3 and 4.3.4**) respectively there is a distinct separation between fractions with very little overlap. This further strengthens that each fraction is enriched for specific proteins.

The trend observed can be said to represent the response of the CHO cells to temperature shift relative to the enriched fractions as there is very little overlap between each fraction. The number of proteins changing in each fraction is then strongly associated with each enrichment.

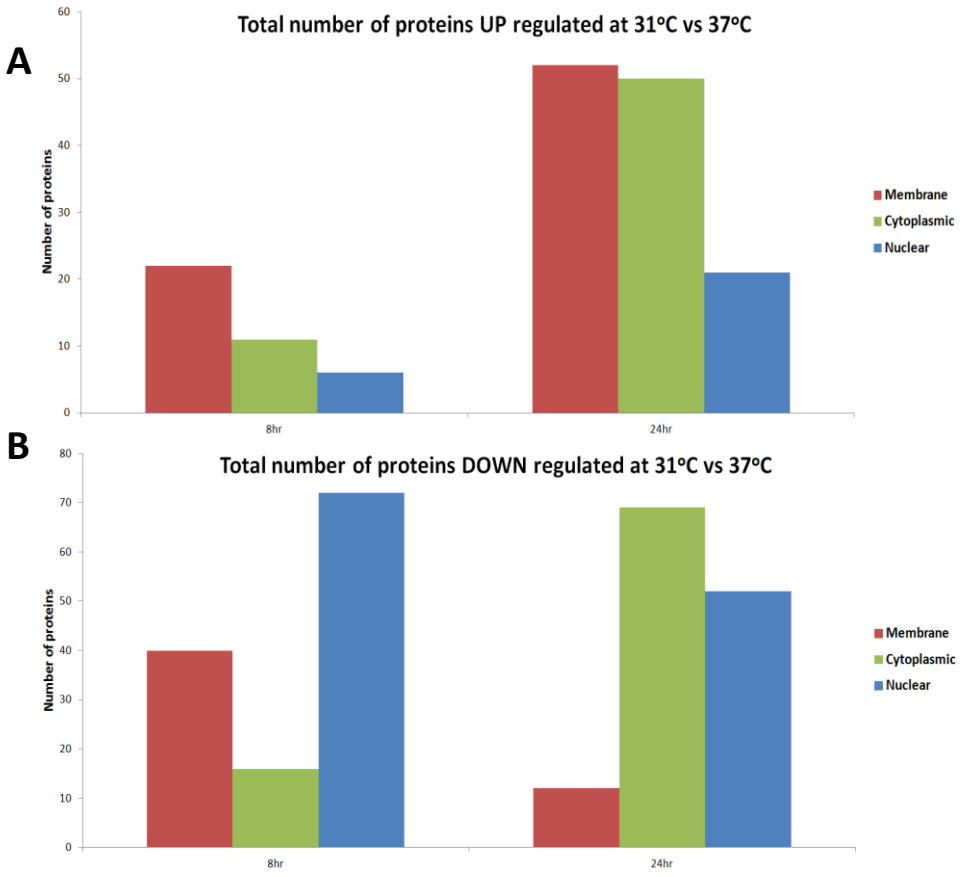


Figure 4.3.3 Up-regulated proteins (A) and down-regulated proteins (B) in CHO-K1 SEAP cells at 31 °C after 8 and 24 hr in membrane, cytoplasmic and nuclear fractions. Up-regulated proteins can be seen to have an increase in all fractions between 8 and 24 hr while down regulation shows a larger number in the membrane and nuclear fraction at 8 hr compared to 24 hr. The large degree of differential regulation (up and down) within the cytoplasm may indicate a shift a re-organising of the cytosolic proteome in response to temperature shift e.g. Heatshock proteins which are localised in the cytoplasm.

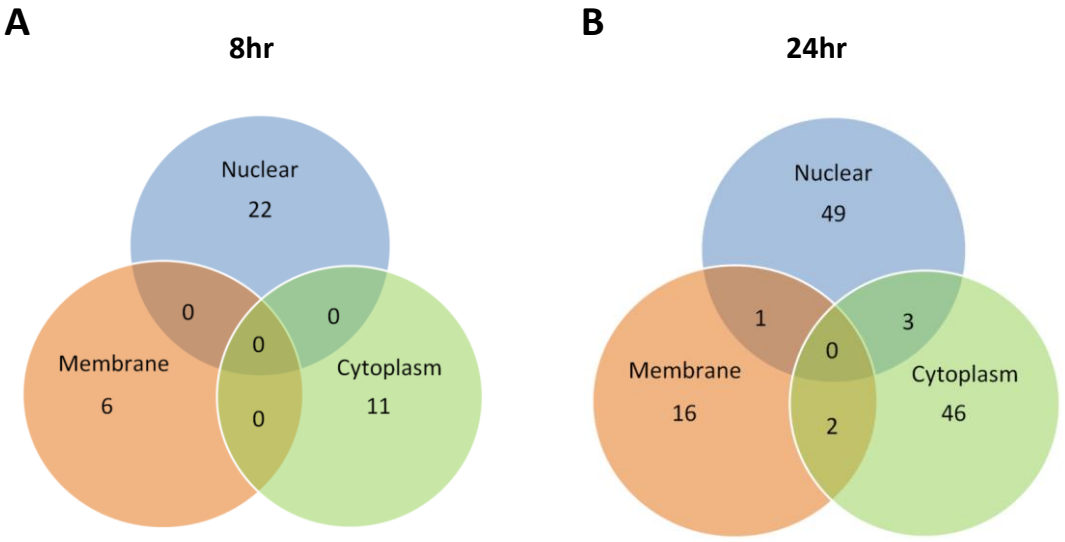


Figure 4.3.4 Overlapping up-regulated proteins in CHO-K1 SEAP cells at 31 °C after 8 (A) and 24 hr (B) between membrane, cytoplasmic and nuclear fractions. This shows that distinct proteins are differentially expressed in each enriched fraction, further indicating that as well as the qualitative distribution the differential distribution also shows the strong specificity of the enrichment.

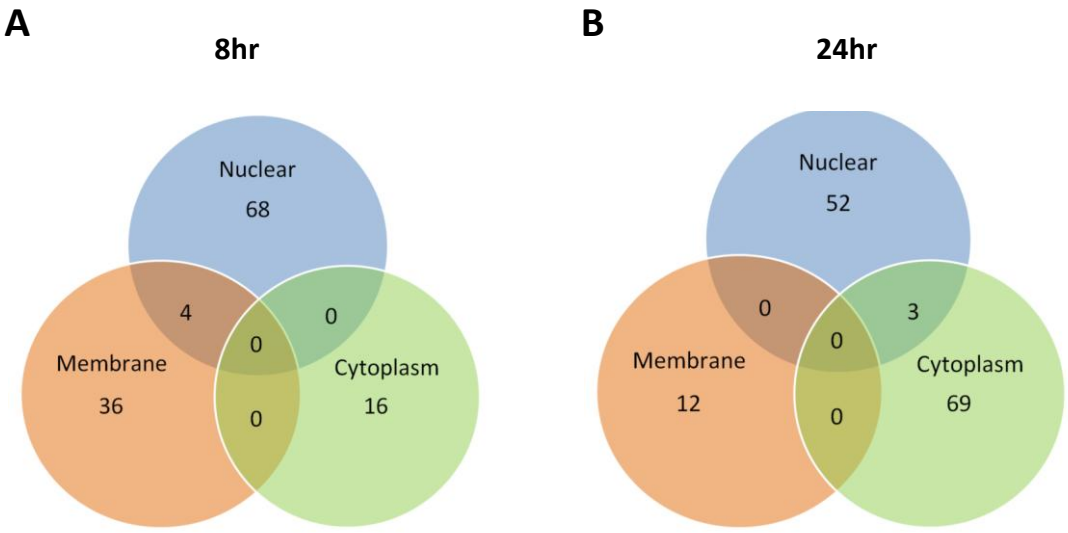


Figure 4.3.5 Overlapping down-regulated proteins in CHO-K1 SEAP cells at 31 °C after 8 (A) and 24 hr (B) between membrane, cytoplasmic and nuclear fractions. This shows that, as well as the qualitative distribution, the differential distribution also displays the strong specificity of the enrichment.

4.3.3 Pathway analysis of temperature shift

In order to determine significant enrichment of biological processes, molecular function, cellular components associated with the differential lists GO analysis through DAVID, PANTHER analysis and KEGG analysis were used as in **Chapter 3**. Significant enrichment was considered for adjusted Bonferroni $p \leq 0.05$. From the temperature shift experiment combined with membrane, cytoplasmic and nuclear enrichment there were multiple comparisons possible and resulted in multiple differentially expressed lists of proteins associated with temperature shift (time point analysis) and associated with the growth of CHO cells over time (time course analysis) at 31 and 37 ^{o}C. In the following three sections (**Section 4.3.3.1 - 4.3.3.3**) will deal with the time point analysis, differentially regulated proteins 8 and 24 hr after temperature shift. The top five terms that are deemed significant are included in the following results (see **Appendix iii.-viii.**) for shortened tables).

4.3.3.1 DAVID temperature shift

As in **Chapter 3** pathway analysis the three categorisations in DAVID for the GO analysis of Biological Process (BP), Molecular Function (MF) and Cellular Component (CC) were used. Taking each of these categories into account we can see what functions of the CHO cells are being affected 8 and 24 hr after temperature shift in the 31 ^{o}C cells. The significant (adjusted Bonferroni $p \leq 0.05$) key findings using DAVID are shown in **Table 4.3.2** (more expanded table found in **Appendix iii.-v.)**.

Taking BP, MF and CC into account together the down regulation of proteins is associated with protein localisation, protein transport and translation processes with these proteins functionally associated with the ribosome, RNA binding and structural activity. These down-regulated proteins are largely associated with the cytosol and ribosome components. While up-regulated proteins were significantly associated with specific CC they were not significantly associated with any BP or MF. This suggest that metabolic activity and differential translation is occurring but also that structural activities are also being significantly affected. The fractionation also indicates that these changes start at the 8 hr time point in the nuclear enriched sample and by 24 hr there are proteins associated with various components (CC) down-regulated in the membrane fraction. The fractionation therefore allows for some insight into localisation events related to temperature shift.

Enrichment category	↑ at 8 hr	↑ at 24 hr	↓ at 8 hr	↓ at 24 hr
Biological Process	-	-	Protein localization, protein transport (**Mem**), translation and translational elongation (**Nuc**)	Translation and translational elongation (**Nuc**)
Molecular Function	-	-	RNA binding, ribosomal constituents, actin binding and cytoskeletal binding (**Nuc**)	Heat shock protein binding (**Cyto**) , RNA binding and structural activity (**Nuc**)
Cellular Component	-	Organelle envelope, organelle inner membrane, mitochondrial envelope and mitochondrial envelope (**Mem**)	Organelle membrane, endoplasmic reticulum part (**Mem**), cytosolic part, non membrane bound cell organelles, and ribosomal subunit (**Nuc**)	Endoplasmic reticulum (**Cyto**), intracellular non-membrane bound organelles, cytosolic ribosome and ribonucleoprotein complex (**Nuc**)

Table 4.3.2 Summary of significantly (adjusted Bonferroni $p \leq 0.05$) enriched GO biological process, molecular function and cellular component terms using DAVID associated with proteins up-regulated (↑) and down-regulated (↓) 8 and 24 hr after 31 ^{o}C temperature shift. Also labelled is the enriched fraction (**Mem, Cyto, Nuc**) for which the terms are associated with. These terms mostly are associated with translational, metabolic and structural events. Additionally this differential expression can be inferred from the enriched fractions with most significant change occurring in the nuclear fraction at 8 hr and resulting in significant changes in the membrane by 24 hr.

4.3.3.2 PANTHER temperature shift

Using over representation analysis with PANTHER it was possible to compare the differentially regulated proteins in each fraction against the reference human database. Each differential list from each of the three fractions membrane, cytoplasmic and nuclear at both 8 hr and 24 hr after temperature shift were analysed for enriched biological process (BP), molecular function (MF) and cellular component (CC).

The overall PANTHER output is consistent with that of GO through DAVID in **Section 4.3.3.1** showing a small number of BP and MF linked to translational and structural protein down regulation and an even larger emphasis on cytoskeletal elements, translational regulation and metabolic activity at 24 hr post temperature shift (**Table 4.3.3**). The same trend is also observed with enriched terms in each fraction. At 8 hr there is a large emphasis on protein down regulation in the nucleus related to translation and cytoskeletal elements. At 24 hr a larger number of enriched terms were found overall with significant up regulation of specific terms now taking place in the membrane and cytoplasmic fraction. Interestingly many of the same terms enriched in the nucleus at 8 hr are also enriched at 24 hr indicating that the cells are still being driven by the same process of metabolic, translational and cytoskeletal regulation from the nuclear enriched fraction. It is also worth noting that these up regulated terms at 8 hr are not enriched for using DAVID potentially showing the merit of using different pathway analysis tools to identify significant cellular functions (see **Appendix vi.-viii.** for expanded tables).

Enrichment category	↑ at 8 hr	↑ at 24 hr	↓ at 8 hr	↓ at 24 hr
Biological Process	Protein metabolic process (**Mem**)	RNA splicing, DNA replication (**Mem**) , metabolic processing (**Cyto**) and the tricarboxylic acid cycle (**Nuc**)	Translation and cellular component organisation (**Nuc**)	Protein metabolic process, mRNA processing (**Cyto**), translation and cellular component organisation (**Nuc**)
Molecular Function	RNA helicase activity (**Nuc**)	RNA, mRNA and DNA binding, structural constituent of the ribosome (**Mem**), helicase and translational initiation factor (**Cyto**) oxidoreductase activity (**Nuc**)	RNA binding, actin binding, cytoskeletal protein binding and ribosomal constituents (**Nuc**)	Isomerase activity (**Cyto**), Structural molecule activity, cytoskeletal binding and structural ribosomal constituent (**Nuc**)
Cellular Component	-	Macromolecular complex and intracellular component (**Mem**)	Mitochondrial (**Mem**) and cytoskeletal (**Nuc**)	Membrane (**Mem**) and cytoskeleton, ribonucleoprotein complex (**Nuc**)

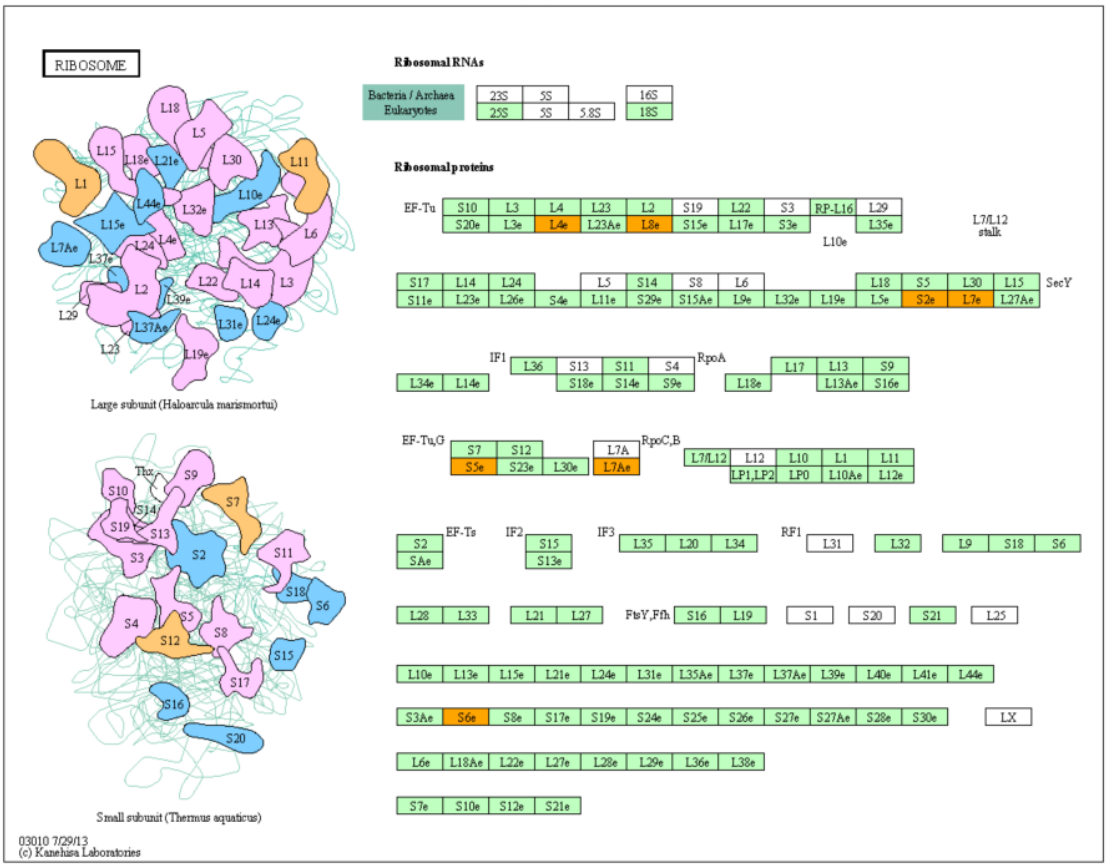
Table 4.3.3 Summary of significantly (adjusted Bonferroni p≤0.05) enriched PANTHER biological process, molecular function and cellular component terms using PANTHER associated with proteins up-regulated (↑) and down-regulated (↓) 8 and 24 hr after 31 ^{o}C temperature shift. Also labelled is the enriched fraction (**Mem, Cyto, Nuc**) for which the terms are associated with. These terms mostly are associated with translational, metabolic and structural events. Further to this differential expression can be inferred from the enriched fractions with consistent enrichment of terms in the nuclear fraction at 8 and 24 hr resulting in cross fraction changes at 24 hr.

4.3.3.3 KEGG temperature shift

Enrichment of specific pathways was assessed using KEGG. Unlike the KEGG analysis in **Chapter 3** where unfractionated samples were used, here it was possible to assess in membrane, cytoplasmic and nuclear enriched fractions. Significance of an adjusted Bonferroni $p \leq 0.05$ was chosen and showed one pathway to be affected by temperature shift at 8 hr and 2 affected at 24 hr.

At 8 hr there were 7 ribosomal proteins significantly ($p=1.64x10^{-4}$) associated with the ribosomal pathway (**Figure 4.3.6**) . These were linked with down-regulated proteins in the nuclear enriched fraction implying that ribosomal elements were being differentially regulated in response to temperature shift. At 24 hr 5 proteins were significantly ($p=7.72 \times 10^{-4}$) associated with DNA replication (**Figure 4.3.7**) in the up-regulated proteins of the cytoplasmic enriched fraction and again 8 ribosomal proteins significantly ($p=1.02 \times 10^{-6}$) associated with the ribosomal pathway (**Figure 4.3.8**) in the down-regulated proteins of the nuclear enriched fraction.

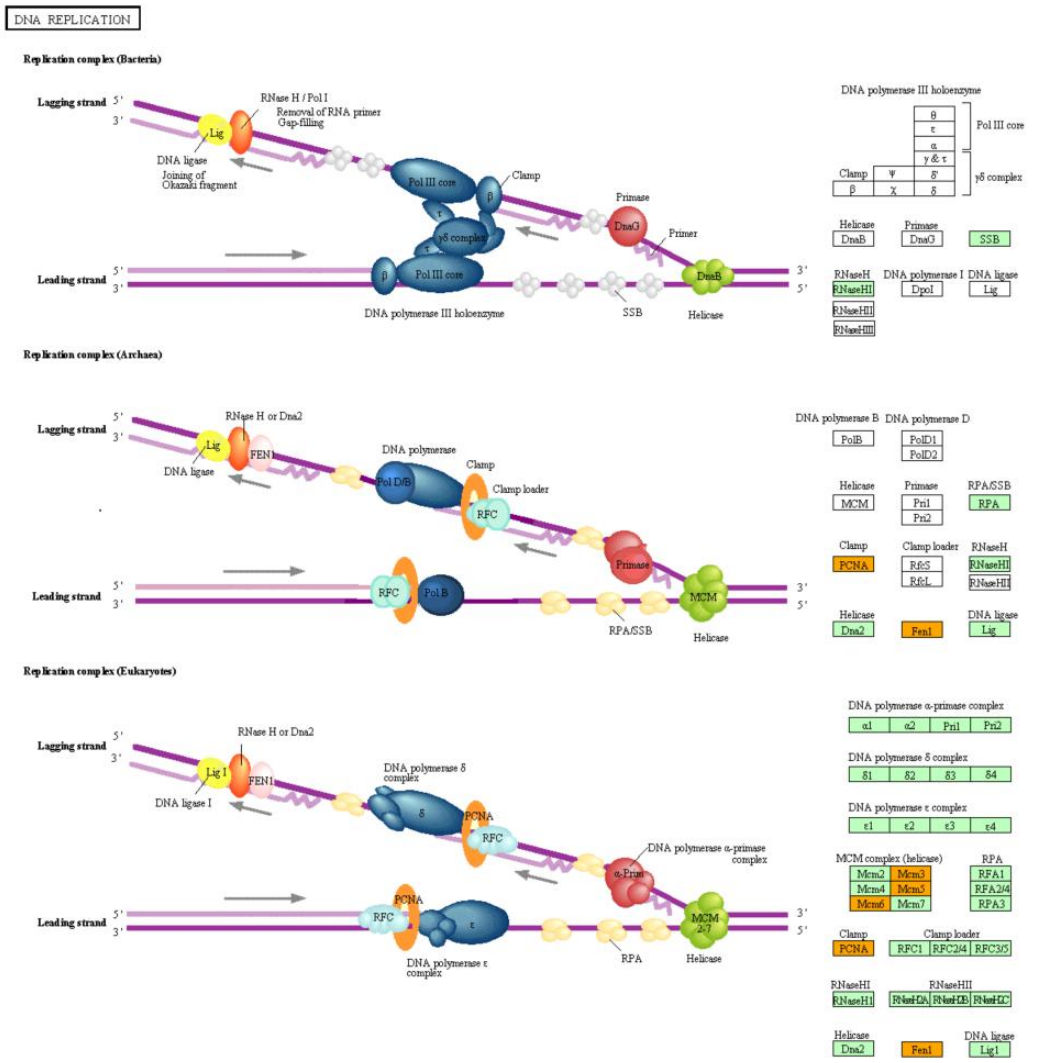
Overall, when the reduced proliferative phenotype at 31 ^{o}C is taken into consideration, the same ribosomal proteins differentially regulated at each time point may imply an overall down regulation of ribosomal activity or specific ribosomal activities rather than different ribosomal activities being altered at each time point. Looking at specific up-regulated proteins in the DNA replication pathway such as PCNA, FEN1 and MCM proteins there is a strong implication that DNA replication initiation is surprisingly increased which may be in response to proteins required by other BP and MF seen in pathway analysis in **Section 4.3.3.1 and 4.3.3.2.**



8hr Nuclear DOWN

Term	Count	Genes	Adjusted p-value
Ribosome	7	RPL7, RPL8, RPL7A, RPL4, RPS2, RPS6, RPS5	1.64E-04

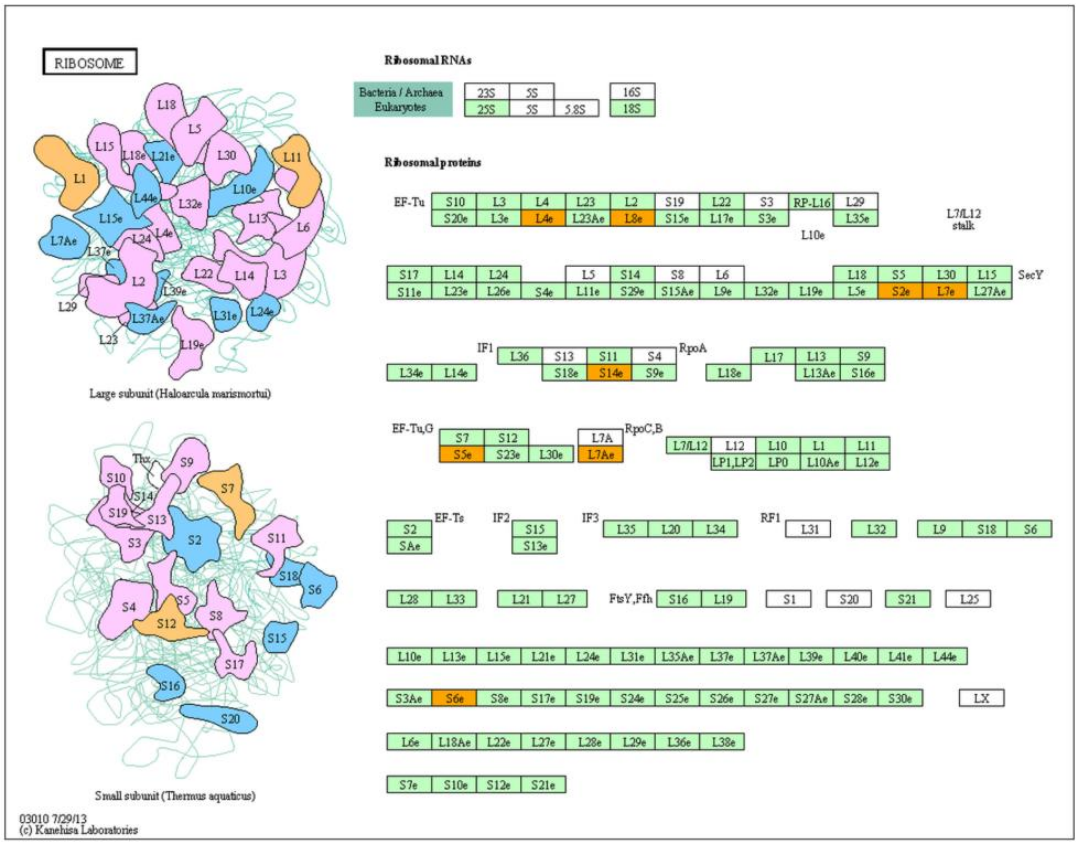
Figure 4.3.6 KEGG pathway diagram representing nuclear enriched protein down-regulation in the ribosome pathway 8 hr after temperature shift. Seven of the ribosomal proteins (highlighted in orange above) found in the nuclear protein enriched fraction that were down-regulated at 31 ^{o}C in CHO cells 8 hr after temperature shift are labelled in the KEGG diagram representing the large ribosomal (50S) and small (30S) ribosomal subunit. Similar proteins are seen at 24 hr in the same pathway indicating that the same processes within the ribosome or related to the ribosome are affected at both time points.



24hr Cytoplasmic UP

Term	Count	Genes	Adjusted p-value
DNA replication	5	PCNA, MCM3, FEN1, MCM5, MCM6	7.72E-04

Figure 4.3.7 KEGG pathway diagram representing cytoplasm enriched protein up-regulation in the DNA replication pathway 24 hr after temperature shift. Five proteins (7 highlighted in orange above with enrichment of PCNA and FEN1 in Archaea replication complex and all 5 enriched in Eukaryote replication complex) were found to be significantly associated with DNA replication from the list of proteins which were found significantly up implying that DNA replication initiation maybe increased to compensate for processes and functions highlighted by the pathway analysis.



24hr Nuclear DOWN

Term	Count	Genes	Adjusted p-value
Ribosome	8	HNRNPH2, RPL7, RPL8, RPL7A, RPL4, RPS2, RPS6, RPS5	1.02E-06

Figure 4.3.8 KEGG pathway diagram representing nuclear enriched protein down-regulation in the ribosome pathway 24 hr after temperature shift. Eight of the ribosomal proteins (highlighted in orange above) found in the nuclear protein enriched fraction that were down-regulated at 31 ^{o}C in CHO cells 8 hr after temperature shift are labelled in the KEGG diagram representing the large ribosomal (50S) and small (30S) ribosomal subunit. Similar proteins are seen at 8 hr in the same pathway indicating that the same processes within the ribosome or related to the ribosome are affected at both time points.

4.3.4 Differentially expressed/abundant proteins over time at 37 and 31 ^{o}C

As well as comparing temperature at each time point the proteins differentially regulated over time were also determined by comparing the 8 hr time point to the 24 hr time point at each temperature. Time point analysis, while powerful for determining differential expression, are often limited in scope. Analysing change over time can show a more dynamic profile of change in the cell. By looking at the cellular functions changed over a 16 hr time period time (8 hr vs 24 hr) it may be possible to observe changes to the cells over time at 37 and 31 ^{o}C instead of at a single moment in time. As with the previous CHO database identification lists 1 peptide hit differential identifications were included. Proteins were considered differentially expressed with a fold change ≥1.2 fold and a significance of $p \leq 0.05$.

In total there were 12 individual lists summarised in the table below. These were split between the two temperatures 37 ^{o}C and 31 ^{o}C (6 each) with down-regulated at 24 hr and up-regulated at 24 hr lists for each of the three fractions membrane, cytoplasmic and nuclear (**Table 4.3.4**) . The aim of this comparison was to identify proteins that potential didn't yield a significant result with 31 ^{o}C vs 37 ^{o}C comparison but were significantly differentially regulated over time from 8 to 24 hr at 37 or 31 ^{o}C only. Such a protein could be said to be associated with the proliferation related to a specific temperature over time as opposed to at a specific time point as with the 31 ^{o}C vs 37 ^{o}C comparison.

Temperature	Fraction	↑ over 16hr	↓over 16hr
31 ^{o}C	**Membrane**	82	55
	Cytoplasm	80	130
	Nuclear	52	27
37 ^{o}C	**Membrane**	59	65
	Cytoplasm	78	98
	Nuclear	57	62

Table 4.3.4 Number of proteins differentially expressed at 8 and 24 hr after 31 ^{o}C temperature shift and in each of the three enriched fractions membrane, cytoplasm and nuclear. Up-regulated proteins are denoted by "↑" and down-regulated proteins by "↓" (See **Appendix ix.** and **x.** for expanded tables).

To assess the number of identifications in the lists the total numbers of proteins differentially regulated over time between 8 and 24 hr (16 hr) were graphed for 31 and 37 ^{o}C (**Figure 4.3.9**) and shows that there are a greater number of membrane enriched fraction proteins up-regulated over time in the 31 ^{o}C and there are also a greater number of proteins associated with the cytoplasmic enriched fraction and a reduced number of proteins associated with the nuclear fraction down-regulated at 31 ^{o}C. These profiles show that there is a clear difference in the localisation of protein differential regulation occurring over time between the two temperatures. By assessing if the trend was represented by unique IDs per fraction (**Figure 4.3.10 and 4.3.11**) it was found that there is very little overlap between fractions and that the trend observed in each fraction is representative of a trend associated with localisation.

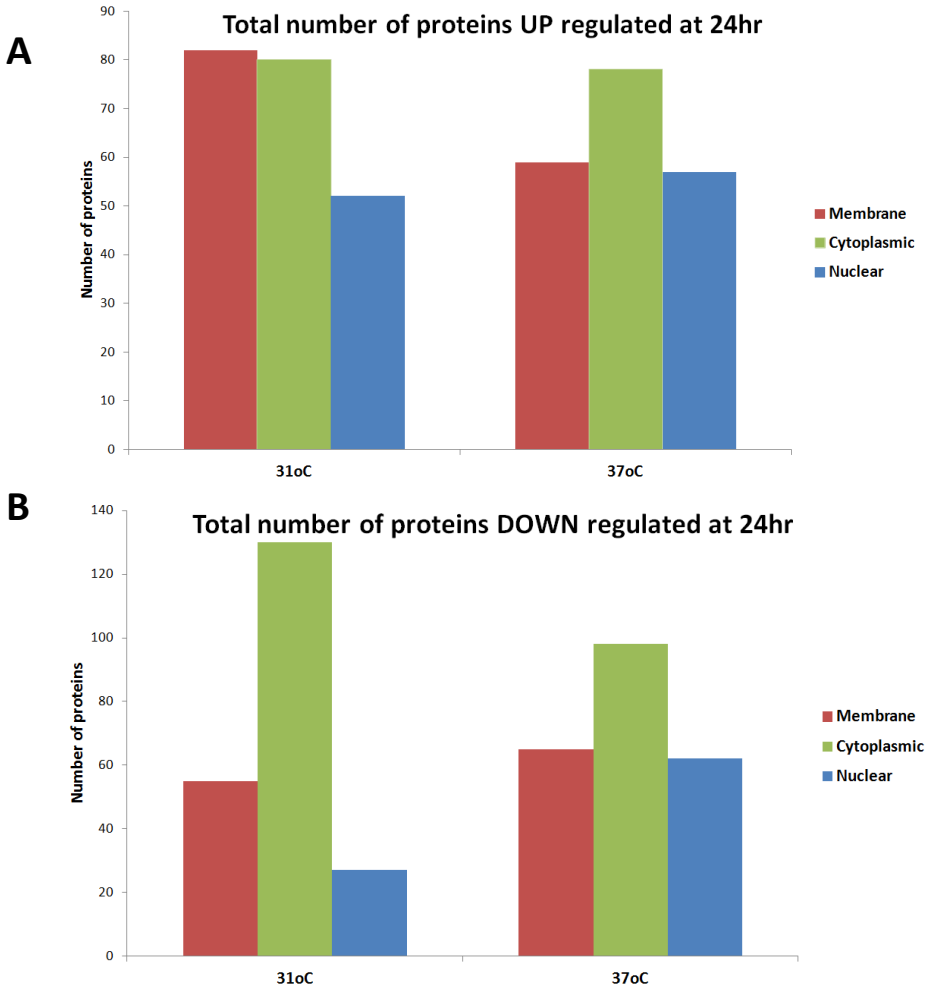


Figure 4.3.9 Up (A) and down (B) regulated proteins in CHO-K1 SEAP cells at 24 hr compared to 8 hr time point at 31 and 37 ^{o}C in membrane, cytoplasmic and nuclear fractions. This shows that the total number of up-regulated proteins overtime at each temperature is particularly different in the membrane fraction while down-regulation is particularly dissimilar in the cytoplasmic and nuclear fractions.

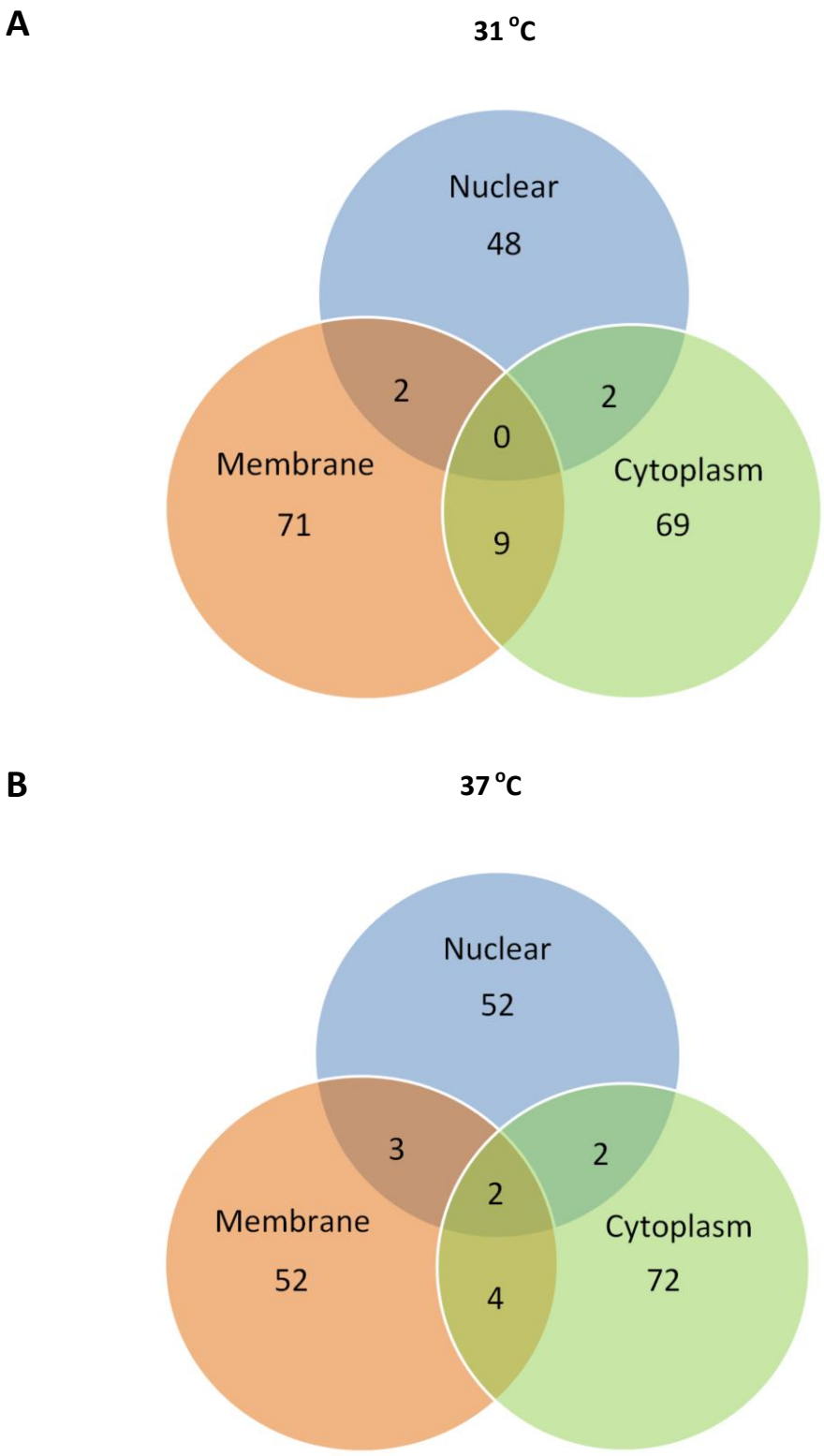


Figure 4.3.10 Up-regulated proteins overlap in CHO-K1 SEAP cells over 16hr at 24 hr compared to 8 hr time point at 31 (A) and 37 °C (B) between membrane, cytoplasmic and nuclear fractions. This shows that differentially expressed identifications are largely associated with each specific enriched fraction.

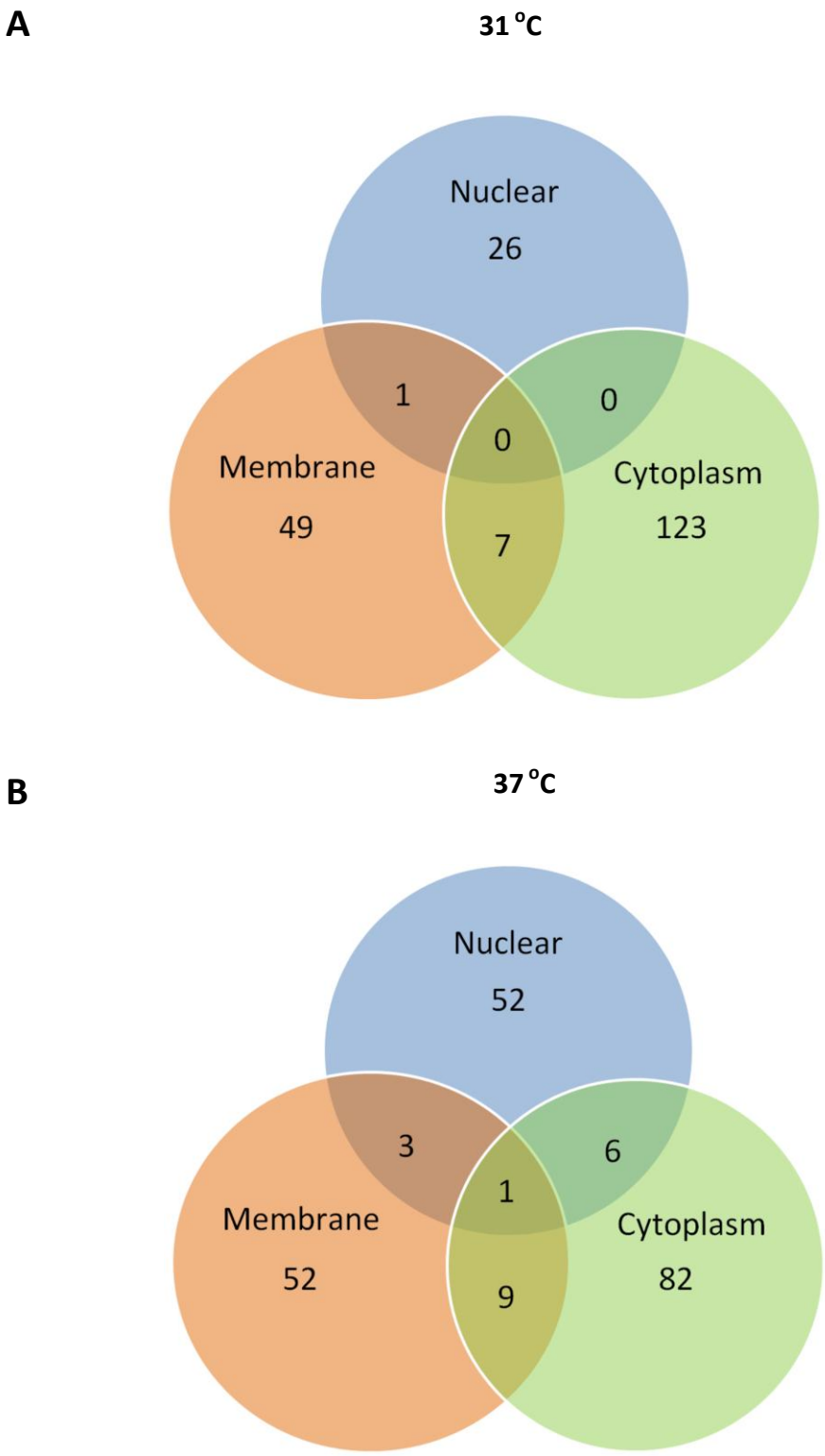


Figure 4.3.11 Down-regulated proteins overlap in CHO-K1 SEAP cells over 16hr at 24 hr compared to 8 hr time point at 31 (A) and 37 °C (B) between membrane, cytoplasmic and nuclear fractions. This shows that differentially expressed identifications are largely associated with each specific enriched fraction.

Further to this the uniqueness of IDs between the two temperatures was also assessed (**Figure 4.3.12**). This showed that while there are a similar number of proteins differentially regulated in certain fractions between the two temperatures over time the proteins that are differentially regulated are distinct from each other.

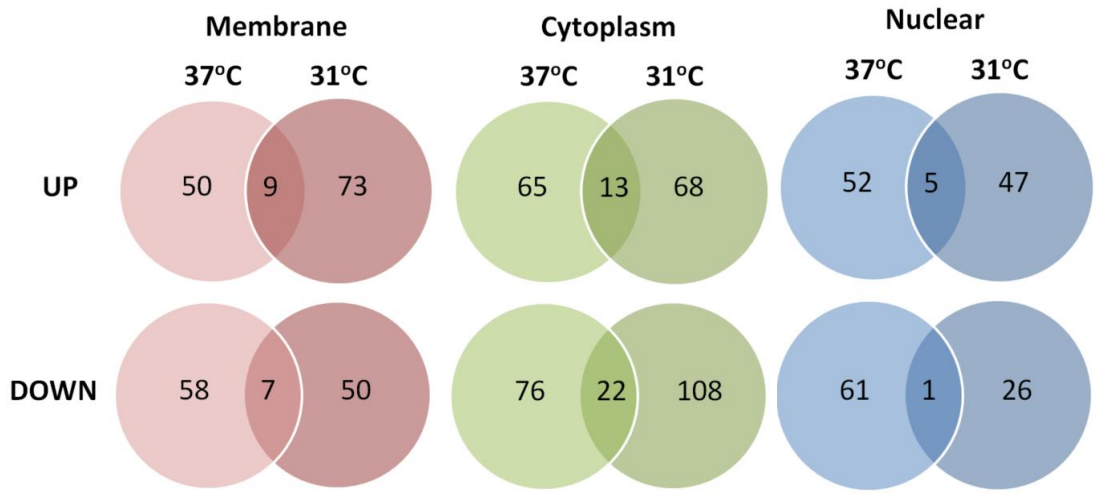


Figure 4.3.12 Overlapping identifications up or down-regulated over time between 31 and 37 °C. This shows that the population of proteins differentially regulated at a given temperature (37 or 31 °C) in a given direction (UP or DOWN) are distinct between the two temperatures over time.

In summary this analyses showed that there were a large pool of potential targets to choose from for functional validation. The overlap between fractions at each temperature also showed that differentially regulated proteins were distinctly associated with each subcellular fraction (**Figure 4.3.10 and 4.3.11**). The overlap between each fraction at both temperatures over time also showed that proteins dysregulated at each temperature had very little overlap (**Figure 4.3.12**). This indicates that the dysregulation of protein expression over time at both temperatures relates to different subsets of proteins. More specifically the comparison managed to highlight several proteins that were not significantly differentially regulated due to temperature shift but were differentially regulated over time at one temperature. Cyclon (CCDC86) was one such protein chosen for functional analysis. Cyclon was down-regulated 25 fold at 31 °C temperature shift in the Nuclear fraction of CHO-K1-SEAP cells over 16 hr between the 8 and 24 hr time point.

4.3.5 Pathway analysis of temperature shift over time at 31 and 37 ^{o}C

Three pathway tools DAVID, PANTHER and KEGG were used to assesses significantly enriched (Bonferroni $p \leq 0.05$) enriched biological process (BP), molecular functions (MF), cellular components (CC) and pathways over time between 8 and 24 hr after temperature shift in CHO-K1-SEAP cells grown at 37 ^{o}C and at 31 ^{o}C. This is different from **Section 4.3.3** where different temperatures were compared at a given time point. From this analysis the differential effect on BP, MF, CC and pathways can be determined over 16hr time period at 31 ^{o}C (temperature shift) and at 37 ^{o}C. As with **Section 4.3.3** the membrane, cytoplasmic and nuclear enrichment resulted in multiple differentially expressed lists of proteins for analysis associated with the 16hr growth.

4.3.5.1 DAVID analysis time course

DAVID analysis for CHO cells grown over 16 hr at 31 and 37 ^{o}C was conducted similarly to previous sections with differentially regulated protein lists being submitted and significant (Bonferroni $p \leq 0.05$) GO terms for biological process (BP), molecular function (MF). The three enriched fractions at both the 31 and 37 ^{o}C temperatures for up and down-regulated proteins over 16hr were processed through these enrichment categories. This resulted in enriched terms associated with up-regulated and down-regulated proteins at 31 ^{o}C and at 37 ^{o}C in the three enriched fractions. Those terms that were enriched specifically to either 31 or 37 ^{o}C were of most interest to determine the effects of temperature shift (**Table 4.3.5).** Key findings included:

- At 31 ^{o}C protein up regulation is associated with MF related to cytoskeletal binding and actin binding while down-regulated proteins are associated with BP related to both posttranscriptional regulation and initiation of translation with MF linked to binding of ATP, mRNA binding and initiation of translation.

- At 37 ^{o}C MF linked to up regulation included RNA-dependant ATPase and purine NTP-dependant helicase activity and MF linked to down regulation at 37 ^{o}C of purine binding and actin binding are not enriched for in the corresponding differential lists for 31 ^{o}C.

Temperature shift therefore is shown to potentially have the opposite effect on structural elements of CHO cells grown at 37 ^{o}C, suppress proteins involved in translation while also stabilising helicase activities related to proteins that are normally differentially expressed over time at 37 ^{o}C. (See **Appendix xi.-xiii.** for expanded tables).

Enrichment category	↑ over 16hr at 31oC	↑ over 16hr at 37oC	↓ over 16hr at 31oC	↓ over 16hr at 37oC
Biological Process	None unique to 31oC	None unique to 37oC	Regulation of translation, posttranscriptional regulation of gene expression (**Mem**), translational initiation (**Cyto**), cell cycle process (**Nuc**)	None unique to 37oC
Molecular Function	Actin binding and cytoskeletal binding (**Nuc**)	Purine NTP-dependant helicase activity, RNA-dependant ATPase activity (**Membrane**) and structural constituent of the ribosome (**Nuc**)	Translation initiation factor activity(**Mem**), mRNA binding (**Cyto**), ATP binding and adenyl ribonucleotide binding (**Nuc**)	Actin binding (**Mem**), purine ribonucleotide binding and purine nucleotide binding (**Cyto**)

Table 4.3.5 Summary of significantly (adjusted Bonferroni p≤0.05) enriched GO biological process and molecular function terms using DAVID that are uniquely associated with proteins up-regulated (↑) and down-regulated (↓) over 16hr at 31 ^{o}C temperature shift or 37 ^{o}C. Also labelled is the enriched fraction (**Mem, Cyto, Nuc**) for which the terms are associated with. Over time cells grown in temperature shift conditions show a dysregulation in cytoskeletal processes and translation factor activity compared to dysregulation of helicase activity, structural ribosomal activity and nucleotide binding activity.

4.3.5.2 PANTHER analysis time course

As with previous temperature shift analysis using PANTHER in **Section 5.4.3.2** the analysis of significantly (Bonferroni p≤0.05) enriched biological process (BP), molecular function (MF) were analysed across three enriched fractions membrane, cytoplasmic and nuclear. Unlike the time point analysis however differential lists between 8 and 24 hr for each temperature of 31 and 37 ^{o}C were analysed. This showed BP and MF that were enriched over the course of 16hr at both temperatures.

Panther analysis for CHO cells grown over 16 hr at 31 and 37 ^{o}C was conducted similarly to previous sections with differentially regulated protein lists being submitted and significant (Bonferroni p≤0.05) GO terms for biological process (BP), molecular function (MF) and cellular component (CC) being retrieved. The analysis was conducted for each of the three enriched fractions at both the 31 and 37 ^{o}C temperatures for up and down-regulated proteins over 16hr (**Table 4.3.6**) . This resulted in enriched terms associated with up-regulated proteins at 31 and 37 ^{o}C in membrane, cytoplasmic and nuclear enriched fractions and enriched terms associated with down-regulated proteins at 31 and 37 ^{o}C in membrane, cytoplasmic and nuclear enriched fractions. Key findings are listed below (**Appendix xiv.-xvi.** for expanded tables).

- Up regulation at 31 ^{o}C is observed to be linked to MF and CC oxidoreductase activity and cytoskeletal structure and other structural elements respectively.
- Down regulation at 31 ^{o}C is associated with BP linked to chromosome segregation, translation and CC linked to ribonucleoprotein and macromolecular complex affected.
- Actin cytoskeleton is the only term, which is found in CC down-regulated proteins, that doesn't appear with down-regulated proteins at 31 ^{o}C
- Most of the terms enriched at 37 ^{o}C are also enriched at 31 ^{o}C with only cytoskeleton CC being uniquely associated with protein down regulation at 37 ^{o}C.

The results indicate that at 31 ^{o}C protein up regulation may be associated with reappropriation of MF within the same BP as at 37 ^{o}C while down-regulated proteins at 31 ^{o}C related to the same MF as 37 ^{o}C but within different BP. All of these differential changes found at 31 ^{o}C relate to the cytoskeleton with additional dysregulation in the chromosomes and oxidoreductase activity and point to structural changes in the cell.

Enrichment category	↑ over 16hr at 31oC	↑ over 16hr at 37oC	↓ over 16hr at 31oC	↓ over 16hr at 37oC
Biological Process	Cellular component organization, Cellular component organisation or biogenesis, cell cycle (**Nuc**)	None unique to 37oC	Nuclear transport and chromosome segregation (**Nuc**)	None unique to 37oC
Molecular Function	Oxidoreductase activity (**Cyto**), structural constituent of the cytoskeleton and single stranded DNA binding (**Nuc**)	None unique to 37oC	None unique to 31oC	None unique to 37oC
Cellular Component	Vesicle coat (**Mem**), organelle, intracellular, actin cytoskeleton, cytoskeleton (**Nuc**)	None unique to 37oC	Ribonucleoprotein complex (**Mem,Cyto,Nuc**) and the macromolecule complex CC (**Mem,Nuc**)	Cytoskeleton (**Mem**)

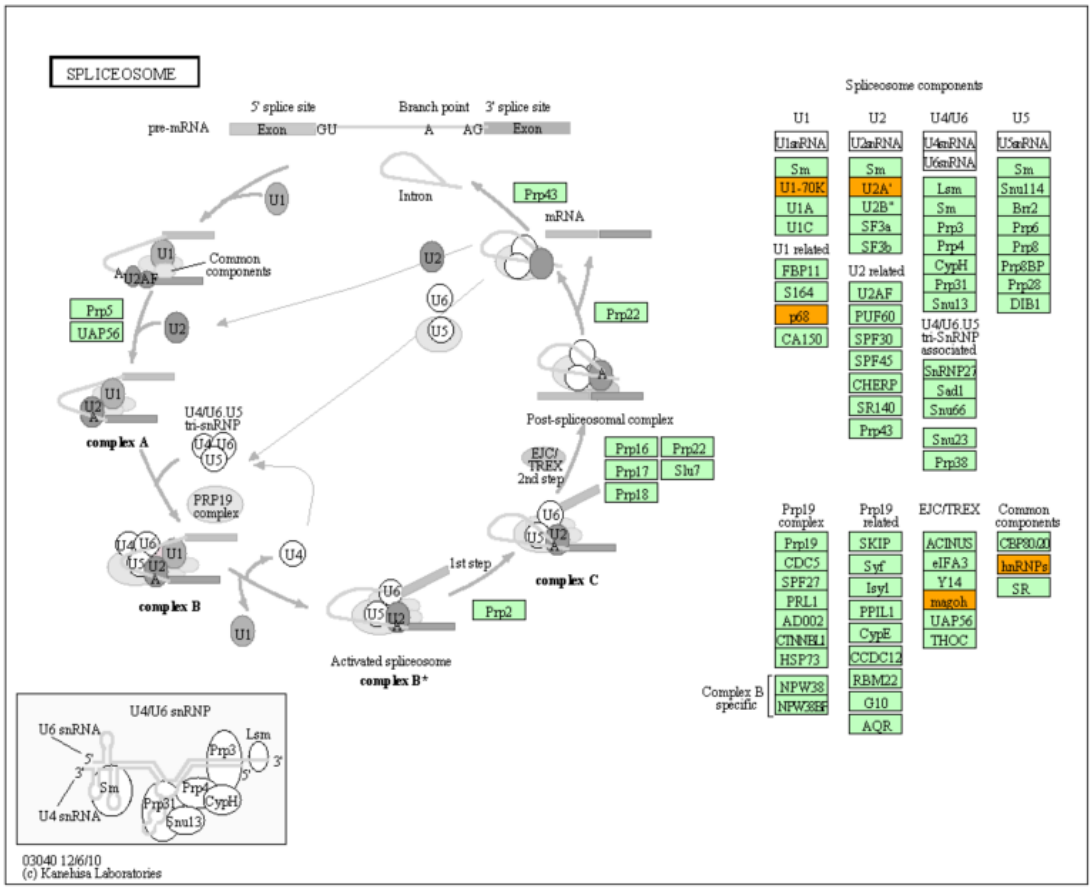
Table 4.3.6 Summary of significantly (adjusted Bonferroni p≤0.05) enriched PANTHER biological process, molecular function terms using DAVID that are uniquely associated with proteins up-regulated (↑) and down-regulated (↓) over 16hr at 31 ^{o}C temperature shift or 37 ^{o}C. Also labelled is the enriched fraction (**Mem, Cyto, Nuc**) for which the terms are associated with. Over time cells grown in temperature shift conditions show a dysregulation in cytoskeletal and translational processes compared to 37 ^{o}C.

4.3.5.3 KEGG analysis time course

Enrichment of specific pathways at 31 and 37 ^{o}C over 16hr was assessed using KEGG. As with the KEGG analysis in Section 5.4.3.3 involving comparing the same time points enrichment of KEGG pathways was deemed to be significant with Bonferonni $p<0.05$ and assessed across the membrane, cytoplasmic and nuclear fractions.

KEGG shows that the same pathways are significantly dysregulated in the same fractions at 31 and 37 ^{o}C. Up-regulated proteins in the membrane enriched fraction are associated with the spliceosome (**Figure 4.3.13 and 4.3.17**), down-regulated proteins in the cytoplasm are associated with the ribosome pathway (**Figure 4.3.14 and 4.3.18**) and conversely up-regulated proteins in the nuclear enriched fraction were linked to the ribosome pathway also (F**igure 4.3.16 and 4.3.19**). The only exception was the ribosomal pathway being significantly linked to the up-regulated in the membrane fraction at 37 ^{o}C only.

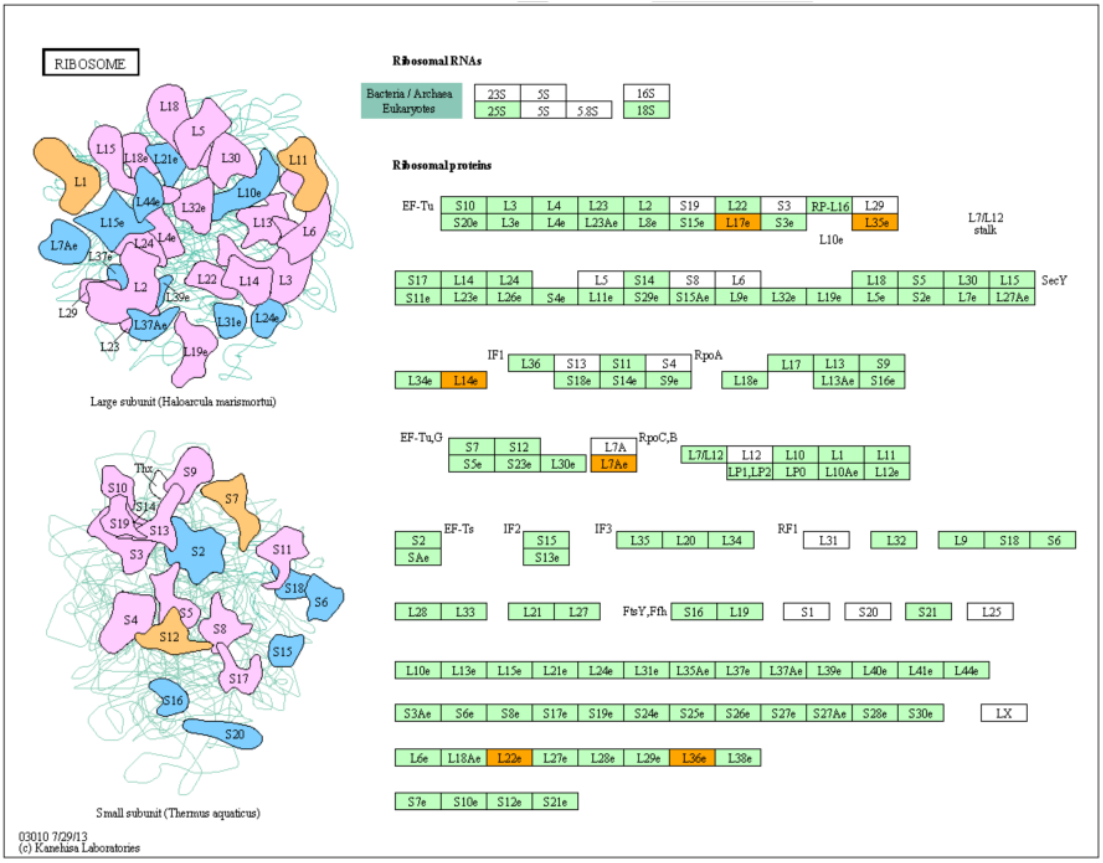
Interestingly, despite the same processes being dysregulated at both temperature points over time, the proteins involved in the enrichment of each process are different at each temperature. For example up regulation of proteins are associated with the spliceosome pathway which in turn includes up regulation of SNRAP1, MAGOH, HNPNPC, SNRNP70, DDX5, HNRNPA1 at 37 ^{o}C and up regulation of EIFA3, HNRNPK, SNRPD2, SNRNP70, RBMX, PUF60, HNRNPU, PRPF6. This would suggest that while the same pathways are dysregulated over time as the pathways are altered between 31 ^{o}C temperature shift. This is also consistent with findings from DAVID (**Section 4.3.5.1**) and PANTHER (**Section 4.3.5.2**) where molecular functions were seen to be dysregulated more than biological processes and highlights the subtle effect of temperature shift on CHO cells over time in culture. Rather than differential regulation of large bioprocesses temperature shift is controlled by multiple smaller changes in molecular functions. Another interesting observation is the increase in ribosomal proteins in the nuclear enriched fraction. While ribosomes are found in the cytoplasm the importation of ribosomal proteins from the cytoplasm back into the nucleus, as well as assembly of other ribosomal proteins in the nucleolus, is a key step in ribosomal subunit maturation in eukaryotes and extensively studied in yeast (Moy and Silver 1999). It has also been reported that this can be a temperature dependant process in protozoa (Moy and Silver 1999, Giese and Wunderlich 1983) and may mark the first reports of occurring in CHO.



37oC Membrane UP

Term	Count	Genes	Adjusted p-value
Spliceosome	7	SNRPA1, MAGOH, HNRNPC, SNRNP70, DDX5, HNRNPA1, HNRNPU	1.16E-03

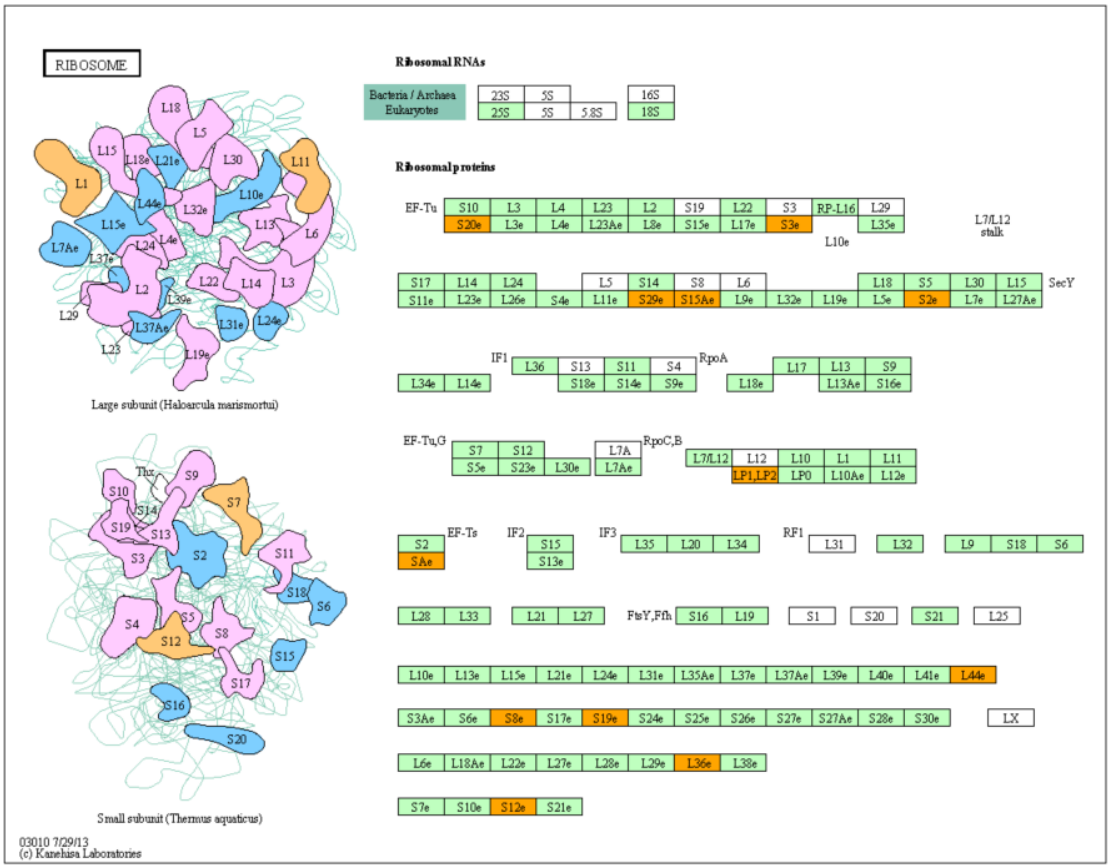
Figure 4.3.13 KEGG pathway diagram representing membrane enriched protein up-regulation in the spliceosome pathway over 16 hr at 37 ^{o}C. By comparing 37 ^{o}C 8 vs 24 hr membrane enriched up-regulated proteins (16hr), highlighted in orange above, the spliceosome pathway was shown to be significantly ($p=1.16 \times 10^{-3}$) differentially regulated containing 7 identifications SNRP70 (U1-70K), SNRPA1 (U2A'), DDX5 (p68), Magoh (magoh) and HNRNPC, HNRNPA1 and HNRNPU (hnRNPs).



37oC Membrane DOWN

Term	Count	Genes	Adjusted p-value
Ribosome	6	RPL17, RPL14, RPL22, RPL35, RPL36, RPL7A	5.17E-03

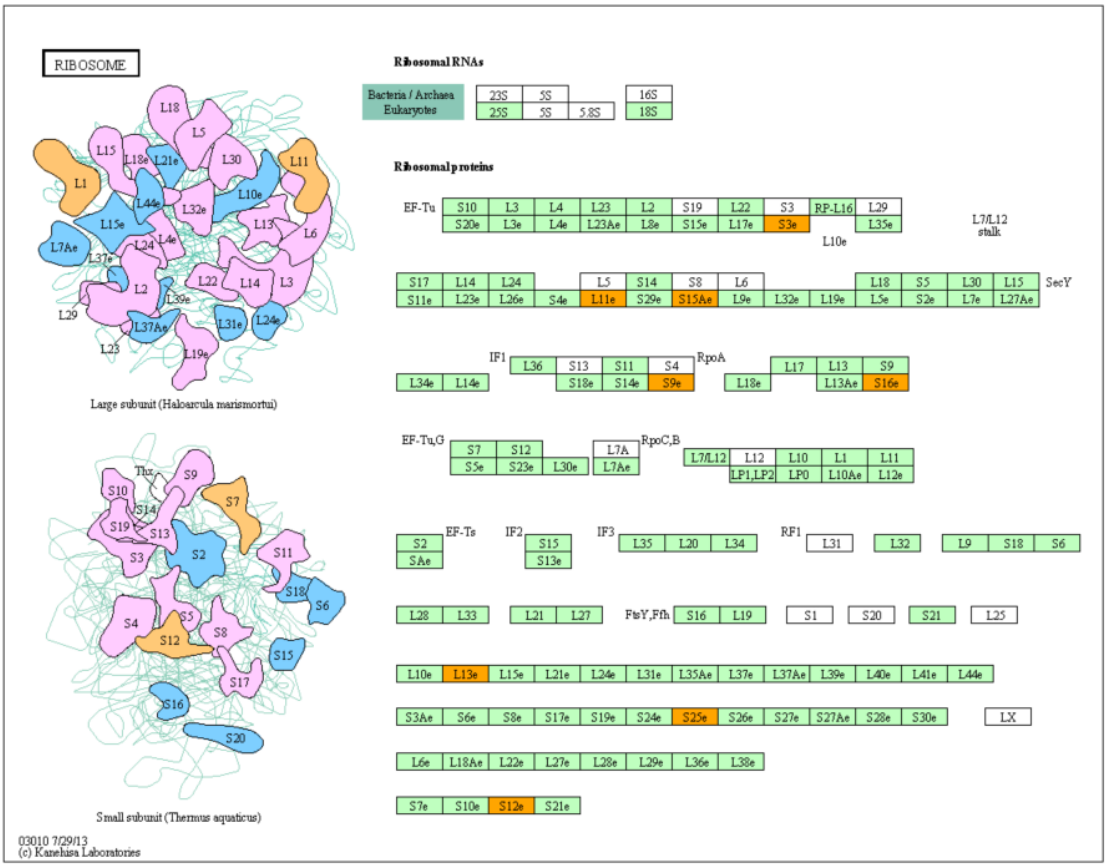
Figure 4.3.14 KEGG pathway diagram representing membrane enriched protein down-regulation in the ribosome pathway over 16hr at 37 ^{o}C. By comparing 37 ^{o}C 8 vs 24 hr membrane enriched down-regulated proteins (16 hr), highlighted in orange above, the ribosome pathway was shown to be significantly ($p=5.17 \times 10^{-3}$) differentially regulated containing 6 identifications.



37oC Cytoplasmic DOWN

Term	Count	Genes	Adjusted p-value
Ribosome	12	RPSA, RPL36A, RPS19, RPS29, RPS12, RPLP2, RPL36, RPS15A, RPS20, RPS2, RPS8, RPS3	7.31E-09

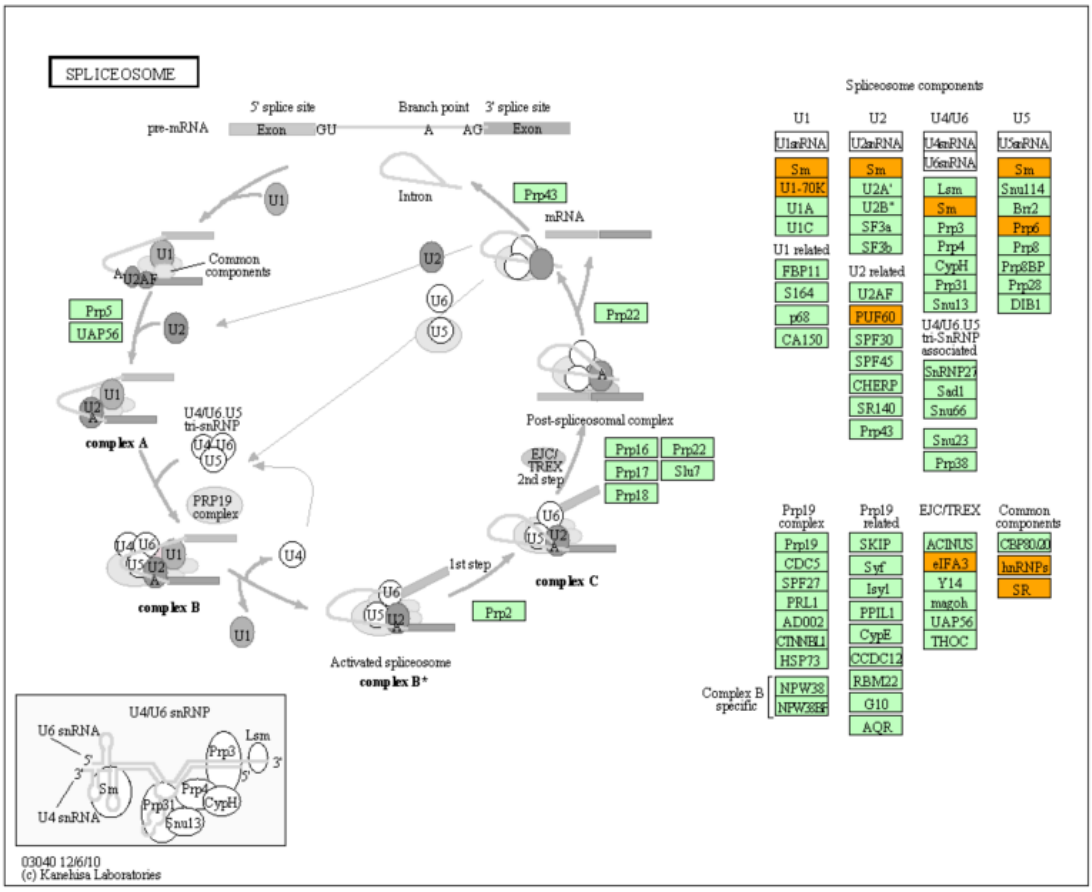
Figure 4.3.15 KEGG pathway diagram representing cytoplasm enriched protein down-regulation in the ribosome pathway over 16hr at 37 ^{o}C. By comparing 37 ^{o}C 8 vs 24 hr cytoplasmic enriched down-regulated proteins (16 hr), highlighted in orange above, the ribosome pathway was shown to be significantly ($p=7.31 \times 10^{-9}$) differentially regulated containing 12 identifications.



37oC Nuclear UP

Term	Count	Genes	Adjusted p-value
Ribosome	8	RPS25, RPS16, RPL13, RPS12, RPS15A, RPS9, RPL11, RPS3	1.25E-05

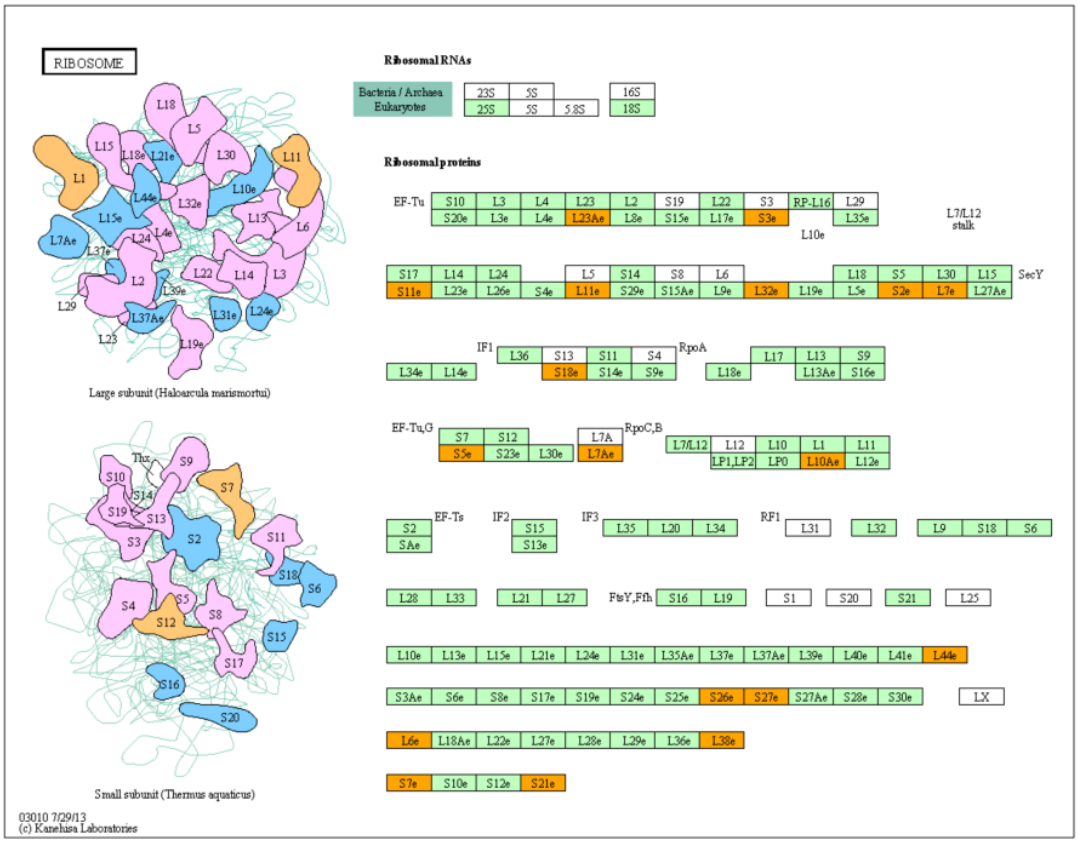
Figure 4.3.16 KEGG pathway diagram representing nuclear enriched protein up-regulation in the ribosome pathway over 16 hr at 37 ^{o}C. By comparing 37 ^{o}C 8 vs 24 hr nuclear enriched up-regulated proteins (16 hr), highlighted in orange above, the ribosomal pathway was shown to be significantly ($p=1.25 \times 10^{-5}$) differentially regulated containing 8 identifications.



31oC Membrane UP

Term	Count	Genes	Adjusted p-value
Spliceosome	8	EIF4A3, HNRNPK, SNRPD2, SNRNP70, RBMX, PUF60, HNRNPU, PRPF6	3.39E-03

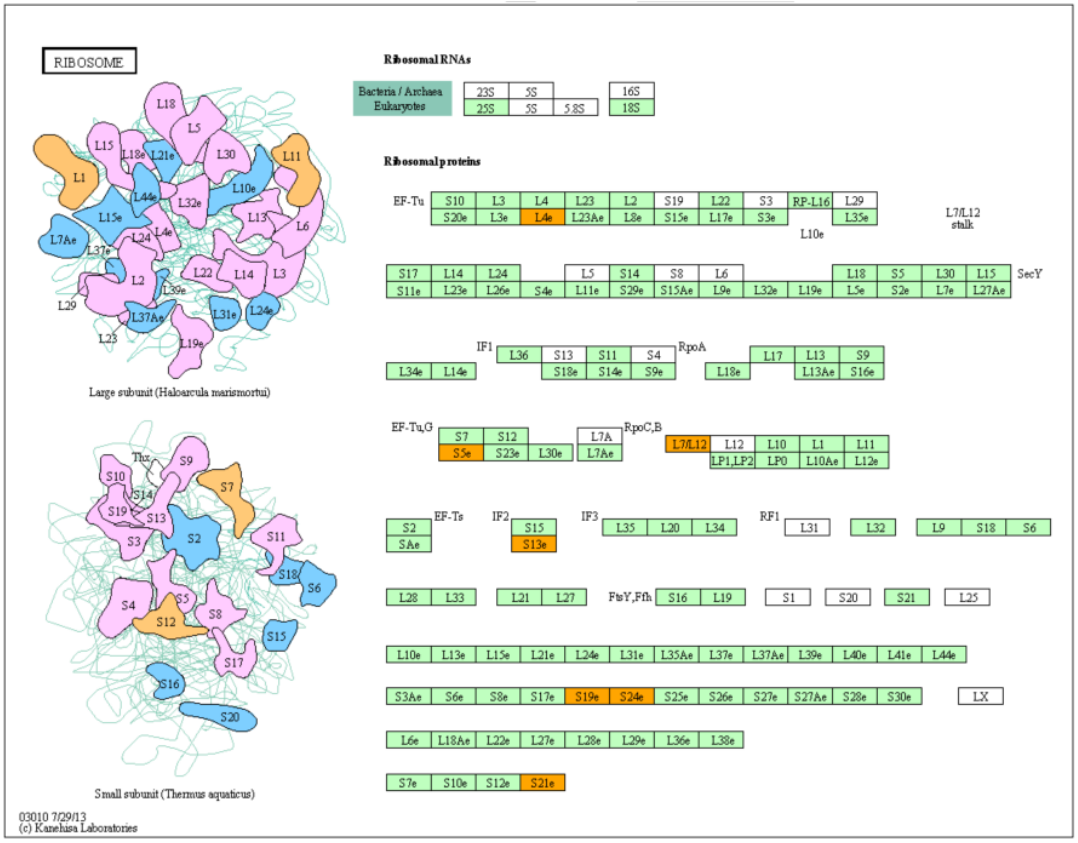
Figure 4.3.17 KEGG pathway diagram representing membrane enriched protein up-regulation in the spliceosome pathway over 16 hr at 37 ^{o}C. By comparing 31 ^{o}C 8 vs 24 hr membrane enriched up-regulated proteins (16 hr), highlighted in orange above, the spliceosome pathway was shown to be significantly ($p=3.39 \times 10^{-3}$) differentially regulated containing 8 identifications SNRPD1, SNRPD2 (Sm), PRPF6 (Prp6), PUF60 (PUF60), EIF4A3 (eIFA3), RBMX, HNRNPK and HNRNPK (hnRNPs).



31oC Cytoplasmic DOWN

Term	Count	Genes	Adjusted p-value
Ribosome	18	RPL36A, RPL23A, RPL38, RPS2, RPS5, RPS3, RPS7, RPS26, RPS27, RPS18, RPL32, RPL7, RPL6, RPL11, RPS11, RPL10A, RPL7A, RPS21	7.39E-16

Figure 4.3.18 KEGG pathway diagram representing cytoplasm enriched protein down-regulation in the ribosome pathway over 16 hr at 31 ^{o}C. By comparing 31 ^{o}C 18 vs 24 hr cytoplasmic enriched down-regulated proteins (16 hr), highlighted in orange above, the ribosome pathway was shown to be significantly ($p=7.39 \times 10^{-16}$) differentially regulated containing 18 identifications.



31oC Nuclear UP

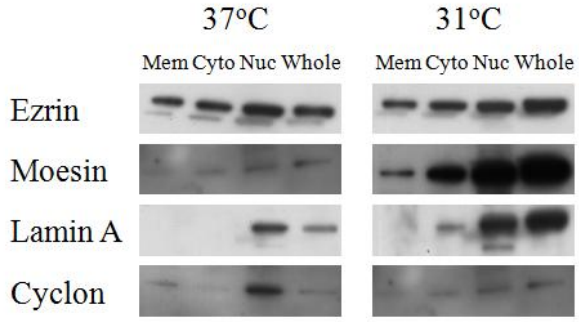
Term	Count	Genes	Adjusted p-value
Ribosome	6	RPS19, RPS13, RPL4, RPS21, RPS5, RPS24	6.85E-03

Figure 4.3.19 KEGG pathway diagram representing nuclear enriched protein up-regulation in the ribosome pathway over 16 hr at 31 ^{o}C. 31 ^{o}C 6 vs 24 hr nuclear enriched up-regulated proteins highlighted in orange above show the ribosome pathway to be significantly ($p=6.85 \times 10^{-3}$) differentially regulated containing 8 identifications.

4.4 Functional Validation

From the differential lists, four proteins were chosen for functional validation Cyclon (CCDC86 [coiled-coil domain containing protein]), Ezrin (EZR), Moesin (MSN) and Lamin A/C (LMNA). EZR, MSN and LMNA were highly differentially expressed at each time point at 31 ^{o}C in the temperature shifted comparison (**Section 4.3.2**) with EZR and LMNA being differentially regulated over time also. CCDC86 was also highly differentially expressed being 25.96 fold down-regulated at 31 ^{o}C over a period of 16hr between 8 and 24 hr after temperature shift but not at 37 ^{o}C in the time course analysis (**Section 4.3.4**). Western blot analysis of samples grown for 24 hr after temperature shift confirmed that enrichment of the targets had occurred at both temperatures with Ezrin, Moesin and Lamin A up-regulated at 31 ^{o}C and Cyclon down-regulated at 31 ^{o}C (**Figure 4.4.1)**.

A



B

Time point analysis (31 vs 37°C)

		Membrane		Cytoplasmic		Nuclear	
Gene name	**Description**	**8hr**	**24hr**	**8hr**	**24hr**	**8hr**	**24hr**
EZR	Ezrin				↑1.68		↑1.9
MSN	Moesin		↑Infinity			↓1.68	
LMNA	Lamin-A/C		↑3.03		↑7.29		

Time course analysis (8 vs 24hr)

		Membrane		Cytoplasmic		Nuclear	
Gene name	**Description**	**31°C**	**37°C**	**31°C**	**37°C**	**31°C**	**37°C**
CCDC86	Coiled-coil domain-containing protein 86					↓25.96	
EZR	Ezrin		↓1.69	↑1.87			
LMNA	Lamin-A/C		↓1.95	↑6.21			↓2.39

Figure 4.4.1 Western blot analysis showing the expression intensity of targets in CHO-K1-SEAP cells grown over 24 hr (A) with corresponding label-free LC-MS/MS data (B). Targets were picked for functional validation which displayed high fold change dysregulation between 31 and 37 ^{o}C and over a 16 hr time period.

As well as fold change, each target was also picked for re-occurrence in previous profiling studies in our lab and novelty in relation to CHO phenotypes. Specific reasons also included the close relationship between Ezrin and Moesin in the Ezrin-Moesin-

Radixin complex of proteins (Adada et al. 2014) and the strong association of LMNA and CCDC86 to pathology research related to proliferation and cell survival(Lamberti et al. 2004, Politano et al. 2013, Emadali et al. 2013, Shishkov et al. 2013b). CCDC86 in particular was chosen as it only appeared differentially regulated in the time course analysis.

To assess the functional role of these proteins custom siRNA against the CHO sequence (http://CHOgenome.org) were supplied by Integrated DNA Technologies (http://eu.idtdna.com). Knockdown was assessed by Western blot using the negative siRNA control and three siRNA to each target. Cyclon showed almost complete knockdown for all three siRNA (**Figure 4.4.2 A**), Ezrin showed knockdown in 2 of 3 (siRNA1 and siRNA3) of the siRNA used (**Figure 4.4.3 A**), Moesin knockdown was not confirmed by Western blot but a phenotypic effect was observed (**Figure 4.4.4**) and Lamin A/C showed reduced expression in all three siRNA molecules used (**Figure 4.4.5 A**). From previous pathway analysis data in section 5.2 and 5.4 enriched terms suggested that cytoskeletal elements were differentially regulated as well as proliferation from temperature shift therefore the cell size and cell number was assessed after knockdown of each target (**Figure 4.4.2-4.4.5**). The following table (**Table 4.4.1)** lists the key findings and the significance of each of the three siRNA using 2 tailed students-t test compared to siRNA negative control (See **Appendix xviii.** for values).

Target	Total viable cells/ml	Cell viability	Cell average diameter	Cell average area	Cell average perimeter
Cyclon	- / * / *	- / - / -	- / - / -	- / * / *	- / * / *
Ezrin	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
Moesin	- / * / *	- / - / -	- / - / -	* / * / ***	* / * / ***
Lamin A/C	- / * / -	- / - / -	- / - / -	- / */ -	- / - / -
- = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001					

Table 4.4.1 Cyclon and Moesin were shown to have a significant effect on cell proliferation as well as cell perimeter and cell area. Moesin knockdown was not confirmed by Western blot but may be complicated by cross reactivity between Ezrin, Radixin and Moesin. It is also worth noting that all the commercial antibodies used are against human analogues of these proteins which may complicate Western blot results.

Cyclon

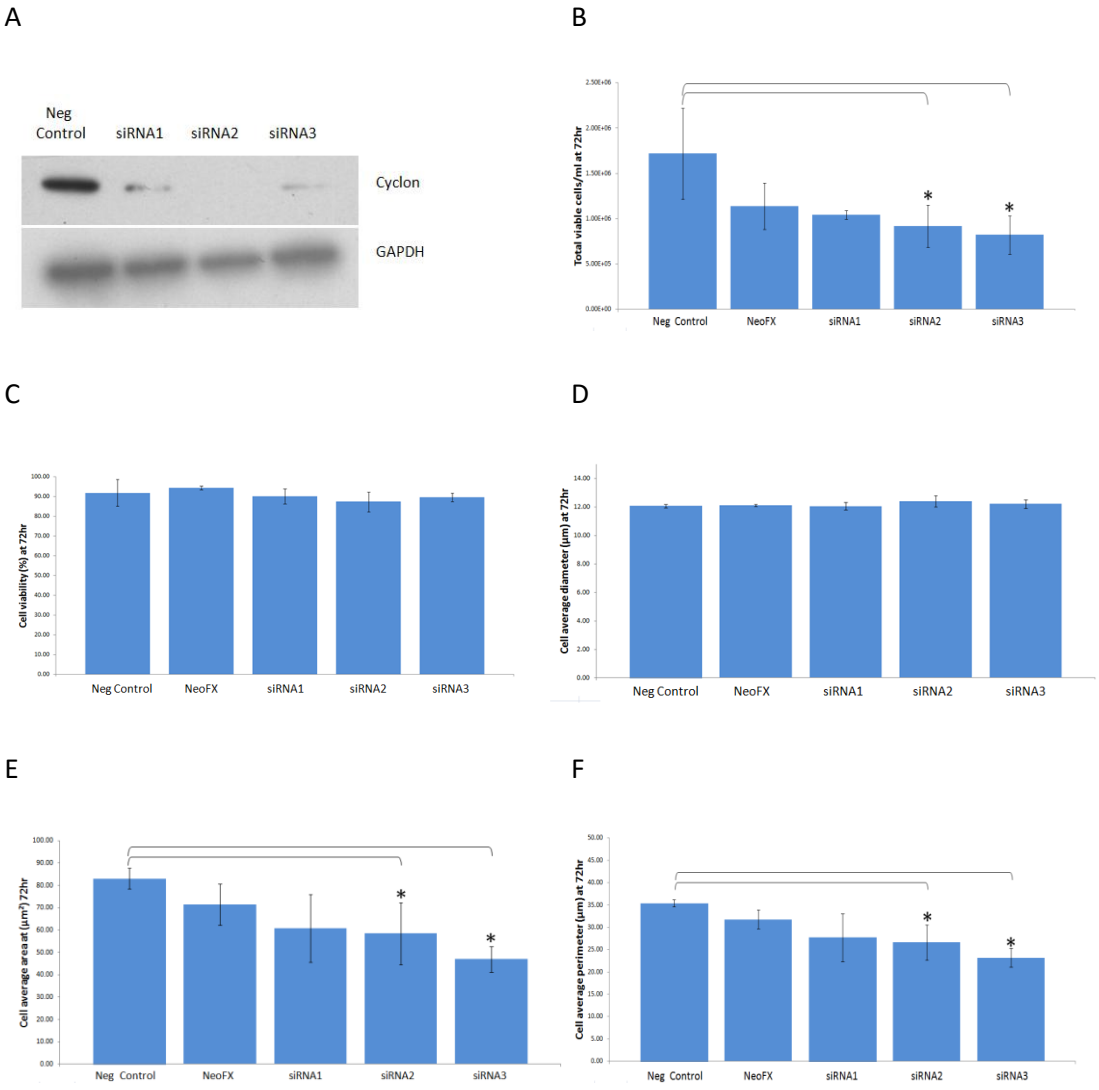


Figure 4.4.2 Effect of Cyclon knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). * = p≤0.05 compared to neg control.

Ezrin

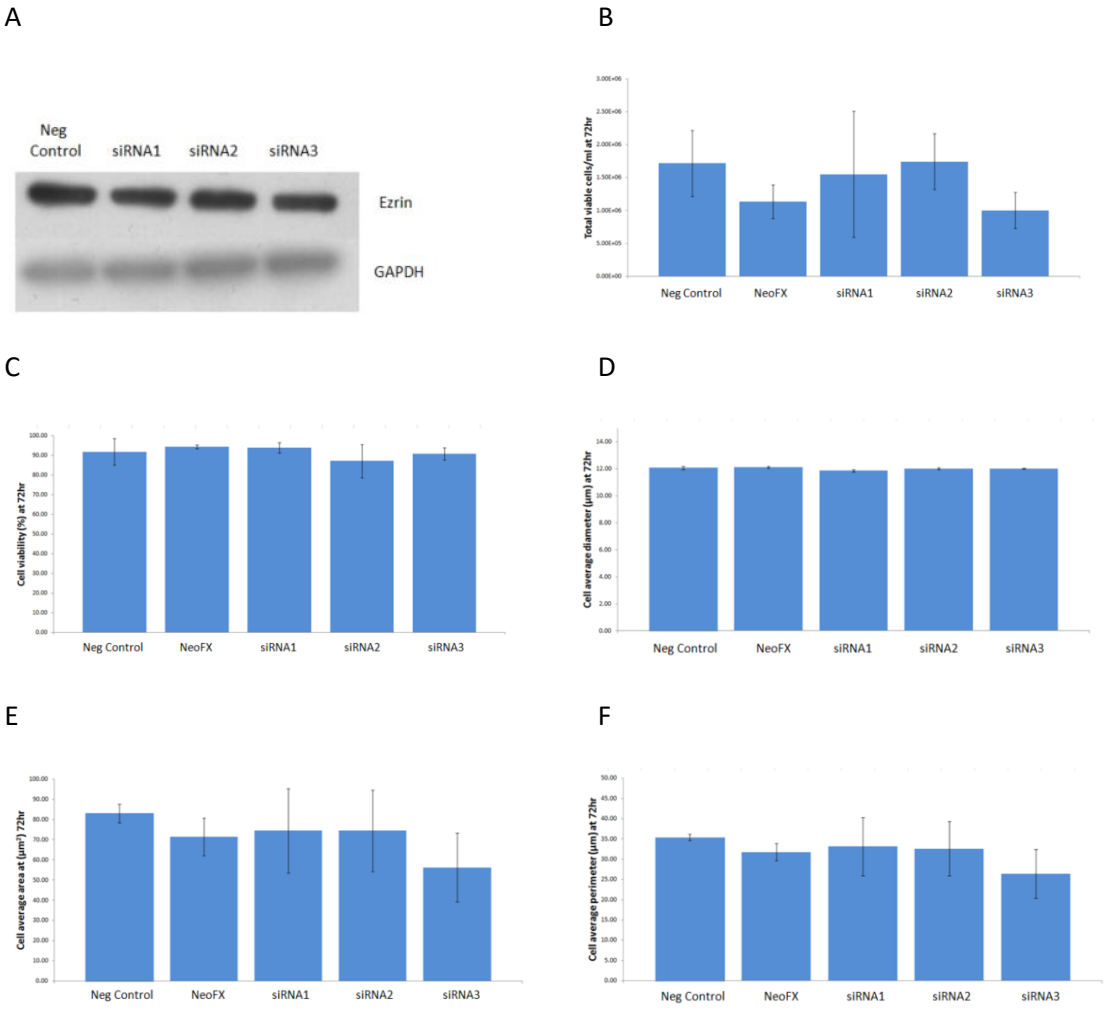


Figure 4.4.3 Effect of Ezrin knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). Knockdown of Ezrin was demonstrated by Western blot anlaysis but no significant effect was observed.

Moesin

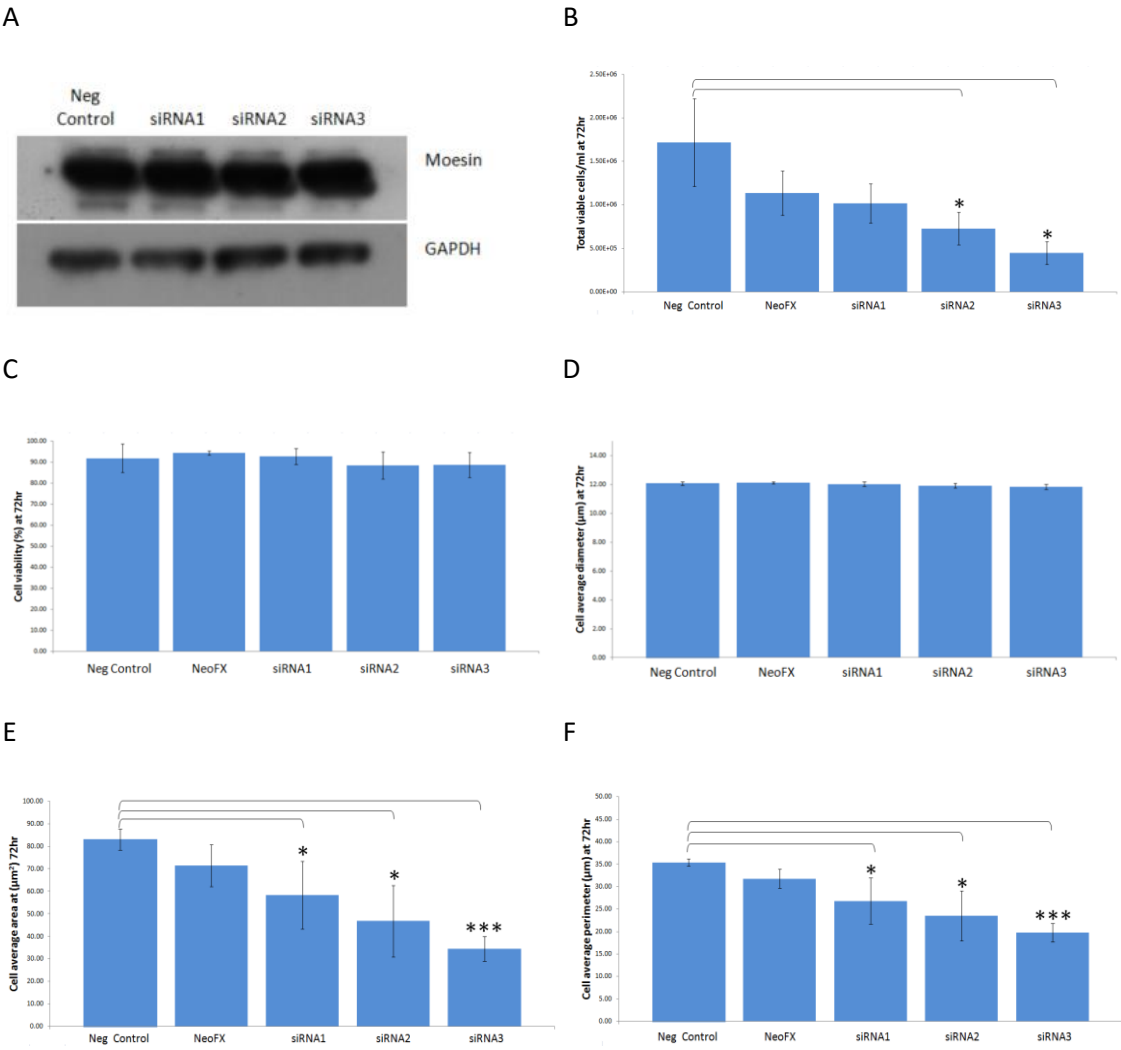


Figure 4.4.4 Effect of Moesin knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). Due to inconclusive Western blot results the observed significant parameters are likely to be independent of Moesin. * = $p \leq 0.05$, *** = $p \leq 0.05$ compared to neg control.

Lamin A/C

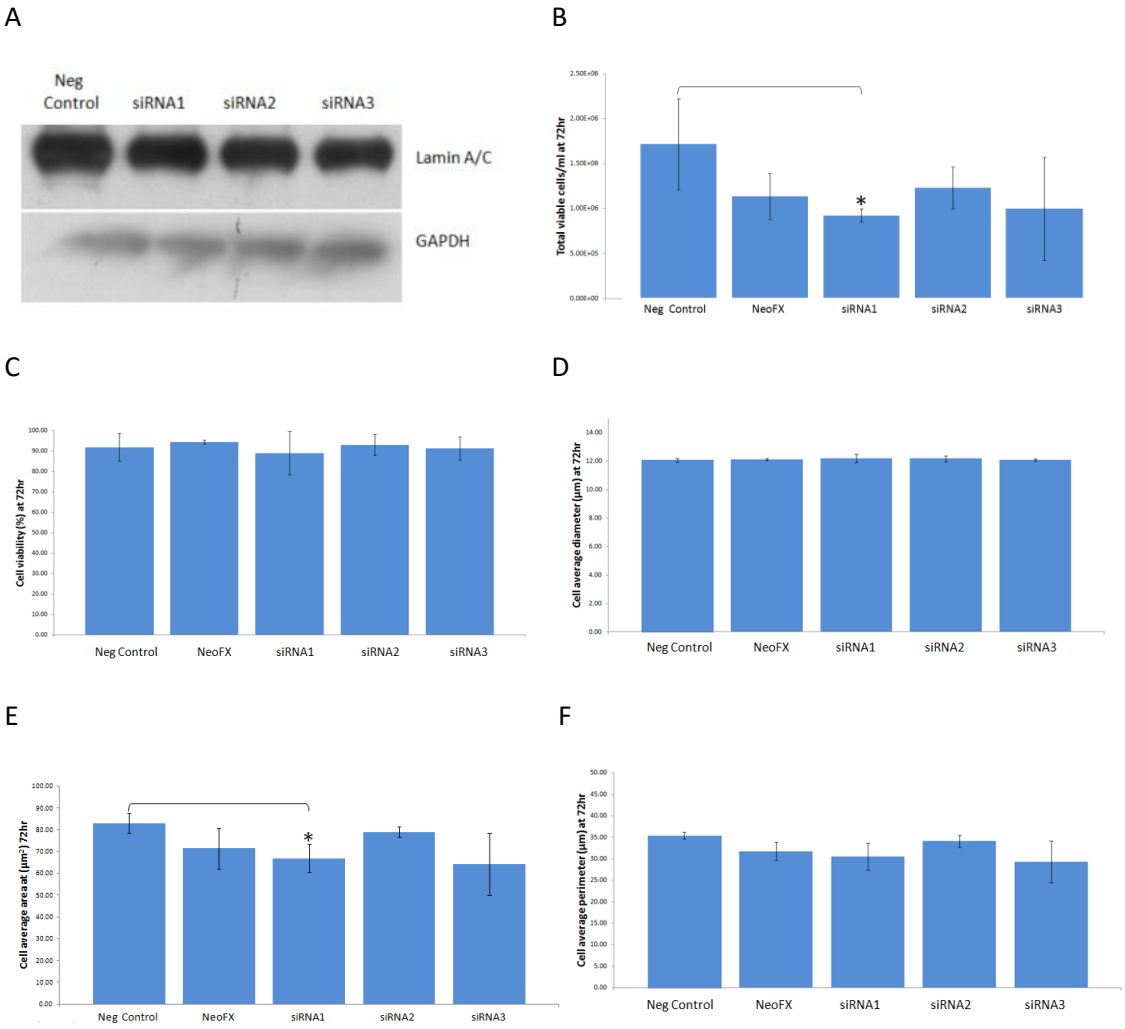


Figure 4.4.5 Effect of Lamin A/C knockdown after 72 hr on target expression using Western blot (A), total cells/ml (B), cell viability (C), cell average diameter (D), cell average area (E), cell average perimeter (F). * = p≤0.05 compared to neg control.

4.5 Results Summary

Quantitative LC-MS/MS analysis was combined with subcellular fractionation to determine the deeper mechanisms behind 31 ^{o}C temperature shifted CHO-K1-SEAP cells. Using commercially available subcellular fractionation kits we obtained a membrane, cytoplasmic and nuclear fraction for CHO cells. We confirmed that through overlapping IDs, pathway analysis and Western blot analysis that each fraction had enrichment for the expected sub cellular components and that a large number of IDs were unique to each fraction compared to an unfractionated sample. CHO-K1-SEAP cells were grown for 72 hr then one set was shift to 31 ^{o}C while one set remained at 37 ^{o}C. Cell culture was continued for a further 8 and 24 hr at which point samples were taken and subcellularly fractionated..

The first analysis compared cells grown at 31 ^{o}C to 37 ^{o}C at 8 hr and at 24 hr after temperature shift. This was known as the time point analysis. Abundance of IDs showed the expression pattern of temperature shifted CHO cells across sub cellular fractions. Pathway analysis confirmed the high degree of translational repression and anti-apoptotic process but also surprisingly showed detailed information on the DNA replication pathway being up-regulated. The second analysis compared 8 and 24 hr time points within each temperature allowing for proteins differentially regulated over time at each temperature to be identified. This revealed that over time a similar total of number of proteins are differentially regulated at each temperature and that there is little overlap between these protein IDs. Pathway analysis shows at the biological process and molecular function level translation and structural proteins are down-regulated over time in temperature shifted cells but surprisingly at the higher pathway level ribosomal activity is dysregulated in both directions and the spliceosome pathway is up-regulated at both temperatures but with different proteins enriched at each temperature.

From the both analysis 4 targets were chosen for functional knockdown follow up Ezrin, Moesin, Cyclon and Lamin 1 experiments. Cyclon knockdown significantly reduced cell number by a mean of 66% (n=2) while Moesin knockdown significantly reduced cell number by a mean of 50% (n=2). Knockdown also reduced cell size and had no negative impact on cell viability. While this effect is promising more work will be required to assess their ability to induce a temperature shift phenotype - reduced proliferation and increased productivity over time which is the one of the goals for selecting these targets for functional validation.

CHAPTER 5
Identification of novel membrane protein targets in breast cancer

5.1 Identifying potential membrane proteins over expressed in breast cancer

Breast cancer is a complex heterogeneous disease with an ever-increasing number of subtypes emerging. In this study we aim to identify membrane proteins that are over-expressed in aggressive breast cancer subtypes ER+, LN+, HER2+ and TNBC breast cancer that could be validated as potential candidate targets for therapeutic targeting using ADCs (see introduction **Section 1.8.3**).

To achieve this we combined a number of publicly available transcriptomic datasets (Sircoulomb et al. 2010, Pau Ni et al. 2010, Tollet-Egnell et al. 2001, Skrzypczak et al. 2010, Neve et al. 2006, LaBreche, Nevins and Huang 2011). These were of only a handful of suitable datasets chosen from 29,357 publically available breast cancer gene lists. Datasets below the top 150 dropped off dramatically below 25 samples. The datasets in the literture presented above were suitable due to their experimental design, they contained tissue rather than cell line data, they contained gene expression data and the most limiting factor was that these lists contained normal breast tissue data.We also had an in house dataset consisting of 62, 25, 60 and 8 samples of ER+, HER2+, LN+ and TNBC subtypes respectively with 17 normal samples for comparison.

Datasets were processed using ebayes function of the limma package (Smyth 2004) from Bioconductor (See materials and methods **Section 2.7.3**). A shortlist of differrentially expressed genes was generated between normal breast tissue and breast cancer subtypes of HER2+, TNBC, ER+ and LN+. The criteria used to generate the gene candidates for further protein validation and investigation as potential ADC targets can be summarised as follows:

- Over expressed in breast cancer compared to normal breast
- High fold change in breast cancer compared to normal breast
- Predicted membrane localisation in breast cancer

Using the above criteria there were 238 differentially expressed proteins across 4 different breast cancer subtypes (**Table 5.1.1**) which were all reported to be membrane localised according to corresponding protein data on the UniProt database.

	No of identifications
ER+ vs normal	20
HER2+ vs normal	123
LN+ vs normal	37
TNBC vs normal	58

Table 5.1.1 Summary of number of potential membrane expressed targets for follow up. The above numbers relate to transcripts that were ≥1.2 fold over expressed significantly (adj p≤0.05) over expressed in breast cancer subtypes ER+, HER2+, LN+ or TNBC.

Manual searching of targets within these lists for evidence of functional analysis in breast cancer and other cancers further reduced the number of candidates down to 40 proteins (**Table 5.1.2**).

- High fold change in breast cancer compared to normal breast
- Previously unexploited (to our knowledge) in breast cancer
- Availability of commercial antibodies

This ultimately resulted in the selection of 5 candidate targets for follow up namely Immunoglobulin superfamily, member 9 (IGSF9), Killer cell lectin-like receptor subfamily G, member 2 (KLRG2), SLAM family member 8 (SLAMF8), Tetraspanin 13 (TSPAN13), Low density lipoprotein receptor-related protein 8, apolipoprotein e receptor (LRP8) (**Table 5.1.3**).

ER+ vs normal		
Gene Name	**Full name**	**Fold change**
CA12	Carbonic anhydrase XII	7.21
CD24	CD24 molecule	9.82
CD24	CD24 molecule	8.76
EFR3A	EFR3 homolog A (S. cerevisiae)	1.38
SPINT2	Serine peptidase inhibitor, Kunitz type, 2	10.72
TACSTD2	Tumor-associated calcium signal transducer 2	11.15
TSPAN13	**Tetraspanin 13**	**2.20**

LN+ vs normal		
Gene Name	**Full name**	**Fold change**
CA12	Carbonic anhydrase XII	6.49
PTPRK	Protein tyrosine phosphatase, receptor type, K	2.59
SLC39A6	Solute carrier family 39 (zinc transporter), member 6	2.90
SLC39A6	Solute carrier family 39 (zinc transporter), member 6	2.44
SPINT2	Serine peptidase inhibitor, Kunitz type, 2	11.62
TACSTD2	Tumor-associated calcium signal transducer 2	11.74
TSPAN13	**Tetraspanin 13**	**2.09**

HER2+ vs normal		
Gene Name	**Full name**	**Fold change**
CELSR1	Cadherin, EGF LAG seven-pass G-type receptor 1	9.22
FCGR1B	Fc fragment of IgG, high affinity Ib, receptor (CD64)	5.59
GPRC5A	G protein-coupled receptor, family C, group 5, member A	11.68
IGSF9	**Immunoglobulin superfamily, member 9**	**18.87**
LSR	Lipolysis stimulated lipoprotein receptor	8.36
RAB15	RAB15, member RAS onocogene family	7.24
SLAMF8	**SLAM family member 8**	**2.71**
SLC4A8	Solute carrier family 4, sodium bicarbonate cotransporter, member 8	16.26
STRA6	Stimulated by retinoic acid gene 6 homolog (mouse)	8.48

TNBC vs normal		
Gene Name	**Full name**	**Fold change**
ADAM19	A disintegrin and metalloproteinase domain 19 (meltrin beta)	1.65
ATP13A3	ATPase type 13A3	17.92
CXADR	Coxsackie virus and adenovirus receptor	12.37
DSC2	Desmocollin 2	7.82
EPT1	Ethanolaminephosphotransferase 1	5.24
FCGR1B	Fc fragment of IgG, high affinity Ia, receptor (CD64)	6.83
FLVCR1	Feline leukemia virus subgroup C cellular receptor	5.41
GJB2	Gap junction protein, beta 2, 26kDa (connexin 26)	11.60
HMMR	Hyaluronan-mediated motility receptor (RHAMM)	9.40
KLRG2	**Killer cell lectin-like receptor subfamily G, member 2**	**5.10**
KRTCAP3	Keratinocyte associated protein 3	26.45
LRP8	**Low density lipoprotein receptor-related protein 8, apolipoprotein e receptor**	**9.82**
LRRC15	Leucine rich repeat containing 15	24.66
LSR	Liver-specific bHLH-Zip transcription factor	7.74
PCDH17	Procadherin 17	1.72

Table 5.1.2 Candidate targets satisfying criteria for protein validation follow up. All targets were reported as being ***membrane expressed*** and ***>1.2 fold over expressed*** in breast cancer. The final targets chosen for validation are highlighted TSPAN13, IGSF9, SLAMF8, KLRG2 and LRP8

Gene name	Full name	Subtype with over expression vs Normal	Fold change
IGSF9	Immunoglobulin superfamily, member 9	HER2+	18.87
KLRG2	Killer cell lectin-like receptor subfamily G, member 2	TNBC	5.10
SLAMF8	SLAM family member 8	HER2+	2.71
TSPAN13	Tetraspanin 13	ER+	2.20
TSPAN13	Tetraspanin 13	LN+	2.09
LRP8	Low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	TNBC	9.82

Table 5.1.3 Summary of 5 potential membrane expressed breast cancer candidate targets for validation. The above 5 gene transcripts were ≥1.2 fold over expressed significantly (adj $p \leq 0.05$) in breast cancer subtypes ER+, HER2+, LN+ or TNBC compared to normal breast, did not have functional work related to breast cancer in the literature and commercial antibodies were available for validation by IHC analysis and Western blot analysis.

5.2 Validation of expression in cell lines

Having chosen 5 targets for validation from the bioinformatics data in **Section 5.1** namely KLRG2, LRP8, TSPAN13, IGSF9 and SLAMF8 (**Table 5.1.2**) it was necessary to validate their expression in breast cancer cell lines representative of breast cancer subtypes as these targets were chosen based on micro array data and their expression at the protein level required confirmation. Furthermore the localisation of th**ese targets to the membrane was initia**lly predicted from UniProt. For this reason representative cell lines were also enriched for membrane proteins using a commercial kit method involving differential solubility and centrifugation to determine if these 5 target proteins were associated with the membrane (see materials and methods **Section 2.4.3**).

A panel of 9 cell lines (n=2) (**Table 5.2.1**) along with their corresponding membrane enriched fractions (n=2) were chosen to investigate the expression of each of the five targets. Initial Western blot analysis showed the LRP8 expression to be unsuitable for Western blot analysis resulting in smeared non-specific binding. TSPAN13 and SLAMF8 did not show expression in membrane enriched fractions of the cell lines used. As membrane expression was predicted using UniProt Western blot analysis conformation of membrane expression was not guaranteed. The remaining targets IGSF9 and KLRG2 showed expression in whole cell lysates of the cell lines (**Figure 5.2.1 A**) and also showed expression in the membrane enriched fractions (**Figure 5.2.1 B**).

Specifically IGSF9 was expressed in the MDA-MB-157, T47D and BT474 cell lines and in the membrane enriched fraction of MDA-MB-157 and BT474. KLRG2 was expressed in a wider number of cell lines tested appearing in all except MDA-MB-361 in the whole lysate preparations and also not being expressed in the membrane fraction of the MDA-MB-361, MDA-MB-157 and BT20 cell lines. While these did not totally align to the subtypes we expected from the bioinformatic data the targets did at least show higher expression in a select number of breast cancer cell lines and were also associated with the membrane enriched fractions of a subset of these breast cancer cell lines with KLRG2 showing expression across a greater number of the cell lines tested than IGSF9.

			IGSF9		KLRG2	
Cell line	**Subtype**	**Hormone status*** (Chavez, Garimella and Lipkowitz 2010)(Chavez, Garimella and Lipkowitz 2010)(Chavez, Garimella and Lipkowitz 2010)	**Whole cell lysate**	**Membrane enriched fraction**	**Whole cell lysate**	**Membrane enriched fraction**
BT20	Basal A	TNBC	NO	NO	YES	NO
MDA-MB-468		TNBC	NO	NO	YES	YES
HCC-1954		HER2+	-	-	YES	YES
MDA-MB-231	Basal B	TNBC	NO	NO	YES	
MDA-MB-157		TNBC	YES	YES	YES	NO
HS578T		TNBC	NO	-	YES	YES
T47D	Luminal	ER+, PR+	YES	NO	YES	YES
MCF7		ER+, PR+	NO	NO	YES	YES
MDA-MB-361		ER+, HER2+	NO	NO	NO	NO
BT474		ER+, PR+, HER2+	YES	YES	YES	YES
SKBR3		HER2+	NO	-	YES	YES

Table 5.2.1 Cell lines used for validation of target expression with IGSF9 and KLRG2 expression indicated. 11 cell lines representing the TNBC, HER2+, ER+ and PR+ subtypes were chosen to validate the expression of the 5 targets chosen KLRG2, LRP8, TSPAN13, IGSF9 and SLAMF8. A whole cell lysate (n=2) and membrane enriched lysate (n=2) were analysed for each cell line. Cell lines are outlined above together with their reported subtype. (Neve et al. 2006, Chavez, Garimella and Lipkowitz 2010)*

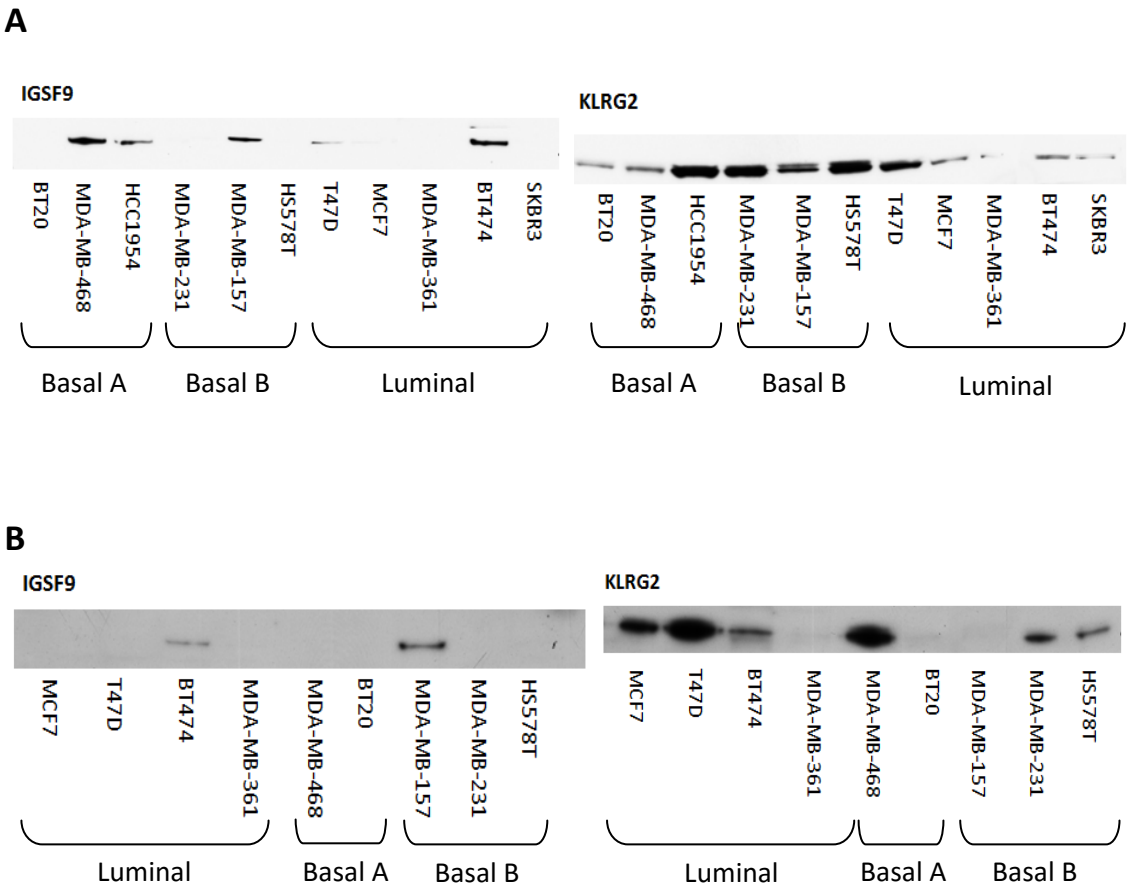


Figure 5.2.1 Western blot analysis showing differential protein expression of IGSF9 and KLRG2 across a panel different breast cancer cell lines (A) whole cell preparation and (B) membrane enriched fractions. IGSF9 was shown to be expressed in fewer cell lines in both whole cell lysates (5/11) and membrane enriched fractions (2/9) compared to KLRG2 which showed expression in almost all the whole cell preparations (10/11) and over half the membrane enriched fractions (6/9).

5.3 Preliminary IHC analysis

Having confirmed the expression of two potential over expressed breast cancer targets IGSF9 and KLRG2 in a panel of breast cancer cell lines it was necessary to determine their expression in normal breast compared to breast cancer tissues. Optimisation of initial five targets KLRG2, LRP8, TSPAN13, IGSF9 and SLAMF8 using invasive breast tissue samples indicated that IGSF9 and KLRG2 produced strong reproducible staining, SLAMF8 produced inconsistent staining and TSPAN13 and LRP8 did not produce strong staining in the small number of tissues used for optimisation. As tissue samples were limited and due to the consistent and highly prevalent staining observed with IGSF9 and KLRG2 it was decided to carry out further IHC analysis of these two targets. Note that all samples were obtained from Prof. Susan Kennedy with ethical approval.

5.3.1 Normal breast vs breast tumor

Using invasive (HER2 and TNBC status unknown) breast tumor (n=4) sections and normal breast tissue (n=3) sections we showed that IGSF9 produced membrane staining as well as granular cytoplasmic staining in these tumors which compared observed weak immunoreactivity in the normal breast tissue stained (**Figure 5.3.1**). KLRG2 was found to produce membrane (and some nuclear staining) in tumor sections compared to weak to negative staining in normal breast tissue (**Figure 5.3.2**). In summary the invasive breast tumors stained produced strong positive staining for both targets compared to normal breast tissue.

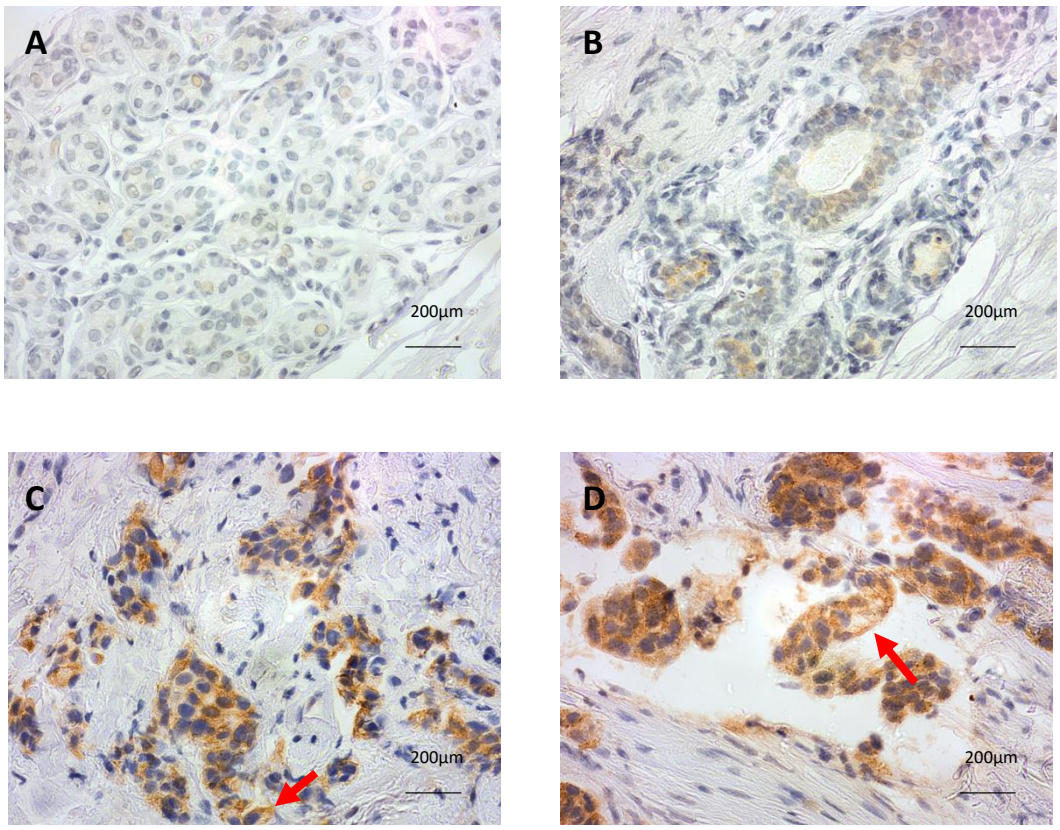


Figure 5.3.1 Representative normal (A, B) and invasive breast tissue (C, D) (subtype unknown) images of IHC analysis of IGSF9. Strong cytoplasmic and possible membrane immunoreactivity (red arrows) can be observed in both invasive breast cancer sections compared to the normal breast. (original magnification 40x).

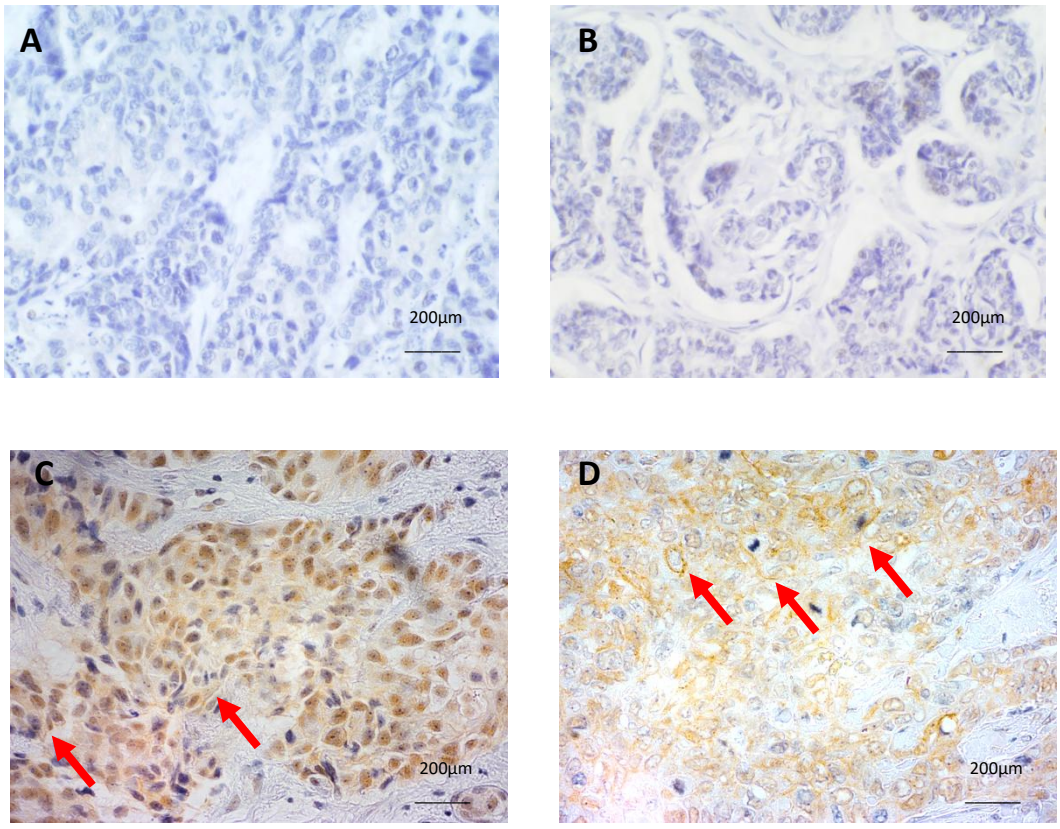


Figure 5.3.2 Represenative normal breast tissue (A, B) and invasive breast cancer tissue (C, D) (subtype unknown) images of IHC analysis of KLRG2.. Some weak KLRG2 immunractivity in some areas can be seen in normal breast (40x original magnification). Clear KLRG2 membrane immunoreactivity is observed in images C & D (red arrow).

5.3.2 Candidate target expression in proliferative tissues

Comparing the immunoreactivity of IGSF9 and KLRG2 in normal breast to invasive breast we found that that breast cancer tissue exhibited intense staining while normal tissue was found to show very limited staining. To further test the specific immunoreactivity of IGSF9 and KLRG2 we stained normal colon sections for Ki67, a widely used IHC marker for proliferating cells. Ideally a target used as an antibody drug conjugate (ADC) or for any targeted treatment will not target proliferative cells and potentially cause unwanted side effects. By comparing the same area of the section stained with IGSF9 or KLRG2 we could determine if either target had an affinity for proliferative cells. **An ideal ADC target will have limited expression in highly proliferating cells.**

Using normal colon cells tissue sections (n=3) we showed that both KLRG2 and IGSF9 were not expressed in Ki-67 proliferating positive cells (i.e. highly proliferating cells) in the photographed location of the tissue section (**Figure 5.3.3**).

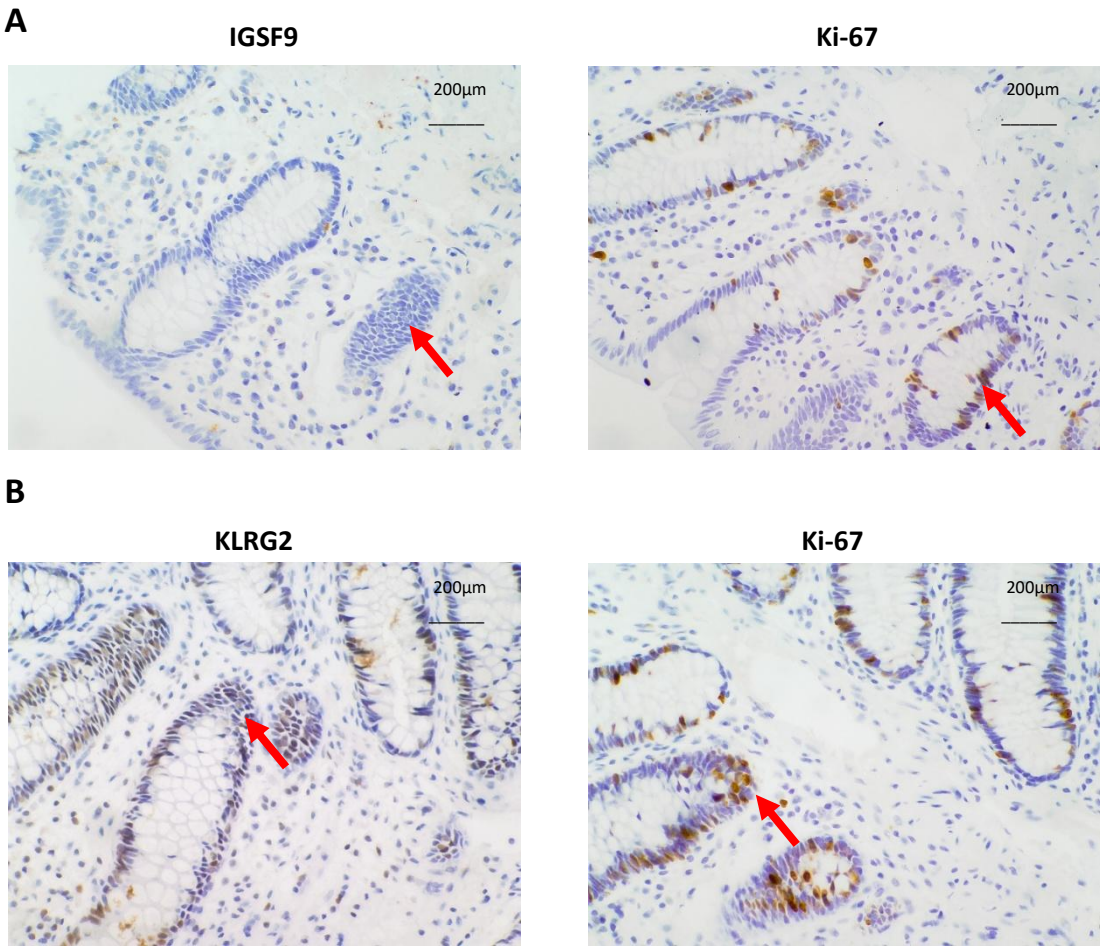


Figure 5.3.3 Normal colon images showing IHC analysis of IGSF9 (A) and KLRG2(B) in normal colon tissue sections stained with the Ki-67 proliferation marker. IGSF9 and KLRG2 show negligible immunoreactivity in Ki-67 positive cells. (original magnification 40x).

5.3.3 Normal tissue expression

We next validated the expression of IGSF9 and KLRG2 in a panel of *available* normal tissues. A panel of normal tissues including normal breast (See **Figure 5.3.1 and 5.3.2** compared to invasive breast) cervix, colon, duodenum, gastric, lung, oesophagus, prostate, salivary gland, spleen, tonsil were stained with anti-IGSF9 (**Figure 5.2.4 and 5.2.5**) or anti-KLRG2 (**Figure 5.2.6 and 5.2.7**). Overall, weak to no immunoreactivity was observed in all the of the normal tissues stained for both IGSF9 and KLRG2 (**Table 5.3.1 and 5.3.2**). Strong IGSF9 immunoreactivity was observed in duodenum but in <10% of the tissue. Only analysis of IGSF9 in the digestive tissues (duodenum, oesophagus and salivary gland) produced moderate staining similar to that seen in **Figure 5.2.2** with invasive breast.

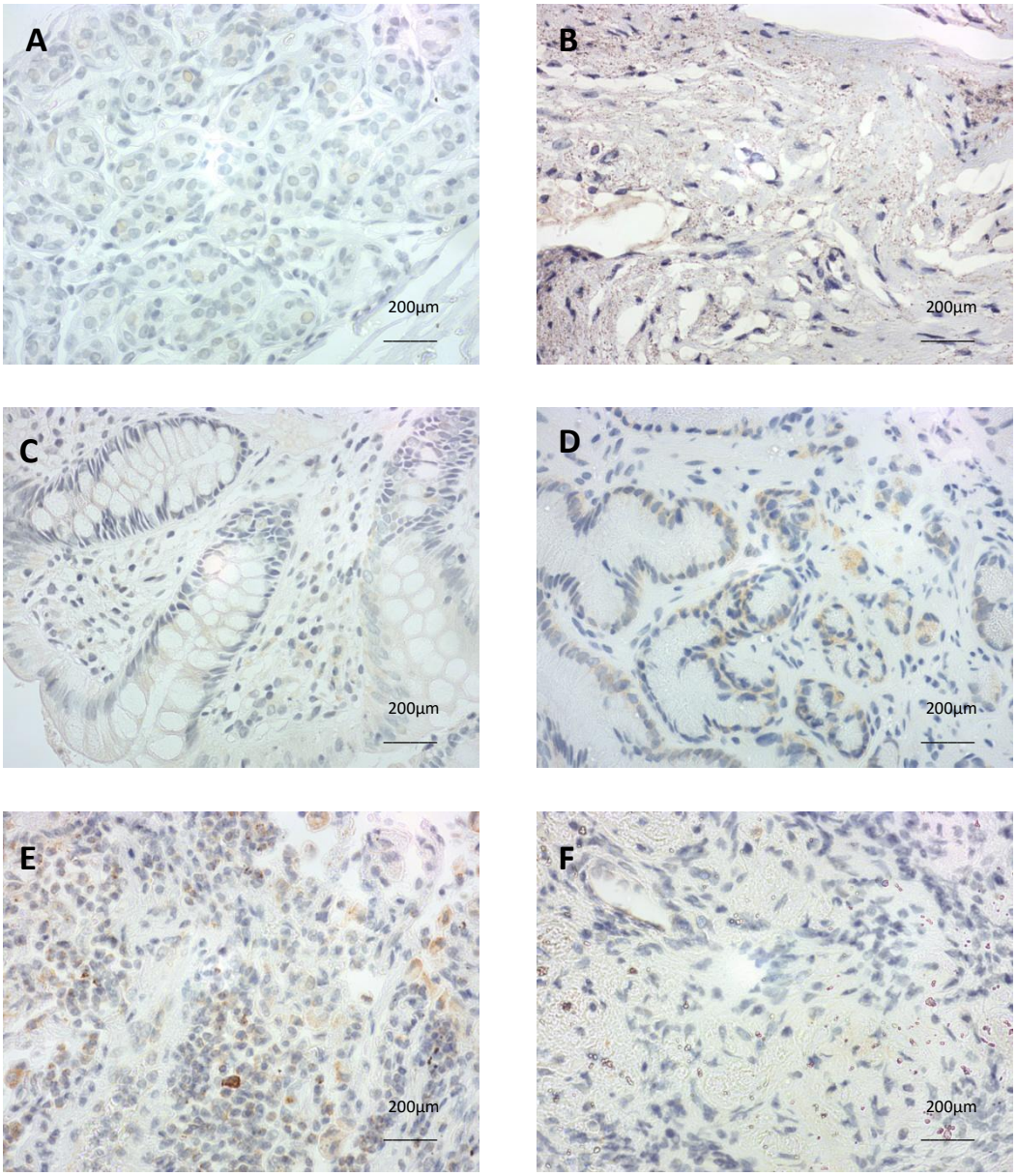


Figure 5.3.4 IGSF9 immunoreactivity across representative normal tissues. Breast (A) cervix (B), colon (C), prostate (F) all show negligible immunoreactivity with gastric (D) and lung (E) showing weak immunoreactivity. (Original magnification 40x).

Normal Tissue	KLRG-2 membrane immunoreactivity
Breast	Negative (4/4)
Colon	Weak 20% of tissue (2/2)
Gastric	Negative (1/1)
Tonsil	Negative (2/2)
Prostate	Negative (2/2)
Spleen	Weak 20% of tissue (2/2)
Duodenum	Strong staining <10% of tissue (1/1)
Lung	Moderate staining 10% of tissue (3/3)
Kidney	Negative (1/1)
Liver	Negative (1/1)

Table 5.3.1 IGSF9 immunoreactivity across a cohort of 18 available normal tissues. Overall observed immunreactivity was mostly negative, weak with low coverage (Colon, Spleen). Moderate and strong immunoreactivity was observed in lung and duodenum at 10% and <10% coverage respectively. This data demonstrates overall that IGSF9 is not highly expressed in thenormal tissue.

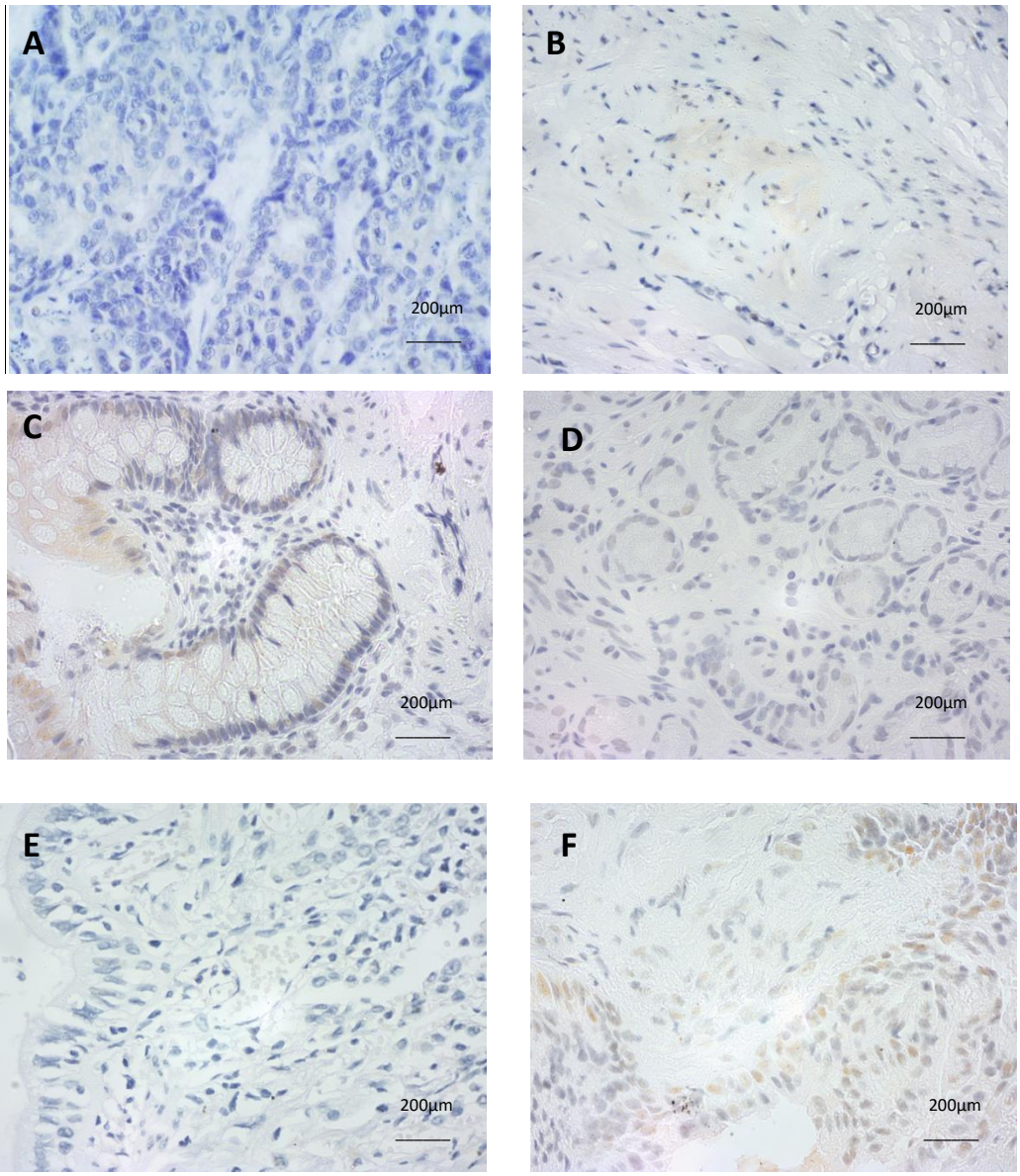


Figure 5.3.5 KLRG2 immunoreactivity across representative normal tissues. Breast (A) cervix (B), gastric (D), lung (E), all show negligible immunoreactivity. Colon (C) and prostate (F) show weak immunreactivity. (Original magnification 40x).

Normal Tissue	KLRG-2 membrane immunoreactivity
Breast	Negative (4/4)
Gastric	Weak <10% of tissue (1/1)
Tonsil	Negative (2/2)
Prostate	Negative (5/5)
Spleen	Negative (1/1)
Duodenum	Moderate 10% of tissue (1/1)
Lung	Negative (2/2)
Kidney	Weak 60% of tissue (1/2)
Liver	Negative (2/2)
Cerebellum	Negative (1/1)
Cerebrum	Negative (1/1)
Colon	Weak 20% of tissue (4/4)
Muscle skeletal	Negative (1/1)
Ovary	Negative (2/2)
Pancreas	Negative (1/1)
Skin	Negative (2/2)
Small inrestine	Negative (1/1)
Stomach	Weak 60% of tissue (1/1)
Spleen	Negative (1/1)
Thyroid gland	Negative (1/1)

Table 5.3.2 KLRG2 immunoreeactivity across a cohort of 36 available normal tissues. The highest coverage of staining was observed in stomach and kideny at 60% but consisted of weak immunoreactivity. Moderate immunoreactivity was only seen in duodenum in 10% of the tissue. From these numbers KLRG2 was observed to not be highly associated with normal tissue.

5.4 IGSF9

As one of the five short listed potential membrane targets from the bioinformatics transcript lists IGSF9 was found to be 18 fold up-regulated in HER2+ breast cancer compared to normal breast tissue (**Table 6.1.2**). Using a panel of breast cancer cell lines we confirmed IGSF9 expression by Western blot analysis in the membrane enriched fractions and unfractionated whole cell lysates of MDA-MB-157 and BT474 cell lines (**Figure 5.2.1**). Initial IHC results revealed that IGSF9 showed strong cytoplasmic and membrane expression in invasive breast cancer (**Figure 5.3.1**), showed overall low immunoreactivity across normal tissues (**Figure 5.3.3**) with moderate staining staining in only 3 of the normal tissues investigated (**Figure 5.3.4 and 5.3.5**). Negligible staining was observed in normal breasttissues analysed. This differential cancer/normal tissue expression i.e. weak immunoreactivity in normal breast (and other normal tissues) with strong specificity for breast cancer tissue and visible membrane expression validated IGSF9 for further IHC analysis using a larger cohort of breast cancer tissues wiht clincopathological information available (courtesy of our clinical collaborators).

This preliminary study showed that IGSF9 had moderate to strong cytoplasmic and membrane immunoreactivity in a panel of HER2+ breast cancer sections (n=6) (**Figure 5.4.1 A,B,C**). We also tested IGSF9 expression in triple negative breast cancer (TNBC) (n=11) and found that it also produced moderate to strong cytoplasmic and membrane staining (**Figure 5.4.1 D,E,F**). We also stained several other cancer tissues (**Figure 5.4.2**) and observed moderate staining in primary melanoma (n=1) compared to metastatic melanoma (n=1) and showed weak to moderate immunoreactivity in one pancreatic tumor (n=1).

As IGSF9 appeared to be expressed in all of the HER2+ sections teasted, analysisof a larger patient cohort was conducted using a tissue matrix array (TMA) (**Figure 5.4.3 - 5.4.6**) containing 102 sample cores and a separate TMA containing a cohort of 70 HER2+ patient sample cores (**Figure 5.4.7 - 5.4.11**). (Prof. Susan Kennedy and Prof. Joe Duffy, St Vincents University Hospital)

5.4.1 IGSF9 expression in HER2+ and TNBC tumors

IGSF9 was found to produce moderate to strong staining in both HER2+ breast cancer tissues and TNBC tissues (**Figure 5.3.1**). Membrane staining was observed in both HER2+ and TNBC tissues but punctuate cytoplasmic staining was also observed. Furthermore increased staining intensity was observed in tumors with a Her2+ score of 0 (A) up to score of +2 (B) and +3 (C). This is consistent with IGSF9 being up-regulated in HER2+ breast cancer indicated by the initial bioinformatics list it was derived from. It did not however appear as up-regulated in the TNBC bioinformatics list despite IGSF9 showing very strong immunoreactivity in TNBC. The following table (**Table 5.3.1**) summarises the number of tissues stained and the intensity of staining produced:

Breast Cancer Subtype	**No. producing staining**	**Localisation of staining**
HER2+	**4/6** (moderate-strong)	Membrane, Cytoplasmic
TNBC	**9/11** (moderate-strong)	Membrane, Cytoplasmic

Table 5.4.1 Summary of numbers of HER2+ and TNBC tissues producing staining with IGSF9 and their localisation. While IGSF9 was identified as over expressed in HER2+ breast cancer compared to normal tissue from the initial bioinformatics data it also showed similar staining in 9/11 TNBC tissues stained.

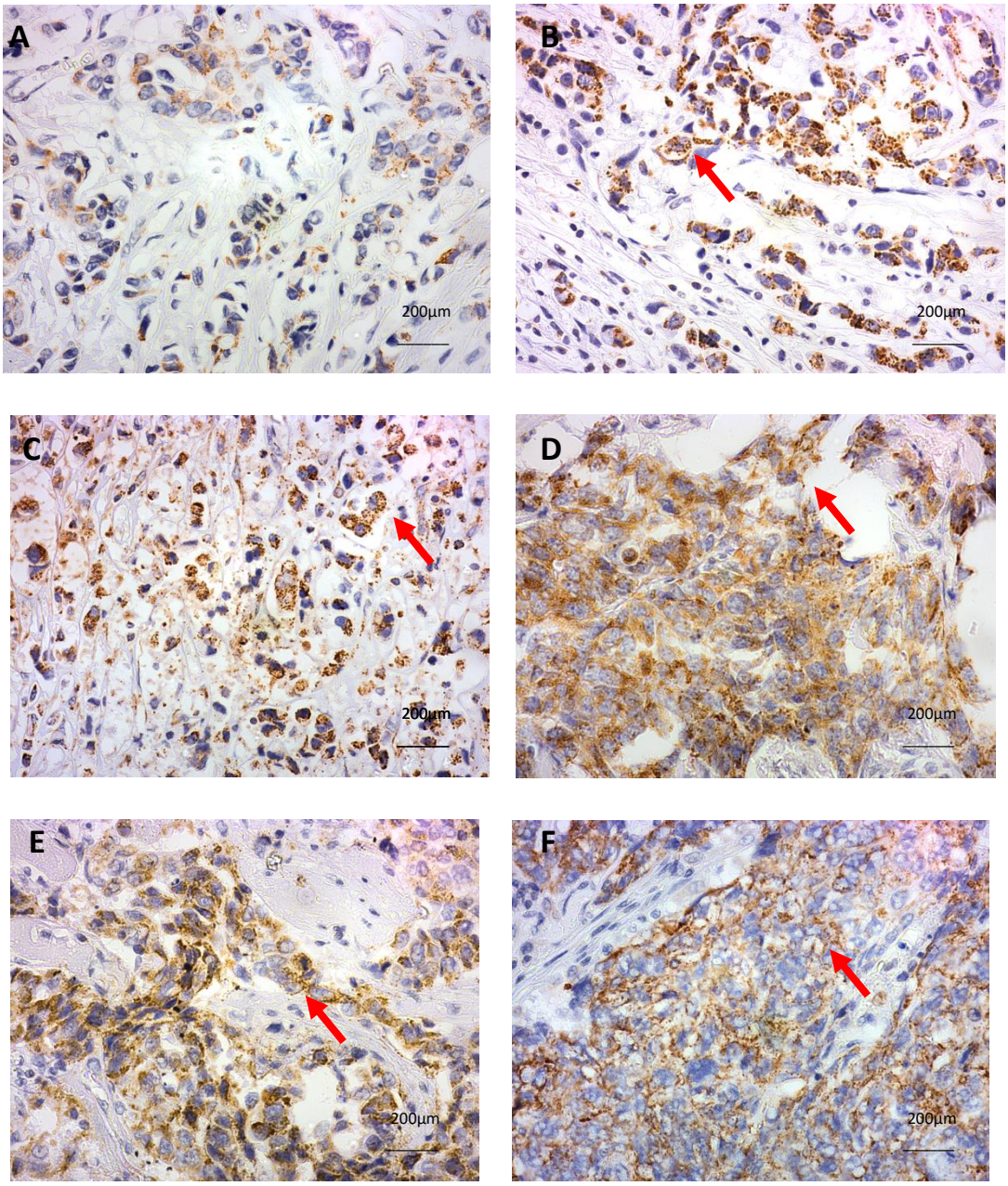


Figure 5.4.1 IGSF9 staining of HER2+ breast tumors. The top image shows a tissue section with HER2+ score of 0 (A) showing and weak IGSF9 immunoreactivity with while tumors with a HER2+ score of +2 (B) and +3 (C) respectively show strong IGSF9 immunoreactivity. IGSF9 staining of TNBC breast cancer tissue samples (D,E,F) show strong "punctuate" like) immunoreactivity in the cytoplasm as well as membrane straining (examples shown) in all samples stained. (original magnification 40x).

5.4.2 IGSF9 expression in other cancer tissues

IGSF9 was observed to be expressed in TNBC tissues despite only being predicted to be over expressed in HER2+ breast cancer. For this reason and it's novel status in the literature in relation to cancer functional work we tested a small number of available cancer tissues (**Figure 5.4.2**).

We observed that IGSF9 showed strong membrane staining in primary melanoma (B) compared to paired metastatic melanoma. We also found that it had weak to moderate membrane and cytoplasmic immunoreactivity in pancreatic cancer (C,D). IGSF9 has been identified in mice to play a positive role in cell adhesion through synapse maturation and outgrowth (Mishra et al. 2008, Mishra et al. 2014). This may explain why the metastatic melanoma shows no staining for IGSF9 as metastatic cells are known to have reduced adhesion properties (Bendas and Borsig 2012).

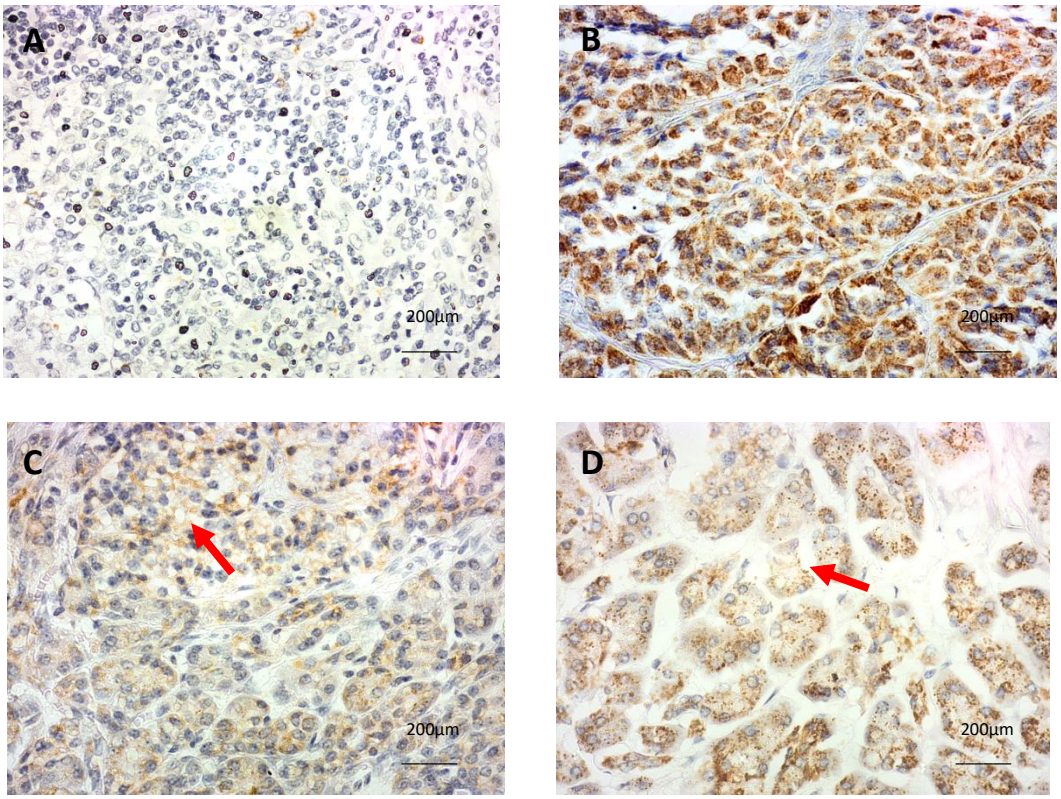


Figure 5.4.2 IGSF9 staining of metastatic (A), primary melanoma (B) and pancreatic cancers sections (C, D). The metastatic melanoma section shows negligible immunoreactivity compared to the strong membrane and cytoplasmic staining observed in the primary section melanoma. Pancreatic cancer sections show varying immunoreactivity from weak (C) to moderate membrane staining (D). (Original magnification 40x)

5.4.3 TMA IHC analysis of IGSF9

TMAs contain many cores of separate paraffin fixed tissue blocks allowing for multiple samples to be analysed at once. TMAs (constructed at St Vincents University Hospital with respective cores from HER2+ patients and TNBC patients) were used for further IHC analysis of IGSF9 with ethical approval from St Vincents University hospital. The first TMA consisted of a selection of breast cancer types including HER2 positive, TNBC and normal breast tissue (**Table 5.4.2**), both TMAs containing clinical data such as tumour size and receptor status. The following table summarises the constituents of each TMA:

No. Patients	Normal Breast	HER2+	ER+	PR+	TNBC
102	2	23	59	48	16

Table 5.4.2 Tissue matrix array sample numbers for normal breast and HER2+, ER+, PR+ and TNBC tumor tissue.

The TMA was stained using an IGSF9 antibody and resulted in the characteristic membrane and cytoplasmic staining observed in the preliminary IHC analysis. The stained TMA who scored according to intensity (+1 to +3 from weak to strong) and for coverage of staining (% area of tumor tissue) graphed in as x y z plots (**Figure 5.4.3 to 5.4.6**) the HER2+ and TNBC group showed the highest relative number of stained samples.

The second TMA consisted of a large cohort (n=70) of HER2+ archival breast cancer samples (**Table 5.4.3**) with ER+/ER- and LN+/LN- status known for each sample.

No. Patients	HER2+	HER2+, ER+	HER2+, ER-	HER2+, LN+	HER2+, LN-
69	69	43	26	33	35

Table 5.4.3 HER2+ breast tumor tissue matrix array sample numbers with ER2+, ER-, LN+ and LN- tumor status.

Both TMAs allowed us to determine if IGSF9 expression is biased towards a particular breast cancer subtype. The multi breast cancer TMA suggests that IGSF9 exhibits higher staining intensity associated with HER2+ and TNBC tumors ompared to ER+ and PR+ tumors which show a higher frequency of a negative staining score. The HER2+ TMA shows that a similar trend of mostly moderate staining associated with HER2+ tumors regardless of ER or LN expression. The association between IGSF9 and

strong staining in HER2+ tissues and TNBC tissues was in agreement with that observed in preliminary IHC analysis (**Table 5.4.1**).

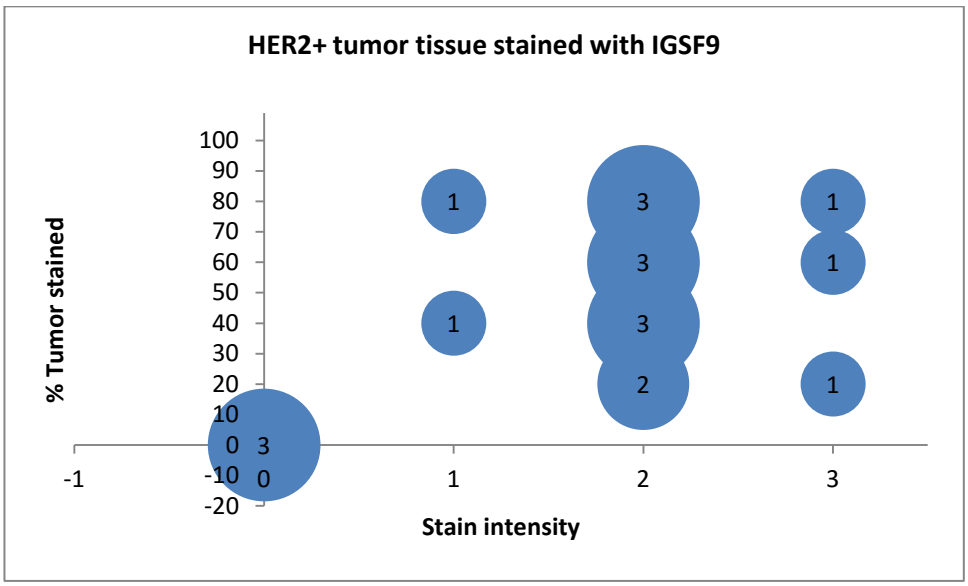


Figure 5.4.3 HER2+ TMA tissue cores (n=19) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. Low numbers of negatively stained tissues (n=3) and a high number of moderately stained tissues suggest that IGSF9 expression may be associated with HER2+ breast cancer.

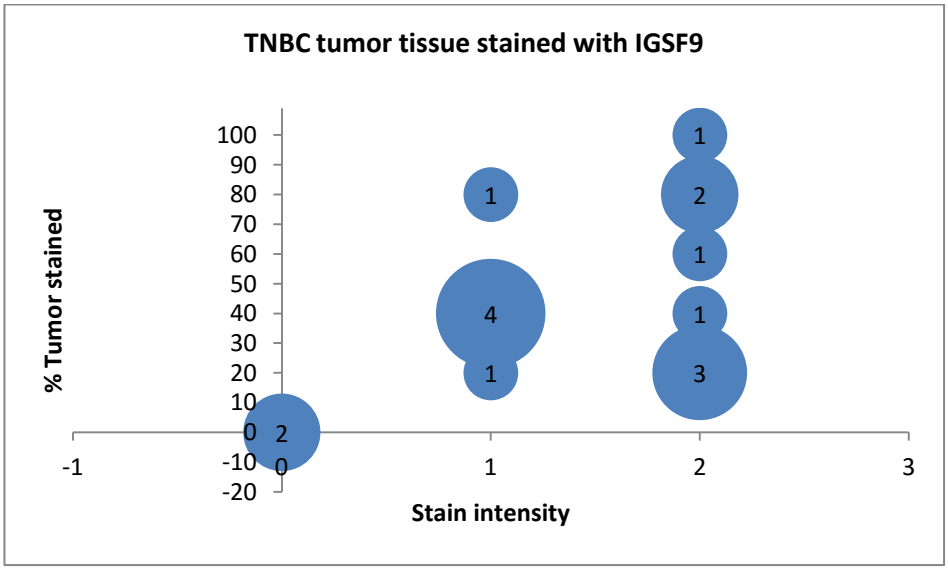


Figure 5.4.4 TNBC TMA tumor cores (n=16) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. Low numbers of negatively stained tissues (n=2) and a high number of moderately stained tissues suggest IGSF9 expression may be associated with TNBC breast cancer.

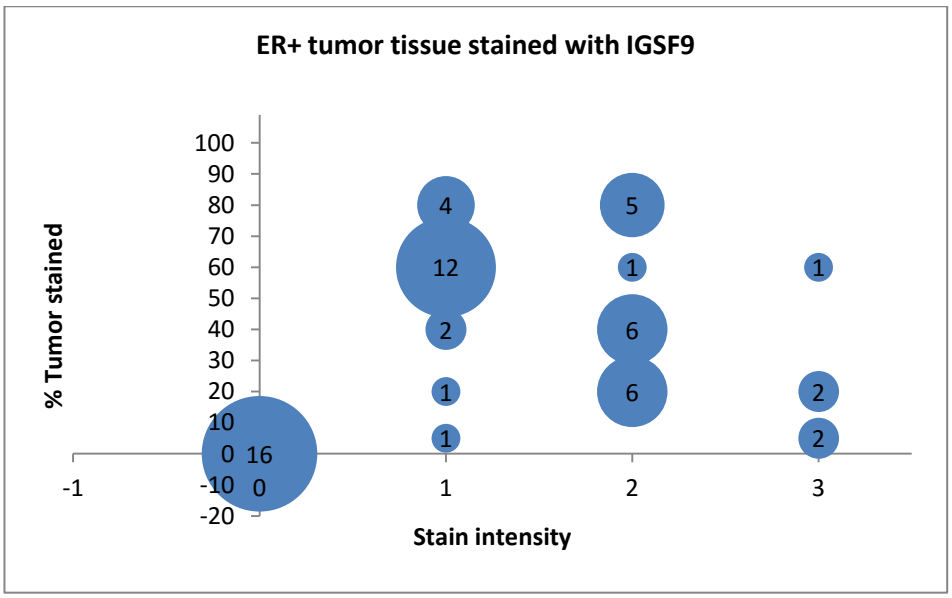


Figure 5.4.5 ER+ TMA tumor cores (n=59) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. High numbers of negative stained tissues (n=16) and a bias toward weak stained tissues suggest IGSF9 expression may not be associated with ER2+ breast cancers.

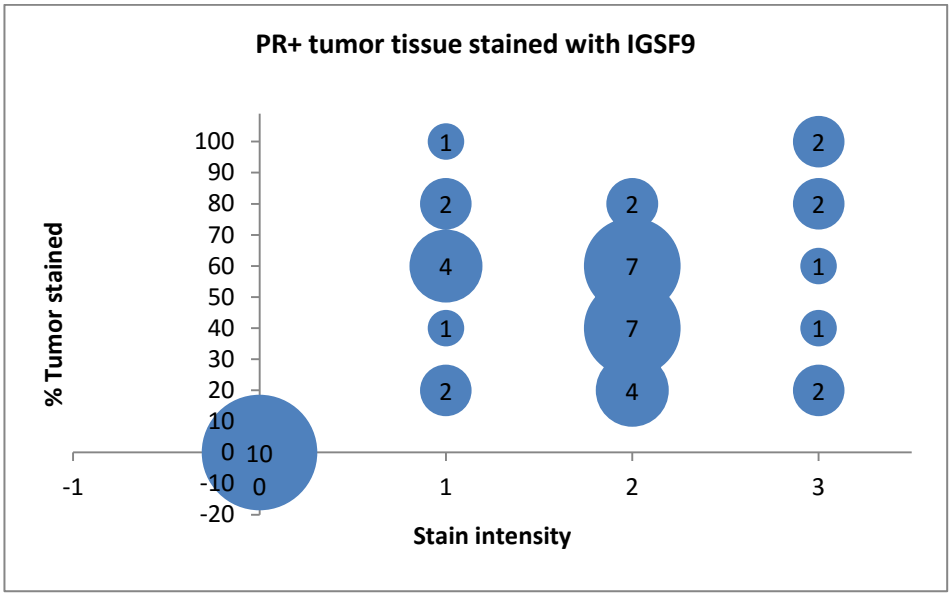


Figure 5.4.6 PR+ TMA tumor cores (n=48) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. High numbers of negative stained tissues (n=10) and with only a bias toward moderate stained tissues suggest IGSF9 expression may be weakly associated with PR+ breast cancers.

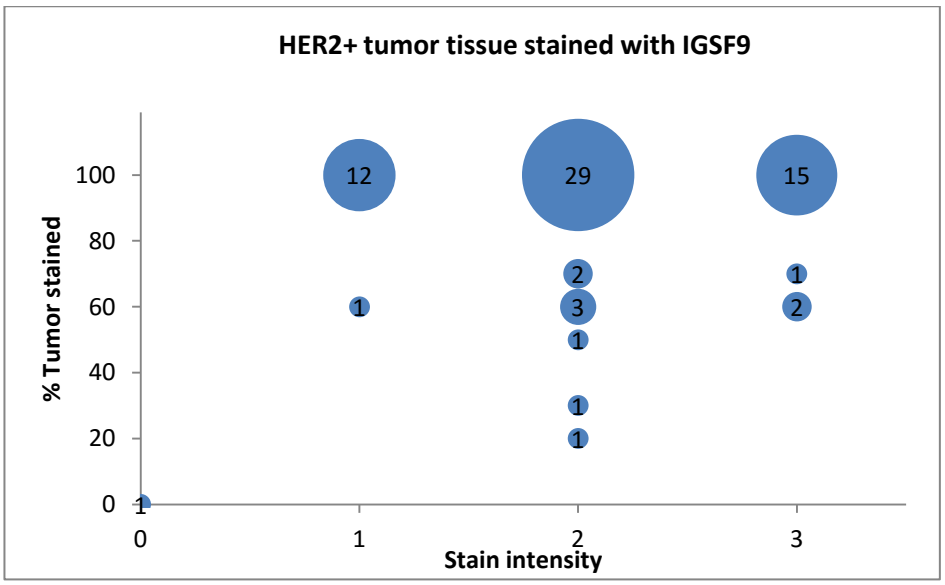


Figure 5.4.7 HER2+ TMA tumor cores (n=69) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. High moderately stained tissues and high percentage of staining coverage suggest IGSF9 expression may be associated with HER2+ breast cancers.

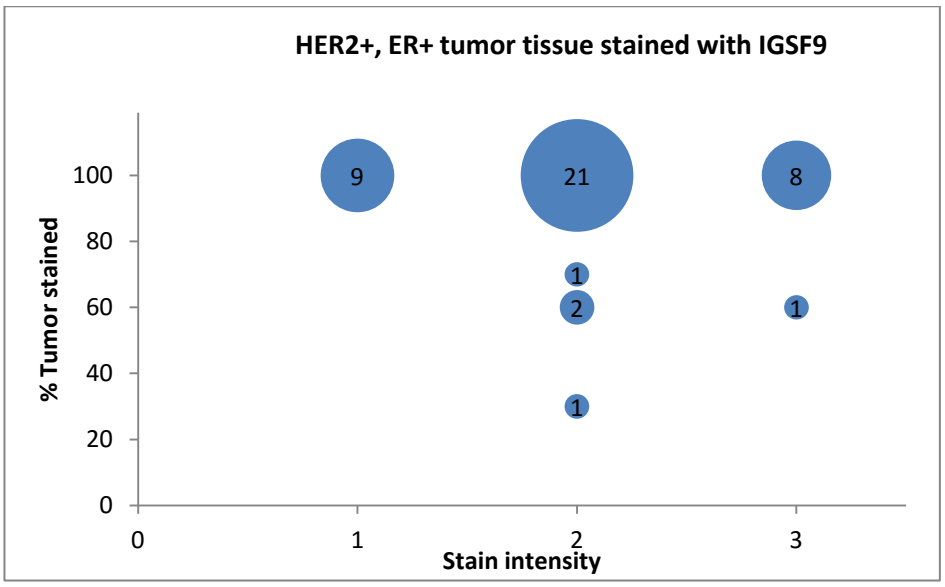


Figure 5.4.8 HER2+, ER+ TMA tumor cores (n=43) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. Exhibiting a similar distribution in staining intensity and staining coverage to the HER2+ mixed population (**Figure 5.4.7**) suggest IGSF9 is not specifically associated with ER+ expression in HER2+ breast cancer.

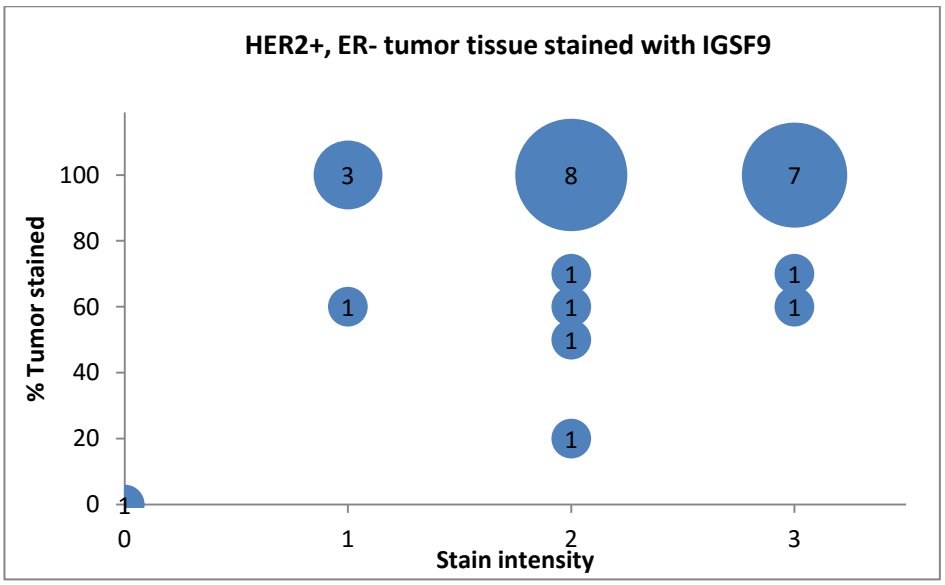


Figure 5.4.9 HER2+, ER- TMA tumor cores (n=26) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. Exhibiting a similar distribution in staining intensity and staining coverage to the HER2+ mixed population (**Figure 5.4.7**) suggest IGSF9 is not specifically associated with ER- expression in HER2+ breast cancer.

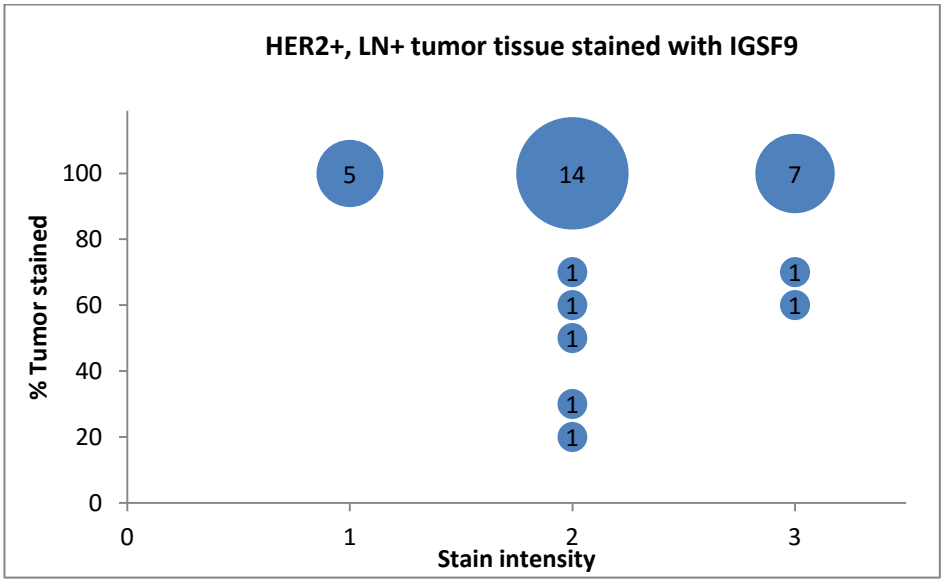


Figure 5.4.10 HER2+, LN+ TMA tumor cores (n=33) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. Exhibiting a similar distribution in staining intensity and staining coverage to the HER2+ mixed population (**Figure 5.4.7**) suggest IGSF9 is not specifically associated with LN+ expression in HER2+ breast cancer.

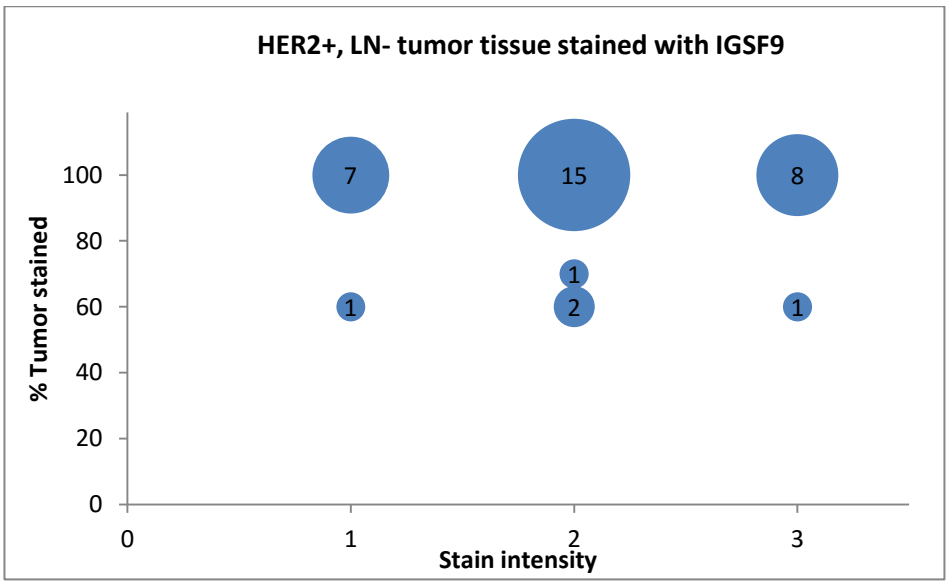


Figure 5.4.11 HER2+, LN- TMA tumor cores (n=35) stained for IGSF9 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane and cytoplasmic staining was observed in all IGSF9 positive tissues. Exhibiting a similar distribution in staining intensity and staining coverage to the HER2+ mixed population (**Figure 5.4.7**) suggest IGSF9 is not specifically associated with LN- expression in HER2+ breast cancer.

5.5 KLRG2

The second target chosen for follow up from the five short listed potential membrane targets from the bioinformatic transcript lists was KLRG2. It was found to be 5 fold up-regulated in TNBC compared to normal breast tissue (**Table 5.1.2**). Our panel of breast cancer cell lines confirmed IGSF9 expression by Western blot analysis in whole cell lysates as well as in membrane enriched fractions of MCF7, T47D, BT474, MDA-MB-468, MDA-MB-231 and HS578T cell lines (**Figure 5.2.1**). Initial IHC results showed that KLRG2 had strong cytoplasmic and membrane expression in two out of four invasive breast cancer tissues tested (**Figure 5.2.3**), was not preferentially associated with binding to proliferative tissue (**Figure 5.2.4**) and only showed low/negligible staining in all normal tissue types tested (**Figure 5.2.6**) with negligible staining in normal breast. From these results KLRG2 exhibited promising characteristics as a molecular ADC target exhibiting with low immunoreactivity in normal tissues and strong specificity for breast cancer tissue with visible membrane expression.

Using available breast cancer tissues (subtype known) we analysed the expression of KLRG2 in TNBC and HER2+ breast cancer. As KLRG2 was derived from the TNBC vs normal list we anticipated that it's expression would by particularly high in TNBC tissue sections. We found that KLRG2 displayed strong immunoreactivity in TNBC tissue sections with distinct membrane and nuclear staining visible (**Figure 5.5.1 D,E,F**). We also found that KLRG2 had low to negligible membrane and nuclear immunoreactivity to HER2+ breast cancer sections (**Figure 5.5.1 A,B,C**). We also stained a small no. of other cancer tissues (**Figure 5.4.2**) and found it showed low to negligible staining in primary melanoma (n=1) compared to metastatic melanoma (n=1) and it also showed weak to moderate immunoreactivity in pancreatic cancer (n=1).

As in **Section 5.4** with IGSF9 we conducted a larger sample analysis with a multiple breast cancer subtype TMA (**Figure 5.5.4**) containing 102 sample cores.

5.5.1 KLRG2 expression inHER2+ and TNBC tumors

KLRG2 was shown to produce negative to low staining in HER2+ breast cancer tissues from the tumors with a HER2+ score of 0 (A) up to score of +2 (B) and +3 (C) (**Figure 5.5.1 A, B, C**). As expected we observed strong staining in TNBC tissues (**Figure 5.5.1 D, E, F)**. Membrane staining was clearly observed in TNBC tissues and often accompanied by strong nuclear staining. The individual clinical tissue samples illustrated that KLRG2 showed an affinity toward strong membrane staining of TNBC tissues which is in agreement with the bioinformatics list that the target was derived from. The following table (**Table 5.5.1**) summarises the tissues stained and the intensity of staining produced.

Breast Cancer Subtype	**No. producing staining**	**Localisation of staining**
HER2+	**2/9** (moderate)	Membrane, Nuclear
TNBC	**8/11** (strong)	Membrane, Nuclear

Table 5.5.1 Summary of numbers of HER2+ and TNBC tissues producing staining with KLRG2 and their localisation. KLRG2 shows stronger immunoreactivity overall for TNBC from these results which is to be expected as KLRG2 was over expressed in TNBC from the bioinformatics data.

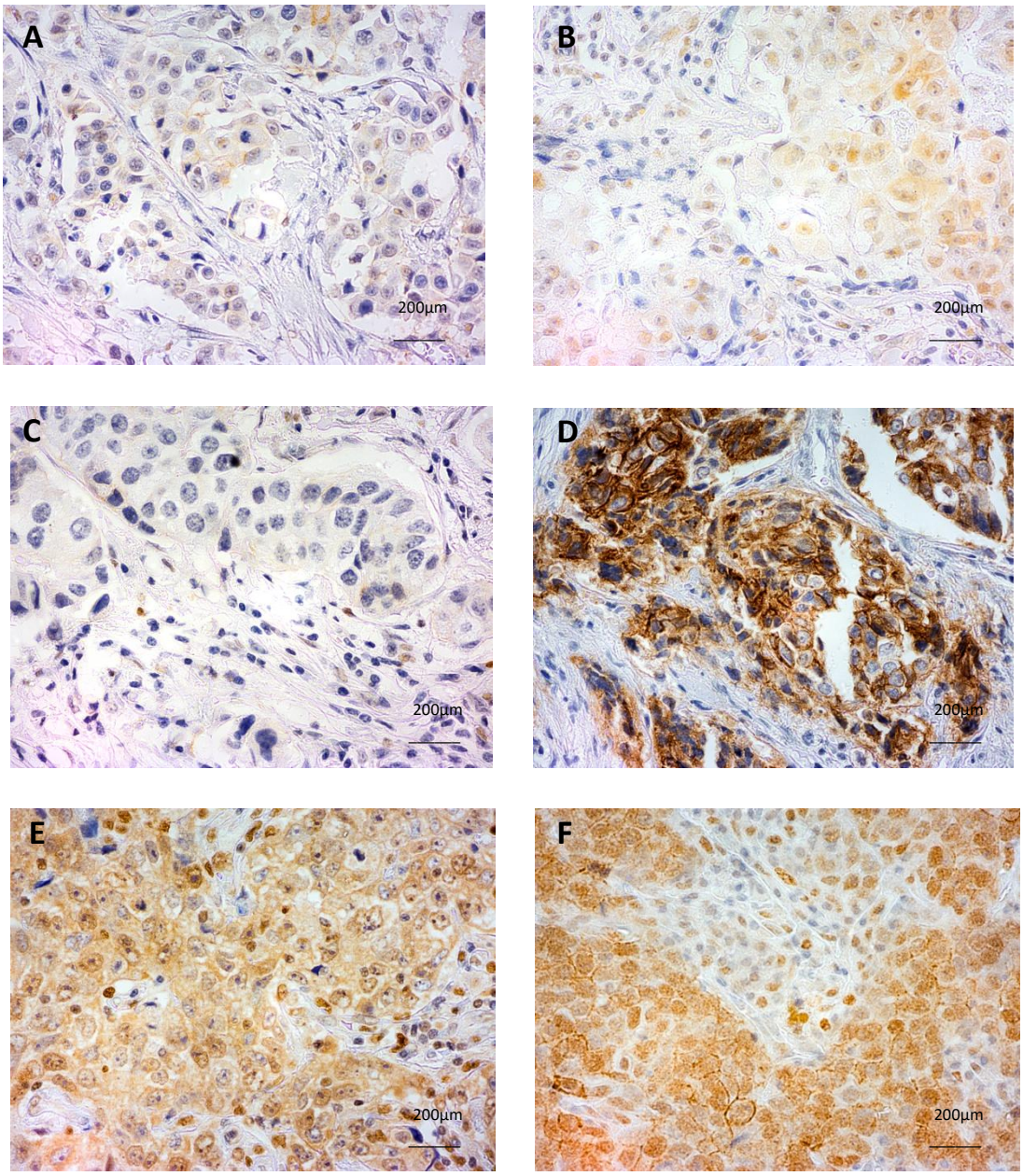


Figure 5.5.1 KLRG2 expression in HER2+ breast cancer tissues. Weak membrane and nuclear immunoreactivity can be seen in all three samples (A, B, C). KLRG2 staining on TNBC (D, E, F) tissue samples showing strong clear membrane immunoreactivity accompanied by nuclear staining (Original magnification 40x).

5.5.2 KLRG2 expression in other cancer tissue

Taking into account the results from staining melanoma and pancreatic cancer tissue with IGSF9 in **Section 5.4.2** we analysed the immunoreactivity of KLRG2 in the same tissue sections. We found that there was negligible staining in the primary and metastatic melanoma sections (A and B) and very weak to negligible staining found in the pancreatic cancer tissue (D, C). This again reflects the observation found in staining HER2+ breast and TNBC tissue, that KLRG2 is potentially more selective than IGSF9, specifically selective to high expression in TNBC.

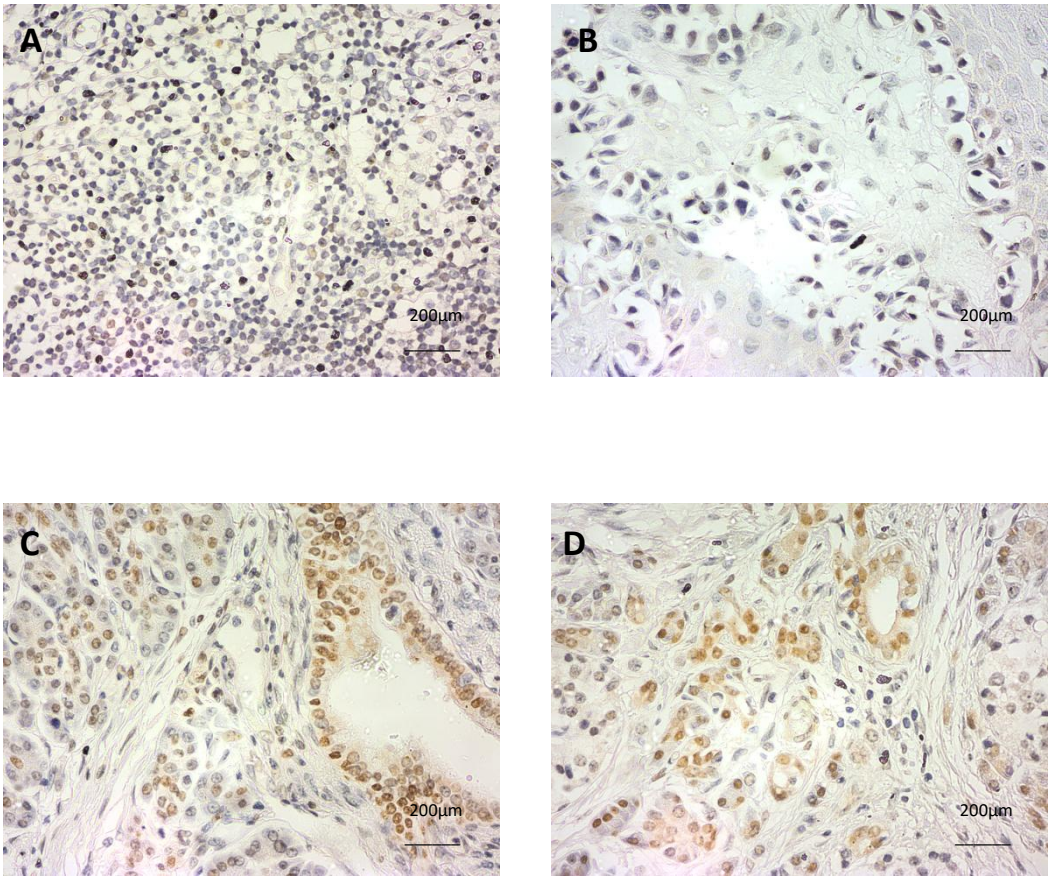


Figure 5.5.2 KLRG2 staining of metastatic (A) and primary melanoma (B) and pancreatic cancer sections (C, D). The melanoma section shows negligible immunoreactivity. Pancreatic cancers show all show weak nuclear immunoreactivity. (Original magnification 40x).

5.5.3 TMA IHC analysis of KLRG2

As in **Section 5.3.3** with IGSF9 we used the same TMA to investigate the immunoreactivity in a large cohort of breast cancer samples. These were obtained with ethical approval from St Vincents University Hospital the contents of which are summarised in the previous **Table 5.4.2.**

IHC analysis of the TMA cores for KLRG2 expression resulted in strong membrane immunoreactivity together with moderate nuclear staining in some cases; this was in agreement with the preliminary IHC analysis. The KLRG2 stained TMA was divided into four subtypes scored according to the intensity of the immunoreactivity (+1 to +3 from weak to strong) and for coverage of staining (% area of tissue core stained) and graphed as x y z plots (**Figure 5.5.3 to 5.5.6**) the HER2+ and TNBC group showed the highest relative number of tissue cores showing KLRG2 immunoreactivity.

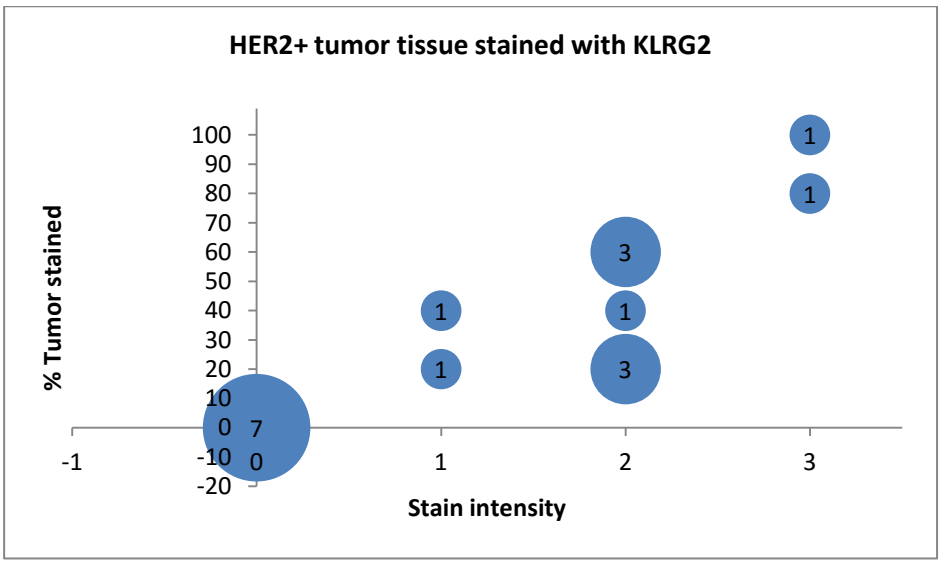


Figure 5.5.3 HER2+ TMA tumor cores (n=18) stained for KLRG2 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane staining was observed in all KLRG2 positive tissues. The observed relatively high number of negatively stained tissues (n=7) and the a low percentage of each tumor core exhibiting KLRG2 immunoreactivity suggests that KLRG2 expression may not *strongly* associated with the HER2+ breast cancer subtype.

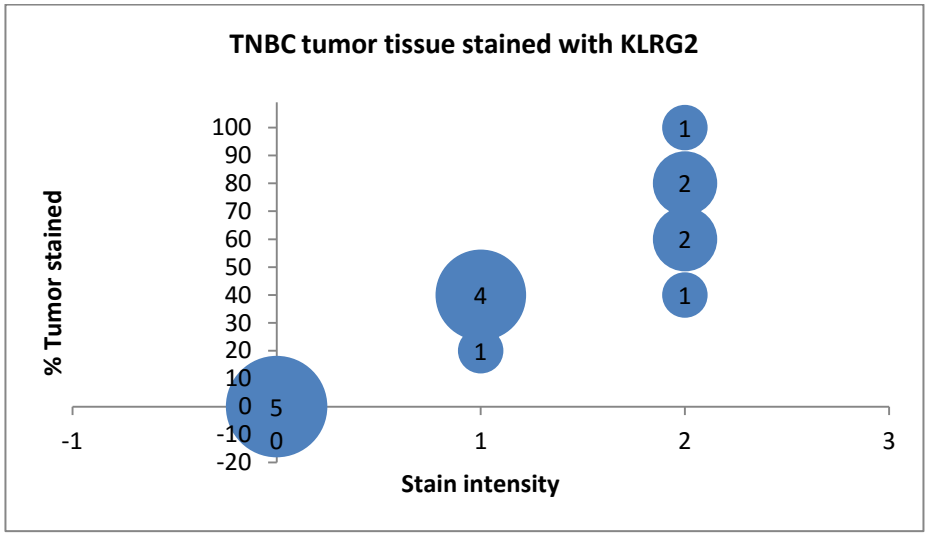


Figure 5.5.4 TNBC TMA tumor cores (n=16) stained for KLRG2 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane staining was observed in all KLRG2 positive tissues. The observed relatively high number of negatively stained tissues (n=5) which showed KLRG2 immunoreactivity together with a high percentage of stained cores showing, KLRG2 expression suggests that KLRG2 expression is associated with the TNBC breast cancer subtype.

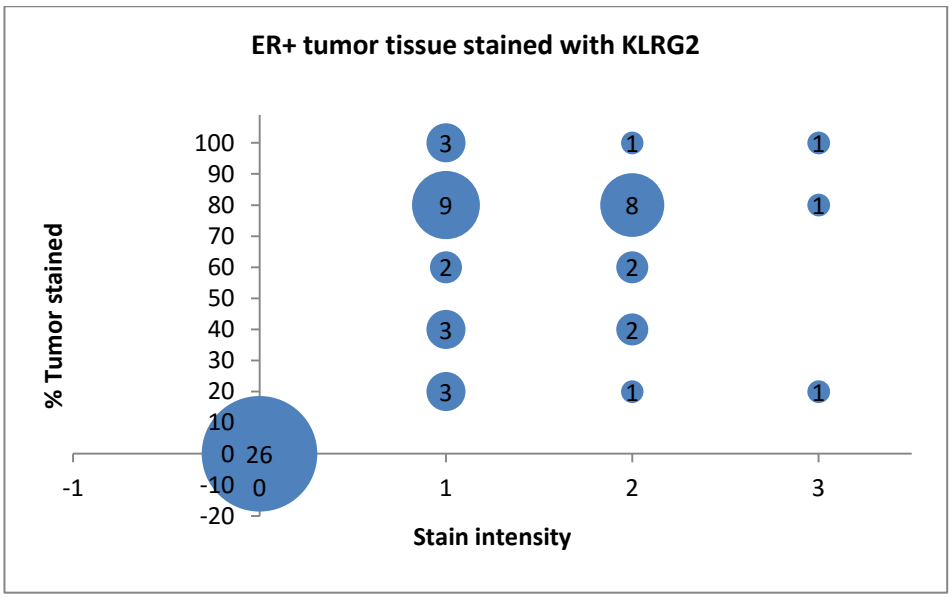


Figure 5.5.5 ER+ TMA tumor tissue (n=63) stained using KLRG2. Stain intensity was measured as weak (1), moderate (2) or high (3). Characteristic membrane staining was observed in all KLRG2 positive tissues. A relatively high number of negative stained tissues (n=26) and a wide range of percentage of each tumor stained, KLRG2 expression may not strongly be associated with ER2+ breast cancer.

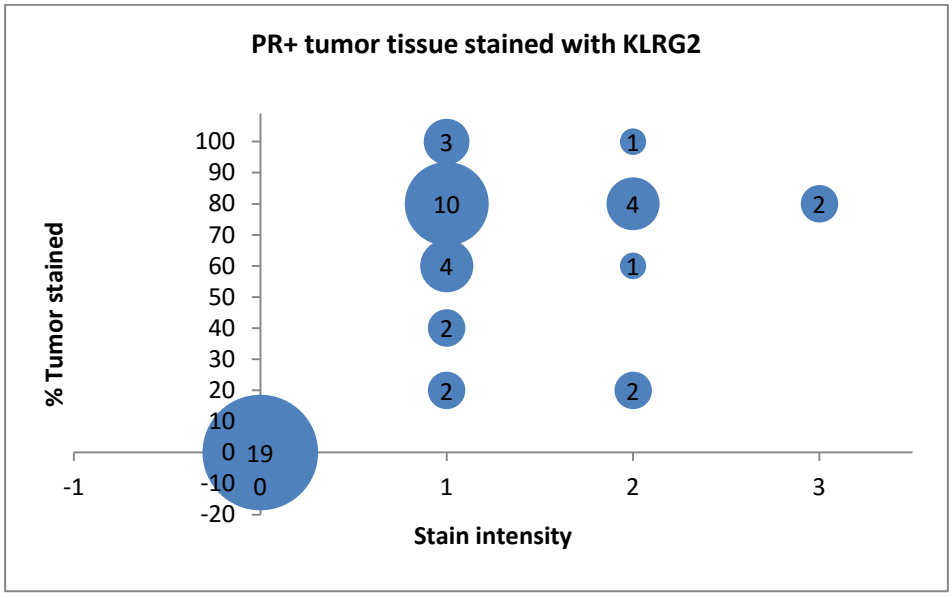


Figure 5.5.6 PR+ TMA tumor cores (n=63) stained for KLRG2 expression. Staining intensity was scored as weak (1), moderate (2) or high (3). Characteristic membrane staining was observed in all KLRG2 positive tissues. The relatively high number of negative stained tissues (n=19) together with, for the most part, weak KLRG2 immunoreactivity suggests that KLRG2 expression may not be associated strongly with the ER2+ breast cancer subtype.

5.5.4 Survival outcome

To investigate any possible correlation between the expression of KLRG2 and patient outcome/survival we used BreastMark, an algorithm available online that pools multiple gene expression microarrays to calculate and correlate the survival rate associated with differential gene expression of a specific gene (Madden et al. 2013).

The results indicate that there is a significant ($p=0.72 \times 10^{-3}$) association between high expression of KLRG2 in the basal breast cancer subtype and poor rate of survival (**Figure 5.5.7**). BreastMark has assigned a hazard ratio of 2.65 to the association between patients expressing low levels of KLRG2 and high levels of KLRG2. A hazard ratio of 2.65 equates to an over two fold increase in the risk of a poor outcome in patients expressing high levels of KLRG2 compared to those expressing low levels of KLRG2. Overall the graph indicates that patients presenting a basal type breast cancer tumor are 2.65 times more likely to have a poor outcome over a ten year period.

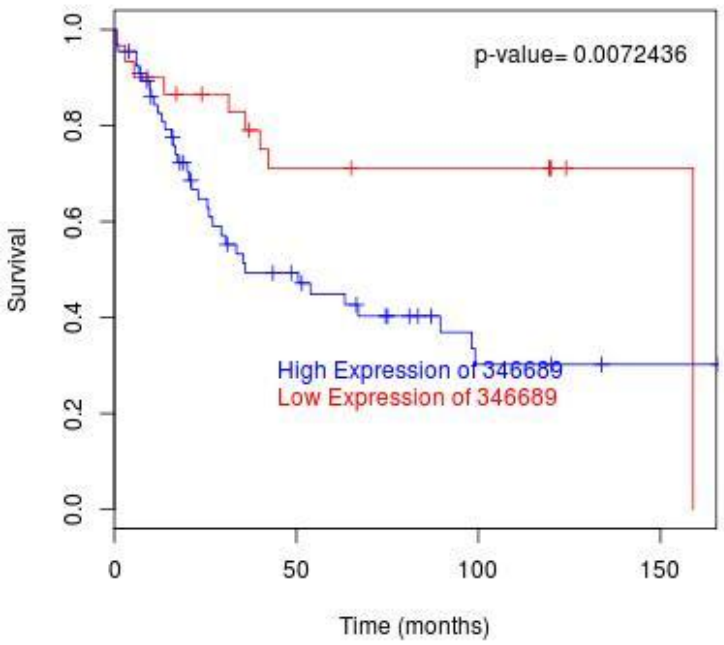


Figure 5.5.7 Kaplan-Meier cumulative survival curve showing poor patient outcome associated with KLRG2 (3466689) over expression in basal type breast cancer. The data shows a significant difference ($p=0.72 \times 10^{-3}$) in the poor prognosis of patients (n=96) with basal type breast cancer over expressing (blue) KLRG2 compared to those with low KLRG2 (red) expression. A hazard ratio of 2.65 indicates an over two fold likely hood of a poorer prognosis in these patients (Madden et al. 2013).

CHAPTER 6
Discussion

6.1 Effect of miR-7 on CHO cells introduction

To date our laboratory has established miR-7 as having a significant effect on the phenotype of CHO cells in culture. It was first identified as part of a micro-RNA screen on temperature shifted CHO cells and was proposed as a molecular mechanism to induce a temperature shifted phenotype (Gammell et al. 2007). This phenotype of reduced cell proliferation without affecting cell viability, resulting in greater productivity due to this extended viability in culture, was confirmed by up regulation of miR-7 through transient transfection in CHO-K1 cells (Barron et al. 2011a). While this phenotypic affect of temperature shift is desirable for industrial applications, prolonging time in culture and increasing productivity over time, the financial feasibility of reducing the temperature of a typical large scale bioreactor with several thousands of litres in culture volume may not be desirable.

Alternatively we proposed identifying targets that induce this temperature shift like phenotype of miR-7 over expressing CHO cells. Additionally, with little information publicly available on the molecular mechanisms of CHO cells, this study also revealed many as of yet unreported effects of miR-7 on the CHO cell proteome. This was met with several challenges including the acquisition of sample with miR-7 arresting cell growth, the acquisition and proofing of protein identifications from multiple species databases without a CHO protein database available and finally the annotation and curation of protein identifications using CHO databases acquired after the initial multi-species database method. With extensive analysis, protein ID curation and validation of protein expression it was possible to identify a large number of proteins differentially expressed in response to miR-7 over expression compared to a miR scramble negative. Not only were we able to identify and validate differentially expressed proteins but we were also able to identify and validate the expression of predicted direct targets of miR-7 in CHO-K1-SEAP cells, overlap differentially expressed proteins with previously reported differentially expressed genes in miR-7 over expressing CHO cells and also able to tie these associations into three types of pathway analysis using the differential expressed identifications generated.

6.1.1 Quantitative label-free LC-MS/MS

Quantitative abel-free LC-MS/MS proteomic technology has been a new analysis within our lab in the last 5 years and a relatively new method of analysis within the omics sciences. The methodology in sample processing is relatively fast and high throughput.

Working with Chinese hamster ovary material there are particular limitations with the availability of Chinese hamster protein databases and the annotation of available Chinese hamster protein databases. Initially we relied on the merger of output from closely related species output with far more sequence information in human, mouse and rat databases as Chinese hamster protein databases were not available. While this does produce IDs it introduces a limitation of certainty for the resulting IDs and if they are truly the same proteins that can be found in Chinese hamster proteome. It also introduces the possibility of not identifying unique Chinese hamster proteins or peptides that are not homologous to respective human, mouse or rat proteins or are simply not present in any of the three species.

6.1.2 Species homology identifications

Our initial published study relied on sequence homology between the CHO peptide sequence data with human mouse and rat sequence IDs. This technique has been used to similar effect in previous CHO sequencing studies where gene transcript databases specific to CHO were not available (Yee et al. 2008). Mouse and rat transcript data has been reported to have 92% homology to CHO transcript sequences (Ernst et al. 2006). Using Progenesis^{TM} QI (http://www.nonlinear.com/) it was possible to combine human, mouse and rat database IDs and remove duplicate IDs arising from species homology. This resulted in a sizable number of proteins identified as up or down-regulated 48 and 96 hr after transient up regulation of miR-7 compared to a scramble control.

As the IDs were obtained through sequence homology a panel of proteins were validated as differentially regulated by Western blot demonstrating that the use of different species databases was effective in identifying protein differential regulation. These findings remain the only reported ones on the effect of miR-7 on the CHO cell proteome profile and of any large scale mammalian proteome profile. While using different species databases to identify CHO proteins does introduce constraints in maximising IDs and validity of differential protein expression, multi-species analysis

also allows potential inferences to be made about the role of miR-7 in other mammalian systems with the high degree of sequence homology shown across species.

6.1.3 Chinese hamster ovary identifications

As described in the introduction we later gained access to two CHO specific protein databases the first being generated from the curated Bielfeld-BOKU-CHO (BBCHO) database (Becker et al. 2011) and the second being compiled from non redundant NCBI CHO IDs based on genomic data by Xu *et al.* (Xu et al. 2011). Both were used in parallel as in a previous study in our lab to maximise the number of IDs obtained (Meleady et al. 2012a).

Using the CHO database resulted in many 1 peptide protein IDs. By overlapping the IDs from the multi-species database search of >1 peptide and the CHO database with 1 peptide IDs included it was found that over half the IDs in the multi-species list were accounted for. We also verified using miRWalk that two predicted direct targets of miR-7, stathmin and catalase were present in the multi-species list with greater than 1 peptide each list and were each identified as 1 peptide IDs in the CHO database. For this reason 1 peptide IDs were included using the CHO database. It is possible that the overlapped IDs between the two search methods which included the two predicted direct targets of miR-7 had less sequence coverage for the peptide fragments identifying the proteins in the CHO database. Using a CHO specific database also means we could have greater certainty as to the validity of the IDs compared to the reliance on species homology. Furthermore for pathway analysis using the CHO database derived IDs each protein name required manual curation. This involved UniProt BLAST searching the peptides of each ID to verify the correct gene names. Consequentially each differential protein ID was manually verified from the CHO database output.

6.1.4 Predicted direct targets of miR-7

Using miRWalk, an online tool that aggregates several miR target prediction databases (Dweep et al. 2011), the two strongest predicted direct targets of miR-7 were catalase and stathmin in the down-regulated proteins of the multi-species list. As these databases do not contain alignment to Chinese hamster species we used the multi-species database differential IDs which contain human, mouse and rat species for the predicted target

search. Only the down-regulated list was used as microRNA negatively regulated protein expression.

It was later confirmed that catalase and stathmin were also down-regulated in the CHO database list. Catalase and stathmin were also confirmed to be down-regulated using Western blot analysis. A number of other potential targets of miR-7 were also predicted within the multi-species list. Discussed below are the predicted targets that were associated with cellular processes that were enriched in pathway analysis.

6.1.4.1 Catalase

Catalase is a well studied 60kDa antioxidant enzyme found in the peroxisome which converts toxic hydrogen peroxide and reactive oxygen species (ROS) through numerous reactions into harmless products and is expressed across many species (Zamocky, FurtmÃ¼ller and Obinger 2008)

Catalase was found to be down-regulated at 48 and 96 hr after exogenous up regulation of miR-7 compared to scramble control. This was further verified by Western blot. The down regulation of catalase in response to increased miR-7, and being a predicted direct target of miR-7, is consistent with how microRNA negatively regulate protein expression.

Catalase inhibition has been reported to reduce the cell proliferation of human immortalised myelogenous leukaemia cells (K562) (Takeuchi et al. 1995), human myeloid cells (U937), human melanoma cells (A375-C6), human B cells (Daudi) (Miyamoto et al. 1996) and in human promyelocytic cells (HP60) (Hachiya and Akashi 2005). Being such a widely studied protein catalase has also been implicated in a wide variety of cellular dysfunctions with reduction in catalase and its associated antioxidant properties being noted in renal failure (Aziz et al. 2015), diabetes and heart failure (Hayden and Tyagi 2003).

With respect to the over expression of miR-7 in CHO the resulting reduction in catalase expression would appear to correlate with reducing cell proliferation. As disease related studies suggest however catalase has a strong anti-oxidative effect. It was noted that Superoxide Dismutase 1 (SOD1) and Superoxide Dismutase 2 (SOD2) were both up-regulated in miR-7 over expressing CHO cells. Both these proteins convert superoxide to hydrogen peroxide. The resulting Hydrogen peroxide can then be acted on by catalase

or the other members of the anti-oxidant proteins. This would suggest that there is a possible increase of hydrogen peroxide. This highly suggests that other members of the ROS reducing family are compensating for the reduced catalase expression as the viability of CHO cells with increased miR-7 do not have reduced viability. Indeed looking for other ROS mitigating proteins in the differential protein lists we found several proteins related to glutathione synthesis and glutathione transferases that may play a role in compensating for the loss of catalase due to increased miR-7.

6.1.4.2 Stathmin

Stathmin is a 17kDa protein involved in preventing and promoting microtubule disassembly (Belmont and Mitchison 1996). Stathmin was found to be down-regulated at 96 hr after exogenous up regulation of miR-7 compared to a scramble control and was confirmed by Western blot analysis as down-regulated at 48 and 96 hr after miR-7 over expression.

Stathmin plays a key role in cell cycle progression by destabilising microtubules involving tubule to allow cell division. The inhibition of stathmin has been well documented to significantly reduce cell proliferation by halting the microtubule dynamics in a variety of cell types including liver cancer (Malz et al. 2009), melanoma (Chen et al. 2013) and leukaemia (Machado-Neto, Saad and Traina 2014). Due to its strong association with proliferation and cell cycle progression in cancer it is also known as Oncoprotein 18. It also interacts with other signalling pathways as part of cell cycle progression, being activated by phosphorylation from Cyclin Dependant Kinases (CDK) CDK1, CDK2 (Brattsand et al. 1994) and CDK5 (Beretta, Dobransky and Sobel 1993). These CDK proteins are in turn negatively regulated by p27 (Berton et al. 2014).

Increase in miR-7 in CHO cells resulting in decreased stathmin expression and arrested cell proliferation correlates with stathmin's reported functions in the literature. Looking at additional proteins in the list many are linked to structural changes in the cell and specifically at 96 hr are associated with structural molecule activity, specifically Tubulin alpha-1C chain (TPM4), Tubulin alpha-1C chain (TUBA1C), Tubulin beta (TUBB) and Tubulin beta-2C chain (TUBB2C) and Tubulin beta-4 chain (TUBB4) all down-regulated at 48 hr and Microtubule-associated protein 6 (MAP6) and

Microtubule-associated protein RP/EB family member 1 (MAPRE1) both down-regulated at 96 hr.

In a previous microarray study by Dr. Noelia Sanchez on the affect of miR-7 on CHO cells it was reported that p27 mediated cell cycle arrest with miR-7 directly interacting with S-Phase Kinase-Associated Protein 2 (SKP2) thereby increasing p27 (Sanchez et al. 2013). This further points to the involvement of stathmin in cell cycle arrest considering the relationship between p27, CDKs and stathmin. It is possible that both events are not mutually exclusive with increased p27 inhibiting CDK halting G1 to S phase transition, as well as miR-7 over expression simultaneously and directly reducing stathmin expression, resulting in arrested microtubule dynamics and an accumulation of the aforementioned structural tubulin proteins.

6.1.4.3 CCT3

The protein Chaperonin containing TCP1, Subunit 3 gamma (CCT3) is a 60kDa protein part of the T-Complex 1 ring complex (TRiC). Through the TRiC this protein is involved in the folding of various proteins including actin and tubulins (Nadler-Holly et al. 2012). We found that CCT3 was down-regulated in miR-7 over-expressing CHO cells compared to the transfected negative control.

CCT3 has been observed to have increased expression in hepatocellular carcinoma compared to adjacent normal tissue and was also confirmed to suppress cell proliferation in hepatocellular carcinoma cell lines (Cui et al. 2015). While CCT3 has also been identified as a potential biomarker in ovarian carcinoma (Peters et al. 2005) its specific functions are largely underreported with more focus on the overall function of TRiC which consists of all 8 of the CCT chaperonin proteins in a double stacked ring formation (Lopez, Dalton and Frydman 2015). It has been reported however that these chaperones are co-regulated with the translational apparatus (Albanèse et al. 2006)

.

As stathmin was a direct target of miR-7 and with its close association to tubulin CCT3 may also be involved in the process of arresting tubulin processing. The down regulation of CCT3 in parallel with the observed reduction in proliferation in miR-7 over expressing CHO cells also correlates with its reported affect on proliferation. Additionally its role in translational as part of TRiC may point to the arrest of further

processes in CHO cells with increased miR-7. CCT3 down regulation by miR-7 then may be contributing to cell cycle arrest by inhibition of protein translation and the processing of tubulins and actin.

6.1.4.4 PA2G4

Proliferation-Associated 2G4 (PA2G4) is a 38 kDa protein involved in growth regulation. PA2G4 was found down-regulated 48 and 96 hr after miR-7 over expression compared a negative control in CHO cells. Apart from its obvious reference to growth regulation and its down regulation correlating with reduced proliferation in miR-7 over expressing cells it also has a number of specific associations with other proteins.

PA2G4, also known as ErbB3-Binding Protein 1 (EBP1), has been strongly associated with ribosomal inhibition in conjunction with transcription initiation factor 1 (TIFIA) and guanosine triphosphate in T-cells (Nguyen le et al. 2015). Transcription has also been reported to be repressed by EBP1 involving E2F1 transcription factor and histone deacetylases (Zhang et al. 2003). PA2G4/EBP1 has also been shown to interact with p53 tumour suppressor in glioblastoma with EBP1 over expression leading to p53 degradation and subsequently promoting cancer growth (Kim et al. 2010). It also interacts with Bcl-2 mRNA leading to increased Bcl-2 mRNA stability and increased Bcl-2 expression leading to apoptosis resistance (Bose et al. 2006).

Numerous transcription and RNA processing proteins were down-regulated 48 and 96 hr after up regulation of miR-7 compared to negative control in CHO cells accompanied with simultaneous up regulation of anti-apoptotic proteins. Anti apoptotic proteins such as SOD1, SOD2 and GSTP1 are more strongly associated with catalase and anti-ROS activity. The up regulation of Eukaryotic Translation Initiation Factor 5A (EIF5A) and Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase is, however, associated with p53 up regulation (Li et al. 2004). From previous findings in our lab p53 was observed to be down-regulated in miR-7 over expressing CHO cells (Sanchez et al. 2013). Consistent with p53 down regulation a cell survival protein Activation Protein Zeta (YWHAZ) was up-regulated 96 hr after miR-7 up regulation and has been noted to increase cell survival by destabilising p53 in various breast cancers (Bergamaschi et al. 2013). Our previous results from microarray analysis and Western blot combined with our reported down regulation of and prediction of PA2G4/EBP1 as a direct target of miR-7 would imply that p53 is simultaneously depleted but stabilised by the reduced

expression of PA2G4/EBP1. PA2G4/EBP1 also potentially playing a role in transcriptional regulation.

6.1.4.5 RAN

Ras-related nuclear protein (RAN) is a 24Da protein involved in GTP binding and the translocation of RNA and proteins via the nuclear pore complex (Stewart 2007). RAN was found to be down-regulated at 48 hr after miR-7 over expression compared to negative control in CHO cells.

RAN is a member of the Ras oncogene superfamily and has been implicated in many processes involving cell cycle progression. This has been reported to be mediated by RAN-GTP, Regulator of chromosome condensation (RCC1) and Ran-binding protein 1(RANBP1) activity in spindle formation (Ciciarello et al. 2010, Clarke and Zhang 2008). As a transport protein it has been associated with increasing nuclear pore permeability with Importin-β (IMB1) (Lowe et al. 2015).

Interestingly we found that RANBP1 was down-regulated at 48 and 96 hr and IMB1 was down-regulated at 48 hr after exogenous miR-7 over expression. This may indicate that reduced RAN expression is functioning to both reduce spindle assembly and reorganisation with RANBP1 and is also reducing the translocation of proteins through the nuclear pore assembly with IMB1.

6.1.4.6 EEF1A1

Elongation factor 1-alpha 1 (EEF1A1) is 50kDa translational control protein that ezymatically delivers tRNAs to the ribosome. EEF1A1 was found to be down-regulated at 48 and 96 hr after miR-7 over expression compared to a negative control in CHO cells. This would suggest a reduction in translation.

EEF1A1 has been implicated in many processes involved in cell survival and stability. It was found to reduce translation and proliferation in a several cancers (Lin and Souchelnytskyi 2011). It has also been specifically linked to being up-regulated by p53, initiating microtubule severing and initiation of cell death in erythroleukemic cells (Kato et al. 1997). EEF1A1 has been shown to increase the rate of apoptosis when up-

regulated and decrease apoptosis rate when down-regulated in nutrient starved conditions (Duttaroy et al. 1998).

These reported phenotypes correlate with the miR-7 over expression phenotype observed in CHO cells with reduced proliferation with no affect on viability. Further to this it has also been noted that EEF1A1 is necessary for the initiation of the heat shock response in heat shock conditions (Vera et al. 2014). This correlates with a number of heat shock proteins down and up-regulated in miR-7 over expressing CHO cells. EEF1A1 therefore may play a role in several processes contributing to the miR-7 over expression phenotype including reduction of translation, reduction of proliferation, passive microtubule preservation and anti-apoptotic processes that may include dysregulation of heat shock proteins

6.1.4.7 RPL15

The 60S ribosomal protein L15 (RPL15) is a 24kDa member of the large 60S ribosomal subunit. It was observed to be down-regulated along with a host of other ribosomal proteins 96 hr after miR-7 over expression in CHO cells compared to a negative control.

Ribosomal proteins, as well as having a role in translation as part of the ribosome, are increasingly emerging as having multiple cellular functions including ribosome independent regulators of translation (Xue and Barna 2012). RPS15 has been shown to be over expressed in gastric cancer tissue and cell lines and was confirmed to suppress growth in gastric cancer cell lines when knocked down (Wang et al. 2006). RPS15 was also noted to be over expressed in oesophageal cancer (Zhang et al. 2004).

This reported reduction in proliferation associated with reduction of RPS15 may also be at play in miR-7 over expressing CHO cells. With so many other ribosomal proteins dysregulated in response to miR-7 over expression it may be that there are other effectors of translational machinery in operation.

6.1.5 Pathway analysis of differentially expressed proteins in miR-7 over expressing CHO cells

While differentially expressed protein expression provides an insight into the affect of miR-7 on the CHO cell proteome it can be difficult to determine the larger functions and processes being affected. We used 3 types of pathway analysis to determine the molecular functions (MF), biological processes (BP) and cellular components (CC) associated with differential protein expression in miR-7 up-regulated CHO cells compared to a scramble control at 48 and 96 hr after transfection.

Pathway analysis is often used as a proof of concept to confirm the effect of an experimental condition resulting in either MF, BP or CC only being used. Using MF, BP and CC together to assess the affect of miR-7 we have shown that all three are highly complementary to each other and reveal specific information on the global affect of miR-7 in CHO cells.

Using three pathway analysis tools it was shown that each revealed unique MF, BP and CC terms which contribute to explaining the specific effects of miR-7 in CHO cells. A key example of this is Glutathione metabolism which was significantly associated with up-regulated proteins at 48 and 96 hr in KEGG and glutathione transferase activity which was a significant MF associated with up-regulated proteins at 96 hr using DAVID. Using PANTHER there were no significantly enriched terms directly associated with Glutathione.

Additionally we used the differential IDs from the multi-species list and the CHO database to assess enriched BP and observed that similar terms were associated with each list. Terms such "cell redox homeostasis" and "cellular homeostasis" associated with protein up and down regulation respectively in miR-7 over expressing cells at 48 hr compared to scramble control CHO cells were associated with both database lists. Moving forward, only the CHO database output was used for pathway analysis.

6.1.5.1 Global effect of miR-7 on protein in CHO cells

Overall we observed that there were a smaller number of significantly enriched terms associated with the up-regulated proteins and an overall greater number of dysregulated terms associated with the 96 hr after miR-7 over expression time point compared to 48 hr. This shows the negative regulation affect of microRNA and shows that even after 48

hr after exogenous over expression of miR-7 there is little impact on MF and BP with respect to protein up regulation in particular.

A general trend observed was a greater number of enriched activities associated with down regulation in both DAVID and PANTHER tools. This is also to be expected taking into account that the up regulation of miR-7 at 48 and 96 hr resulted in a larger number of proteins down-regulated than up-regulated, consistent with the negative regulation effect of microRNA.

The one exception to this negative regulation bias in the pathway analysis is at 48 hr after miR-7 up regulation DAVID pathway analysis results show only 2 BP associated with protein down regulation and 6 BP associated with up regulation potentially revealing the specific early 48 hr effect of miR-7 in CHO cells. Specifically these were translation and translational elongation. The down regulation of proteins within these pathways may explain the larger number of dysregulated proteins at 96 hr and potentially point to a cascade of dysregulation leading to the more numerous enriched BP and MF seen at 96 hr.

It was also noted, looking at the output of each pathway analysis tool, that the large number of differentially regulated proteins that were identified as being associated with miR-7 over expression in CHO affect a wide variety of cellular activities. This is to be expected as microRNAs, initially thought to have highly specific effects, are now well understood to have multiple targets and effects including miR-7 which is thought to target multiple oncogenes (Liu et al. 2014).

Generally, however, pathway analysis reveals 4 major groups of dysregulated cellular activities with repeated terms to translation, cellular structure, homeostasis, apoptosis and glutathione metabolism. It was also observed that the predicted direct targets of miR-7 CAT, STMN1, PA2G4, RAN, CCT3 and EEF1A1 all feature across these enriched pathways.

6.1.5.2 miR-7 negatively regulates proteins associated with translation

Among the terms associated with protein down regulation in miR-7 over expressing CHO cells were several BP associated with translation and RNA processing, MF associated with translational activity and CC associated with the ribosome. These BP

include many of the ribosomal proteins found in the down-regulated protein lists but also includes several elongation factors.

Most notably it includes EEF1A1 one of the predicted direct targets of miR-7. Similarly enriched MF associated with down-regulated proteins shows RNA binding to be enriched which again includes 2 of the predicted miR-7 direct targets PA24G and RPS15. CC also confirms the large enrichment of ribosomal localised proteins with frequent reference to ribosomal gene ontology terms.

As stated previously the inhibition of specific chaperones such as CCT3 or the direct inhibition of ribosomal proteins such as RPS15 by miR-7 over expression may play a role in translational repression. At the early time point of 48 hr after miR-7 over expression protein down regulation is associated with translation. This may imply that miR-7, at least initially, acts on translational regulation, directly or indirectly, involving predicted direct targets RPL15, EEF1A1 and PA2G4 which are all flagged by pathway analysis as being associated with these pathways.

6.1.5.3 miR-7 negatively regulates proteins associated with cellular structure

Pathway analysis indicates the enrichment of terms associated with structural events in miR-7 over expressing CHO cells. Down regulation of proteins is associated with late stage enrichment of macromolecule complex organization and assembly at 96 hr after miR-7 over expression. Enriched MF also shows structural molecule activity associated with down-regulated proteins at both 48 and 96 hr after miR-7 over expression.

These enriched terms also contain the predicted direct targets of miR-7 CAT and STMN1. Interestingly, predicted direct target RPS15 only appears in structural molecule MF at 96 hr.

Taking into account the predicted direct targets this suggests that microtubule disassembly is down-regulated, with stathmin playing a key role in tubulin degradation in order to trigger cell proliferation. This is further supported by the presence of TUBB4 being implicated in the same enriched terms.

6.1.5.4 miR-7 positively regulates homeostasis and anti-apoptotic proteins

As microRNAs directly inhibit mRNAs leading to negative regulation of protein expression it is worth noting that they can give rise to numerous positive regulation events as a result. Pathway analysis indicates that many anti-apoptotic and homeostasis BP and MF are associated with protein up regulation in miR-7 over expressing CHO cells compared to negative control.

At 48 hr after miR-7 over expression BP terms such as cellular homeostasis, regulation of apoptosis and oxidoreductase activity as well as MF such as antioxidant activity and oxidoreductase activity suggest the promotion of cell survival related proteins. At 96 hr the same terms are enriched with additional specific references to anti-apoptosis and negative regulation of apoptosis and programmed cell death BP and similarly for MF the same terms are enriched with the addition of glutathione transferase activity. Homeostasis in this context is associated with antioxidant proteins such as thioredoxin and combined with the other enriched terms suggests that the cells are working to maintain pro antioxidant processes.

All of these terms would support our observed miR-7 over expression phenotype in CHO cells with prolonged life in culture with no affect on viability. In the context of predicted direct targets it was observed that many antioxidant proteins are up-regulated such as SOD1, SOD2, GSTP1, TXN and TXNRD1. This points towards increased oxidative stress in miR-7 over expressing CHO cells compared to negative control transfected cells. This in part may be explained by the down regulation of predicted miR-7 target catalase as a key anti-oxidant protein leading to the over compensating antioxidant activity of other proteins. It has been reported that knockdown of Peroxiredoxin 1 and 2 leads to increased catalase expression in podocytes (Hsu et al. 2011) showing that this type of antioxidant compensation does occur.

6.1.5.5 miR-7 positively regulates proteins involved in glutathione metabolism

As noted from the predicted target analysis of miR-7 we found that catalase was down-regulated in miR-7 over expressing CHO cells. As a powerful antioxidant protein its reduced expression pointed toward the possibility of increased hydrogen peroxide in miR-7 over expressing cells. We noted that several proteins related to glutathione were also up-regulated in miR-7 over expressing cells compared to the negative control at 48 and 96 hr. As members of the anti-ROS family of proteins both catalase and Glutathione

(GSH) neutralises hydrogen peroxide converting it to oxygen and water. GSH mediates this through the Glutathione peroxidase (GPX) (Dan Dunn et al. 2015). Although GPX itself wasn't found to be up-regulated in response to miR-7 up regulation using KEGG analysis it is possible to identify the specific reactions within the Glutathione metabolism pathway that are up-regulated.

Proteins up-regulated in the Glutathione metabolism pathway in response to miR-7 over expression were Glutathione S-transferase Mu 1 (GSTM1), Glutathione synthetase (GSS), Glutamate cysteine ligase regulatory subunit (GCLM), Glutathione S-transferase P (GSTP1), Glutathione S-transferase A2 (GSTA2), Glutathione S-transferase A3 (GSTA3) and Isocitrate dehydrogenase 1 (NADP+), soluble (IDH1).

Looking at the location of these proteins in the pathway in KEGG the 4 proteins GSTM1, GSS, GCLM and GSTP1 were identified as up-regulated 48 hr after miR-7 over expression. The up regulation of these 4 proteins would imply that GSH is being consumed by GSTM1 and GSTP1 to produce L-Glutamate. The up regulation of GSS shows that L-Glutamate is converted into L-γ-Glutamylcysteine which ultimately becomes GSH via the up regulation of GCLM. A GSH assay however would be required to determine the if it is being consumed within the cell. With catalase down-regulated and numerous other antioxidant proteins up-regulated it could be that GSH is being recruited as an antioxidant. This occurs via the reduction of GSH to $GS^{\bullet}$ radical which when accumulated to harmful levels is in turn rendered inactive to Glutathione disulphide (GSSG) (Franco et al. 2007).

At 96 hr after miR-7 up regulation in CHO cells GSTPM1, GSTP1 and GCLM as well as GSTA3 and GSTA2 were up-regulated in the same fashion as they were at 48 hr suggesting again that a feedback loop involving the production of L-Glutamate is being driven by the consumption of GSH. Additionally IDH1 was also shown to be up-regulated in the GSH pathway which is shown by KEGG to convert NADP+ to NADPH. While NADPH can be implicated across many cellular processes it may indicate that it is being used in the GSH pathway to produce GSSG. This may be driven by the production of $GS^{\bullet}$ from GSH being produced and consumed as an antioxidant at 48 and 96 hr. If GSH is indeed being used as an antioxidant then this may also indicate that GSSG is up-regulated to prevent $GS^{\bullet}$ radical damage. GSSG can then be fed back into the production of GSH to perpetuate the cycle.

The reduction of GSH to $GS^{\bullet}$ as an antioxidant makes sense from the other pathway analysis results. Anti-apoptotic and anti-oxidant BP and MF activities are up-regulated at 48 and 96 hr as well as GSH metabolism MF being up-regulated at 96 hr after miR-7 over expression in CHO cells. KEGG illustrates that GSH production should be increased 48 and 96 hr in a self perpetuating.. By 96 hr the up regulation of IDH is also implicated in GSH metabolism and may point toward the mopping up of $GS^{\bullet}$. This GSSG could in turn feed back into GSH production. It is therefore proposed that with the loss of catalase, a predicted direct target of miR-7, the glutathione metabolism pathway acts as an anti-oxidant sink replacing the role of catalase to remove ROS species in the CHO cells over expressing miR-7. This also explains how CHO cells survive with reduced catalase and may also explain the increased longevity and productivity of miR-7 over expressing CHO cells as it has previously been reported that GSH is an up-regulated feature in high producer CHO cell lines (Orellana et al. 2015). Such a self perpetuating sink for oxidative damage fuelled by L-Glutamate and GSH combined with reduced proliferation may explain miR-7 effect of increasing productivity over time in CHO cells.

6.1.6 Transcriptomic and proteomic overlap of differentially expressed targets in miR-7 over expressing CHO cells

Previously in our lab a study was carried out by Dr. Noelia Sanchez which identified potential miR-7 targets in CHO cells using micro array analysis (Sanchez et al. 2013). The targets from this analysis were overlapped with the differential protein IDs that were obtained in our data. We found that once again catalase which was previously discussed (Section 7.1.4.1) was a predicted direct target at 72 hr after miR-7 over expression compared to negative control in the microarray study further strengthening the evidence that catalase is a direct target of miR-7.

One target in particular Histone deacetylase 1 (HDAC1) was of particular interest as it was previously noted by Western blot analysis that Histone acetylation was lower 48 and 96 hr after miR-7 over expression in CHO cells compared to negative control. Therefore we decided to investigate the expression of HDAC1.

6.1.6.1 HDAC1

Histone Deacetylase 1 is a 55 kDa protein involved in histone deacetylation and many

other proteins in the BRAF-histone deacetylase complex (BHC) and the Nucleosome remodelling and deacetylating complex (NuRD). In a previous analysis by Dr. Noleia Sanchez in our lab it was shown that HDAC1 was down-regulated at the transcriptional level 72 hr after up regulation of miR-7 in CHO cells (**Appendix xviii**). We also confirmed by Western blot that HDAC1 protein was down-regulated 48 and 96 hr after miR-7 up regulation in CHO cells compared negative control. Contradicting we found that Acetyl-histone H3 and Acetyl-histone H4 were down-regulated in miR-7 over expressing CHO cells which would imply that HDAC1 would be up-regulated.

Searching for potential binding partners of HDAC1 in the differentially expressed lists revealed no members of the BHC or NuRD complexes. As part of the these complexes, HDAC1 functions to deacetylate histones allowing Chromatin to unfold and allow DNA to be transcribed. Without histone acetylation translation is repressed (Zhang et al. 1998). EBP1 was also reported to recruit HDAC2 but not HDAC1 and was proposed to partly be responsible for observed translational repression (Zhang et al. 2003) as it has also been reported that HDAC1 complexes can recruit HDAC2 (Humphrey et al. 2001). HDAC has also been reported to acetylate hundreds of proteins through acetylation of Lysine, including p53 (Choudhary et al. 2009).

What is known about HDAC is that it has many functions outside of deacetylation of Histones. We also know that miR-7 over expressing CHO cells from pathway analysis results have highly reduced transcription, with up-regulated anti-apoptotic processes and from the previous microarray study in our lab, we also know that p53 is down-regulated in miR-7 up-regulated cells potentially by directly targeting PSME3. The observed reduction of Acetyl-histones may be as a result of translational repression preventing Histone acetylation in part driven by HDAC1 being recruited by EBP1. EBP1 or PA24G is a predicted direct target of miR-7 and was also down-regulated in our quantitative label-free proteomic analysis.

Taking a deeper look at the differential protein lists it was noticed that a key protein in the Citric Acid cycle (TCA) was down-regulated. Citrate Synthase (CS) catalyses condensation of Acetyl-CoA and oxaloacetate to CoA-SH and citrate which initiates the TCA cycle (Berg, Tymoczko and Stryer 2002). It was also observed in the previously discussed up-regulated Glutathione metabolism pathway (**Section 6.1.5.5**) that Acetyl-CoA is also implicated in a terminal off branch which produces L-Glutamate and creating a possible feedback loop. This may suggest that Glutathione metabolism is

rapidly consuming Acetyl-CoA as it may not be becoming consumed in the TCA cycle. The absence of acetyl groups on Histone H3 and H4 in miR-7 over expressing cells also indicates that acetylation may be being redirected to other processes in the cell. The reduced acetylation of histones then may contribute to translational repression but also allows for acetylation to occur in other cellular process such as Glutathione metabolism.

6.1.7 Proposed effect of miR-7 on CHO cell proteome

Overall using a combination of label-free quantitative mass spectrometry, multiple pathway analysis and overlapping with proteomic data with transcriptomic data we found that miR-7 regulates several process that can largely be divided into translational repression, arrest of structural modifications and alteration of antioxidant production.

As previously reported miR-7 directly targets SKP2 and PSME3 in CHO cells 72 hr after miR-7 over expression. With our proteomic findings we would propose multiple new targets which all at least contribute to the phenotype of reduced proliferation and increased longevity in culture.

Translational repression is potentially mediated by a direct interaction of miR-7 with RPS15 and elongation factor EEF1A1, indirectly through the direct targeting of CCT3 ribosomal protein chaperone and indirectly through redirection of acetylation events reducing histone acetylation.

Inhibition of structural activity is implied by the confirmed down regulation of Stathmin as well as the down regulation of several tubulins in the differential expression data. RAN also a predicted target of miR-7 was noted to be down-regulated in the label-free data and is also involved in micro tubule assembly.

Terms associated with antioxidant activity were associated with protein up-regulation as well as the Glutathione metabolism pathway which, considering the specific proteins that are up-regulated in the pathway, may lead to increased Glutathione and L-Glutamate. Glutathione metabolism then is potentially acting as an sponge for reactive oxygen species following the reduced expression of catalase which was confirmed to be down-regulated in miR-7 over expressing CHO cells by label-free quantitative proteomics, Western blot analysis, transcriptional microarray and a predicted direct target of miR-7. These interactions are represented in **Figure 6.1.1** below.

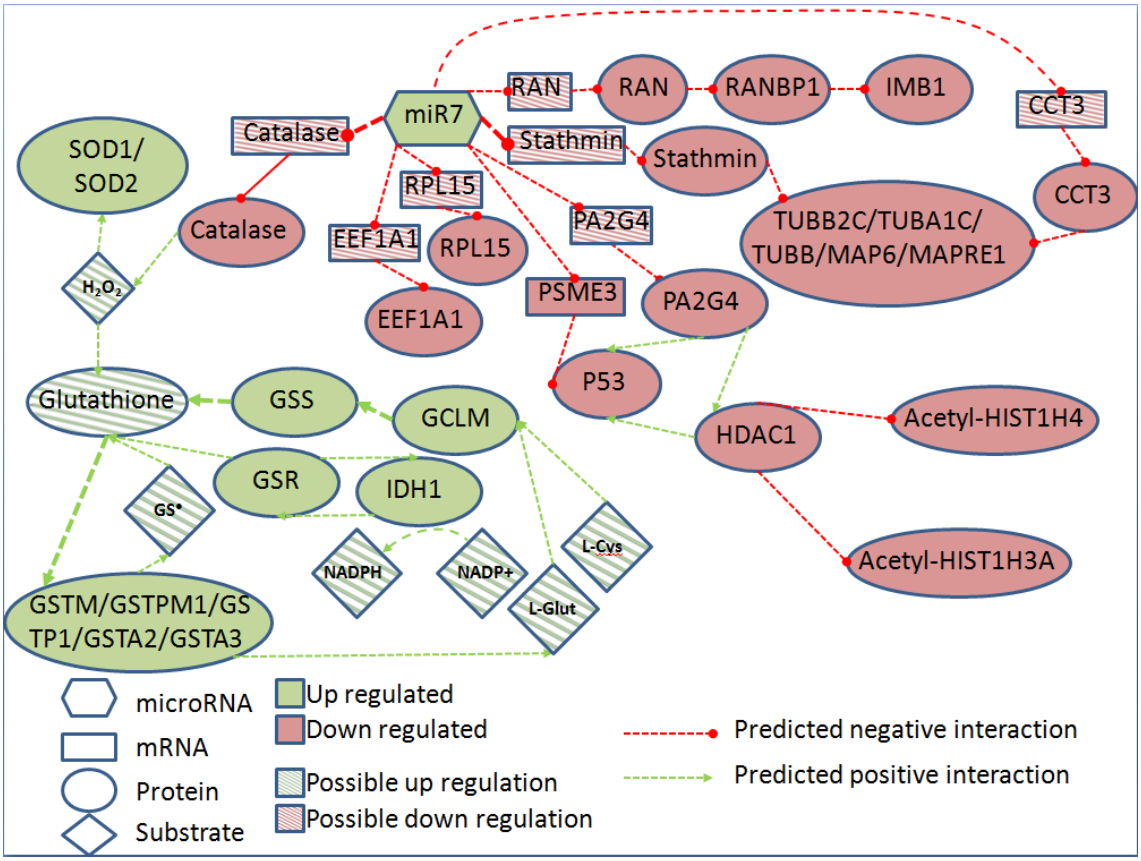


Figure 6.1.1 Proposed effect of miR-7 on the CHO cell proteome. Antioxidant up-regulation may be occuring due to direct inhibition of antioxidant catalse by miR-7 while negative regulation of stathmin and tubulins arrests cell cycle progression together with reduced transcription activity through ribosomal proteins.

6.2 Effect of temperature shift on CHO cells introduction

As discussed in the introduction **(Section 1.3)** temperature shift is a technique used to prolong the life of cells in culture thereby increasing their productivity over time. We have already demonstrated this in previous studies in our lab providing us with differentially expressed microRNAs related to the temperature shift phenotype (Gammell et al. 2007). It was from this study which miR-7 was identified as a potential mechanism to induce a temperature shift like state, the affect of which is discussed in the previous section.

Previous to our current study Dr. Niraj Kumar in our lab had conducted a 2D-DIGE analysis of 31 ^{o}C temperature shifted CHO cells in which identified 23 differentially expressed proteins compared to CHO cells grown at 37 ^{o}C (Kumar et al. 2008) 7 of which were greater than two fold . For this study we used quantitative label-free LC-MS/MS to identify differentially regulated proteins combined with subcellular fractionation to gain a deeper insight into temperature shifted CHO cells with an aim to identify new potential targets to manipulate the CHO cell phenotype.

6.2.1 Subcellular fractionation of CHO-K1-SEAP cells undergoing temperature shift

The fractionation methods used to investigate the effect of temperature shift on the CHO proteome were benchtop enrichment kits. These were used in an attempt to allow a greater number of IDs to be obtained and potentially greater sequence coverage of proteins by reducing the complexity of the sample and analysing each fraction.

Using commercial kit based fractionation methods there are a number of limitations to be considered. While "fractionation" is a term that is frequently used for these methods strictly it is "enrichment" and the degree of enrichment required verification. Other CHO fractionation studies CHO studies have used more complex chromatographic and cell culture based protien labelling methods and produced a similar number of identifications with Orellana et al identifying 2000 proteins (Orellana et al. 2015). Larger studies have used far more complex but time consuming sample pre fractionation. Baycin-Hizal et al conducted 120 mass spectrometry analysis combining a large number of fractionations per sample such as glycoprotein enrichment, secretome analysis, 8 fraction strong cation exchange and 27 gel bands per sample (Baycin-Hizal et al. 2012). This resulted in the identification of 6164 CHO proteins. Considering that

the use of three commercially available subcellular fractionation kits produced over 2298 CHO protein IDs with 12 mass spectrometry analyses (3 samples for unfractionated, nuclear, cytoplsmic and membrane, pooled and repeat IDs removed) it would appear that there may be dimishing returns in pre fractionation of sample. This of course could only be confirmed with a properly controlled study.

6.2.1.1 Overlap between fractionated samples

To assess enrichment of nuclear, cytoplasmic and membrane enriched CHO cell fractions we used pathway analysis of all the proteins identified in each fraction by LC-MS compared to un-fractionated lysate.

By overlapping the unique gene index (GI) identifier of each protein we were able to show that a greater number of proteins were associated with being only in one of the three fractionated samples than with the un-fractionated sample. That is to say there were specifically 241, 412 and 309 protein IDs that were uniquely associated with the nuclear, membrane and cytoplasmic samples respectively compared to 137 uniquely associated with the un-fractionated sample.

This indicated that the large pool of proteins that were only accessible through fractionation were identified. Furthermore, pooling the total IDs between fractions and the un-fractionated sample 2298 unique IDs were obtained.

6.2.1.2 Pathway analysis validation of enrichment

To assess enrichment of nuclear, cytoplasmic and membrane enriched CHO cell fractions we used pathway analysis of the proteins identified in each fraction by LC-MS compared to un-fractionated lysate. Pathway analysis is often used to determine the degree of fractionation of fractionated samples (Orellana et al. 2015, Baycin-Hizal et al. 2012).(Orellana et al. 2015)

We showed that each fraction, as well as the un-fractionated sample, displayed a different profile of enrichment while each fraction showed a bias for its expected enriched subcellular terms. The top 10 cellular component terms showed that while the enrichment did not achieve perfect fractionation, there was a different subset of proteins in each fraction indicating that the method reduced sample complexity.

6.2.1.3 Western blot validation of enrichment

Using proteins that are typically associated with nuclear, cytoplasmic and membrane localisations we showed that each corresponding subcellular fraction contained the highest amount of associated nuclear, cytoplasmic or membrane protein target.

Namely these were PDIA3 which is known to be associated with the nucleus (Wu et al. 2010b), HSP90 which is known to be associated with the cytoplasm (Langer, Rosmus and Fasold 2003) and IGF1Rβ which is known to be a membrane localised receptor (Guvakova and Surmacz 1997, Hellawell et al. 2002). The associated localisation of each protein was also confirmed with an online literature mining search tool called COMPARTMENTS (Binder et al. 2014).

6.2.2 CHO database annotation

As explained in the introduction and results section (**Section 3.4.1**) the CHO protein database used for this analysis was composed of BBCHO (Becker et al. 2011)(Meleady et al. 2012a), Meleady et al. 2012a) and NCBI (Xu et al. 2011) databases derived from transcript and gene data respectively. While these databases do hold certain advantages over using species homology they do have the disadvantage of not containing usable identifiers for pathways analysis. Furthermore the identifiers can also contain several names which make it difficult to assign a valid protein or gene name through automatic parsing.

Essentially this required manual BLAST searching of peptide identifications to assign the correct gene name for pathway analysis purposes. Combining the differential protein identification output from the multiple lists from within the temperature shift analysis and from the miR-7 analysis (**Chapter 3**) we were able construct a reference list of 1135 unique gene names and 746 unique protein names to the pooled differential IDs.

Despite the recent use of the now publicly available CHO genome in several proteomic studies this issue has not been addressed fully despite many of these studies using pathway analysis tools in conjunction with these databases. The annotation step represents a significant amount of data mining and it may be useful for CHO proteomics in the future to streamline this process with the construction of correctly annotated databases. Currently the CHO genome collaboration is adding to this and may benefit

from submissions of this nature to their proteomic reference information (Kremkow et al. 2015).

6.2.3 Quantitative Label-free LC-MS/MS analysis of temperature shifted sub cellular enriched CHO-K1-SEAP cells

To investigate the affect of temperature shift on CHO cells two groups of CHO-K1-SEAP cells were grown for 72 hr at 37 ^{o}C before one group was transferred to 31 ^{o}C temperature shift. Cells at both temperatures were taken down 8 and 24 hr after this temperature shift and an aliquot was taken for each membrane, cytoplasmic and nuclear fractionation.

This allowed two separate comparisons to be made using Progenesis label free software. The first comparison was between 31 ^{o}C and 37 ^{o}C at each time point of 8 and 24 hr. This time point analysis essentially provides a snapshot of changes between CHO cells grown at different temperatures. The second analysis involved comparing the 8 hr vs 24 hr time point for each temperature. This time course analysis identifies differential protein regulation over 16hr between 8 and 24 hr after temperature shift allowing changes over time at 31 ^{o}C and 37 ^{o}C to be observed separately and compared. This arguably highlights dynamic process in the cell than a time point analysis (Chong et al. 2012, Parnell et al. 2011, Nambu et al. 2015).

6.2.3.1 Differentially regulated proteins at 31 ^{o}C, 8 and 24 hr after temperature shift

Comparing sub cellular fractions of CHO-K1-SEAP cells grown for 8 and 24 hr at 31 ^{o}C compared to 37 ^{o}C we observed several patterns with regard protein differential regulation in 31 ^{o}C temperature shifted cells.

Graphing the total number of differentially regulated proteins in 31 ^{o}C temperature shifted CHO cells there was a clear bias toward down regulation at the early time point of 8 hr after temperature shift. This was particularly noticeable in the large number of proteins down-regulated in the nucleus (72) compared to the highest number of up-regulated proteins in at 8 hr in the membrane (22). A relatively large number of proteins were also down-regulated at 8 hr in the membrane (40). This shifts drastically at 24 hr

where protein up regulation more than doubles and accompanied by a more than 5 fold increase in down-regulated proteins in the cytoplasm.

This at least shows that temperature shift has a much more dramatic affect on protein down regulation rather than up regulation at 8 hr in CHO cells. It has been reported that CHO cells produce heat inducible proteins in response to low temperature (Kaufmann et al. 1999, Gammell et al. 2007) and the process of heat shock itself being well known to instigate the up regulation of heat shock proteins (Voellmy and Boellmann 2007). It may be that while early down regulation is seen more than up regulation the most significant effectors are in the up-regulated proteins that reduce transcription such as heat shock proteins. It has also been reported that lipid membrane triggers in the membrane may in fact trigger heat shock response (Saidi et al. 2009) which may also explain the large membrane response observed in CHO.

There are also a number of similar proteins differentially regulated in a previous study in our lab using 2D-DIGE to identify differentially regulated proteins associated with temperature shift in CHO (Kumar et al. 2008). In this study 23 proteins were identified as differentially regulated 144 hr after reducing temperature. These proteins functionaly align with the results of pathway analysis from the label-free study with a reduction in translational initiation via eukaryotic initiation factors and mRNA binding proteins such as HNRPC. Also interesting is that despite temperature shift and miR-7 up-regulation being entirely different stimuli to promote a similar phenotype, these protein classes of initiation factor inhibition and RNA binding up-regulation can be seen in both studies. It may also not be that surprising considering miR-7 was differentially regulated in temperature shifted CHO cells and therefore may affect many proteins involved in temperature shift. A similar 2D-PAGE study by Baik et al identified 7 proteins that were over two fold increased in expression at 33 ^{o}C in recombinant EPO producing CHO one of which was PDIA3 (ERp57) (Baik et al. 2006).

The proteomic mechanism behind cold shock response has been investigated in several other cell types. HeLa cells were found to respond to cold shock with differential regulation of many proteins but this response was not associated with some of the typical CHO cold response proteins such as vimentin and protein disulphide isomerase (Underhill and Smales 2007). More often though the same proteins are repeatedly identified with temperature shift ond other stress response phenotypes (Petrak et al. 2008) hence the need for deeper proteome studies. For that reason targets not normally associated with tempersture shift or CHO, as well as being strongly associated with

influencing cell phenotypes, were selected for validation (**Section 6.2.5**). Namely these were ezrin, moesin, lamin A/C and cyclon.

6.2.3.2 Pathway analysis of differentially regulated proteins at 31 ^{o}C, 8 and 24 hr after temperature shift

Differentially regulated proteins at 8 and 24 hr in temperature shifted CHO-K1-SEAP cells were analysed for enriched bioprocesses (BP), molecular functions (MF), Cellular components (CC) and pathways using DAVID, PANTHER and KEGG.

After 8 hr at 31 ^{o}C CHO cells were shown to have a large degree of down-regulated BP and MF these were largely related to protein localisation, translation and cytoskeleton. At 24 hr these down-regulated processes carry over with further emphasis on reduced structural activity and also included terms involving metabolic process and heat shock protein binding. Protein up regulation at 24 hr also shows that the TCA cycle BP and oxidoreductase activity MF are potentially up-regulated. Dysregulation of these process in temperature shift CHO cells were previously inferred in our lab from a 2D DIGE (Kumar et al. 2008). Looking at pathway results in KEGG however we are able to add to this.

KEGG shows that translation is down-regulated at 8 and 24 hr after temperature shift as well as an up regulation of DNA replication. The comparatively large amount of down regulation observed compared to up regulation in temperature shifted CHO cells was then likely driven by translational repression. It is interesting then that at 24 hr proteins in the DNA replication pathway are up-regulated.

Looking at these proteins there are several DNA replication licensing factors (MCM3, MCM5 and MCM6), Proliferating cellular nuclear antigen 1 (PCNA) and Flap end nuclease 1 (FEN1). PCNA and FEN1 are known to interact with each other PCNA acting as an enhancer (Tom, Henricksen and Bambara 2000). FEN1 is involved in the removal of RNA primers from DNA and trimming Okazaki fragment over hangs (Wu et al. 1996). This repair process may point toward replication damage occurring in temperature shifted CHO cells as the down regulation of PCNA has been linked to multiple DNA replication defects in mice (Zheng et al. 2007). As translation is down-regulated the role of PCNA/FEN1 may point to the fidelity of the DNA replication and

eventual translation being even increased in order to maintain the up-regulated cell process mentioned.

The inclusion of 3 MCM family members however implies that DNA replication initiation may be increased in the cell. The MCM2-7 complex is responsible for the once per cell cycle DNA replication event in eukaryotes (Tye 1999). From the temperature shift phenotype we know that cell proliferation is reduced and that translation is also reduced. The synthesis of MCM proteins in the cell is associated with G1 phase in the cell (Sclafani and Holzen 2007). The resulting increase in MCM in temperature shifted CHO cells may indicate a cell cycle related characteristic specifically that temperature shifted CHO cells are arrested at G1 phase combined with translational repression compared to CHO cells at 37 ^{o}C.

6.2.4 Differentially regulated proteins over 16hr at 31 and 37 ^{o}C

For the time course analysis over 16hr we compared CHO-K1-SEAP cells grown at 31 and 37 ^{o}C. Prior to this they were grown for 72 hr at 37 ^{o}C with the temperature shift group transferred to 31 ^{o}C. Sample taken 8 and 24 hr later at both temperatures were fractionated into membrane, cytoplasmic and nuclear fractions.

The resulting differential IDs over this 16hr growth period revealed several differences between CHO cells grown over time at 31 and 37 ^{o}C. Graphing total IDs for each fraction there was a similar profile of up regulation between both temperatures in all three fractions. Down-regulated proteins however were shown to be over 50% more down-regulated in the cytoplasm at 31 ^{o}C compared to 37 ^{o}C over time and down regulation of nuclear proteins was seen to be over twice as down-regulated at 37 ^{o}C compared to 31 ^{o}C.

This may lend itself to that fact that from the time course analysis there was a large amount of down regulation related to translation and to the ribosomal pathway suggesting that protein down regulation may also be occurring over time due to this translational repression.

Looking further at overlaps between the IDs in each fraction at each temperature over the 16hr period it was also observed that each temperature had distinct differentially expressed proteins in each fraction. This may lead us to believe that, and considering the

dramatic temperature shift phenotype, there was completely different cellular process active over time between the two temperatures.

6.2.4.1 Pathway analysis of differentially regulated proteins over 16hr at 31 °C and 37 °C

From the small amount of overlap in differential proteins between 31 and 37 °C CHO cells over the 16hr culture period we expected that enriched bioprocess (BP), molecular functions (MF) and KEGG pathways would be dissimilar. This was only partly true.

Pathway analysis showed that over time down regulation of protein expression at 31 °C in CHO cells was associated with translation and cell cycle progression BP and MF while up regulation was associated with cytoskeleton and cellular organisation BP and MF. CHO cells grow at 37 °C however only showed enriched MF over time related to actin binding.

The absence of enriched BP and MF at 37 °C is somewhat misleading. Considering that the comparison spans differential regulation over 16hr it is possible that the spread of functionality over this time period resulted in both DAVID and PANTHER not flagging any BP as specifically enriched.

KEGG pathway enrichment does not exhibit the same pattern and shows that down-regulated proteins are associated with the ribosomal pathway while up-regulated proteins are associated with the ribosomal pathway and spilceosome. What is interesting is that these pathways show the enrichment of completely different proteins.

At 31 °C there are a number of proteins of interest that are implicated in these pathways. RPL36a for example is known to promote proliferation acute myeloid leukaemia cells (Wu et al. 2010a) and hepatocellular carcinoma (Kim et al. 2004) and is down-regulated at 31 °C. It is well known however that translational repression occurs in temperature shifted CHO cells (Masterton et al. 2010). What is interesting in these findings is that there are also up-regulated ribosomal proteins. This may point to translation of specific proteins at 31 °C or to the increasing reports of ribosomal proteins having additional functions outside of translation (Lindstrom 2009) such as with RPL36a above.

The up regulation of proteins in the spliceosome at 31 °C also points towards translational related control. Proteins specifically up-regulated in the spliceosome at 31 °C include EIF4A3, HNRNPK, SNRPD2, RBMX, PUF60, HNRNPU and PRPF6.

EIF4A3 (Eukaryotic Translation Initiation Factor 4A3) depletion has been shown to increase pro-apoptotic splice variants of Bcl-x in HeLa cells (Michelle et al. 2012). HNRNP (Heterogenous nuclear ribonucleoprotein) family members are not all well understood but HNRNPK has been reported to be over expressed across various cancers (Barboro, Ferrari and Balbi 2014). SNRPD2, RBMX, PUF60 and PRPF6 are all poorly studied but PUF60 has been linked to apoptotic regulation (Ren et al. 2015).

Taking all of this into consideration the pathway analysis showed that despite the completely dissimilar differential regulated protein IDs at each temperature over 16hr CHO-K1-SEAP cells had similar pathways dysregulated. From the individual BP and MF analysis there was a great difference between both temperature conditions but KEGG confirmed that the over arching dysregulated processes were different, also showing that the subset of proteins responsible for this at each temperature are different. Novel and poorly understood spliceosome proteins were identified as playing a potential role in this functionality in CHO cells grown at 31 ^{o}C over time compared to 37 ^{o}C.

6.2.5 Selection of targets for functional validation

We identified several proteins of interest that were highly dysregulated in temperature shifted CHO-K1-SEAP cells and novel in CHO functional studies. We selected 4 to assess their functional effect on the CHO cell phenotype. Knockdown of two of these proteins Cyclon and Moesin showed a significant affect on cell proliferation and cell size without affecting viability.

6.2.5.1 Ezrin (EZR) and Moesin (MSN)

Ezrin and Moesin are part of the Ezrin-Radixin-Moesin (ERM) group of proteins which are part of the lager ezrin/radixin/moesin (FERM) family (Moleirinho et al. 2013). Ezrin and Moesin while part of the same group of proteins have different expression levels across different tissues. Ezrin was found to be up-regulated 24 hr after temperature shift in CHO cells grown at 31 ^{o}C while Moesin was found to be down-regulated.

Increasingly Ezrin has been shown to be linked to poor prognosis in a variety of cancers including cervical (Kong et al. 2013)(Kong et al. 2013), head and neck squamous carcinoma (Schlecht et al. 2012), the transformation of breast cancer from benign to

malignant (Gschwantler-Kaulich et al. 2013) and also the localisation of Ezrin to the cytoplasm also seems to correlate to its role in the development of cancer. It has also been reported that localisation of Ezrin and Moesin to the cytoplasm correlates with a number of genes related to proliferation and cell death processes in cancer cells (Schlecht et al. 2012), (Kong et al. 2013)(Kong et al. 2013)Kong et al. 2013). The ERM proteins are thought to also play a key role in actin and tubulin dynamics in oocyte development (Namgoong and Kim 2016), structural proteins that are well known to play a role in cell cycle arrest.

No significant affect was observed with knockdown of Ezrin after 72 hr. Moesin knockdown produced a significant ($p \leq 0.05$) reduction in cell proliferation and cell perimeter with all three siRNA 72 hr after transfection. Cell number was also significantly decreased in 2 out of 3 of the siRNA used. Cell viability was unaffected. This established that Moesin plays a role in altering cell proliferation in CHO-K1-SEAP cells.

The expression of Moesin was not confirmed to be knocked down using Western blot analysis. This may be the result of an antibody compatability issue using human humanised antibodies which can commonly occuring analysing CHO. The FERM family of proteins including the ERM like protein Merlin all contain a FERM domain which has been proven to exist in sorting nexus proteins also (Ghai et al. 2011). This may pose a problem for antibody specificity in particular considering possible CHO squence variations. Although the anti-Moesin siRNA were custom designed to known CHO sequences off target effects may resulted in the effect seen. What ever the cause it is likely to be a consistent cause accross samples therefore the differing oligo variations used for the Moesin may not be the likely cause of the observed effect. While Ezrin, Moesin and Radixin are struturally similar and usually coexpressed they have been reported in knockout mice to be functually distinct (Doi et al. 1999, Kikuchi et al. 2002). This may further be support by Ezrin being increased and Moesin being decreased 24 hr after temperature shift. It would seem unlikely then that compensating expression of one of the other ERM family members potentially lead to Ezrins lack of effect and Moesins lack of observed knowdown.

6.2.5.2 Cyclon

Cyclon (CCDC86) is 40kDa T-Cell response regulating protein (Hoshino and Fujii 2007) but to date there are very few studies on its function to date. Cyclon was found to be down-regulated in temperature shifted CHO-K1-SEAP cells over 16hr but not at 37 ^{o}C. This would suggest that Cyclon was potentially involved in the temperature shift phenotype over time.

Although little is known about Cyclon a recent study does report it's over-expression being associated with poor outcome for lymphoma patients and that knockdown of Cyclon in Raji BL haematopoietic cells significantly reduced tumour size (Ren et al. 2015). It has been mentioned as a potential novel marker for normal physological function such as hypocampus development (Shishkov et al. 2013a), as a potential disease marker and treatment predictor in schizophrenia (Chervenkov, Shishkov and Tonchev 2013) and up-regulation has been proposed as a marker for poor chemotherapy response in testicular cancer (Knapp 2013). In total there are only 10 papers listed on PubMed all of which surround varying topics.

Using siRNA knockdown we showed that Cyclon in CHO-K1-SEAP cells significantly ($p \leq 0.05$) reduced cell number more than 2 fold and cell area and perimeter more than 20% 72 hr after transfection in 2 out of 3 siRNA used. Viability was unaffected.

This adds to the currently lacking functional information on Cyclon and identifies Cyclon as potential novel target for the manipulation of the CHO-K1-SEAP phenotype.

6.2.5.3 Lamin A/C

Lamin A/C (LMNA) is a very well studied 74kDa member of the Lamin family of proteins. It was found to be up-regulated in temperature shifted CHO-K1-SEAP cells at 24 hr after temperature shift.

Lamin A/C is known to be associated with many diseases (Politano et al. 2013). It has been identified as a potential deficiency marker in cervical cancer (Capo-Chichi et al. 2015). The migration, proliferation and invasion of prostate cancer is also increased in prostate cancer with increased Lamin A/C (Kong et al. 2012). It has also been shown to contribute to telomerase homeostasis leading partly explaining its role in aging disorders like progeria (Das et al. 2013). As far as phenotypic roles are concerned it has been known to be reduced in expression in senescent in dermal fibroblasts and

keratinocyts (Dreesen et al. 2013). This indicates that knockdown would reduce proliferation.

Using 3 siRNAs we found no significant affect of Lamin A/C on CHO-K1-SEAP cells after 72 hr. It may still have an affect on antioxidant process however as it is noted to alter ROS activity (Sieprath et al. 2015) although this would require further confirmation.

6.2.6 Conclusion

Fractionation allowed the identification of a large subset of proteins that were not identified in the unfractionated sample. It allowed patterns of expression to be seen in subcellular enriched fractions in CHO-K1-SEAP cells and identify the early effects of temperature shift. Label-free LC-MS/MS combined with pathway analysis built upon previous findings on the effect of temperature shift on CHO cells within our lab and identified a larger cohort of proteins associated with translational repression. The time course analysis also revealed differentially regulated proteins associated with the spliceosome in temperature shifted CHO cells.

Finally two targets chosen from the analysis Moesin and Cyclon had a significant effect on cell size and proliferation without affecting viability and may potentially be used to activate a temperature shift like phenotype in CHO cells.

6.3 Breast Cancer target discovery

As discussed in **Section 1.7** breast cancer is a complex and heterogeneous disease which is currently defined by multiple levels of subtyping (Inic et al. 2014, Goldhirsch et al. 2007). While there are better clinical outcomes for some types of breast cancer such as chemo resistant HER2+ and the constant revision and further stratification of breast cancer subtypes means that even within these prognosticly favourable classifications there is a need for more effective treatment options. Still TNBC remains one of the subtypes with the poorest outcome and therefore has the greatest requirement for new therapeutic options. For this reason we were interested in assessing TNBC, HER2+, ER+ and LN+ breast cancer gene profiling datasets for novel potential therapeutic membrane protein targets for use as ADCs like T-DM1 used for HER2+ breast tumors.

To achieve this, four publicly available transcriptomic datasets (one in-house) comparing breast cancer tissue to normal breast tissue were overlapped and searched for membrane localised proteins. By careful literature searching a follow-up list of 5 proteins was chosen - SLAMF8, LRP8, TSPAN13, IGSF9, KLRG2 All of these proteins had not associated with breast cancer functional or protein validation studies in the literature. The novelty of these targets meant that commercially available antibodies for follow-up investigations were limited. Therefore, of those proteins shortlisted, only two could be clearly confirmed by Western blot analysis and preliminary IHC experiments. The analysis of a larger cohort of samples showed that both IGSF9 and KLRG2 preferentially stained the cell membrane of breast cancer tissue with very little immunoreactivity staining in normal breast and other normal tissues, that may represent a future validated molecular targetusing an ADC therapeutic approach.

6.3.1 SLAMF8

Signalling lymphocytic activation molecule family member 8 (SLAMF8) also known as B-lymphocyte activator macrophage expressed (BLAME) is a cell surface protein 285 amino acids long with a molecular weight of 32 kDa. It also has a 176 amino acid 20 kDa isoform that shares 100% homology with its longer form. SLAMF8 is part of the SLAM family of proteins which are part of the CD2 subset in the immuno globulin super family (IGSF) of proteins. It has been reported to be expressed across many lymphocytes and monocytes including T cells, B cells and NK cells as well as dendrite

cells. Tissues associated with these cell types such as spleen, thymus, lymph nodes and bone marrow all express high amounts (Kingsbury et al. 2001). Much more is known about the other members 1-7 of the SLAM family. What is known is that SLAMF8 lacks a characteristic tyrosine containing motif in it's cytoplasmic tail that binds SLAM associated protein (SAP) (Fraser et al. 2002).

This motif is known to mediate the signalling response in 6 of the other SLAM family members 1,3,4,5,6 and7 family members. It has been shown that the presence of the SAP binding cytoplasmic motif in SLAM mediates CD4+ T cell binding with B cells allowing for presented antigen and T cell signals to stimulate B cell immune response (Ma and Deenick 2011). The role of SAP and SLAM family molecules in the humoral immune response, New York Academy of Sciences, 2010). In particular this relates to X-linked linked lymphoproliferative disease (XLP). XLP is linked to a mutation in the SAP gene leading to a non functional SAP protein. Consequently the SAP binding SLAM members do not become activated leading to sufferers of this disease to be highly susceptible to other diseases and present with immune system related conditions. As SLAMF8 does not contain the cytoplasmic domain to bind SAP, SLAMF8 function has not been a focus of research as XLP related SLAM members have been. Kingsbury et al observe that SLAMF8 transduced mice show increased levels of specific B cell lineages. It remained unclear what the mechanism was behind this.

SLAMF8 IHC did not prduce staining and as a result it was not used for the larger sample cohorts. As this protein is involved in immune response it may be more useful to test for SLAMF8 on a large nuber of lymph node associated breast cancers. Using $HER2_{+,}$ TNBC and the very limited lymphnode (status known) breast cancer tissue samples we could not confirm SLAMF8 expression.

6.3.2 LRP8

Low density lipoprotein receptor protein 8 (LRP8) is a 106 kDa member of the low density lipoproteins (LDL). LRP8 binds to it's ligand called reelin, triggering numerous neuronal ehancers (Riddell et al. 1999). Tissue localisation occurs mostly around the brain but in has a broader envolvement in the nervous system. The presence of LRP8 in patients with congenital early onset coranary artery disease and mycardial infarction had significantly poorer outcome than those who did not express LRP8 (Shen et al. 2014). Being involved in brain function also means that LRP8 is associated with a number of

brain pathologies such as brain cholesterol homeostasis (Gosselet et al. 2009) and dementia (Jaeger and Pietrzik 2008).

Initial Western blot analysis validation of LRP8 expression did not produce bands and could not be determined as expressed in breast cancer cell lysates. LRP8 is part of the low density lipoprotein family and structurally contains a high number of disulphide bridge bonds. As discussed by Bajari et al low density lipoprotein analysis can be achieved using various non-ionic detergents such as n-octyl-β-D-glucopyranoside to increase solubilisation of these proteinsfrom membrane enriched fractions (Bajari et al. 2005). An alternative approach to quantify these proteins is a ligand binding assay. LRP8 has been reported to have several ligands including apolipoprotein E, F-spondin and reelinn (Divekar et al. 2014) which could be used for this type of quantification. The analysis of LRP8 in breast cancer may require further optimisation in sample preparation based on these studies.

6.3.3 TSAN13

Tetraspanin 13 (TSPAN13) is a 22 kDa protein from the transmembrane 4 superfamily. TSPAN13 contain s4 hydrophobic domains and mediates growth and motility. Literature availability is very limited with only 13 publications according to PubMed. Most of these publications are related to cancer, TSPAN13 being proposed as a cancer supressor in one of it's first publications (Huang et al. 2007). Subsequent publications however show that high TSPAN13 expression results in a poorer outcome in prostate cancer directly correlating with epithelial cell abnormality (prostatic intrapithelial neoplasia) (Arencibia et al. 2009).

The limited number of TSPAN13 studies use PCR to quantify expression (Iwai et al. 2007). The role of TSPAN8 in cancer has been well studied in the carcinoma (Kanetaka et al. 2001) and melanoma metastasis (Berthier-Vergnes et al. 2011). TSPAN12 has been implicated in breast cnacer development in xenogrpht mouse models (Knoblich et al. 2014). As these other tetraspanin family members have direct associations with cancer progression and have large functional studies associated with them TSPAN13 remains a very interesting target for further study.

6.3.4 IGSF9

Immuno globulin superfamily member 9 (IGSF9) also known as dendrite arborisation and maturation 1 (DASM1) and originally known as Protein Turtle homolog is a relatively large transmembrane protein 1179 amino acids long with a mass of 127 kDa containing 5 extracellular immunoglobulin domains and 2 fibronectin type III repeats. Discovered in 2002 IGSF9 was found to be involved in nervous system development in mice (Doudney et al. 2002). More specifically it is involved in the development of neural dendrite outgrowth and synapse maturation. Previous to this an IGSF protein was found in *Drosophila* named Turtle protein (Bodily et al. 2001). It was later found that Turtle protein was a homolog of IGSF9 suggesting that it has an important role due to it being highly conserved across species. In a bioinformatics review on IGSF9 Hansen and Walmod show that it is conserved across several vertebrates and invertebrates with varying numbers of genetic repeats and isoforms between species (Hansen and Walmod 2013). Furthermore a close homolog called IGSF9b has been reported to be coexpressed in the developing hippocampus (Mishra et al. 2008). IGSF9b is 1349 amino acids long with a mass of 147 kDa and has a 34% homology to IGSF9. Little or nothing is known about IGSF9b but it may have a role in substituting the function of IGSF9a.

A study involving *in vitro* siRNA knockdown has shown IGSF9 to be specifically involved in dendrite outgrowth (Shi et al. 2004)while axon outgrowth was unaffected. Mishra *et al* however report that IGSF9 is not specifically involved in dendrite outgrowth. This study uses DASM1 null mice to show that dendrite arborisation is uninhibited. Furthermore they repeat siRNA mediated knockdown of DASM1 in hippocampus cells isolated from the DASM1 null mice and show that dendrite development is inhibited as was seen with Shi *et al*. Mishra *et al* conclude that results of initial work be Shi *et al* were due to off target effects caused by DASM1 siRNA but they also mention the possibility that IGSF9b might compensate for the knockdown of IGSF9a. The sequence against which the IGSF9 antibody used in our work is shown below in full. The antigen sequence of the antibody matches human IGSF9a 100% and matches human IGSF9b 50%. It may be worth noting that in all of the literature mentioned above IGSF9 has been implicated in the hippocampus and neural cell development especially in the foetal and pupal stages of development stages of development. The role of IGSF9 in tumorigenesis has not been a focus of study in the literature. The potential for IGSF9 to be involved in cell signalling related to invasiveness has been suggest by Al-Anzi and Wyman through their work on the turtle

gene in Drosophila (Al-Anzi and Wyman 2009). Turtle protein in Drosophila has 5 different isoforms 2 of which are diffusible isoforms lacking the hydrophobic transmembrane region of membrane bound isoforms. In this study they conclude that the diffusible isoforms cause axonal invasion into regions not normally associated with axon branching. This characteristic of diffusible Turtle protein may suggest that the IGSF9 homolog in mammals could potentially be involved in similar signalling pathways.

As can be seen from the findings the role of IGSF9 in vertebrates and its Turtle protein homolog in *Drosophila* contradict one another. It is clear that it is involved in neural development but to what extent is not certain. Using mouse models Shi *et al* have determined that it is involved, and necessary for the dendrite outgrowth but not axon outgrowth. Using mouse models and cells isolated from the same mouse models Mishra *et al* conclude that IGSF1/DASM1 is not necessary for dendrite outgrowth. Also conversely Al-Anzi and Wyam show that the IGSF9 homologue in *Drosophila* Turtle protein is a key attractant in axon branching. This is however a different in vivo model.

6.3.4.1 IGSF9 expression in cell lines

Initial confirmation of cell line expression was conducted using a panel of representative breast cancer subtypes. By preparing a membrane enriched sample of each cell line we wished to determine the degree to which IGSF9 was associated with the cell membrane.

Based on the results IGSF9 was shown to be expressed in 5/10 of the unfractionated breast cancer cell lysates and 2/10 of the membrane protein enriched lysates. Namely these two cell lines showing membrane expression of IGSF9 were the luminal MDA-BT474 (HER2+) and MDA-MB-157 (TNBC) cell lines. According to cell line characterisation by Chavez et al. and Neve et al. (Neve et al. 2006, Chavez, Garimella and Lipkowitz 2010) the BT474 cell line is ER+, PR+ and HER2+ while the MDA-MB-157 is ER-, PR- and HER-. This would suggest, that despite IGSF9 over expression being associated with HER2+ gene array data for breast cancer tissue, IGSF9 expression is associated with markers other than ER, PR or HER2. IGSF9 is also not associated with proliferative cells which is an ADC target suitability requiremnet.

These results, while limited in their interpretation using well established subtypes, did confirm membrane expression in some cell lines. The membrane enrichment in itself may require further optimisation with specific cell lines as IGSF9 appears absent in 3 of the cell lines where there was expression in the whole un-enriched samples. In a comparison between 5 different membrane enrichment methods, the method used for this experiment, it was shown to provide the greatest degree of enrichment based on band resolution and yield (Bunger, Roblick and Habermann 2009).

6.3.4.2 IGSF9 has low expression in normal tissue

18 tissues normal tissue types were analysed for IGSF9 immunoreactivity. Namely these were normal breast, cervix, colon, gastric, lung, prostate. As with any potential target for therapeutic applications, immunoreactivity should ideally be as low as possible to reduce cytotoxicity. Results from this cohort of tissues show that IGSF9 has ideal characteristics for use as a therapeutic agent with regard to potential side effects

6.3.4.3 IGSF9 expression is not specifically associated with HER2+ breast tumor tissue

As IGSF9 was initially derived from a differential expression list of HER2+ breast tissue vs. normal breast tissue it was expected that IHC analysis of breast cancer tissue would confirm this specificity. We found that IGSF9 displayed moderate (+2) to strong (+3) immuno reactivity to 4/6 HER2+ breast tumors and 9/11 TNBC tumors. This does not show a clear bias toward HER2+ over TNBC. A larger cohort of samples in the form of two TMAs confirmed the lack of a trend related to HER2+ (n=19) and TNBC (n=16), with both subtypes displaying largely moderate (+2) staining. Stratification of this TMA based on ER+ or PR+ status showed a higher degree of negative immunoreactivity than in the HER2+ and TNBC groups. While this may indicate that ER+ and PR+ do not predict the expression of IGSF9, as shown in **Section 1.8** the percentage of ER+ and PR+ can determine predictive outcome. This data in future TMAs may be useful in determining if IGSF9 is preferentially associated with these subtypes.

Using a HER2+ TMA it was possible to shed light on this observation. All samples in this cohort were HER2+ (n=69) however ER+ and LN+ varied. Results indicated that

combinations of HER2+/ER+, HER2+/ER-, HER+/LN+ and HER2+ LN- all displayed largely moderate staining in a similar pattern to that of the HER2+ total population . This may indicate that IGSF9 immunoreactivity is not directly associated with ER+/- or LN+/- in HER2+ breast cancer

This taken together with its moderate to strong staining of TNBC tissue point toward it having high specificity for breast cancer but possibly to subtypes with the main subtypes of HER2+ and TNBC which was also mirrored in the cell line

6.3.5 KLRG2

KLRG2 (killer cell lectin-like receptor subfamily G, member 2) is a 42 kDa NK (natural killer cells) lectin-like receptor protein which is closely related to the C-type lectins (CLEC) classification of proteins. It also has a smaller annotated 32 kDa isoform. Not much is known about KLRG2. It has been linked to prostate cancer aggressiveness in a gene array study (Liu et al. 2011b) and expression in the brain (Lysenko et al. 2013) but otherwise there is limited literature for KLRG2. More information is available on the other family member KLRG1. Blast results show that there is 13% homology between the two proteins most of which occurs at the lectin binding domain.

KLRG1 is the only KLRG family member which much is known about. KLRG1 is a 195 kDa transmembrane protein preferentially expressed in NK and T-cells known to inhibit their cytotoxic function. It has been shown that KLRG1+ NK cells bind to E cadherin. Down regulation of E cadherin is known to be associated with cancer invasiveness and migration. KLRG1 then has a role analogous to that of the “missing self” mechanism of MHC-NK cell binding, potentially playing an important part in tumour surveillance. No such functional data is available on KLRG2.

6.3.5.1 KLRG2 expression in cell lines

Breast cancer cell line expression of KLRG2 was confirmed in a panel of 10/11 whole cell lysates and 6/9 membrane enriched fractions.

KLRG2 was chosen from a panel of over expressed transcripts associated with TNBC tumor tissue compared to normal tissue. As such, the Basal category of cell lines were expected to be the most enriched in the panel of cell lines. This in part was observed for

the whole lysates where very clear strong staining can be observed in the Basal A and Basal B cell lines compared to comparatively weak expression overall in the luminal categorised cell lines. For the membrane enriched fractions however 3/4 luminal showed strong KLRG2 expression while 3/6 Basal category cell lines showed strong KLRG2 expression.

While this may be explained as being due to further subtypes within these cell lines (as cell line categorisation can vary between publications) it may be as a result of the membrane fractionation. The fact that the whole cell lysates are stratified in terms of KLRG2 expression as we expected and the membrane enriched fractions are not may point toward further optimisation for particular cell lines with this method. As a proof of concept however the membrane enrichment does confirm that KLRG2 is enriched in 6/9 different breast cancer cell lines.

6.3.5.2 KLRG2 is lowly expressed in normal tissue

Using a cohort of 36 normal tissues KLRG2 was shown to display low immunoreactivity in normal tissue. Normal breast, cervix, colon, gastric, lung and prostate tissue all displayed little or no immunoreactivity compared to breast cancer tissue and in proliferating tissue

In particular normal breast tissue showed the least immunoreactivity. The normal breast tissue shows that KLRG2 has an extremely high affinity for breast cancer tissue with far higher staining intensity compared to normal breast tissue than that observed with IGSF9, which also displayed a high level of breast cancer specific immunoreactivity.

6.3.5.3 KLRG2 expression appears highest in TNBC tissue samples

From the initial whole cell lysates and the bioinformatics profiling KLRG2 was expected to display higher immunoreactivity in TNBC tissue. The membrane enriched cell lysate samples alluded to KLRG2 not being specific for TNBC. Taking a cohort of 9 HER2+ and 11 TNBC tissues it was confirmed that KLRG2 did have higher immunoreactivity in TNBC tissue. Moderate (+2) staining in 2/9 HER2+ tissues was observed while 8/11 TNBC tissues produced strong (+3) staining. This further points to

the possibility that the membrane fractionation technique requires further optimisation for each cell line.

IHC analysis of a cohort of 102 patient samples using a TMA showed that a similar relative number of tumors produced negative immunoreactivity based on HER2+, TNBC, ER+ or PR+ status. PR+ and ER+ status was however shown to produce mostly weak (+1) staining while HER2+ and TNBC status displayed mostly moderate (+2) staining. Further stratification or re classification of this TMA may reveal more about the specificity for some ER+ and PR+ tumors over others and also why HER2+ and TNBC both display a similar trend in immunoreactivity.

As with IGSF9 these results show that KLRG2 has a strong affinity for breast cancer displaying very clear membrane staining. As to what subtypes this strong affinity is associated with is not clear but it maybe be a subtype within the TNBC classification.

CHAPTER 7
Conclusion

7.1 Effect of miR-7 on the CHO cell proteome

Using quantitative label-free LC-MS/MS it was possible to identify a wide range of differentially regulated proteins making thorough pathway analysis. Two predicted direct targets to miR-7 were confirmed down-regulated after miR-7 up regulation - Catalase and Stathmin. Additionally, a previously unavailable reference list of CHO protein IDs were generated for future CHO proteomic studies.

- CHO-K1-SEAP cells were transfected with pre-miR-7 and the effect of this transient over expression was investigated using a statistical cut of ANOVA ≤0.05, fold change >1.2 and peptide identifications >1 peptide.

- No CHO specific protein identification database was available at the time of the initial analysis therefore a multi species approach with combined rat, mouse and human database searching was used to acquire IDs. In total we found 275 proteins to be differentially regulated after pre-miR-7 transfection, 134 proteins 48 hr after miR-7 up regulation and 141 proteins 96 hr after miR-7 up regulation.

- Over expression of miR-7 was validated by rtPCR and label free mass spectrometry differential expression of Histone H3, Histone H4, HSPA8, 14-3-3 epsilon, GRP78 andPDIA6 was validated by Western blot analysis.

- A reference list of 1135 unique manually annotated gene names and 746 unique manually annotated protein names associated with an in house CHO database was generated for use with future CHO proteomic experiments.

- Overlapping CHO database identifications with the multi-species search identifications the overlap between the two showed less peptides in the CHO output therefore the one peptide identifications were not filtered out for the CHO database and subsequently included in all further experiments using the CHO database. 142 proteins from the CHO database including 1 peptide matches were represented in the multi-species output (52%).

- In total we found 386 CHO proteins to be differentially regulated after pre-miR-7 transfection, 186 proteins 48 hr after miR-7 up regulation and 200 proteins 96 hr after miR-7 up regulation.

- Pathway analysis indicated that proteins related to homeostasis and anti apoptotic process were up-regulated with down regulation linked to translation, RNA binding while KEGG pathways of Glutathione metabolism were up-regulated and Ribosome pathway down-regulated which all point toward increased detoxification and reduced protein synthesis and also suggests Glutathione metabolism is functioning in a self perpetuating manner to produce antioxidant Glutathione.

- Catalase and Stathmin were predicted targets of miR-7 and were down-regulated 4.15 and 3.11 fold at 48 and 96 hr and down-regulated 22.2 fold at 96 hr respectively with validation by Western blot.

- HDAC1 was confirmed to be down-regulated both at the proteomic level and the transcriptomic level in miR-7 over expressing CHO cells. Contrary to this the Acetylation of Histone H3 and H4 were decreased suggesting that the acetylation is redirected and preserved in response to miR-7 up regulation

❖ Overall we were able to identify a larger number of proteins differentially regulated in CHO cells using a label free method, compiled a useful database of annotated proteins for downstream analysis. With this we identified potential direct targets of miR-7 in CHO cells and proposed a complex network of protein regulation associated with the reduced proliferation and increased productivity in CHO with elevated mi-R7. These protein targets may potentially be used in the future to produce a CHO cell line with this industrially desirable phenotype.

7.2 Effect of temperature shift on CHO-K1-SEAP cells using sub cellular fractionation

Using sub cellular fractionation kits we were able to reduce the sample complexity of the CHO proteome and identify proteins related to temperature shift that were uniquely identified through the fractionation, highly novel with respect to CHO and with follow up siRNA knockdown showing a significant reduction to cell proliferation, potentially providing a target to induce a temperature shift phenotype cell line for bioprocessing.

- Simple benchtop fractionation kits were used to enrich for membrane, cytoplasmic and nuclear components. Enrichment was confirmed using pathway analysis and via Western blot for membrane, cytoplasmic and nuclear specific proteins IGF1Rβ, HSP90 and catalase respectively

- Sub cellular enrichment was shown to produce 412, 309 and 241 unique to the membrane, cytoplasmic and nuclear enriched fractions when compared to the non enriched sample.

- The differential protein IDs were used to compare the in house CHO database and the non redundant publically available NCBI CHO database showing that each had a small number of protein IDs unique to each database and that both databases have a similar trend in the frequency of peptide number IDs.

- Differential analysis consisted of comparing 31 to 37 ^{o}C at each time point of 8 and 24 hr and a time course analysis identifying differentially regulated proteins at the 24 hr time point compared to the 8 hr time point at each temperature of 31 and 37 ^{o}C.

- Proteins identified as differentially regulated were shown to be mostly unique to each of the three enriched fractions with differentially regulated proteins being mostly unique to each temperature in the time course analysis.

- Pathway analysis for the initial time point comparison showed that down regulation of translation and structural process occurred at 31 ^{o}C accompanied by an increase in metabolic processing and DNA replication. For the time course analysis differentially regulated proteins at 31 ^{o}C were associated with

associated with different molecular functions and biological process but interestingly 31 ^{o}C and 37 ^{o}C pathways of the ribosomal and the spliceosome were dysregulated in both only different proteins were involved at each temperature.

- Four targets were chosen to follow up with transient knockdown. Ezrin, Moesin and Lamin A/C were chosen as highly down-regulated at 31 ^{o} C in the time point analysis and Cyclon was highly differentially regulated at 31 ^{o}C over time but not at 37 ^{o}C and not in the time point analysis.

- We found that siRNA knockdown in two out of three siRNAs significantly ($p \leq 0.05$) for Cyclon reduced viable cell number by an average of 50% and knockdown of Moesin reduced viable cell number by 66% without impairing viability. Cell size (area and perimeter) were also significantly affected with both targets.

❖ Cell size and morphology, as well as the typical metabolic and transcriptional regulation, can be seen to be involved in the temperature shift phenotype which can be seen in great detail using sub cellular fractionation. Using these deeper proteome identifications, cyclon protein was transiently knocked down and shown to significantly reduce cell proliferation and cell size and may be useful to induce a temperature shift phenotype in CHO cells to increase cellular productivity.

7.3 Identification of novel membrane targets in breast cancer

With publicly available transcriptional data sets we were able to construct a list of potential membrane proteins over expressed in TNBC, ER+, LN+ or HER2+ sub types compard to normal breast tissues. We identified a panel of proteins that had not previously been associated with breast cancer functional studies and chose these for validation experiments to determine their potential suitability as ADC molecular targets.

- The criteria used to choose final candidate protein targets (i.e. reducing the large panel of transcriptional targets down) novelty of the target (previously unexploited, no listed patents), no known functional association at the protein level (in any breast cancer or other cancer types) and availability of antibodies for validation studies.

- Five targets IGSF9, KLRG2, SLAMF8, LRP8 and TSPAN13 were chosen for initial follow up all of which had antibodies available against them, had no protein functional studies in breast cancer, were all implicated to have membrane expression in the literature and no patents in the area of breast cancer associated with them (commercially novel).

- Initial Western blots analysis showed expression of four targets TSPAN13, SLAMF8, IGSF9 and KLRG2 across 9 different breast cancer cell lines with expression of IGSF9 and KLRG2 confirmed also in membrane enriched fractions of theses cell lines.

- IHC analysis showed strong expression of IGSF9 and KLRG2 in breast cancer tissue compared to normal breast tissue. IGSF9 was derived from the HER2+ target list and was found to produce strong cytoplasmic and membrane immunoreactivity

❖ Overall IGSF9 and KLRG2 represent novel proteins with no publications relating to breast cancer. Both proteins are present in membrane enriched fractions of breast cancer cell lines. IGSF9 and KLRG2 display membrane immunoreactivity in breast cancer tissues, with moderate to strong immunoreactivity observed in both HER2 and TNBC subtypes.

- Limited expression of IGSF9 and KLRG2 was observed in normal and highly proliferating proliferative tissues compared to IGSF9/KLRG2 immunoreactivity observed in breast cancer.

- While the specific breast cancer subtype affiliation for these two targets is not totally clear,the results presented demonstate that both of these candidate targets show a high prevalence of expression across TNBC and HER2 positive breast cancers with limited normal tissue expression, suggesting that they may represent potential molecular targets that maybe ameanable to therapeutic targeting using MAbs/ ADCs.

Future work

Effect of miR-7 on the CHO cell proteome

- ❖ While catalase and stathmin were predicted to be direct targets of miR-7 and their abundance was confirmed as reduced in miR-7 transiently up-regulated CHO cells, these are indirect methods to determine their association with miR-7. A luciferase assay GFP method or biotin labelled methods are available to conduct this experiment.

- ❖ Knockdown of catalase and stathmin using transient siRNA transfection would establish if they have a role in CHO cell phenotype. Following these results a stable CRISPR/Cas9 mediated catalase and stathmin knockout cell line could be gnerated or a double knockout may be required to achieve a miR-7 over-expressing penotype.

- ❖ The glutathione metabolism pathway was found contain many up-regulated proteins in response to increased miR-7. A glutathione (GSH) assay could be used to establish if GSH production is also increasing. As this also indicated increased anti-oxidant activity, a reactive oxygen species (ROS) kit could also be used to assess this activity in miR-7 up regulated cells.

- ❖ A conflicting negative relationship between decreased histone deacetylase 1 (HDAC1) and increased histone acetylation was observed. Acetylation events could be assessed from the MS raw data as well as further Western blot analysis on additional members of the histone acetylation pathway to determine the cause of this observed histone deacetylation.

Effect of temperature shift on CHO-K1-SEAP cells using subcellular fractionation

- ❖ As we found that using subcellular fractionation doubled the number of qualitative IDs it may be possible to further increase this number with further fractionation steps such as with cation exchange, 1D gel separation, phopho/glyco/ubiquitinated protein enrichment or size exclusion chromtography, all of which can quickly and easily be achieved with commercial kits.

- ❖ The time course analysis as a comparison allowed a functionally validated target (cyclon) to be identified. While this analysis proved useful a true time course analysis has many more time points than two. Seeing as how a minimum amount of

time points in this proved useful, there may be further functional targets identified with a greater number of time points for this type of comparison.

- ❖ Of the 4 targets chosen for functional validation 3 (Ezrin, Moesin and Lamin A/C) were not confirmed to be knocked down via Western blot analysis. Different antibodies could be used to assess knockdown in these samples as humanised antibodies sometimes produce unreliable results on CHO samples. Alternatively RT-PCR could be used to assess knockdown. Moesin in particular should be reanalysed as a phenotypic response was observed with anti-Moesin siRNA transfection.

- ❖ Cyclon was confirmed knocked down via Western blot analysis and significantly reduced cell proliferation and cell size with no affect on percentage viable cells. Other assays such as a SEAP assay would determine if cyclon has an effect on productivity and confirm if it truly does cause a temperature shift phenotype when down-regulated.

- ❖ Stable gene edited clones could be generated with CRISPR/Cas9 and various knockout and insertion combination cell lines generated with Ezrin, Moesin, Lamin A/C and Cyclon with the overall objective to generate a stable cell line with a temperature shifted phenotype - resistance to apoptosis, reduced cell growth and increased productivity. As Cyclon was shown to reduce proliferation it could be more suitable to generate a stable clone with an inducible promoter so that high cell densities

Identification of novel membrane protein targets in breast cancer

- ❖ An extended panel of breast cancer cell lines representing various TNBC subtypes should be analysed for IGSF9 and KLRG2 expression.

- ❖ As TMAs represent a high throughput method of analysis, more of these should be analysed with IGSF9 and KLRG2. In particular TMAs with complete clinicopathological information including patient outcome could be analysed to determine association of the targets with clinical outcome or tumour features. To address if these targets represent potential ADC targets, (i) internalisation studies should be carried out to demonstrate that these target antigens can be internalised on the cell surface of breast cancer cells. (ii) Microscopic cellular localisation studies of cell surface staining of targets should be carried out to demonstrate cell surface localisation of the target antigens.

- Another desirable but not essential characteristic for an ADC target is that the target has a functional effect on cancer cells. Transient knockdown followed by any number of cancer functional assays such as migration, invasion, proliferation and anoikis assays should be carried out.

- IGSF9 and KLRG2 represent highly novel targets with very little functional work available. Analysing there expression in a wider range of cancer tissue may produce more novel applications.

- While the antibodies used in this study were suitable for IGSF9 and KLRG2 the antibodies used for TSPAN13, LRP8 and SLAMF8 were not suitable. Custom antibody design may solve some of these problems and allow the validation of other targets with few commercially available antibodies.

Bibliography

Tamoxifen for early breast cancer: An overview of the randomised trials. early breast cancer trialists' collaborative group. 1998. *Lancet (London, England),* 351(9114), pp.1451-1467.

Adada, M., Canals, D., Hannun, Y.A. and Obeid, L.M. 2014. Sphingolipid regulation of ezrin, radixin, and moesin proteins family: Implications for cell dynamics. *Biochimica Et Biophysica Acta,* 1841(5), pp.727-737.

Ahn, S.M., Byun, K., Cho, K., Kim, J.Y., Yoo, J.S., Kim, D., Paek, S.H., Kim, S.U., Simpson, R.J. and Lee, B. 2008. Human microglial cells synthesize albumin in brain. *PloS One,* 3(7), pp.e2829.

Al-Anzi, B. and Wyman, R.J. 2009. The drosophila immunoglobulin gene turtle encodes guidance molecules involved in axon pathfinding. *Neural Development,* 4pp.31-8104-4-31.

Albanèse, V., Yam, A.Y., Baughman, J., Parnot, C. and Frydman, J. 2006. Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell,* 124(1), pp.75-88.

Al-Fageeh, M.B., Marchant, R.J., Carden, M.J. and Smales, C.M. 2006. The cold-shock response in cultured mammalian cells: Harnessing the response for the improvement of recombinant protein production. *Biotechnology and Bioengineering,* 93(5), pp.829-835.

Al-Fageeh, M.B. and Smales, C.M. 2013. Alternative promoters regulate cold inducible RNA-binding (CIRP) gene expression and enhance transgene expression in mammalian cells. *Molecular Biotechnology,* 54(2), pp.238-249.

Al-Fageeh, M.B. and Smales, C.M. 2006. Control and regulation of the cellular responses to cold shock: The responses in yeast and mammalian systems. *The Biochemical Journal,* 397(2), pp.247-259.

Anderson, W.F., Chatterjee, N., Ershler, W.B. and Brawley, O.W. 2002. Estrogen receptor breast cancer phenotypes in the surveillance, epidemiology, and end results database. *Breast Cancer Research and Treatment,* 76(1), pp.27-36.

Arencibia, J.M., Martin, S., Perez-Rodriguez, F.J. and Bonnin, A. 2009. Gene expression profiling reveals overexpression of TSPAN13 in prostate cancer. *International Journal of Oncology,* 34(2), pp.457-463.

Arpino, G., Weiss, H., Lee, A.V., Schiff, R., De Placido, S., Osborne, C.K. and Elledge, R.M. 2005. Estrogen receptor-positive, progesterone receptor-negative breast cancer: Association with growth factor receptor expression and tamoxifen resistance. *Journal of the National Cancer Institute,* 97(17), pp.1254-1261.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. and Sherlock, G. 2000. Gene ontology: Tool for the unification of biology. the gene ontology consortium. *Nature Genetics,* 25(1), pp.25-29.

Aziz, M.A., Majeed, G.H., Diab, K.S. and Al-Tamimi, R.J. 2015. The association of oxidant-antioxidant status in patients with chronic renal failure. *Renal Failure*, pp.1-7.

Baik, J.Y., Lee, M.S., An, S.R., Yoon, S.K., Joo, E.J., Kim, Y.H., Park, H.W. and Lee, G.M. 2006. Initial transcriptome and proteome analyses of low culture temperature-induced expression in CHO cells producing erythropoietin. *Biotechnology and Bioengineering*, 93(2), pp.361-371.

Bajari, T.M., Strasser, V., Nimpf, J. and Schneider, W.J. 2005. LDL receptor family: Isolation, production, and ligand binding analysis. *Methods (San Diego, Calif.)*, 36(2), pp.109-116.

Barboro, P., Ferrari, N. and Balbi, C. 2014. Emerging roles of heterogeneous nuclear ribonucleoprotein K (hnRNP K) in cancer progression. *Cancer Letters*, 352(2), pp.152-159.

Bardou, V.J., Arpino, G., Elledge, R.M., Osborne, C.K. and Clark, G.M. 2003. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, 21(10), pp.1973-1979.

Barnabe, N. and Butler, M. 1994. Effect of temperature on nucleotide pools and monoclonal antibody production in a mouse hybridoma. *Biotechnology and Bioengineering*, 44(10), pp.1235-1245.

Barok, M., Joensuu, H. and Isola, J. 2014. Trastuzumab emtansine: Mechanisms of action and drug resistance. *Breast Cancer Research : BCR*, 16(2), pp.209.

Barron, N., Kumar, N., Sanchez, N., Doolan, P., Clarke, C., Meleady, P., O'Sullivan, F. and Clynes, M. 2011a. Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. *Journal of Biotechnology*, 151(2), pp.204-211.

Barron, N., Sanchez, N., Kelly, P. and Clynes, M. 2011b. MicroRNAs: Tiny targets for engineering CHO cell phenotypes? *Biotechnology Letters*, 33(1), pp.11-21.

Bauwens, C., Yin, T., Dang, S., Peerani, R. and Zandstra, P.W. 2005. Development of a perfusion fed bioreactor for embryonic stem cell-derived cardiomyocyte generation: Oxygen-mediated enhancement of cardiomyocyte output. *Biotechnology and Bioengineering*, 90(4), pp.452-461.

Baycin-Hizal, D., Tabb, D.L., Chaerkady, R., Chen, L., Lewis, N.E., Nagarajan, H., Sarkaria, V., Kumar, A., Wolozny, D., Colao, J., Jacobson, E., Tian, Y., O'Meally, R.N., Krag, S.S., Cole, R.N., Palsson, B.O., Zhang, H. and Betenbaugh, M. 2012. Proteomic analysis of chinese hamster ovary cells. *Journal of Proteome Research*, 11(11), pp.5265-5276.

Beck, A. and Reichert, J.M. 2013. Approval of the first biosimilar antibodies in europe: A major landmark for the biopharmaceutical industry. *MAbs*, 5(5), pp.621-623.

Becker, J., Hackl, M., Rupp, O., Jakobi, T., Schneider, J., Szczepanowski, R., Bekel, T., Borth, N., Goesmann, A., Grillari, J., Kaltschmidt, C., Noll, T., Pühler, A., Tauch, A. and Brinkrolf, K. 2011. Unraveling the chinese hamster ovary cell line transcriptome by next-generation sequencing. *Journal of Biotechnology,* 156(3), pp.227-235.

Belmont, L.D. and Mitchison, T.J. 1996. Identification of a protein that interacts with tubulin dimers and increases the catastrophe rate of microtubules. *Cell,* 84(4), pp.623-631.

Bendas, G. and Borsig, L. 2012. Cancer cell adhesion and metastasis: Selectins, integrins, and the inhibitory potential of heparins. *International Journal of Cell Biology,* 2012pp.676731.

Beretta, L., Dobransky, T. and Sobel, A. 1993. Multiple phosphorylation of stathmin. identification of four sites phosphorylated in intact cells and in vitro by cyclic AMP-dependent protein kinase and p34cdc2. *The Journal of Biological Chemistry,* 268(27), pp.20076-20084.

Berg, J., Tymoczko, J. and Stryer, L. 2002. The Citric Acid Cycle Oxidizes Two-Carbon Units *IN:* Anonymous*Biochemistry.* 5th edition ed.W H Freeman, pp.Section 17.1.

Bergamaschi, A., Frasor, J., Borgen, K., Stanculescu, A., Johnson, P., Rowland, K., Wiley, E.L. and Katzenellenbogen, B.S. 2013. 14-3-3zeta as a predictor of early time to recurrence and distant metastasis in hormone receptor-positive and -negative breast cancers. *Breast Cancer Research and Treatment,* 137(3), pp.689-696.

Berger, S., Lowe, P. and Tesar, M. 2015. Fusion protein technologies for biopharmaceuticals: Applications and challenges. *MAbs,* 7(3), pp.456-460.

Berthier-Vergnes, O., Kharbili, M.E., de la Fouchardiere, A., Pointecouteau, T., Verrando, P., Wierinckx, A., Lachuer, J., Le Naour, F. and Lamartine, J. 2011. Gene expression profiles of human melanoma cells with different invasive potential reveal TSPAN8 as a novel mediator of invasion. *British Journal of Cancer,* 104(1), pp.155-165.

Berton, S., Pellizzari, I., Fabris, L., D'Andrea, S., Segatto, I., Canzonieri, V., Marconi, D., Schiappacassi, M., Benevol, S., Gattei, V., Colombatti, A., Belletti, B. and Baldassarre, G. 2014. Genetic characterization of p27(kip1) and stathmin in controlling cell proliferation in vivo. *Cell Cycle (Georgetown, Tex.),* 13(19), pp.3100-3111.

Bertucci, F., Finetti, P., Cervera, N., Esterni, B., Hermitte, F., Viens, P. and Birnbaum, D. 2008. How basal are triple-negative breast cancers? *International Journal of Cancer.Journal International Du Cancer,* 123(1), pp.236-240.

Binder, J.X., Pletscher-Frankild, S., Tsafou, K., Stolte, C., O'Donoghue, S.I., Schneider, R. and Jensen, L.J. 2014. COMPARTMENTS: Unification and visualization of protein subcellular localization evidence. *Database : The Journal of Biological Databases and Curation,* 2014pp.bau012.

Bleckwenn, N.A. and Shiloach, J. 2004. Large-scale cell culture. *Current Protocols in Immunology / Edited by John E.Coligan ...[Et Al.],* Appendix 1pp.Appendix 1U.

Bodily, K.D., Morrison, C.M., Renden, R.B. and Broadie, K. 2001. A novel member of the ig superfamily, turtle, is a CNS-specific protein required for coordinated motor control. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience,* 21(9), pp.3113-3125.

Bohnsack, M.T., Czaplinski, K. and Gorlich, D. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, N.Y.),* 10(2), pp.185-191.

Boja, E., Rivers, R., Kinsinger, C., Mesri, M., Hiltke, T., Rahbar, A. and Rodriguez, H. 2010. Restructuring proteomics through verification. *Biomarkers in Medicine,* 4(6), pp.799-803.

Boja, E.S. and Fales, H.M. 2001. Overalkylation of a protein digest with iodoacetamide. *Analytical Chemistry,* 73(15), pp.3576-3582.

Borth, N., Zeyda, M., Kunert, R. and Katinger, H. 2000. Efficient selection of high-producing subclones during gene amplification of recombinant chinese hamster ovary cells by flow cytometry and cell sorting. *Biotechnology and Bioengineering,* 71(4), pp.266-273.

Bose, S.K., Sengupta, T.K., Bandyopadhyay, S. and Spicer, E.K. 2006. Identification of Ebp1 as a component of cytoplasmic bcl-2 mRNP (messenger ribonucleoprotein particle) complexes. *The Biochemical Journal,* 396(1), pp.99-107.

Boyle, P. 2012. Triple-negative breast cancer: Epidemiological considerations and recommendations. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO,* 23 Suppl 6pp.vi7-12.

Brattsand, G., Marklund, U., Nylander, K., Roos, G. and Gullberg, M. 1994. Cell-cycle-regulated phosphorylation of oncoprotein 18 on Ser16, Ser25 and Ser38. *European Journal of Biochemistry / FEBS,* 220(2), pp.359-368.

Bunger, S., Roblick, U.J. and Habermann, J.K. 2009. Comparison of five commercial extraction kits for subsequent membrane protein profiling. *Cytotechnology,* 61(3), pp.153-159.

Burris, H.A.,3rd, Hurwitz, H.I., Dees, E.C., Dowlati, A., Blackwell, K.L., O'Neil, B., Marcom, P.K., Ellis, M.J., Overmoyer, B., Jones, S.F., Harris, J.L., Smith, D.A., Koch, K.M., Stead, A., Mangum, S. and Spector, N.L. 2005. Phase I safety, pharmacokinetics, and clinical activity study of lapatinib (GW572016), a reversible dual inhibitor of epidermal growth factor receptor tyrosine kinases, in heavily pretreated patients with metastatic carcinomas. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 23(23), pp.5305-5313.

Butler, M. 2005. Animal cell cultures: Recent achievements and perspectives in the production of biopharmaceuticals. *Applied Microbiology and Biotechnology,* 68(3), pp.283-291.

Buyse, M., Loi, S., van't Veer, L., Viale, G., Delorenzi, M., Glas, A.M., d'Assignies, M.S., Bergh, J., Lidereau, R., Ellis, P., Harris, A., Bogaerts, J., Therasse, P., Floore, A., Amakrane, M., Piette, F., Rutgers, E., Sotiriou, C., Cardoso, F., Piccart, M.J. and

TRANSBIG Consortium. 2006. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *Journal of the National Cancer Institute,* 98(17), pp.1183-1192.

Cacciatore, J.J., Chasin, L.A. and Leonard, E.F. 2010. Gene amplification and vector engineering to achieve rapid and high-level therapeutic protein production using the dhfr-based CHO cell selection system. *Biotechnology Advances,* 28(6), pp.673-681.

Cai, Y., Yu, X., Hu, S. and Yu, J. 2009. A brief review on the mechanisms of miRNA regulation. *Genomics, Proteomics & Bioinformatics,* 7(4), pp.147-154.

Capo-Chichi, C.D., Aguida, B., Chabi, N.W., Cai, Q.K., Offrin, G., Agossou, V.K., Sanni, A. and Xu, X.X. 2015. Lamin A/C deficiency is an independent risk factor for cervical cancer. *Cellular Oncology (Dordrecht),*

Carey, L.A., Dees, E.C., Sawyer, L., Gatti, L., Moore, D.T., Collichio, F., Ollila, D.W., Sartor, C.I., Graham, M.L. and Perou, C.M. 2007. The triple negative paradox: Primary tumor chemosensitivity of breast cancer subtypes. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 13(8), pp.2329-2334.

Carlage, T., Hincapie, M., Zang, L., Lyubarskaya, Y., Madden, H., Mhatre, R. and Hancock, W.S. 2009. Proteomic profiling of a high-producing chinese hamster ovary cell culture. *Analytical Chemistry,* 81(17), pp.7357-7362.

Carlson, R.W., Allred, D.C., Anderson, B.O., Burstein, H.J., Carter, W.B., Edge, S.B., Erban, J.K., Farrar, W.B., Goldstein, L.J., Gradishar, W.J., Hayes, D.F., Hudis, C.A., Jahanzeb, M., Kiel, K., Ljung, B.M., Marcom, P.K., Mayer, I.A., McCormick, B., Nabell, L.M., Pierce, L.J., Reed, E.C., Smith, M.L., Somlo, G., Theriault, R.L., Topham, N.S., Ward, J.H., Winer, E.P., Wolff, A.C. and NCCN Breast Cancer Clinical Practice Guidelines Panel. 2009. Breast cancer. clinical practice guidelines in oncology. *Journal of the National Comprehensive Cancer Network : JNCCN,* 7(2), pp.122-192.

Castiglia, S., Mareschi, K., Labanca, L., Lucania, G., Leone, M., Sanavio, F., Castello, L., Rustichelli, D., Signorino, E., Gunetti, M., Bergallo, M., Bordiga, A.M., Ferrero, I. and Fagioli, F. 2014. Inactivated human platelet lysate with psoralen: A new perspective for mesenchymal stromal cell production in good manufacturing practice conditions. *Cytotherapy,* 16(6), pp.750-763.

Chandramouli, K. and Qian, P.Y. 2009. Proteomics: Challenges, techniques and possibilities to overcome biological sample complexity. *Human Genomics and Proteomics : HGP,* 2009pp.10.4061/2009/239204.

Chase, L.G., Lakshmipathy, U., Solchaga, L.A., Rao, M.S. and Vemuri, M.C. 2010. A novel serum-free medium for the expansion of human mesenchymal stem cells. *Stem Cell Research & Therapy,* 1(1), pp.8.

Chavez, K.J., Garimella, S.V. and Lipkowitz, S. 2010. Triple negative breast cancer cell lines: One tool in the search for better treatment of triple negative breast cancer. *Breast Disease,* 32(1-2), pp.35-48.

Cheang, M.C., Voduc, D., Bajdik, C., Leung, S., McKinney, S., Chia, S.K., Perou, C.M. and Nielsen, T.O. 2008. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 14(5), pp.1368-1376.

Chen, J., Abi-Daoud, M., Wang, A., Yang, X., Zhang, X., Feilotter, H.E. and Tron, V.A. 2013. Stathmin 1 is a potential novel oncogene in melanoma. *Oncogene,* 32(10), pp.1330-1337.

Cheng, A.M., Byrom, M.W., Shelton, J. and Ford, L.P. 2005. Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Research,* 33(4), pp.1290-1297.

Chervenkov, T., Shishkov, R. and Tonchev, A.B. 2013. Expression and differential response to haloperidol treatment of Cyclon/CCDC86 mRNA in schizophrenia patients. *Neurochemistry International,* 62(6), pp.870-872.

Chi, S.W., Hannon, G.J. and Darnell, R.B. 2012. An alternative mode of microRNA target recognition. *Nature Structural & Molecular Biology,* 19(3), pp.321-327.

Chiou, S.Y., Lee, Y.S., Jeng, M.J., Tsao, P.C. and Soong, W.J. 2013. Moderate hypothermia attenuates oxidative stress injuries in alveolar epithelial A549 cells. *Experimental Lung Research,* 39(6), pp.217-228.

Cho, W. and Stahelin, R.V. 2005. Membrane-protein interactions in cell signaling and membrane trafficking. *Annual Review of Biophysics and Biomolecular Structure,* 34pp.119-151.

Chong, H.S., Campbell, L., Padula, M.P., Hill, C., Harry, E., Li, S.S., Wilkins, M.R., Herbert, B. and Carter, D. 2012. Time-course proteome analysis reveals the dynamic response of cryptococcus gattii cells to fluconazole. *PloS One,* 7(8), pp.e42835.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V. and Mann, M. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science,* 325(5942), pp.834-840.

Christin, C., Bischoff, R. and Horvatovich, P. 2011. Data processing pipelines for comprehensive profiling of proteomics samples by label-free LC–MS for biomarker discovery. *Talanta,* 83(4), pp.1209-1224.

Ciciarello, M., Roscioli, E., Di Fiore, B., Di Francesco, L., Sobrero, F., Bernard, D., Mangiacasale, R., Harel, A., Schinina, M.E. and Lavia, P. 2010. Nuclear reformation after mitosis requires downregulation of the ran GTPase effector RanBP1 in mammalian cells. *Chromosoma,* 119(6), pp.651-668.

Clarke, C., Barron, N., Gallagher, M., Henry, M., Meleady, P. and Clynes, M. 2012. Target Prediction Algorithms and Bioinformatics Resources for miRNA Studies *IN:* Barron, N. (ed.) Springer Netherlands, pp.29-48.

Clarke, P.R. and Zhang, C. 2008. Spatial and temporal coordination of mitosis by ran GTPase. *Nature Reviews.Molecular Cell Biology,* 9(6), pp.464-477.

Colomer, R., Beltran, M., Dorcas, J., Cortes-Funes, H., Hornedo, J., Valentin, V., Vargas, C., Mendiola, C. and Ciruelos, E. 2005. It is not time to stop progesterone receptor testing in breast cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 23(16), pp.3868-9; author reply 3869-70.

Correa-Medina, M., Bravo-Egana, V., Rosero, S., Ricordi, C., Edlund, H., Diez, J. and Pastori, R.L. 2009. MicroRNA miR-7 is preferentially expressed in endocrine cells of the developing and adult human pancreas. *Gene Expression Patterns : GEP,* 9(4), pp.193-199.

Crown, J., O'Shaughnessy, J. and Gullo, G. 2012. Emerging targeted therapies in triple-negative breast cancer. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO,* 23 Suppl 6pp.vi56-65.

Cui, S.M., Zhao, J.X., Liu, X.M., Chen,Y.Q., Zhang,H. and Chen, W. 2016. Maximum-biomass concentration predicition for Bifidobacteria in the pH-contolled fed-batch culture. *Applied Microbiology,* 62(3), pp.256-263.

Cui, X., Hu, Z.P., Li, Z., Gao, P.J. and Zhu, J.Y. 2015. Overexpression of chaperonin containing TCP1, subunit 3 predicts poor prognosis in hepatocellular carcinoma. *World Journal of Gastroenterology,* 21(28), pp.8588-8604.

Cuzick, J., Dowsett, M., Pineda, S., Wale, C., Salter, J., Quinn, E., Zabaglo, L., Mallon, E., Green, A.R., Ellis, I.O., Howell, A., Buzdar, A.U. and Forbes, J.F. 2011. Prognostic value of a combined estrogen receptor, progesterone receptor, ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 29(32), pp.4273-4278.

Dai, X., Chen, A. and Bai, Z. 2014. Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling. *Scientific Reports,* 4pp.6566.

Dan Dunn, J., Alvarez, L.A., Zhang, X. and Soldati, T. 2015. Reactive oxygen species and mitochondria: A nexus of cellular homeostasis. *Redox Biology,* 6pp.472-485.

Dandachi, N., Dietze, O. and Hauser-Kronberger, C. 2002. Chromogenic in situ hybridization: A novel approach to a practical and sensitive method for the detection of HER2 oncogene in archival human breast carcinoma. *Laboratory Investigation; a Journal of Technical Methods and Pathology,* 82(8), pp.1007-1014.

Das, A., Grotsky, D.A., Neumann, M.A., Kreienkamp, R., Gonzalez-Suarez, I., Redwood, A.B., Kennedy, B.K., Stewart, C.L. and Gonzalo, S. 2013. Lamin A Deltaexon9 mutation leads to telomere and chromatin defects but not genomic instability. *Nucleus (Austin, Tex.),* 4(5), pp.410-419.

Dasari, S. and Tchounwou, P.B. 2014. Cisplatin in cancer therapy: Molecular mechanisms of action. *European Journal of Pharmacology,* 740pp.364-378.

Davie, J.R. 2003. Inhibition of histone deacetylase activity by butyrate. *The Journal of Nutrition,* 133(7 Suppl), pp.2485S-2493S.

De Jesus, M. and Wurm, F.M. 2011. Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. *European Journal of Pharmaceutics and Biopharmaceutics : Official Journal of Arbeitsgemeinschaft Fur Pharmazeutische Verfahrenstechnik e.V,* 78(2), pp.184-188.

Devraj, K., Geguchadze, R., Klinger, M.E., Freeman, W.M., Mokashi, A., Hawkins, R.A. and Simpson, I.A. 2009. Improved membrane protein solubilization and clean-up for optimum two-dimensional electrophoresis utilizing GLUT-1 as a classic integral membrane protein. *Journal of Neuroscience Methods,* 184(1), pp.119-123.

Divekar, S.D., Burrell, T.C., Lee, J.E., Weeber, E.J. and Rebeck, G.W. 2014. Ligand-induced homotypic and heterotypic clustering of apolipoprotein E receptor 2. *The Journal of Biological Chemistry,* 289(23), pp.15894-15903.

Doi, Y., Itoh, M., Yonemura, S., Ishihara, S., Takano, H., Noda, T. and Tsukita, S. 1999. Normal development of mice and unimpaired cell adhesion/cell motility/actin-based cytoskeleton without compensatory up-regulation of ezrin or radixin in moesin gene knockout. *The Journal of Biological Chemistry,* 274(4), pp.2315-2321.

Doolan, P., Meleady, P., Barron, N., Henry, M., Gallagher, R., Gammell, P., Melville, M., Sinacore, M., McCarthy, K., Leonard, M., Charlebois, T. and Clynes, M. 2010. Microarray and proteomics expression profiling identifies several candidates, including the valosin-containing protein (VCP), involved in regulating high cellular growth rate in production CHO cell lines. *Biotechnology and Bioengineering,* 106(1), pp.42-56.

Doudney, K., Murdoch, J.N., Braybrook, C., Paternotte, C., Bentley, L., Copp, A.J. and Stanier, P. 2002. Cloning and characterization of Igsf9 in mouse and human: A new member of the immunoglobulin superfamily expressed in the developing nervous system. *Genomics,* 79(5), pp.663-670.

Douglas, D.J. and Konenkov, N.V. 2014. Mass selectivity of dipolar resonant excitation in a linear quadrupole ion trap. *Rapid Communications in Mass Spectrometry : RCM,* 28(5), pp.430-438.

Dowsett, M., Cuzick, J., Wale, C., Howell, T., Houghton, J. and Baum, M. 2005. Retrospective analysis of time to recurrence in the ATAC trial according to hormone receptor status: An hypothesis-generating study. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 23(30), pp.7512-7517.

Dowsett, M., Houghton, J., Iden, C., Salter, J., Farndon, J., A'Hern, R., Sainsbury, R. and Baum, M. 2006. Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, EGF receptor and HER2 status. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO,* 17(5), pp.818-826.

Dreesen, O., Chojnowski, A., Ong, P.F., Zhao, T.Y., Common, J.E., Lunny, D., Lane, E.B., Lee, S.J., Vardy, L.A., Stewart, C.L. and Colman, A. 2013. Lamin B1 fluctuations have differential effects on cellular proliferation and senescence. *The Journal of Cell Biology,* 200(5), pp.605-617.

Druz, A., Betenbaugh, M. and Shiloach, J. 2012. Glucose depletion activates mmu-miR-466h-5p expression through oxidative stress and inhibition of histone deacetylation. *Nucleic Acids Research,* 40(15), pp.7291-7302.

Druz, A., Son, Y.J., Betenbaugh, M. and Shiloach, J. 2013. Stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in CHO cells. *Metabolic Engineering,* 16pp.87-94.

Dunnwald, L.K., Rossing, M.A. and Li, C.I. 2007. Hormone receptor status, tumor characteristics, and prognosis: A prospective cohort of breast cancer patients. *Breast Cancer Research : BCR,* 9(1), pp.R6.

Duquesne, K. and Sturgis, J.N. 2010. Membrane protein solubilization. *Methods in Molecular Biology (Clifton, N.J.),* 601pp.205-217.

Duttaroy, A., Bourbeau, D., Wang, X.L. and Wang, E. 1998. Apoptosis rate can be accelerated or decelerated by overexpression or reduction of the level of elongation factor-1 alpha. *Experimental Cell Research,* 238(1), pp.168-176.

Dweep, H., Sticht, C., Pandey, P. and Gretz, N. 2011. miRWalk--database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. *Journal of Biomedical Informatics,* 44(5), pp.839-847.

Eiteman, M.A. and Altman, E. 2006. Overcoming acetate in escherichia coli recombinant protein fermentations. *Trends in Biotechnology,* 24(11), pp.530-536.

Emadali, A., Rousseaux, S., Bruder-Costa, J., Rome, C., Duley, S., Hamaidia, S., Betton, P., Debernardi, A., Leroux, D., Bernay, B., Kieffer-Jaquinod, S., Combes, F., Ferri, E., McKenna, C.E., Petosa, C., Bruley, C., Garin, J., Ferro, M., Gressin, R., Callanan, M.B. and Khochbin, S. 2013. Identification of a novel BET bromodomain inhibitor-sensitive, gene regulatory circuit that controls rituximab response and tumour growth in aggressive lymphoid cancers. *EMBO Molecular Medicine,* 5(8), pp.1180-1195.

Ernst, W., Trummer, E., Mead, J., Bessant, C., Strelec, H., Katinger, H. and Hesse, F. 2006. Evaluation of a genomics platform for cross-species transcriptome analysis of recombinant CHO cells. *Biotechnology Journal,* 1(6), pp.639-650.

Fan, C., Oh, D.S., Wessels, L., Weigelt, B., Nuyten, D.S., Nobel, A.B., van't Veer, L.J. and Perou, C.M. 2006. Concordance among gene-expression-based predictors for breast cancer. *The New England Journal of Medicine,* 355(6), pp.560-569.

Farmer, P., Bonnefoi, H., Becette, V., Tubiana-Hulin, M., Fumoleau, P., Larsimont, D., Macgrogan, G., Bergh, J., Cameron, D., Goldstein, D., Duss, S., Nicoulaz, A.L., Brisken, C., Fiche, M., Delorenzi, M. and Iggo, R. 2005. Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene,* 24(29), pp.4660-4671.

Filipowicz, W., Bhattacharyya, S.N. and Sonenberg, N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nature Reviews.Genetics,* 9(2), pp.102-114.

Fischer, S., Buck, T., Wagner, A., Ehrhart, C., Giancaterino, J., Mang, S., Schad, M., Mathias, S., Aschrafi, A., Handrick, R. and Otte, K. 2014. A functional high-content miRNA screen identifies miR-30 family to boost recombinant protein production in CHO cells. *Biotechnology Journal,* 9(10), pp.1279-1292.

Fischer, S., Paul, A.J., Wagner, A., Mathias, S., Geiss, M., Schandock, F., Domnowski, M., Zimmermann, J., Handrick, R., Hesse, F. and Otte, K. 2015. miR-2861 as novel HDAC5 inhibitor in CHO cells enhances productivity while maintaining product quality. *Biotechnology and Bioengineering,* 112(10), pp.2142-2153.

Fogolin, M.B., Wagner, R., Etcheverrigaray, M. and Kratje, R. 2004. Impact of temperature reduction and expression of yeast pyruvate carboxylase on hGM-CSF-producing CHO cells. *Journal of Biotechnology,* 109(1-2), pp.179-191.

Fomina-Yadlin, D., Mujacic, M., Maggiora, K., Quesnell, G., Saleem, R. and McGrew, J.T. 2015. Transcriptome analysis of a CHO cell line expressing a recombinant therapeutic protein treated with inducers of protein expression. *Journal of Biotechnology,* 212pp.106-115.

Fountoulakis, M. and Takacs, B. 2001. Effect of strong detergents and chaotropes on the detection of proteins in two-dimensional gels. *Electrophoresis,* 22(9), pp.1593-1602.

Fox, S.R., Patel, U.A., Yap, M.G. and Wang, D.I. 2004. Maximizing interferon-gamma production by chinese hamster ovary cells through temperature shift optimization: Experimental and modeling. *Biotechnology and Bioengineering,* 85(2), pp.177-184.

Francis, P., Martinez, D.M., Taghipour, F., Bowen, B.D. and Haynes, C.A. 2006. Optimizing the rotor design for controlled-shear affinity filtration using computational fluid dynamics. *Biotechnology and Bioengineering,* 95(6), pp.1207-1217.

Franco, R., Schoneveld, O.J., Pappa, A. and Panayiotidis, M.I. 2007. The central role of glutathione in the pathophysiology of human diseases. *Archives of Physiology and Biochemistry,* 113(4-5), pp.234-258.

Fraser, C.C., Howie, D., Morra, M., Qiu, Y., Murphy, C., Shen, Q., Gutierrez-Ramos, J.C., Coyle, A., Kingsbury, G.A. and Terhorst, C. 2002. Identification and characterization of SF2000 and SF2001, two new members of the immune receptor SLAM/CD2 family. *Immunogenetics,* 53(10-11), pp.843-850.

Fredholm, B.B., Battig, K., Holmen, J., Nehlig, A. and Zvartau, E.E. 1999. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews,* 51(1), pp.83-133.

Fulford, L.G., Easton, D.F., Reis-Filho, J.S., Sofronis, A., Gillett, C.E., Lakhani, S.R. and Hanby, A. 2006. Specific morphological features predictive for the basal phenotype in grade 3 invasive ductal carcinoma of breast. *Histopathology,* 49(1), pp.22-34.

Furukawa, K. and Ohsuye, K. 1998. Effect of culture temperature on a recombinant CHO cell line producing a C-terminal alpha-amidating enzyme. *Cytotechnology,* 26(2), pp.153-164.

Gammell, P., Barron, N., Kumar, N. and Clynes, M. 2007. Initial identification of low temperature and culture stage induction of miRNA expression in suspension CHO-K1 cells. *Journal of Biotechnology,* 130(3), pp.213-218.

Gamper, N., Stockand, J.D. and Shapiro, M.S. 2005. The use of chinese hamster ovary (CHO) cells in the study of ion channels. *Journal of Pharmacological and Toxicological Methods,* 51(3), pp.177-185.

Gauci, S., Helbig, A.O., Slijper, M., Krijgsveld, J., Heck, A.J. and Mohammed, S. 2009. Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. *Analytical Chemistry,* 81(11), pp.4493-4501.

Gershon, P.D. 2014. Cleaved and missed sites for trypsin, lys-C, and lys-N can be predicted with high confidence on the basis of sequence context. *Journal of Proteome Research,* 13(2), pp.702-709.

Ghai, R., Mobli, M., Norwood, S.J., Bugarcic, A., Teasdale, R.D., King, G.F. and Collins, B.M. 2011. Phox homology band 4.1/ezrin/radixin/moesin-like proteins function as molecular scaffolds that interact with cargo receptors and ras GTPases. *Proceedings of the National Academy of Sciences of the United States of America,* 108(19), pp.7763-7768.

Giese, G. and Wunderlich, F. 1983. In vitro ribosomal ribonucleoprotein transport. temperature-induced "graded unlocking" of nuclei. *The Journal of Biological Chemistry,* 258(1), pp.131-135.

Goker, E., Waltham, M., Kheradpour, A., Trippett, T., Mazumdar, M., Elisseyeff, Y., Schnieders, B., Steinherz, P., Tan, C. and Berman, E. 1995. Amplification of the dihydrofolate reductase gene is a mechanism of acquired resistance to methotrexate in patients with acute lymphoblastic leukemia and is correlated with p53 gene mutations. *Blood,* 86(2), pp.677-684.

Goldhirsch, A., Wood, W.C., Gelber, R.D., Coates, A.S., Thurlimann, B., Senn, H.J. and 10th St. Gallen conference. 2007. Progress and promise: Highlights of the international expert consensus on the primary therapy of early breast cancer 2007. *Annals of Oncology : Official Journal of the European Society for Medical Oncology / ESMO,* 18(7), pp.1133-1144.

Gomme, P.T., Hunt, B.M., Tatford, O.C., Johnston, A. and Bertolini, J. 2006. Effect of lobe pumping on human albumin: Investigating the underlying mechanisms of aggregate formation. *Biotechnology and Applied Biochemistry,* 43(Pt 2), pp.103-111.

Gosselet, F., Candela, P., Sevin, E., Berezowski, V., Cecchelli, R. and Fenart, L. 2009. Transcriptional profiles of receptors and transporters involved in brain cholesterol homeostasis at the blood-brain barrier: Use of an in vitro model. *Brain Research,* 1249pp.34-42.

Graham, J.M. 2015. Fractionation of subcellular organelles. *Current Protocols in Cell Biology / Editorial Board, Juan S.Bonifacino ...[Et Al.],* 69pp.3.1.1-22.

Griffiths, S.W. and Cooney, C.L. 2002. Development of a peptide mapping procedure to identify and quantify methionine oxidation in recombinant human alpha1-antitrypsin. *Journal of Chromatography.A,* 942(1-2), pp.133-143.

Gruber, R., Karreth, F., Kandler, B., Fuerst, G., Rot, A., Fischer, M.B. and Watzek, G. 2004. Platelet-released supernatants increase migration and proliferation, and decrease osteogenic differentiation of bone marrow-derived mesenchymal progenitor cells under in vitro conditions. *Platelets,* 15(1), pp.29-35.

Gschwantler-Kaulich, D., Natter, C., Steurer, S., Walter, I., Thomas, A., Salama, M. and Singer, C.F. 2013. Increase in ezrin expression from benign to malignant breast tumours. *Cellular Oncology (Dordrecht),* 36(6), pp.485-491.

Guvakova, M.A. and Surmacz, E. 1997. Overexpressed IGF-I receptors reduce estrogen growth requirements, enhance survival, and promote E-cadherin-mediated cell-cell adhesion in human breast cancer cells. *Experimental Cell Research,* 231(1), pp.149-162.

Hachiya, M. and Akashi, M. 2005. Catalase regulates cell growth in HL60 human promyelocytic cells: Evidence for growth regulation by H(2)O(2). *Radiation Research,* 163(3), pp.271-282.

Hansen, M. and Walmod, P.S. 2013. IGSF9 family proteins. *Neurochemical Research,* 38(6), pp.1236-1251.

Harris, L.N., Broadwater, G., Lin, N.U., Miron, A., Schnitt, S.J., Cowan, D., Lara, J., Bleiweiss, I., Berry, D., Ellis, M., Hayes, D.F., Winer, E.P. and Dressler, L. 2006. Molecular subtypes of breast cancer in relation to paclitaxel response and outcomes in women with metastatic disease: Results from CALGB 9342. *Breast Cancer Research : BCR,* 8(6), pp.R66.

Harris, L.N., You, F., Schnitt, S.J., Witkiewicz, A., Lu, X., Sgroi, D., Ryan, P.D., Come, S.E., Burstein, H.J., Lesnikoski, B.A., Kamma, M., Friedman, P.N., Gelman, R., Iglehart, J.D. and Winer, E.P. 2007. Predictors of resistance to preoperative trastuzumab and vinorelbine for HER2-positive early breast cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 13(4), pp.1198-1207.

Hata, A. and Kashima, R. 2015. Dysregulation of microRNA biogenesis machinery in cancer. *Critical Reviews in Biochemistry and Molecular Biology,* pp.1-14.

Hayden, M.R. and Tyagi, S.C. 2003. Myocardial redox stress and remodeling in metabolic syndrome, type 2 diabetes mellitus, and congestive heart failure. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research,* 9(7), pp.SR35-52.

Hellawell, G.O., Turner, G.D., Davies, D.R., Poulsom, R., Brewster, S.F. and Macaulay, V.M. 2002. Expression of the type 1 insulin-like growth factor receptor is up-regulated in primary prostate cancer and commonly persists in metastatic disease. *Cancer Research,* 62(10), pp.2942-2950.

Hendrick, V., Winnepenninckx, P., Abdelkafi, C., Vandeputte, O., Cherlet, M., Marique, T., Renemann, G., Loa, A., Kretzmer, G. and Werenne, J. 2001. Increased

productivity of recombinant tissular plasminogen activator (t-PA) by butyrate and shift of temperature: A cell cycle phases analysis. *Cytotechnology,* 36(1-3), pp.71-83.

Hernandez Bort, J.A., Hackl, M., Hoflmayer, H., Jadhav, V., Harreither, E., Kumar, N., Ernst, W., Grillari, J. and Borth, N. 2012. Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures. *Biotechnology Journal,* 7(4), pp.500-515.

Hieber, L., Beck, H.P. and Lucke-Huhle, C. 1981. G2-delay after irradiation with alpha-particles as studied in synchronized cultures and by the bromodeoxyuridine-33258H technique. *Cytometry,* 2(3), pp.175-178.

Ho, C.S., Lam, C.W., Chan, M.H., Cheung, R.C., Law, L.K., Lit, L.C., Ng, K.F., Suen, M.W. and Tai, H.L. 2003. Electrospray ionisation mass spectrometry: Principles and clinical applications. *The Clinical Biochemist.Reviews / Australian Association of Clinical Biochemists,* 24(1), pp.3-12.

Hodge, K., Have, S.T., Hutton, L. and Lamond, A.I. 2013. Cleaning up the masses: Exclusion lists to reduce contamination with HPLC-MS/MS. *Journal of Proteomics,* 88pp.92-103.

Hoshino, A. and Fujii, H. 2007. Redundant promoter elements mediate IL-3-induced expression of a novel cytokine-inducible gene, cyclon. *FEBS Letters,* 581(5), pp.975-980.

Hsu, H.H., Hoffmann, S., Di Marco, G.S., Endlich, N., Peter-Katalinic, J., Weide, T. and Pavenstadt, H. 2011. Downregulation of the antioxidant protein peroxiredoxin 2 contributes to angiotensin II-mediated podocyte apoptosis. *Kidney International,* 80(9), pp.959-969.

Huang da, W., Sherman, B.T. and Lempicki, R.A. 2009a. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research,* 37(1), pp.1-13.

Huang da, W., Sherman, B.T. and Lempicki, R.A. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols,* 4(1), pp.44-57.

Huang, H., Sossey-Alaoui, K., Beachy, S.H. and Geradts, J. 2007. The tetraspanin superfamily member NET-6 is a new tumor suppressor gene. *Journal of Cancer Research and Clinical Oncology,* 133(10), pp.761-769.

Huggett, B. and Lahteenmaki, R. 2012. Public biotech 2011--the numbers. *Nature Biotechnology,* 30(8), pp.751-757.

Humphrey, G.W., Wang, Y., Russanova, V.R., Hirai, T., Qin, J., Nakatani, Y. and Howard, B.H. 2001. Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and mta-L1. *The Journal of Biological Chemistry,* 276(9), pp.6817-6824.

Inic, Z., Zegarac, M., Inic, M., Markovic, I., Kozomara, Z., Djurisic, I., Inic, I., Pupic, G. and Jancic, S. 2014. Difference between luminal A and luminal B subtypes

according to ki-67, tumor size, and progesterone receptor negativity providing prognostic information. *Clinical Medicine Insights.Oncology,* 8pp.107-111.

Iorio, M.V. and Croce, C.M. 2012. MicroRNA dysregulation in cancer: Diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Molecular Medicine,* 4(3), pp.143-159.

Irvine, G.B. 2001. Determination of molecular size by size-exclusion chromatography (gel filtration). *Current Protocols in Cell Biology / Editorial Board, Juan S.Bonifacino ...[Et Al.],* Chapter 5pp.Unit 5.5.

Ismail, K.S.K., Najafpour, G., Younesi, H., Mohamed, A.R. and Kamaruddin, A.H. 2008. Biological hydrogen production from CO: Bioreactor performance. *Biochemical Engineering Journal,* 39(3), pp.468-477.

Iwai, K., Ishii, M., Ohshima, S., Miyatake, K. and Saeki, Y. 2007. Expression and function of transmembrane-4 superfamily (tetraspanin) proteins in osteoclasts: Reciprocal roles of tspan-5 and NET-6 during osteoclastogenesis. *Allergology International : Official Journal of the Japanese Society of Allergology,* 56(4), pp.457-463.

Jadhav, V., Hackl, M., Klanert, G., Hernandez Bort, J.A., Kunert, R., Grillari, J. and Borth, N. 2014. Stable overexpression of miR-17 enhances recombinant protein production of CHO cells. *Journal of Biotechnology,* 175pp.38-44.

Jaeger, S. and Pietrzik, C.U. 2008. Functional role of lipoprotein receptors in alzheimer's disease. *Current Alzheimer Research,* 5(1), pp.15-25.

Jaffe, J.D., Keshishian, H., Chang, B., Addona, T.A., Gillette, M.A. and Carr, S.A. 2008. Accurate inclusion mass screening: A bridge from unbiased discovery to targeted assay development for biomarker verification. *Molecular & Cellular Proteomics : MCP,* 7(10), pp.1952-1962.

Jaspe, J. and Hagen, S.J. 2006. Do protein molecules unfold in a simple shear flow? *Biophysical Journal,* 91(9), pp.3415-3424.

Jayapal, K.P., Wlaschin, K.F., Hu, W. and Yap, M.G. 2007. Recombinant protein therapeutics from CHO cells-20 years and counting. *Chemical Engineering Progress,* 103(10), pp.40.

Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E. and Forman, D. 2011. Global cancer statistics. *CA: A Cancer Journal for Clinicians,* 61(2), pp.69-90.

Jenkins, N., Parekh, R.B. and James, D.C. 1996. Getting the glycosylation right: Implications for the biotechnology industry. *Nature Biotechnology,* 14(8), pp.975-981.

Jensen, E.B. and Carlsen, S. 1990. Production of recombinant human growth hormone in escherichia coli: Expression of different precursors and physiological effects of glucose, acetate, and salts. *Biotechnology and Bioengineering,* 36(1), pp.1-11.

Jensen, O.N. 2004. Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. *Current Opinion in Chemical Biology,* 8(1), pp.33-41.

Jeon, M.K., Yu, D.Y. and Lee, G.M. 2011. Combinatorial engineering of ldh-a and bcl-2 for reducing lactate production and improving cell growth in dihydrofolate reductase-deficient chinese hamster ovary cells. *Applied Microbiology and Biotechnology,* 92(4), pp.779-790.

Jing, D., Parikh, A. and Tzanakakis, E.S. 2010. Cardiac cell generation from encapsulated embryonic stem cells in static and scalable culture systems. *Cell Transplantation,* 19(11), pp.1397-1412.

Jing, Y., Qian, Y., Ghandi, M., He, A., Borys, M.C., Pan, S.H. and Li, Z.J. 2012. A mechanistic study on the effect of dexamethasone in moderating cell death in chinese hamster ovary cell cultures. *Biotechnology Progress,* 28(2), pp.490-496.

Jing, Y., Qian, Y. and Li, Z.J. 2010. Sialylation enhancement of CTLA4-ig fusion protein in chinese hamster ovary cells by dexamethasone. *Biotechnology and Bioengineering,* 107(3), pp.488-496.

Jun, S.C., Kim, M.S., Baik, J.Y., Hwang, S.O. and Lee, G.M. 2005. Selection strategies for the establishment of recombinant chinese hamster ovary cell line with dihydrofolate reductase-mediated gene amplification. *Applied Microbiology and Biotechnology,* 69(2), pp.162-169.

Kanehisa, M., Goto, S., Kawashima, S. and Nakaya, A. 2002. The KEGG databases at GenomeNet. *Nucleic Acids Research,* 30(1), pp.42-46.

Kanetaka, K., Sakamoto, M., Yamamoto, Y., Yamasaki, S., Lanza, F., Kanematsu, T. and Hirohashi, S. 2001. Overexpression of tetraspanin CO-029 in hepatocellular carcinoma. *Journal of Hepatology,* 35(5), pp.637-642.

Kantardjieff, A., Jacob, N.M., Yee, J.C., Epstein, E., Kok, Y.J., Philp, R., Betenbaugh, M. and Hu, W.S. 2010. Transcriptome and proteome analysis of chinese hamster ovary cells under low temperature and butyrate treatment. *Journal of Biotechnology,* 145(2), pp.143-159.

Kato, M.V., Sato, H., Nagayoshi, M. and Ikawa, Y. 1997. Upregulation of the elongation factor-1alpha gene by p53 in association with death of an erythroleukemic cell line. *Blood,* 90(4), pp.1373-1378.

Katz, E., Fon, M., Eigenheer, R., Phinney, B., Fass, J., Lin, D., Sadka, A. and Blumwald, E. 2010. A label-free differential quantitative mass spectrometry method for the characterization and identification of protein changes during citrus fruit development. *Proteome Science,* 8(1), pp.68.

Kaufmann, H., Mazur, X., Fussenegger, M. and Bailey, J.E. 1999. Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. *Biotechnology and Bioengineering,* 63(5), pp.573-582.

Keane, J.T., Ryan, D. and Gray, P.P. 2003. Effect of shear stress on expression of a recombinant protein by chinese hamster ovary cells. *Biotechnology and Bioengineering*, 81(2), pp.211-220.

Kehoe, D.E., Jing, D., Lock, L.T. and Tzanakakis, E.S. 2010. Scalable stirred-suspension bioreactor culture of human pluripotent stem cells. *Tissue Engineering.Part A*, 16(2), pp.405-421.

Kelly, P.S., Breen, L., Gallagher, C., Kelly, S., Henry, M., Lao, N.T., Meleady, P., O'Gorman, D., Clynes, M. and Barron, N. 2015. Re-programming CHO cell metabolism using miR-23 tips the balance towards a highly productive phenotype. *Biotechnology Journal*, 10(7), pp.1029-1040.

Khelwatty, S.A., Essapen, S., Seddon, A.M. and Modjtahedi, H. 2011. Growth response of human colorectal tumour cell lines to treatment with afatinib (BIBW2992), an irreversible erbB family blocker, and its association with expression of HER family members. *International Journal of Oncology*, 39(2), pp.483-491.

Kikuchi, S., Hata, M., Fukumoto, K., Yamane, Y., Matsui, T., Tamura, A., Yonemura, S., Yamagishi, H., Keppler, D., Tsukita, S. and Tsukita, S. 2002. Radixin deficiency causes conjugated hyperbilirubinemia with loss of Mrp2 from bile canalicular membranes. *Nature Genetics*, 31(3), pp.320-325.

Kildegaard, H.F., Baycin-Hizal, D., Lewis, N.E. and Betenbaugh, M.J. 2013. The emerging CHO systems biology era: Harnessing the 'omics revolution for biotechnology. *Current Opinion in Biotechnology*, 24(6), pp.1102-1107.

Kim, C.K., Nguyen, T.L., Joo, K.M., Nam, D.H., Park, J., Lee, K.H., Cho, S.W. and Ahn, J.Y. 2010. Negative regulation of p53 by the long isoform of ErbB3 binding protein Ebp1 in brain tumors. *Cancer Research*, 70(23), pp.9730-9741.

Kim, D.Y., Lee, J.C., Chang, H.N. and Oh, D.J. 2005. Effects of supplementation of various medium components on chinese hamster ovary cell cultures producing recombinant antibody. *Cytotechnology*, 47(1-3), pp.37-49.

Kim, J.H., You, K.R., Kim, I.H., Cho, B.H., Kim, C.Y. and Kim, D.G. 2004. Over-expression of the ribosomal protein L36a gene is associated with cellular proliferation in hepatocellular carcinoma. *Hepatology (Baltimore, Md.)*, 39(1), pp.129-138.

Kim, S.H. and Lee, G.M. 2007. Down-regulation of lactate dehydrogenase-A by siRNAs for reduced lactic acid formation of chinese hamster ovary cells producing thrombopoietin. *Applied Microbiology and Biotechnology*, 74(1), pp.152-159.

Kim, Y.B., Lee, K.H., Sugita, K., Yoshida, M. and Horinouchi, S. 1999. Oxamflatin is a novel antitumor compound that inhibits mammalian histone deacetylase. *Oncogene*, 18(15), pp.2461-2470.

Kingsbury, G.A., Feeney, L.A., Nong, Y., Calandra, S.A., Murphy, C.J., Corcoran, J.M., Wang, Y., Prabhu Das, M.R., Busfield, S.J., Fraser, C.C. and Villeval, J.L. 2001. Cloning, expression, and function of BLAME, a novel member of the CD2 family. *Journal of Immunology (Baltimore, Md.: 1950)*, 166(9), pp.5675-5680.

Kishishita, S., Nishikawa, T., Shinoda, Y., Nagashima, H., Okamoto, H., Takuma, S. and Aoyagi, H. 2015. Effect of temperature shift on levels of acidic charge variants in IgG monoclonal antibodies in chinese hamster ovary cell culture. *Journal of Bioscience and Bioengineering,* 119(6), pp.700-705.

Knapp, S. 2013. Testis specific gene expression drives disease progression and rituximab resistance in lymphoma. *EMBO Molecular Medicine,* 5(8), pp.1149-1150.

Knoblich, K., Wang, H.X., Sharma, C., Fletcher, A.L., Turley, S.J. and Hemler, M.E. 2014. Tetraspanin TSPAN12 regulates tumor growth and metastasis and inhibits beta-catenin degradation. *Cellular and Molecular Life Sciences : CMLS,* 71(7), pp.1305-1314.

Kong, J., Li, Y., Liu, S., Jin, H., Shang, Y., Quan, C., Li, Y. and Lin, Z. 2013. High expression of ezrin predicts poor prognosis in uterine cervical cancer. *BMC Cancer,* 13(1), pp.520.

Kong, L., Schafer, G., Bu, H., Zhang, Y., Zhang, Y. and Klocker, H. 2012. Lamin A/C protein is overexpressed in tissue-invading prostate cancer and promotes prostate cancer cell growth, migration and invasion through the PI3K/AKT/PTEN pathway. *Carcinogenesis,* 33(4), pp.751-759.

Kozomara, A. and Griffiths-Jones, S. 2014. miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research,* 42(Database issue), pp.D68-73.

Krauss, M., Singer, H. and Hollender, J. 2010. LC-high resolution MS in environmental analysis: From target screening to the identification of unknowns. *Analytical and Bioanalytical Chemistry,* 397(3), pp.943-951.

Kremkow, B.G., Baik, J.Y., MacDonald, M.L. and Lee, K.H. 2015. CHOgenome.org 2.0: Genome resources and website updates. *Biotechnology Journal,* 10(7), pp.931-938.

Kruh, J. 1982. Effects of sodium butyrate, a new pharmacological agent, on cells in culture. *Molecular and Cellular Biochemistry,* 42(2), pp.65-82.

Kumar, N., Gammell, P., Meleady, P., Henry, M. and Clynes, M. 2008. Differential protein expression following low temperature culture of suspension CHO-K1 cells. *BMC Biotechnology,* 8pp.42.

LaBreche, H.G., Nevins, J.R. and Huang, E. 2011. Integrating factor analysis and a transgenic mouse model to reveal a peripheral blood predictor of breast tumors. *BMC Medical Genomics,* 4pp.61-8794-4-61.

Lamberti, A., Caraglia, M., Longo, O., Marra, M., Abbruzzese, A. and Arcari, P. 2004. The translation elongation factor 1A in tumorigenesis, signal transduction and apoptosis: Review article. *Amino Acids,* 26(4), pp.443-448.

Lange, C., Cakiroglu, F., Spiess, A.N., Cappallo-Obermann, H., Dierlamm, J. and Zander, A.R. 2007. Accelerated and safe expansion of human mesenchymal stromal cells in animal serum-free medium for transplantation and regenerative medicine. *Journal of Cellular Physiology,* 213(1), pp.18-26.

Langer, T., Rosmus, S. and Fasold, H. 2003. Intracellular localization of the 90 kDA heat shock protein (HSP90alpha) determined by expression of a EGFP-HSP90alpha-fusion protein in unstressed and heat stressed 3T3 cells. *Cell Biology International,* 27(1), pp.47-52.

Lee, J.S., Ha, T.K., Park, J.H. and Lee, G.M. 2013. Anti-cell death engineering of CHO cells: Co-overexpression of bcl-2 for apoptosis inhibition, beclin-1 for autophagy induction. *Biotechnology and Bioengineering,* 110(8), pp.2195-2207.

Lee, R.C., Feinbaum, R.L. and Ambros, V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* 75(5), pp.843-854.

Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y. and Pietenpol, J.A. 2011. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of Clinical Investigation,* 121(7), pp.2750-2767.

Lehmann, B.D., Pietenpol, J.A. and Tan, A.R. 2015. Triple-negative breast cancer: Molecular subtypes and new targets for therapy. *American Society of Clinical Oncology Educational Book / ASCO.American Society of Clinical Oncology.Meeting,* 35pp.e31-9.

Li, A.L., Li, H.Y., Jin, B.F., Ye, Q.N., Zhou, T., Yu, X.D., Pan, X., Man, J.H., He, K., Yu, M., Hu, M.R., Wang, J., Yang, S.C., Shen, B.F. and Zhang, X.M. 2004. A novel eIF5A complex functions as a regulator of p53 and p53-dependent apoptosis. *The Journal of Biological Chemistry,* 279(47), pp.49251-49258.

Li, Z.Y., Na, H.M., Peng, G., Pu, J. and Liu, P. 2011. Alteration of microRNA expression correlates to fatty acid-mediated insulin resistance in mouse myoblasts. *Molecular bioSystems,* 7(3), pp.871-877.

Liedtke, C., Mazouni, C., Hess, K.R., Andre, F., Tordai, A., Mejia, J.A., Symmans, W.F., Gonzalez-Angulo, A.M., Hennessy, B., Green, M., Cristofanilli, M., Hortobagyi, G.N. and Pusztai, L. 2008. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 26(8), pp.1275-1281.

Lin, K.W. and Souchelnytskyi, S. 2011. Translational connection of TGFbeta signaling: Phosphorylation of eEF1A1 by TbetaR-I inhibits protein synthesis. *Small GTPases,* 2(2), pp.104-108.

Lindstrom, M.S. 2009. Emerging functions of ribosomal proteins in gene-specific transcription and translation. *Biochemical and Biophysical Research Communications,* 379(2), pp.167-170.

Listgarten, J. and Emili, A. 2005. Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. *Molecular & Cellular Proteomics : MCP,* 4(4), pp.419-434.

Liu, B., Qiu, F.H., Voss, C., Xu, Y., Zhao, M.Z., Wu, Y.X., Nie, J. and Wang, Z.L. 2011a. Evaluation of three high abundance protein depletion kits for umbilical cord serum proteomics. *Proteome Science,* 9(1), pp.24-5956-9-24.

Liu, X., Cheng, I., Plummer, S.J., Suarez, B.K., Casey, G., Catalona, W.J. and Witte, J.S. 2011b. Fine-mapping of prostate cancer aggressiveness loci on chromosome 7q22-35. *The Prostate,* 71(7), pp.682-689.

Liu, Z., Jiang, Z., Huang, J., Huang, S., Li, Y., Yu, S., Yu, S. and Liu, X. 2014. miR-7 inhibits glioblastoma growth by simultaneously interfering with the PI3K/ATK and Raf/MEK/ERK pathways. *International Journal of Oncology,* 44(5), pp.1571-1580.

Livasy, C.A., Karaca, G., Nanda, R., Tretiakova, M.S., Olopade, O.I., Moore, D.T. and Perou, C.M. 2006. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc,* 19(2), pp.264-271.

Lopez, T., Dalton, K. and Frydman, J. 2015. The mechanism and function of group II chaperonins. *Journal of Molecular Biology,* 427(18), pp.2919-2930.

Lowe, A.R., Tang, J.H., Yassif, J., Graf, M., Huang, W.Y., Groves, J.T., Weis, K. and Liphardt, J.T. 2015. Importin-beta modulates the permeability of the nuclear pore complex in a ran-dependent manner. *ELife,* 4pp.10.7554/eLife.04052.

Lu, D., Joshi, A., Wang, B., Olsen, S., Yi, J.H., Krop, I.E., Burris, H.A. and Girish, S. 2013. An integrated multiple-analyte pharmacokinetic model to characterize trastuzumab emtansine (T-DM1) clearance pathways and to evaluate reduced pharmacokinetic sampling in patients with HER2-positive metastatic breast cancer. *Clinical Pharmacokinetics,* 52(8), pp.657-672.

Lu, Y. and Leslie, C.S. 2016. Learning to predict miRNA-mRNA interactions from AGO CLIP sequencing and CLASH data. *PLoS Computational Biology,* 12(7), pp.e1005026.

Luo, Q., Tang, K., Yang, F., Elias, A., Shen, Y., Moore, R.J., Zhao, R., Hixson, K.K., Rossie, S.S. and Smith, R.D. 2006. More sensitive and quantitative proteomic measurements using very low flow rate porous silica monolithic LC columns with electrospray ionization-mass spectrometry. *Journal of Proteome Research,* 5(5), pp.1091-1097.

Ly, L. and Wasinger, V.C. 2008. Peptide enrichment and protein fractionation using selective electrophoresis. *Proteomics,* 8(20), pp.4197-4208.

Lysenko, O., Schulte, D., Mittelbronn, M. and Steinle, A. 2013. BACL is a novel brain-associated, non-NKC-encoded mammalian C-type lectin-like receptor of the CLEC2 family. *PloS One,* 8(6), pp.e65345.

Ma, C.S. and Deenick, E.K. 2011. The role of SAP and SLAM family molecules in the humoral immune response. *Annals of the New York Academy of Sciences,* 1217pp.32-44.

Ma, Y., Zhang, W., Wei, J., Niu, M., Lin, H., Qin, W., Zhang, Y. and Qian, X. 2011. Reversed-phase liquid chromatography with double gradient elution for the separation and mass spectrometric analysis of peptides. *Se Pu = Chinese Journal of Chromatography / Zhongguo Hua Xue Hui,* 29(3), pp.205-211.

Maa, Y.F. and Hsu, C.C. 1996. Effect of high shear on proteins. *Biotechnology and Bioengineering,* 51(4), pp.458-465.

Machado-Neto, J.A., Saad, S.T. and Traina, F. 2014. Stathmin 1 in normal and malignant hematopoiesis. *BMB Reports,* 47(12), pp.660-665.

Madden, S.F., Clarke, C., Gaule, P., Aherne, S.T., O'Donovan, N., Clynes, M., Crown, J. and Gallagher, W.M. 2013. BreastMark: An integrated approach to mining publicly available transcriptomic datasets relating to breast cancer outcome. *Breast Cancer Research : BCR,* 15(4), pp.R52.

Maltman, D.J., Brand, S., Belau, E., Paape, R., Suckau, D. and Przyborski, S.A. 2011. Top-down label-free LC-MALDI analysis of the peptidome during neural progenitor cell differentiation reveals complexity in cytoskeletal protein dynamics and identifies progenitor cell markers. *Proteomics,*

Malz, M., Weber, A., Singer, S., Riehmer, V., Bissinger, M., Riener, M.O., Longerich, T., Soll, C., Vogel, A., Angel, P., Schirmacher, P. and Breuhahn, K. 2009. Overexpression of far upstream element binding proteins: A mechanism regulating proliferation and migration in liver cancer cells. *Hepatology (Baltimore, Md.),* 50(4), pp.1130-1139.

Manz, R., Assenmacher, M., Pfluger, E., Miltenyi, S. and Radbruch, A. 1995. Analysis and sorting of live cells according to secreted molecules, relocated to a cell-surface affinity matrix. *Proceedings of the National Academy of Sciences of the United States of America,* 92(6), pp.1921-1925.

Masterton, R.J., Roobol, A., Al-Fageeh, M.B., Carden, M.J. and Smales, C.M. 2010. Post-translational events of a model reporter protein proceed with higher fidelity and accuracy upon mild hypothermic culturing of chinese hamster ovary cells. *Biotechnology and Bioengineering,* 105(1), pp.215-220.

Matejovic, M., Tuma, Z., Moravec, J., Valesova, L., Sykora, R., Chvojka, J., Benes, J. and Mares, J. 2016. Renal proteomic responses to severe sepsis and surgical trauma: Dynamic analysis of porcine tissue biopsies. *Shock (Augusta, Ga.),*

McCarroll, N., Keshava, N., Cimino, M., Chu, M., Dearfield, K., Keshava, C., Kligerman, A., Owen, R., Protzel, A., Putzrath, R. and Schoeny, R. 2008. An evaluation of the mode of action framework for mutagenic carcinogens case study: Cyclophosphamide. *Environmental and Molecular Mutagenesis,* 49(2), pp.117-131.

Meleady, P., Hoffrogge, R., Henry, M., Rupp, O., Bort, J.H., Clarke, C., Brinkrolf, K., Kelly, S., Muller, B., Doolan, P., Hackl, M., Beckmann, T.F., Noll, T., Grillari, J., Barron, N., Puhler, A., Clynes, M. and Borth, N. 2012a. Utilization and evaluation of CHO-specific sequence databases for mass spectrometry based proteomics. *Biotechnology and Bioengineering,* 109(6), pp.1386-1394.

Meleady, P., Gallagher, M., Clarke, C., Henry, M., Sanchez, N., Barron, N. and Clynes, M. 2012b. Impact of miR-7 over-expression on the proteome of chinese hamster ovary cells. *Journal of Biotechnology,* 160(3–4), pp.251-262.

Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabkin, S., Guo, N., Muruganujan, A., Doremieux, O., Campbell, M.J., Kitano, H. and Thomas, P.D. 2005. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Research,* 33(Database issue), pp.D284-8.

Michelle, L., Cloutier, A., Toutant, J., Shkreta, L., Thibault, P., Durand, M., Garneau, D., Gendron, D., Lapointe, E., Couture, S., Le Hir, H., Klinck, R., Elela, S.A., Prinos, P. and Chabot, B. 2012. Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators. *Molecular and Cellular Biology,* 32(5), pp.954-967.

Mishra, A., Knerr, B., Paixao, S., Kramer, E.R. and Klein, R. 2008. The protein dendrite arborization and synapse maturation 1 (dasm-1) is dispensable for dendrite arborization. *Molecular and Cellular Biology,* 28(8), pp.2782-2791.

Mishra, A., Traut, M.H., Becker, L., Klopstock, T., Stein, V. and Klein, R. 2014. Genetic evidence for the adhesion protein IgSF9/Dasm1 to regulate inhibitory synapse development independent of its intracellular domain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience,* 34(12), pp.4187-4199.

Miyamoto, T., Hayashi, M., Takeuchi, A., Okamoto, T., Kawashima, S., Takii, T., Hayashi, H. and Onozaki, K. 1996. Identification of a novel growth-promoting factor with a wide target cell spectrum from various tumor cells as catalase. *Journal of Biochemistry,* 120(4), pp.725-730.

Moleirinho, S., Tilston-Lunel, A., Angus, L., Gunn-Moore, F. and Reynolds, P.A. 2013. The expanding family of FERM proteins. *The Biochemical Journal,* 452(2), pp.183-193.

Monigatti, F. and Berndt, P. 2005. Algorithm for accurate similarity measurements of peptide mass fingerprints and its application. *Journal of the American Society for Mass Spectrometry,* 16(1), pp.13-21.

Moy, T.I. and Silver, P.A. 1999. Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins. *Genes & Development,* 13(16), pp.2118-2133.

Mullard, A. 2013. Maturing antibody-drug conjugate pipeline hits 30. *Nature Reviews.Drug Discovery,* 12(5), pp.329-332.

Muniyappa, M.K., Dowling, P., Henry, M., Meleady, P., Doolan, P., Gammell, P., Clynes, M. and Barron, N. 2009. MiRNA-29a regulates the expression of numerous proteins and reduces the invasiveness and proliferation of human carcinoma cell lines. *European Journal of Cancer (Oxford, England : 1990),* 45(17), pp.3104-3118.

Murray, C.I., Barrett, M. and Van Eyk, J.E. 2009. Assessment of ProteoExtract subcellular fractionation kit reveals limited and incomplete enrichment of nuclear subproteome from frozen liver and heart tissue. *Proteomics,* 9(15), pp.3934-3938.

Nadler-Holly, M., Breker, M., Gruber, R., Azia, A., Gymrek, M., Eisenstein, M., Willison, K.R., Schuldiner, M. and Horovitz, A. 2012. Interactions of subunit CCT3 in the yeast chaperonin CCT/TRiC with Q/N-rich proteins revealed by high-throughput

microscopy analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 109(46), pp.18833-18838.

Nakatani, N., Kozaki, D., Mori, M. and Tanaka, K. 2012. Recent progress and applications of ion-exclusion/ion-exchange chromatography for simultaneous determination of inorganic anions and cations. *Analytical Sciences : The International Journal of the Japan Society for Analytical Chemistry*, 28(9), pp.845-852.

Nam, J.H., Ermonval, M. and Sharfstein, S.T. 2009. The effects of microcarrier culture on recombinant CHO cells under biphasic hypothermic culture conditions. *Cytotechnology*, 59(2), pp.81-91.

Nambu, M., Tatsukami, Y., Morisaka, H., Kuroda, K. and Ueda, M. 2015. Quantitative time-course proteome analysis of mesorhizobium loti during nodule maturation. *Journal of Proteomics*, 125pp.112-120.

Namgoong, S. and Kim, N.H. 2016. Roles of actin binding proteins in mammalian oocyte maturation and beyond. *Cell Cycle (Georgetown, Tex.)*, 15(14), pp.1830-1843.

Natrajan, R., Weigelt, B., Mackay, A., Geyer, F.C., Grigoriadis, A., Tan, D.S., Jones, C., Lord, C.J., Vatcheva, R., Rodriguez-Pinilla, S.M., Palacios, J., Ashworth, A. and Reis-Filho, J.S. 2010. An integrative genomic and transcriptomic analysis reveals molecular pathways and networks regulated by copy number aberrations in basal-like, HER2 and luminal cancers. *Breast Cancer Research and Treatment*, 121(3), pp.575-589.

Neilson, K.A., Ali, N.A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., van Sluyter, S.C. and Haynes, P.A. 2011. Less label, more free: Approaches in label-free quantitative mass spectrometry. *Proteomics*, 11(4), pp.535-553.

Nelson, A.L., Dhimolea, E. and Reichert, J.M. 2010. Development trends for human monoclonal antibody therapeutics. *Nature Reviews.Drug Discovery*, 9(10), pp.767-774.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.P., Tong, F., Speed, T., Spellman, P.T., DeVries, S., Lapuk, A., Wang, N.J., Kuo, W.L., Stilwell, J.L., Pinkel, D., Albertson, D.G., Waldman, F.M., McCormick, F., Dickson, R.B., Johnson, M.D., Lippman, M., Ethier, S., Gazdar, A. and Gray, J.W. 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell*, 10(6), pp.515-527.

Nguyen le, X.T., Lee, Y., Urbani, L., Utz, P.J., Hamburger, A.W., Sunwoo, J.B. and Mitchell, B.S. 2015. Regulation of ribosomal RNA synthesis in T cells: Requirement for GTP and Ebp1. *Blood*, 125(16), pp.2519-2529.

Niebruegge, S., Bauwens, C.L., Peerani, R., Thavandiran, N., Masse, S., Sevaptisidis, E., Nanthakumar, K., Woodhouse, K., Husain, M., Kumacheva, E. and Zandstra, P.W. 2009. Generation of human embryonic stem cell-derived mesoderm and cardiac cells using size-specified aggregates in an oxygen-controlled bioreactor. *Biotechnology and Bioengineering*, 102(2), pp.493-507.

Nielsen, T.O., Hsu, F.D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L., Akslen, L.A., Ragaz, J., Gown, A.M.,

Gilks, C.B., van de Rijn, M. and Perou, C.M. 2004. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 10(16), pp.5367-5374.

Olivotto, I.A., Truong, P.T., Speers, C.H., Bernstein, V., Allan, S.J., Kelly, S.J. and Lesperance, M.L. 2004. Time to stop progesterone receptor testing in breast cancer management. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 22(9), pp.1769-1770.

Orellana, C.A., Marcellin, E., Schulz, B.L., Nouwens, A.S., Gray, P.P. and Nielsen, L.K. 2015. High-antibody-producing chinese hamster ovary cells up-regulate intracellular protein transport and glutathione synthesis. *Journal of Proteome Research,* 14(2), pp.609-618.

Pan, Y., Han, J., Zhang, Y. and Li, X.J. 2010. Role of vimentin in tumor metastasis and drug research. *Sheng Li Ke Xue Jin Zhan [Progress in Physiology],* 41(6), pp.413-416.

Pandey, A.K., Agarwal, P., Kaur, K. and Datta, M. 2009. MicroRNAs in diabetes: Tiny players in big disease. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology,* 23(4-6), pp.221-232.

Panowski, S., Bhakta, S., Raab, H., Polakis, P. and Junutula, J.R. 2014. Site-specific antibody drug conjugates for cancer therapy. *MAbs,* 6(1), pp.34-45.

Parampalli, A., Eskridge, K., Smith, L., Meagher, M.M., Mowry, M.C. and Subramanian, A. 2007. Developement of serum-free media in CHO-DG44 cells using a central composite statistical design. *Cytotechnology,* 54(1), pp.57-68.

Park, J., Lee, E.J., Esau, C. and Schmittgen, T.D. 2009. Antisense inhibition of microRNA-21 or-221 arrests cell cycle, induces apoptosis, and sensitizes the effects of gemcitabine in pancreatic adenocarcinoma. *Pancreas,* 38(7), pp.E190-E199.

Parnell, J.J., Callister, S.J., Rompato, G., Nicora, C.D., Pasa-Tolic, L., Williamson, A. and Pfrender, M.E. 2011. Time-course analysis of the shewanella amazonensis SB2B proteome in response to sodium chloride shock. *Scientific Reports,* 1pp.25.

Patel, V.J., Thalassinos, K., Slade, S.E., Connolly, J.B., Crombie, A., Murrell, J.C. and Scrivens, J.H. 2009. A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *Journal of Proteome Research,* 8(7), pp.3752-3759.

Pau Ni, I.B., Zakaria, Z., Muhammad, R., Abdullah, N., Ibrahim, N., Aina Emran, N., Hisham Abdullah, N. and Syed Hussain, S.N. 2010. Gene expression patterns distinguish breast carcinomas from normal breast tissues: The malaysian context. *Pathology, Research and Practice,* 206(4), pp.223-228.

Perkins, D.N., Pappin, D.J., Creasy, D.M. and Cottrell, J.S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis,* 20(18), pp.3551-3567.

Perou, C.M. 2011. Molecular stratification of triple-negative breast cancers. *The Oncologist,* 16 Suppl 1pp.61-70.

Peters, D.G., Kudla, D.M., Deloia, J.A., Chu, T.J., Fairfull, L., Edwards, R.P. and Ferrell, R.E. 2005. Comparative gene expression analysis of ovarian carcinoma and normal ovarian epithelium by serial analysis of gene expression. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology,* 14(7), pp.1717-1723.

Petersdorf, S.H., Kopecky, K.J., Slovak, M., Willman, C., Nevill, T., Brandwein, J., Larson, R.A., Erba, H.P., Stiff, P.J., Stuart, R.K., Walter, R.B., Tallman, M.S., Stenke, L. and Appelbaum, F.R. 2013. A phase 3 study of gemtuzumab ozogamicin during induction and postconsolidation therapy in younger patients with acute myeloid leukemia. *Blood,* 121(24), pp.4854-4860.

Peterson, S.M., Thompson, J.A., Ufkin, M.L., Sathyanarayana, P., Liaw, L. and Congdon, C.B. 2014. Common features of microRNA target prediction tools. *Frontiers in Genetics,* 5pp.23.

Petrak, J., Ivanek, R., Toman, O., Cmejla, R., Cmejlova, J., Vyoral, D., Zivny, J. and Vulpe, C.D. 2008. Deja vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins. *Proteomics,* 8(9), pp.1744-1749.

Petushkova, N.A., Kuznetsova, G.P., Larina, O.V., Kisrieva, Y.S., Samenkova, N.F., Trifonova, O.P., Miroshnichenko, Y.V., Zolotarev, K.V., Karuzina, I.I., Ipatova, O.M. and Lisitsa, A.V. 2015. One-dimensional proteomic profiling of danio rerio embryo vitellogenin to estimate quantum dot toxicity. *Proteome Science,* 13pp.17-015-0072-7. eCollection 2015.

Piccart-Gebhart, M.J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., Gianni, L., Baselga, J., Bell, R., Jackisch, C., Cameron, D., Dowsett, M., Barrios, C.H., Steger, G., Huang, C.S., Andersson, M., Inbar, M., Lichinitser, M., Lang, I., Nitz, U., Iwata, H., Thomssen, C., Lohrisch, C., Suter, T.M., Ruschoff, J., Suto, T., Greatorex, V., Ward, C., Straehle, C., McFadden, E., Dolci, M.S., Gelber, R.D. and Herceptin Adjuvant (HERA) Trial Study Team. 2005. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *The New England Journal of Medicine,* 353(16), pp.1659-1672.

Piehowski, P.D., Petyuk, V.A., Orton, D.J., Xie, F., Moore, R.J., Ramirez-Restrepo, M., Engel, A., Lieberman, A.P., Albin, R.L., Camp, D.G., Smith, R.D. and Myers, A.J. 2013. Sources of technical variability in quantitative LC-MS proteomics: Human brain tissue sample analysis. *Journal of Proteome Research,* 12(5), pp.2128-2137.

Piovesana, S., Capriotti, A.L., Caruso, G., Cavaliere, C., La Barbera, G., Zenezini Chiozzi, R. and Laganà, A. 2016. Labeling and label free shotgun proteomics approaches to characterize muscle tissue from farmed and wild gilthead sea bream (sparus aurata). *Journal of Chromatography A,* 1428pp.193-201.

Politano, L., Carboni, N., Madej-Pilarczyk, A., Marchel, M., Nigro, G., Fidziaoska, A., Opolski, G. and Hausmanowa-Petrusewicz, I. 2013. Advances in basic and clinical research in laminopathies. *Acta Myologica : Myopathies and Cardiomyopathies :*

Official Journal of the Mediterranean Society of Myology / Edited by the Gaetano Conte Academy for the Study of Striated Muscle Diseases, 32(1), pp.18-22.

Pontiller, J., Gross, S., Thaisuchat, H., Hesse, F. and Ernst, W. 2008. Identification of CHO endogenous promoter elements based on a genomic library approach. *Molecular Biotechnology,* 39(2), pp.135-139.

Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P. and Stoffel, M. 2004. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature,* 432(7014), pp.226-230.

Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X. and Perou, C.M. 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research : BCR,* 12(5), pp.R68.

Printz, C. 2010. New AJCC cancer staging manual reflects changes in cancer knowledge. *Cancer,* 116(1), pp.2-3.

Pritchard, K.I., Messersmith, H., Elavathil, L., Trudeau, M., O'Malley, F. and Dhesy-Thind, B. 2008. HER-2 and topoisomerase II as predictors of response to chemotherapy. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 26(5), pp.736-744.

PUCK, T.T., CIECIURA, S.J. and ROBINSON, A. 1958. Genetics of somatic mammalian cells. III. long-term cultivation of euploid cells from human and animal subjects. *The Journal of Experimental Medicine,* 108(6), pp.945-956.

Putti, T.C., El-Rehim, D.M., Rakha, E.A., Paish, C.E., Lee, A.H., Pinder, S.E. and Ellis, I.O. 2005. Estrogen receptor-negative breast carcinomas: A review of morphology and immunophenotypical analysis. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc,* 18(1), pp.26-35.

Quenel, N., Wafflart, J., Bonichon, F., de Mascarel, I., Trojani, M., Durand, M., Avril, A. and Coindre, J.M. 1995. The prognostic value of c-erbB2 in primary breast carcinomas: A study on 942 cases. *Breast Cancer Research and Treatment,* 35(3), pp.283-291.

Rajaraman, R. and Faulkner, G. 1984. Reverse transformation of chinese hamster ovary cells by methyl xanthines. structure-function relationships. *Experimental Cell Research,* 154(2), pp.342-356.

Rajendra, Y., Kiseljak, D., Baldi, L., Hacker, D.L. and Wurm, F.M. 2011. Influence of glutamine on transient and stable recombinant protein production in CHO and HEK-293 cells. *BMC Proceedings,* 5 Suppl 8pp.P35-6561-5-S8-P35. eCollection 2011.

Rakha, E.A., El-Sayed, M.E., Green, A.R., Paish, E.C., Lee, A.H. and Ellis, I.O. 2007a. Breast carcinoma with basal differentiation: A proposal for pathology definition based on basal cytokeratin expression. *Histopathology,* 50(4), pp.434-438.

Rakha, E.A., El-Sayed, M.E., Green, A.R., Paish, E.C., Powe, D.G., Gee, J., Nicholson, R.I., Lee, A.H., Robertson, J.F. and Ellis, I.O. 2007b. Biologic and clinical characteristics of breast cancer with single hormone receptor positive phenotype.

Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 25(30), pp.4772-4778.

Rakha, E.A., Elsheikh, S.E., Aleskandarany, M.A., Habashi, H.O., Green, A.R., Powe, D.G., El-Sayed, M.E., Benhasouna, A., Brunet, J.S., Akslen, L.A., Evans, A.J., Blamey, R., Reis-Filho, J.S., Foulkes, W.D. and Ellis, I.O. 2009. Triple-negative breast cancer: Distinguishing between basal and nonbasal subtypes. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 15(7), pp.2302-2310.

Rakha, E.A., Reis-Filho, J.S. and Ellis, I.O. 2010. Combinatorial biomarker expression in breast cancer. *Breast Cancer Research and Treatment,* 120(2), pp.293-308.

Rasmussen, B.B., Regan, M.M., Lykkesfeldt, A.E., Dell'Orto, P., Del Curto, B., Henriksen, K.L., Mastropasqua, M.G., Price, K.N., Mery, E., Lacroix-Triki, M., Braye, S., Altermatt, H.J., Gelber, R.D., Castiglione-Gertsch, M., Goldhirsch, A., Gusterson, B.A., Thurlimann, B., Coates, A.S., Viale, G. and BIG 1-98 Collaborative and International Breast Cancer Study Groups. 2008. Adjuvant letrozole versus tamoxifen according to centrally-assessed ERBB2 status for postmenopausal women with endocrine-responsive early breast cancer: Supplementary results from the BIG 1-98 randomised trial. *The Lancet.Oncology,* 9(1), pp.23-28.

Regitnig, P., Reiner, A., Dinges, H.P., Hofler, G., Muller-Holzner, E., Lax, S.F., Obrist, P., Rudas, M. and Quehenberger, F. 2002. Quality assurance for detection of estrogen and progesterone receptors by immunohistochemistry in austrian pathology laboratories. *Virchows Archiv : An International Journal of Pathology,* 441(4), pp.328-334.

Reichert, J.M. 2013. Antibodies to watch in 2013: Mid-year update. *MAbs,* 5(4), pp.513-517.

Ren, C., Chen, T., Sun, H., Jiang, X., Hu, C., Qian, J. and Wang, Y. 2015. The first echinoderm poly-U-binding factor 60 kDa (PUF60) from sea cucumber (stichopus monotuberculatus): Molecular characterization, inducible expression and involvement of apoptosis. *Fish & Shellfish Immunology,* 47(1), pp.196-204.

Resing, K.A. and Ahn, N.G. 2005. Proteomics strategies for protein identification. *FEBS Letters,* 579(4), pp.885-889.

Riddell, D.R., Vinogradov, D.V., Stannard, A.K., Chadwick, N. and Owen, J.S. 1999. Identification and characterization of LRP8 (apoER2) in human blood platelets. *Journal of Lipid Research,* 40(10), pp.1925-1930.

Ruiterkamp, J. and Ernst, M.F. 2011. The role of surgery in metastatic breast cancer. *European Journal of Cancer (Oxford, England : 1990),* 47 Suppl 3pp.S6-22.

Rumachik, N.G., McAlister, G.C., Russell, J.D., Bailey, D.J., Wenger, C.D. and Coon, J.J. 2012. Characterizing peptide neutral losses induced by negative electron-transfer dissociation (NETD). *Journal of the American Society for Mass Spectrometry,* 23(4), pp.718-727.

Ruprecht, B., Koch, H., Medard, G., Mundt, M., Kuster, B. and Lemeer, S. 2015. Comprehensive and reproducible phosphopeptide enrichment using iron immobilized

metal ion affinity chromatography (fe-IMAC) columns. *Molecular & Cellular Proteomics : MCP,* 14(1), pp.205-215.

Sadygov, R.G. 2015. Using SEQUEST with theoretically complete sequence databases. *Journal of the American Society for Mass Spectrometry,* 26(11), pp.1858-1864.

Sadygov, R.G., Cociorva, D. and Yates, J.R.,3rd. 2004. Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. *Nature Methods,* 1(3), pp.195-202.

Saidi, Y., Finka, A., Muriset, M., Bromberg, Z., Weiss, Y.G., Maathuis, F.J. and Goloubinoff, P. 2009. The heat shock response in moss plants is regulated by specific calcium-permeable channels in the plasma membrane. *The Plant Cell,* 21(9), pp.2829-2843.

Sanchez, N., Gallagher, M., Lao, N., Gallagher, C., Clarke, C., Doolan, P., Aherne, S., Blanco, A., Meleady, P., Clynes, M. and Barron, N. 2013. MiR-7 triggers cell cycle arrest at the G1/S transition by targeting multiple genes including Skp2 and Psme3. *PloS One,* 8(6), pp.e65671.

Sapra, P. and Shor, B. 2013. Monoclonal antibody-based therapies in cancer: Advances and challenges. *Pharmacology & Therapeutics,* 138(3), pp.452-469.

Sassoon, I. and Blanc, V. 2013. Antibody-drug conjugate (ADC) clinical pipeline: A review. *Methods in Molecular Biology (Clifton, N.J.),* 1045pp.1-27.

Schatz, S.M., Kerschbaumer, R.J., Gerstenbauer, G., Kral, M., Dorner, F. and Scheiflinger, F. 2003. Higher expression of fab antibody fragments in a CHO cell line at reduced temperature. *Biotechnology and Bioengineering,* 84(4), pp.433-438.

Schechter, I. and Berger, A. 1967. On the size of the active site in proteases. I. papain. *Biochemical and Biophysical Research Communications,* 27(2), pp.157-162.

Scheffler, K. 2014. Top-down proteomics by means of orbitrap mass spectrometry. *Methods in Molecular Biology (Clifton, N.J.),* 1156pp.465-487.

Schepens, B., Tinton, S.A., Bruynooghe, Y., Parthoens, E., Haegman, M., Beyaert, R. and Cornelis, S. 2007. A role for hnRNP C1/C2 and unr in internal initiation of translation during mitosis. *The EMBO Journal,* 26(1), pp.158-169.

Schlecht, N.F., Brandwein-Gensler, M., Smith, R.V., Kawachi, N., Broughel, D., Lin, J., Keller, C.E., Reynolds, P.A., Gunn-Moore, F.J., Harris, T., Childs, G., Belbin, T.J. and Prystowsky, M.B. 2012. Cytoplasmic ezrin and moesin correlate with poor survival in head and neck squamous cell carcinoma. *Head and Neck Pathology,* 6(2), pp.232-243.

Schmidt, M., Hafner, M. and Frech, C. 2014. Modeling of salt and pH gradient elution in ion-exchange chromatography. *Journal of Separation Science,* 37(1-2), pp.5-13.

Schneiderman, M.H., Dewey, W.C. and Highfield, D.P. 1971. Inhibition of DNA synthesis in synchronized chinese hamster cells treated in G1 with cycloheximide. *Experimental Cell Research,* 67(1), pp.147-155.

Schultheiss, M., Schnichels, S., Hermann, T., Hurst, J., Feldkaemper, M., Arango-Gonzalez, B., Ueffing, M., Bartz-Schmidt, K.U., Zeck, G. and Spitzer, M.S. 2016. Hypothermia protects and prolongs the tolerance time of retinal ganglion cells against ischemia. *PloS One,* 11(2), pp.e0148616.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell,* 115(2), pp.199-208.

Sclafani, R.A. and Holzen, T.M. 2007. Cell cycle regulation of DNA replication. *Annual Review of Genetics,* 41pp.237-280.

Senger, R.S. and Karim, M.N. 2003. Effect of shear stress on intrinsic CHO culture state and glycosylation of recombinant tissue-type plasminogen activator protein. *Biotechnology Progress,* 19(4), pp.1199-1209.

Sengupta, D. and Kundu, S. 2012. Role of long- and short-range hydrophobic, hydrophilic and charged residues contact network in protein's structural organization. *BMC Bioinformatics,* 13pp.142-2105-13-142.

Severson, T.M., Peeters, J., Majewski, I., Michaut, M., Bosma, A., Schouten, P.C., Chin, S., Pereira, B., Goldgraben, M.A., Bismeijer, T., Kluin, R.J.C., Muris, J.J.F., Jirström, K., Kerkhoven, R.M., Wessels, L., Caldas, C., Bernards, R., Simon, I.M. and Linn, S. 2015. BRCA1-like signature in triple negative breast cancer: Molecular and clinical characterization reveals subgroups with therapeutic potential. *Molecular Oncology,* 9(8), pp.1528-1538.

Shaw, C.R. and Prasad, R. 1970. Starch gel electrophoresis of enzymes--a compilation of recipes. *Biochemical Genetics,* 4(2), pp.297-320.

Shefet-Carasso, L. and Benhar, I. 2015. Antibody-targeted drugs and drug resistance--challenges and solutions. *Drug Resistance Updates : Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy,* 18pp.36-46.

Shen, G.Q., Girelli, D., Li, L., Rao, S., Archacki, S., Olivieri, O., Martinelli, N., Park, J.E., Chen, Q., Topol, E.J. and Wang, Q.K. 2014. A novel molecular diagnostic marker for familial and early-onset coronary artery disease and myocardial infarction in the LRP8 gene. *Circulation.Cardiovascular Genetics,* 7(4), pp.514-520.

Shen, Y., Tolic, N., Zhao, R., Pasa-Tolic, L., Li, L., Berger, S.J., Harkewicz, R., Anderson, G.A., Belov, M.E. and Smith, R.D. 2001. High-throughput proteomics using high-efficiency multiple-capillary liquid chromatography with on-line high-performance ESI FTICR mass spectrometry. *Analytical Chemistry,* 73(13), pp.3011-3021.

Shi, M., Xie, Z., Yu, M., Shen, B. and Guo, N. 2005. Controlled growth of chinese hamster ovary cells and high expression of antibody-IL-2 fusion proteins by temperature manipulation. *Biotechnology Letters,* 27(23-24), pp.1879-1884.

Shi, S.H., Cox, D.N., Wang, D., Jan, L.Y. and Jan, Y.N. 2004. Control of dendrite arborization by an ig family member, dendrite arborization and synapse maturation 1 (Dasm1). *Proceedings of the National Academy of Sciences of the United States of America,* 101(36), pp.13341-13345.

Shishkov, R., Chervenkov, T., Yamashima, T. and Tonchev, A.B. 2013a. Expression of Cyclon/CCDC86, a novel nuclear protein, in the hippocampus of adult non-human primates. *Journal of Neuroimmunology,* 258(1-2), pp.96-99.

Shishkov, R., Chervenkov, T., Yamashima, T. and Tonchev, A.B. 2013b. Expression of Cyclon/CCDC86, a novel nuclear protein, in the hippocampus of adult non-human primates. *Journal of Neuroimmunology,* 258(1–2), pp.96-99.

Shukla, A.A. and Thömmes, J. 2010. Recent advances in large-scale production of monoclonal antibodies and related proteins. *Trends in Biotechnology,* 28(5), pp.253-261.

Siepen, J.A., Keevil, E.J., Knight, D. and Hubbard, S.J. 2007. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *Journal of Proteome Research,* 6(1), pp.399-408.

Sieprath, T., Corne, T.D., Nooteboom, M., Grootaert, C., Rajkovic, A., Buysschaert, B., Robijns, J., Broers, J.L., Ramaekers, F.C., Koopman, W.J., Willems, P.H. and De Vos, W.H. 2015. Sustained accumulation of prelamin A and depletion of lamin A/C both cause oxidative stress and mitochondrial dysfunction but induce different cell fates. *Nucleus (Austin, Tex.),* 6(3), pp.236-246.

Simpson, P.T., Reis-Filho, J.S., Gale, T. and Lakhani, S.R. 2005. Molecular evolution of breast cancer. *The Journal of Pathology,* 205(2), pp.248-254.

Sircoulomb, F., Bekhouche, I., Finetti, P., Adelaide, J., Ben Hamida, A., Bonansea, J., Raynaud, S., Innocenti, C., Charafe-Jauffret, E., Tarpin, C., Ben Ayed, F., Viens, P., Jacquemier, J., Bertucci, F., Birnbaum, D. and Chaffanet, M. 2010. Genome profiling of ERBB2-amplified breast cancers. *BMC Cancer,* 10pp.539-2407-10-539.

Sissi, C., Moro, S. and Crothers, D.M. 2015. Novel insights on the DNA interaction of calicheamicin gamma(1)(I). *Biopolymers,* 103(8), pp.449-459.

Skottman, H., Stromberg, A.M., Matilainen, E., Inzunza, J., Hovatta, O. and Lahesmaa, R. 2006. Unique gene expression signature by human embryonic stem cells cultured under serum-free conditions correlates with their enhanced and prolonged growth in an undifferentiated stage. *Stem Cells (Dayton, Ohio),* 24(1), pp.151-167.

Skrzypczak, M., Goryca, K., Rubel, T., Paziewska, A., Mikula, M., Jarosz, D., Pachlewski, J., Oledzki, J. and Ostrowski, J. 2010. Modeling oncogenic signaling in colon tumors by multidirectional analyses of microarray data directed for maximization of analytical reliability. *PloS One,* 5(10), pp.10.1371/journal.pone.0013091.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A. and McGuire, W.L. 1987. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science (New York, N.Y.),* 235(4785), pp.177-182.

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J. and Norton, L. 2001. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *The New England Journal of Medicine,* 344(11), pp.783-792.

Smyth, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology,* 3pp.Article3.

Sonna, L.A., Fujita, J., Gaffin, S.L. and Lilly, C.M. 2002. Invited review: Effects of heat and cold stress on mammalian gene expression. *Journal of Applied Physiology (Bethesda, Md.: 1985),* 92(4), pp.1725-1742.

Sou, S.N., Sellick, C., Lee, K., Mason, A., Kyriakopoulos, S., Polizzi, K.M. and Kontoravdi, C. 2015. How does mild hypothermia affect monoclonal antibody glycosylation? *Biotechnology and Bioengineering,* 112(6), pp.1165-1176.

Stendahl, M., Ryden, L., Nordenskjold, B., Jonsson, P.E., Landberg, G. and Jirstrom, K. 2006. High progesterone receptor expression correlates to the effect of adjuvant tamoxifen in premenopausal breast cancer patients. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 12(15), pp.4614-4618.

Stewart, M. 2007. Molecular mechanism of the nuclear protein import cycle. *Nature Reviews.Molecular Cell Biology,* 8(3), pp.195-208.

Stiefel, F., Fischer, S., Sczyrba, A., Otte, K. and Hesse, F. 2016. miRNA profiling of high, low and non-producing CHO cells during biphasic fed-batch cultivation reveals process relevant targets for host cell engineering. *Journal of Biotechnology,* 225pp.31-43.

Strotbek, M., Florin, L., Koenitzer, J., Tolstrup, A., Kaufmann, H., Hausser, A. and Olayioye, M.A. 2013. Stable microRNA expression enhances therapeutic antibody productivity of chinese hamster ovary cells. *Metabolic Engineering,* 20pp.157-166.

Sunley, K. and Butler, M. 2010. Strategies for the enhancement of recombinant protein production from mammalian cells by growth arrest. *Biotechnology Advances,* 28(3), pp.385-394.

Szulawska, A. and Czyz, M. 2006. Molecular mechanisms of anthracyclines action. *Postepy Higieny i Medycyny Doswiadczalnej (Online),* 60pp.78-100.

Szymczak, P. and Cieplak, M. 2007. Proteins in a shear flow. *The Journal of Chemical Physics,* 127(15), pp.155106.

Takeuchi, A., Miyamoto, T., Yamaji, K., Masuho, Y., Hayashi, M., Hayashi, H. and Onozaki, K. 1995. A human erythrocyte-derived growth-promoting factor with a wide target cell spectrum: Identification as catalase. *Cancer Research,* 55(7), pp.1586-1589.

Tao, J.J., Visvanathan, K. and Wolff, A.C. 2015. Long term side effects of adjuvant chemotherapy in patients with early breast cancer. *Breast (Edinburgh, Scotland),* 24 Suppl 2pp.S149-53.

Thaisuchat, H., Baumann, M., Pontiller, J., Hesse, F. and Ernst, W. 2011. Identification of a novel temperature sensitive promoter in CHO cells. *BMC Biotechnology,* 11pp.51-6750-11-51.

Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A. and Narechania, A. 2003. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Research,* 13(9), pp.2129-2141.

Tian, H. and Cronstein, B.N. 2007. Understanding the mechanisms of action of methotrexate: Implications for the treatment of rheumatoid arthritis. *Bulletin of the NYU Hospital for Joint Diseases,* 65(3), pp.168-173.

Tollet-Egnell, P., Flores-Morales, A., Stahlberg, N., Malek, R.L., Lee, N. and Norstedt, G. 2001. Gene expression profile of the aging process in rat liver: Normalizing effects of growth hormone replacement. *Molecular Endocrinology (Baltimore, Md.),* 15(2), pp.308-318.

Tom, S., Henricksen, L.A. and Bambara, R.A. 2000. Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *The Journal of Biological Chemistry,* 275(14), pp.10498-10505.

Tsujimoto, Y. 2003. Cell death regulation by the bcl-2 protein family in the mitochondria. *Journal of Cellular Physiology,* 195(2), pp.158-167.

Turner, N.C., Reis-Filho, J.S., Russell, A.M., Springall, R.J., Ryder, K., Steele, D., Savage, K., Gillett, C.E., Schmitt, F.C., Ashworth, A. and Tutt, A.N. 2007. BRCA1 dysfunction in sporadic basal-like breast cancer. *Oncogene,* 26(14), pp.2126-2132.

Tye, B.K. 1999. MCM proteins in DNA replication. *Annual Review of Biochemistry,* 68pp.649-686.

Underhill, M.F. and Smales, C.M. 2007. The cold-shock response in mammalian cells: Investigating the HeLa cell cold-shock proteome. *Cytotechnology,* 53(1-3), pp.47-53.

Van Breukelen, F. and Martin, S.L. 2002. Invited review: Molecular adaptations in mammalian hibernators: Unique adaptations or generalized responses? *Journal of Applied Physiology (Bethesda, Md.: 1985),* 92(6), pp.2640-2647.

Van Dyk, D.D., Misztal, D.R., Wilkins, M.R., Mackintosh, J.A., Poljak, A., Varnai, J.C., Teber, E., Walsh, B.J. and Gray, P.P. 2003. Identification of cellular changes associated with increased production of human growth hormone in a recombinant chinese hamster ovary cell line. *Proteomics,* 3(2), pp.147-156.

Vanneman, M. and Dranoff, G. 2012. Combining immunotherapy and targeted therapies in cancer treatment. *Nature Reviews.Cancer,* 12(4), pp.237-251.

Vera, M., Pani, B., Griffiths, L.A., Muchardt, C., Abbott, C.M., Singer, R.H. and Nudler, E. 2014. The translation elongation factor eEF1A1 couples transcription to translation during heat shock response. *ELife,* 3pp.e03164.

Voellmy, R. and Boellmann, F. 2007. Chaperone regulation of the heat shock protein response. *Advances in Experimental Medicine and Biology,* 594pp.89-99.

Volinia, S., Calin, G.A., Liu, C.G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R.L., Yanaihara, N., Lanza, G., Scarpa, A.,

Vecchione, A., Negrini, M., Harris, C.C. and Croce, C.M. 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America,* 103(7), pp.2257-2261.

Vrbic, S., Pejcic, I., Filipovic, S., Kocic, B. and Vrbic, M. 2013. Current and future anti-HER2 therapy in breast cancer. *Journal of B.U.ON.: Official Journal of the Balkan Union of Oncology,* 18(1), pp.4-16.

Walsh, G. 2014. Biopharmaceutical benchmarks 2014. *Nature Biotechnology,* 32(10), pp.992-1000.

Walsh, G. and Jefferis, R. 2006. Post-translational modifications in the context of therapeutic proteins. *Nature Biotechnology,* 24(10), pp.1241-1252.

Wang, H. 2014. Lapatinib for the treatment of breast cancer in the people's republic of china. *OncoTargets and Therapy,* 7pp.1367-1373.

Wang, H., Wang, F., Wang, W., Yao, X., Wei, D., Cheng, H. and Deng, Z. 2014. Improving the expression of recombinant proteins in E. coli BL21 (DE3) under acetate stress: An alkaline pH shift approach. *PloS One,* 9(11), pp.e112777.

Wang, H., Zhao, L.N., Li, K.Z., Ling, R., Li, X.J. and Wang, L. 2006. Overexpression of ribosomal protein L15 is associated with cell proliferation in gastric cancer. *BMC Cancer,* 6pp.91.

Winkler, C., Denker, K., Wortelkamp, S. and Sickmann, A. 2007. Silver- and coomassie-staining protocols: Detection limits and compatibility with ESI MS. *Electrophoresis,* 28(12), pp.2095-2099.

Wirapati, P., Sotiriou, C., Kunkel, S., Farmer, P., Pradervand, S., Haibe-Kains, B., Desmedt, C., Ignatiadis, M., Sengstag, T., Schutz, F., Goldstein, D.R., Piccart, M. and Delorenzi, M. 2008. Meta-analysis of gene expression profiles in breast cancer: Toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Research : BCR,* 10(4), pp.R65.

Wolff, A.C., Hammond, M.E., Schwartz, J.N., Hagerty, K.L., Allred, D.C., Cote, R.J., Dowsett, M., Fitzgibbons, P.L., Hanna, W.M., Langer, A., McShane, L.M., Paik, S., Pegram, M.D., Perez, E.A., Press, M.F., Rhodes, A., Sturgeon, C., Taube, S.E., Tubbs, R., Vance, G.H., van de Vijver, M., Wheeler, T.M., Hayes, D.F. and American Society of Clinical Oncology/College of American Pathologists. 2007. American society of clinical Oncology/College of american pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Archives of Pathology & Laboratory Medicine,* 131(1), pp.18-43.

Wong, J.W., Sullivan, M.J. and Cagney, G. 2008. Computational methods for the comparative quantification of proteins in label-free LCn-MS experiments. *Briefings in Bioinformatics,* 9(2), pp.156-165.

Wu, L.M., Wang, S.Y., Wang, S.X., Huang, Y.M. and Li, J.G. 2010a. Proliferation promotion and apoptotic inhibition effects of ribosomal protein RPL36A small interference RNA on U937 cells. *Zhongguo Shi Yan Xue Ye Xue Za Zhi / Zhongguo*

Bing Li Sheng Li Xue Hui = Journal of Experimental Hematology / Chinese Association of Pathophysiology, 18(2), pp.344-349.

Wu, W., Beilhartz, G., Roy, Y., Richard, C.L., Curtin, M., Brown, L., Cadieux, D., Coppolino, M., Farach-Carson, M.C., Nemere, I. and Meckling, K.A. 2010b. Nuclear translocation of the 1,25D3-MARRS (membrane associated rapid response to steroids) receptor protein and NFkappaB in differentiating NB4 leukemia cells. *Experimental Cell Research,* 316(7), pp.1101-1108.

Wu, X., Li, J., Li, X., Hsieh, C.L., Burgers, P.M. and Lieber, M.R. 1996. Processing of branched DNA intermediates by a complex of human FEN-1 and PCNA. *Nucleic Acids Research,* 24(11), pp.2036-2043.

Wurm, F.M. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology,* 22(11), pp.1393-1398.

Xu, P., Dai, X.P., Graf, E., Martel, R. and Russell, R. 2014. Effects of glutamine and asparagine on recombinant antibody production using CHO-GS cell lines. *Biotechnology Progress,* 30(6), pp.1457-1468.

Xu, Q. and Keiderling, T.A. 2004. Effect of sodium dodecyl sulfate on folding and thermal stability of acid-denatured cytochrome c: A spectroscopic approach. *Protein Science : A Publication of the Protein Society,* 13(11), pp.2949-2959.

Xu, X., Nagarajan, H., Lewis, N.E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S., Andersen, M.R., Neff, N., Passarelli, B., Koh, W., Fan, H.C., Wang, J., Gui, Y., Lee, K.H., Betenbaugh, M.J., Quake, S.R., Famili, I., Palsson, B.O. and Wang, J. 2011. The genomic sequence of the chinese hamster ovary (CHO)-K1 cell line. *Nature Biotechnology,* 29(8), pp.735-741.

Xue, S. and Barna, M. 2012. Specialized ribosomes: A new frontier in gene regulation and organismal biology. *Nature Reviews.Molecular Cell Biology,* 13(6), pp.355-369.

Yee, J.C., Gerdtzen, Z.P. and Hu, W.S. 2009. Comparative transcriptome analysis to unveil genes affecting recombinant protein productivity in mammalian cells. *Biotechnology and Bioengineering,* 102(1), pp.246-263.

Yee, J.C., Wlaschin, K.F., Chuah, S.H., Nissom, P.M. and Hu, W.S. 2008. Quality assessment of cross-species hybridization of CHO transcriptome on a mouse DNA oligo microarray. *Biotechnology and Bioengineering,* 101(6), pp.1359-1365.

Yen, C.Y., Russell, S., Mendoza, A.M., Meyer-Arendt, K., Sun, S., Cios, K.J., Ahn, N.G. and Resing, K.A. 2006. Improving sensitivity in shotgun proteomics using a peptide-centric database with reduced complexity: Protease cleavage and SCX elution rules from data mining of MS/MS spectra. *Analytical Chemistry,* 78(4), pp.1071-1084.

Yoon, S.K., Song, J.Y. and Lee, G.M. 2003. Effect of low culture temperature on specific productivity, transcription level, and heterogeneity of erythropoietin in chinese hamster ovary cells. *Biotechnology and Bioengineering,* 82(3), pp.289-298.

Yoshikawa, T., Nakanishi, F., Ogura, Y., Oi, D., Omasa, T., Katakura, Y., Kishimoto, M. and Suga, K.I. 2001. Flow cytometry: An improved method for the selection of

highly productive gene-amplified CHO cells using flow cytometry. *Biotechnology and Bioengineering,* 74(5), pp.435-442.

Zamocky, M., FurtmÃ¼ller, P.,G. and Obinger, C. 2008. Evolution of catalases from bacteria to humans. *Antioxidants & Redox Signaling,* 10(9), pp.1527-1548.

Zhang, J., Gao, F.L., Zhi, H.Y., Luo, A.P., Ding, F., Wu, M. and Liu, Z.H. 2004. Expression patterns of esophageal cancer deregulated genes in C57BL/6J mouse embryogenesis. *World Journal of Gastroenterology,* 10(8), pp.1088-1092.

Zhang, Y., Woodford, N., Xia, X. and Hamburger, A.W. 2003. Repression of E2F1-mediated transcription by the ErbB3 binding protein Ebp1 involves histone deacetylases. *Nucleic Acids Research,* 31(8), pp.2168-2177.

Zhang, Y., LeRoy, G., Seelig, H., Lane, W.S. and Reinberg, D. 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell,* 95(2), pp.279-289.

Zheng, L., Dai, H., Qiu, J., Huang, Q. and Shen, B. 2007. Disruption of the FEN-1/PCNA interaction results in DNA replication defects, pulmonary hypoplasia, pancytopenia, and newborn lethality in mice. *Molecular and Cellular Biology,* 27(8), pp.3176-3186.

Zhou, H., Purdie, J., Wang, T. and Ouyang, A. 2010. pH measurement and a rational and practical pH control strategy for high throughput cell culture system. *Biotechnology Progress,* 26(3), pp.872-880.

Zhu, X., Buhrer, C. and Wellmann, S. 2016. Cold-inducible proteins CIRP and RBM3, a unique couple with activities far beyond the cold. *Cellular and Molecular Life Sciences : CMLS,*

Zimmer, A., Mueller, R., Wehsling, M., Schnellbaecher, A. and von Hagen, J. 2014. Improvement and simplification of fed-batch bioprocesses with a highly soluble phosphotyrosine sodium salt. *Journal of Biotechnology,* 186pp.110-118.

Appendix

Temperature Shift tables time point analysis

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
PSMC1	26S protease regulatory subunit 4	1.31					
PSMD2	26S proteasome non-ATPase regulatory subunit 2			1.26			
PSMD4	26S proteasome non-ATPase regulatory subunit 4						2.03
HACL1	2-hydroxyacyl-CoA lyase 1		1.84				
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase		2.30				
RPS15	40S ribosomal protein S15			72.61			
RPS15A	40S ribosomal protein S15a				1.40		
RPS19	40S ribosomal protein S19						Infinity
RPS4	40S ribosomal protein S4	25.21					
RPL19	60S ribosomal protein L19		2.61				
RPL6	60S ribosomal protein L6	Infinity					
ACTB	Actin, cytoplasmic 2						Infinity
SLC25A5	ADP/ATP translocase 2						1.31
AKR1E2	Aldo-Keto Reductase Family 1, Member E2				3443.91		
GANAB	Alpha 1,3-glucosidase			1.69			
AP2A1	AP-2 complex subunit alpha-1				1.44		
AIFM2	Apoptosis-inducing factor, mitochondrion-associated			1.26			
ATP5Ea	ATP synthase subunit epsilon, mitochondrial		1.57				
ATPT2A2	ATPase, Ca++ transporting, cardiac muscle, fast twitch 1		1.82				
ABCE1	ATP-binding cassette sub-family E member 1				1.23		
BRD2	Bromodomain-containing protein 2						1.49
CILP2a	Cartilage intermediate layer protein 2		1.59				
CSNK2B	Casein kinase 2, beta polypeptide	1.21					
IGF2R	Cation-independent mannose-6-phosphate receptor		30.90				
ARCN1	Coatomer subunit delta		1.41				
CSDE1	Cold shock domain-containing protein E1				1.37		
CIRBPa	Cold-inducible RNA-binding protein		2.43				
NDUFAF1	Complex I intermediate-associated protein 30, mitochondrial		2.71				
CND2	Condensin complex subunit 2	1.40					
EMC8	COX4 neighbor				1.75		
CUL1	Cullin-1				1.82		
CDK2	Cyclin-dependent kinase 2						1.43
NFS1	Cysteine Desulfurase				1.37		
CX6B1a	Cytochrome c oxidase subunit 6B1		1.49				1.89
COX5A	Cytochrome c oxidase subunit Va		1.44				
COX6C	Cytochrome c oxidase subunit Vic		1.36				
DARS	DARS aspartyl-tRNA synthetase	1.43					
DDX19B	DEAD (Asp-Glu-Ala-As) box polypeptide 19B		1.66				
DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17		1.66			2.73	
DDX23	DEAD (Asp-Glu-Ala-Asp) box polypeptide 23						1.52
DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 50	Infinity					
DHX15	DEAH (Asp-Glu-Ala-His) box polypeptide 15					1.54	
DIS3	DIS3 mitotic control homolog (S. cerevisiae)		1.70				
MCM3	DNA replication licensing factor MCM3				1.60		
MCM5	DNA replication licensing factor MCM5				2.13		
MCM6	DNA replication licensing factor MCM6				1.35		
OPA1	Dynamin-like 120 kDa protein, mitochondrial				1.39		
ELAV1	ELAV-like protein 1				1.40		
HADHA	Enoyl-CoA hydratase / long-chain 3-hydroxyacyl-CoA dehydrogenase		Infinity				
IF2G	Eukaryotic translation initiation factor 2 subunit 3				1.33		
EIF3I	Eukaryotic translation initiation factor 3 subunit I				1.22		
EIF4A1	Eukaryotic Translation Initiation Factor 4A1				1.42		
EIF4A3	Eukaryotic Translation Initiation Factor 4A3				2.59		
EIF6	Eukaryotic translation initiation factor 6			1.33			
EZR	Ezrin				1.68		1.90
FASN	Fatty Acid Synthase, Animal Type				2.01		

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
FKBP1	FK506-binding protein 1				Infinity		
FEN1	Flap Structure-Specific Endonuclease 1				1.41		
FLOT1	Flotillin-1		3.16				
ALDOA[a]	Fructose-bisphosphate aldolase A		1.61				
FUNDC2	FUN14 domain-containing protein 2	1.76					
LG3BP	Galectin-3-binding protein	1.82					
GNAI1	Guanine nucleotide-binding protein G(i), alpha-2 subunit		80.37				
GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1		Infinity				
HMOX1	Heme Oxygenase (Decycling) 1				1.84		
ROA3	Heterogeneous nuclear ribonucleoprotein A3		2096.34				
HNRNPF	Heterogeneous nuclear ribonucleoprotein F					1.47	
HNRH3	Heterogeneous nuclear ribonucleoprotein H3		1.45				
HNRPK	Heterogeneous nuclear ribonucleoprotein K		1.42				
HK1	Hexokinase 1		3.34				
HMGB2	High mobility group protein B2	1.71					
HIST1H1C	Histone H1.2				3.41		
HIST1H2AG[a]	Histone H2A type 1				2.79		
H2AZ	Histone H2A.Z				2.44		
IMMT	Inner Membrane Protein, Mitochondrial (Mitofilin)				1.25		
ITGA5	Integrin alpha-5		2.39				
ICAM1	Intercellular adhesion molecule 1			1.23			
IDH3A	Isocitrate dehydrogenase 3 (NAD+) alpha						1.26
LDHA	Lactate dehydrogenase A						1.64
LMNA	Lamin-A/C		3.03		7.29		
TMPO	Lamina-associated polypeptide 2, isoforms alpha/zeta				3.06		1.77
LMNB1	Lamin-B1		3.11				
LAP3	Leucine aminopeptidase 3		1.30				
MTHFD1L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like		Infinity				
TIMM8B[a]	Mitochondrial import inner membrane translocase subunit Tim8 B						1.30
TOM70[a]	Mitochondrial import receptor subunit TOM70		1.22				
MSN	Moesin		Infinity				
MRTO4	mRNA turnover protein 4 homolog		1.68				
NDUFA13	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 13		1.62				
NDUFV1	NADH dehydrogenase (ubiquinone) flavoprotein 1		1.50				1.84
NDUFA3[a]	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 3						1.50
AHNAK	Neuroblast differentiation-associated protein AHNAK	1.85	11.35				
NUDC	Nuclear migration protein nudC			2.05			
NUCL	Nucleolin			1.24			
TPR	Nucleoprotein TPR		1.99				
PPIH	Peptidylprolyl Isomerase H (Cyclophilin H)				1.30		
PLS3	Plastin-3		6.80				
PLEC	Plectin		1.61				
NOMO3	PM5 NODAL modulator 3				1.31		
PABP1	Polyadenylate-binding protein 4			141.03			
PTRF	Polymerase I and transcript release factor						1.39
SNRNP200	Pre-Mrna-Splicing Helicase Brr2				2.31		
PHB	Prohibitin		Infinity				
PCNA	Proliferating cell nuclear antigen				1.34		
PA2G4	Proliferation-associated protein 2G4	1.26					1.11
SCL1	Proteasome (prosome, macropain) subunit, alpha type, 1	1.47					
PSMB6	Proteasome (prosome, macropain) subunit, beta type, 6	1.31					
AHNAK2	Protein AHNAK2	1.79					
PRMT1	Protein Arginine Methyltransferase 1				1.47		
CYR61	Protein CYR61	2.77	2.69				
PDIA6	Protein disulfide isomerase family A, member 6	37.04					
PPP1CB	Protein Phosphatase 1, Catalytic Subunit				Infinity		

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
RBM3	Putative RNA-binding protein 3	1.44			4.23		
LUC7L	Putative RNA-binding protein Luc7-like 2	1.60					
PDXDC1	Pyridoxal-dependent decarboxylase domain containing 1		2.69E+00				
RAGP1	Ran GTPase-activating protein 1					2.97E+00	
RCC1	Regulator of chromosome condensation						1.62
RCF4	Replication factor C subunit 4				1.44		
RPN1	Ribophorin I				1.44		
RPN2	Ribophorin II				1.47		
RBM12	RNA-binding protein 12				1.66		
PPPC1B[a]	serine/threonine-protein phosphatase PP1-beta catalytic subunit				Infinity		
Tf[a]	Serotransferrin				2.33		
SND1	SND1 staphylococcal nuclease and tudor domain containing 1			1.14			
SPTAN1	Spectrin alpha		2.32E+00				
SF3B6	Splicing factor 3B subunit 3	3.55					
SRSF6	Splicing Factor, Arginine/Serine-Rich 6				1.47		
STIP1	Stress-induced-phosphoprotein 1				134.04		
UBE21a	SUMO-conjugating enzyme UBC9		1.39E+00				
SKIV2L2	Superkiller viralicidic activity 2-like 2 (S. cerevisiae)		1.43E+00				
CCT4	T-complex protein 1 subunit delta				1.84		
CCT8	T-complex protein 1 subunit theta				1.30		
TXNDC12[a]	Thioredoxin domain-containing protein 12		Infinity				
THTR	Thiosulfate sulfurtransferase		1.51E+00				
THOC2	THO complex 2					2.75E+00	
TSTA3	Tissue Specific Transplantation Antigen P35B				1.88		
TECR	Trans-2,3-Enoyl-Coa Reductase				2.81		
TCEB2	Transcription elongation factor B polypeptide 2		1.75E+00				
BTF3	Transcription factor BTF3				14.13		
VCP	Transitional endoplasmic reticulum ATPase			1.20			
TUBA1B	Tubulin alpha-1B chain		Infinity				
TBB6	Tubulin beta-6 chain	213.83					
TUBB6	Tubulin beta-6 chain				3.51		2.17
SNRNP70	U1 small nuclear ribonucleoprotein 70 kDa		1.32E+00				
SNRPA	U1 small nuclear ribonucleoprotein A					1.86E+00	
SNRPA1	U2 small nuclear ribonucleoprotein A'		1.23E+00				
UBE2D2	Ubiquitin-conjugating enzyme E2 D/E						1.43
UBE2I	Ubiquitin-conjugating enzyme E2 I		1.39E+00				
UGDH	UDP-glucose dehydrogenase						1.96
C14orf166	UPF0568 protein C14orf166	2.04					
HDLBP	Vigilin				1.51		

[a]

Protein identification obtained from NCBI nonredudant database which was not present in BBCHO or had a higher confidence score compared to BBCHO database

[b]

Fold change showing increased protein expression in CHO cells grown at 31°C compared to 37°C at 8 and 24h

i. Up-regulated proteins in CHO-K1 SEAP cells at 31 °C after 8 and 24 hr in membrane, cytoplasmic and nuclear fractions.

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
RPS14	40S ribosomal protein S14					1.97	
RPS16	40S ribosomal protein S16	3.04					
RPS2	40S ribosomal protein S2					2.02	2.02
RPS5	40S ribosomal protein S5					1.65	1.65
RPS6	40S ribosomal protein S6					1.91	1.91
SLC3A2	4F2 cell-surface antigen heavy chain				1.22		
CH60	60 kDa heat shock protein, mitochondrial					2.50	
RPL11[a]	60S ribosomal protein L11				2.17		
RPL13	60S ribosomal protein L23	Infinity					
RPL26	60S ribosomal protein L26				Infinity		
RPL28	60S ribosomal protein L28				1.81		
RPL36	60S ribosomal protein L36	64.98					
RPL36A	60S ribosomal protein L36a			Infinity			
RPL4	60S ribosomal protein L4					1.50	1.50
RPL7	60S ribosomal protein L7					1.80	1.80
RPL7A	60S ribosomal protein L7a					1.60	1.60
RPL8	60S ribosomal protein L8					1.80	1.80
GRP78	78 kDa glucose-regulated protein				1.76		
AK3	Adenylate kinase 3			1.25			
SLC25A5	ADP/ATP translocase 2	1.29					
ARFGAP2	ADP-ribosylation factor GTPase-activating protein 2					1.98	
ARFGAP3	ADP-ribosylation factor GTPase-activating protein 3						2.24
AARS	Alanyl-tRNA synthetase			1.20			
AKR7A2	Aldo-keto reductase family 7, member A2 (aflatoxin aldehyde reductase)			1.35			
ACTN4	Alpha-actinin-4				2.03		
AMPD1	AMP deaminase				1.48		
RSRC2	Arginine/serine-rich coiled-coil protein 2					1.90	
ASNS	ASNS asparagine synthetase				2.45		
ATXN2L	Ataxin-2-like protein				1.62		
ATP5H	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit d	1.32					
DDX42	ATP-dependent RNA helicase DDX42					1.57	1.57
HSPG2	Basement membrane-specific heparan sulfate proteoglycan core protein	9.89					
BOD1L1	Biorientation of chromosomes in cell division 1-like						27.29
BRD4	Bromodomain containing 4					1.92	
CAMK2D	Calcium/Calmodulin-Dependent Protein Kinase II Delta				1.64		
SLC25A12	Calcium-binding mitochondrial carrier protein Aralar1	1.41					
CNN2	Calponin-2					1.54	1.54
CALR	CALR Calreticulin	5.28			164.79		
CHST11	Carbohydrate (chondroitin 4) sulfotransferase 11					1.47	1.47
CSNK2A1	Casein kinase 2, alpha 1 polypeptide					1.68	1.68
CSNK2B	Casein Kinase 2, Beta Polypeptide				1.44		
CD2BP2	CD2 antigen cytoplasmic tail-binding protein 2					1.67	
CEP170	Centrosomal protein of 170 kDa					1.63	
CHMP6	Charged multivesicular body protein 6	1.28					
CBX1	Chromobox protein homolog 1				1.60		
CHD5	Chromodomain helicase DNA binding protein 5						7.41
ARCN1	Coatomer subunit delta		1.41				
CFL2	Cofilin-2				1.96		
CCDC43	Coiled-coil domain-containing protein 43				1.47		
CCDC50	Coiled-coil domain-containing protein 50				2.99		
CSDE1	Cold shock domain-containing protein E1					1.61	1.61
NDUFAF1	Complex I intermediate-associated protein 30, mitochondrial		2.71				
COR1B	Coronin-1B					4.24	4.24
COR1C	Coronin-1C					1.89	
CTPS1	CTP synthase				1.49		
CREL2	Cysteine-rich with EGF-like domain protein 2				1.65		

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
COX5A	Cytochrome c oxidase subunit Va					1.66	
CYFIP1	Cytoplasmic FMR1-interacting protein 1					1.78	
CKAP4	Cytoskeleton-associated protein 4					1.96	
DDX19B	DEAD (Asp-Glu-Ala-As) box polypeptide 19B		1.66				
DNTTIP2	Deoxynucleotidyltransferase terminal-interacting protein 2					2.04	
DPYSL2	Dihydropyrimidinase-related protein 2					1.39	
DLG1	Disks large homolog 1	3.47					
MCM5	DNA replication licensing factor MCM5	1.42					
POLRMT	DNA-directed RNA polymerase, mitochondrial	1.45					
DNAJB11	DnaJ homolog subfamily B member 11				1.33		
RPN1	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase	1.43				1.51	
STAU1	Double-stranded RNA-binding protein Staufen homolog 1				1.60		
DBNL	Drebrin-like protein			1.47			
DYNC1H1	Dynein heavy chain 1, cytosolic						1.46
EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	3.54					
ENG	Endoglin	1.73					
ESRRA	Estrogen-related receptor alpha	1.27					
EEF1A1	Eukaryotic Translation Elongation Factor 1 Alpha 1				1.39		
EEF1D	Eukaryotic translation elongation factor 1 delta					2.16	
EIF3C	Eukaryotic translation initiation factor 3, subunit C					2.30	
IF4G1	Eukaryotic translation initiation factor 4 gamma 1				1.38		
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1					1.40	
EIF4A2	Eukaryotic translation initiation factor 4A2					1.60	
EIF4B	Eukaryotic translation initiation factor 4B					2.02	2.02
IF5A2	Eukaryotic translation initiation factor 5A-2				3.28		
EIF5A2	Eukaryotic translation initiation factor 5A-2						9.26
FLNB	Filamin-B	1.31				2.09	2.09
FKBP3	FK506 binding protein 3				1.28		
FKBP4	FK506 binding protein 4, 59kDa				1.75		
FKBP5	FK506-binding protein 4/5				3.94		
FLOT1	Flotillin-1		3.16				
GAG	Gag polyprotein					2.23	2.23
GALK1	Galactokinase 1			1.30			
GTF2I	General transcription factor II-I						3.49
GAA	Glucosidase, Alpha				1.31		
GLUD1	Glutamate dehydrogenase 1					1.51	
DHE3[a]	Glutamate dehydrogenase 1, mitochondrial					1.51	
GPX8	Glutathione peroxidase 8 (putative)	1.20					
GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1		Infinity				
HS105	Heat shock protein 105 kDa					2.12	
HNRNPH2	Heterogeneous nuclear ribonucleoprotein A/B						Infinity
HNRH3	Heterogeneous nuclear ribonucleoprotein H3		1.45				
HNRNPK	Heterogeneous nuclear ribonucleoprotein K						Infinity
HNRPM	Heterogeneous nuclear ribonucleoprotein M				1.78		
HTATSF1	HIV Tat-specific factor 1 homolog			1.25			
HSDL2	Hydroxysteroid dehydrogenase-like protein 2	1.48					
HYOU1	Hypoxia up-regulated protein 1				1.47		
IMPDH1	IMP dehydrogenase				1.45		
IGF2BP2	Insulin-like growth factor 2 mRNA-binding protein 2				1.35		
ITGA5	Integrin alpha-5		2.39				
ITB1	Integrin beta-1				1.55		
ILK	Integrin-linked kinase						5.96
KRT10[a]	Keratin, type I cytoskeletal 10					3.07	
KRT6A[a]	Keratin, type II cytoskeletal 6A		6.20				
KRT75[a]	Keratin, type II cytoskeletal 75					2.37	
KLC1	Kinesin light chain 1				1.22		

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
SPC24	Kinetochore protein Spc24				7.08		
LAMC1	Laminin subunit gamma-1				1.29		
LGMN	Legumain				2.05		
LETM1	LETM1 and EF-hand domain-containing protein 1, mitochondrial	11.26					
LAP3	Leucine aminopeptidase 3		1.30				
LEPRE1	Leucine proline-enriched proteoglycan (leprecan) 1			1.57			
LASP1	LIM and SH3 domain protein 1				1.55		
ABLIM1	LIM domain and actin-binding protein 1					1.56	1.56
MVP	Major vault protein				2.57		
MDH1	Malate dehydrogenase 1, NAD (soluble)					1.37	
MESDC2	Mesoderm Development Candidate 2				1.55		
MET	MET proto-oncogene (hepatocyte growth factor receptor)				2.52		
MAP1B	Microtubule-associated protein 1B				2.63		1.96
MAP6	Microtubule-associated protein 6					1.39	1.39
SLC25A20	Mitochondrial carnitine/acylcarnitine carrier protein	1.34					
DIC1	Mitochondrial dicarboxylate carrier	1.34					
TOMM34	Mitochondrial import receptor subunit				2.32		
MSN	Moesin					1.68	
MYBBP1A	MYB binding protein (P160) 1a						2.67
MYH9	Myh9 myosin, heavy chain 11, smooth muscle	9.33			1.73		
MHY11	Myosin, heavy chain 11, smooth muscle					1.75	1.75
SLC9A3R1[a]	Na(+)/H(+) exchange regulatory cofactor NHE-RF1				1.55		
NSF	N-ethylmaleimide-sensitive factor	1.39					
AHNAK	Neuroblast differentiation-associated protein AHNAK				1.49	1.51	1.51
NCLN	Nicalin	1.21					
CALD1	Non-muscle caldesmon					1.68	1.68
NASP	Nuclear autoantigenic sperm protein				1.70		
NUDC	Nuclear migration protein				1.77		
NUMA1	Nuclear mitotic apparatus protein 1						2.41
NOG1	Nucleolar GTP-binding protein 1					1.69	
PLIN4	Perilipin-4				2.29		
PRDX1	Peroxiredoxin-1			Infinity			
PCNP	PEST proteolytic signal-containing nuclear protein				1.86		
SACM1	Phosphatidylinositide phosphatase SAC1	1.35					
PHGDH	Phosphoglycerate Dehydrogenase				1.29		
PSAT1	Phosphoserine Aminotransferase 1				1.44		
PLS3	Plastin-3		6.80				
PABP1	Polyadenylate-binding protein 4	2388.98				1.83	1.83
VBP1	Prefoldin subunit 3			1.29			
PCYOX1	Prenylcysteine oxidase 1	1.26					
PHB	Prohibitin		Infinity				
PHB2	Prohibitin-2					1.98	1.98
CNPY3[a]	Protein canopy 3				1.66		
PDIA3	Protein Disulfide Isomerase Family A, Member 3				9.08		
PDIA4	Protein Disulfide Isomerase Family A, Member 4				1.75		
CCAR2	Protein KIAA1967 homolog					1.92	1.92
PRKCSH	Protein Kinase C Substrate 80K-H				1.31		
PPP1CB	Protein phosphatase 1, catalytic subunit						Infinity
PPP1CA	Protein phosphatase 1, catalytic subunit, gamma isozyme			1.37			
PP1R2P3	Protein Phosphatase 1, Regulatory (Inhibitor) Subunit 2				10.90		
PPP1R7	Protein Phosphatase 1, Regulatory (Inhibitor) Subunit 7				1.32		
RCC2	Protein RCC2						1.28
PTPN6	Protein Tyrosine Phosphatase, Non-Receptor Type 6				1.37		
WIBG	Protein wibg homolog				1.68		
PDXDC1	Pyridoxal-dependent decarboxylase domain containing 1		2.69				
PDHA1	Pyruvate dehydrogenase (lipoamide) alpha 1	1.27					

		31°C v 37°C Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	8hr	24hr	8hr	24hr	8hr	24hr
RBM3	Putative RNA-binding protein 3	1.44			4.23		
LUC7L	Putative RNA-binding protein Luc7-like 2	1.60					
PDXDC1	Pyridoxal-dependent decarboxylase domain containing 1		2.69				
RAGP1	Ran GTPase-activating protein 1					2.97	
RCC1	Regulator of chromosome condensation						1.62
RCF4	Replication factor C subunit 4				1.44		
RPN1	Ribophorin I				1.44		
RPN2	Ribophorin II				1.47		
RBM12	RNA-binding protein 12				1.66		
PPPC1B[a]	serine/threonine-protein phosphatase PP1-beta catalytic subunit				Infinity		
Tf[a]	Serotransferrin				2.33		
SND1	SND1 staphylococcal nuclease and tudor domain containing 1			1.14			
SPTAN1	Spectrin alpha		2.32				
SF3B6	Splicing factor 3B subunit 3	3.55					
SRSF6	Splicing Factor, Arginine/Serine-Rich 6				1.47		
STIP1	Stress-induced-phosphoprotein 1				134.04		
UBE21a	SUMO-conjugating enzyme UBC9		1.39				
SKIV2L2	Superkiller viralicidic activity 2-like 2 (S. cerevisiae)		1.43				
CCT4	T-complex protein 1 subunit delta				1.84		
CCT8	T-complex protein 1 subunit theta				1.30		
TXNDC12[a]	Thioredoxin domain-containing protein 12		#######				
THTR	Thiosulfate sulfurtransferase		1.51				
THOC2	THO complex 2					2.75	
TSTA3	Tissue Specific Transplantation Antigen P35B				1.88		
TECR	Trans-2,3-Enoyl-Coa Reductase				2.81		
TCEB2	Transcription elongation factor B polypeptide 2		1.75				
BTF3	Transcription factor BTF3				14.13		
VCP	Transitional endoplasmic reticulum ATPase			1.20			
TUBA1B	Tubulin alpha-1B chain		Infinity				
TBB6	Tubulin beta-6 chain	213.83					
TUBB6	Tubulin beta-6 chain				3.51		2.17
SNRNP70	U1 small nuclear ribonucleoprotein 70 kDa		1.32				
SNRPA	U1 small nuclear ribonucleoprotein A					1.86	
SNRPA1	U2 small nuclear ribonucleoprotein A'		1.23				
UBE2D2	Ubiquitin-conjugating enzyme E2 D/E						1.43
UBE2I	Ubiquitin-conjugating enzyme E2 I		1.39				
UGDH	UDP-glucose dehydrogenase						1.96
C14orf166	UPF0568 protein C14orf166	2.04					
HDLBP	Vigilin				1.51		

[a] Protein identification obtained from NCBI nonredudant database which was not present in BBCHO or had a higher confidence score compared to BBCHO database

[b] Fold change showing increased protein expression in CHO cells grown at 31°C compared to 37°C at 8 and 24h

ii. Down-regulated proteins in CHO-K1 SEAP cells at 31 °C after 8 and 24 hr in membrane, cytoplasmic and nuclear fractions.

DAVID analysis timepoint comparison with fractionation and temperature shift

Biological Process 8hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	DOWN at 31oC	Count	Adjusted p-value
Membrane	Protein localization	12	3.2E-03
	Protein transport	10	4.2E-02
	Establishment of protein localization	10	4.5E-02
Cytoplasmic	NO SIG		
Nuclear	Translation	14	1.3E-08
	Translational elongation	9	5.1E-07

Biological Process 24hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	DOWN at 31oC	Count	Adjusted p-value
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	Translational elongation	8	4.6E-06
	Translation	10	9.9E-05

iii. Biological processes differentially regulated in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using GO analysis through DAVID showing translational and protein localization dis regulation at 8 hr with translational dysregulation also occurring at 24 hr

Molecular Function 8hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	RNA binding	17	1.2E-07
	Structural constituent of ribosome	8	2.0E-04
	Structural molecule activity	12	9.5E-04
	Actin binding	8	1.5E-02
	Cytoskeletal protein binding	9	3.6E-02

Molecular Function 24hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	Heat shock protein binding	5	2.7E-02
Nuclear	RNA binding	17	7.9E-09
	Structural constituent of ribosome	8	6.2E-05
	Structural molecule activity	11	1.3E-03

iv. Molecular functions differentially regulated in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using GO analysis through DAVID showing structural dysregulation and RNA binding differential regulation at 8 hr and structural, RNA binding and heat shock protein binding regulation at 24 hr.

Cellular Component 8hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	Organelle membrane	13	1.6E-03
	Endoplasmic reticulum part	7	3.8E-02
Cytoplasmic	NO SIG		
Nuclear	Cytosolic ribosome	8	4.1E-06
	Intracellular non-membrane-bounded organelle	29	4.7E-06
	Non-membrane-bounded organelle	29	4.7E-06
	Cytosolic part	9	1.9E-05
	Ribosomal subunit	8	9.8E-05

Cellular Component 24hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	Organelle envelope	15	2.1E-07
	Envelope	15	2.2E-07
	Organelle inner membrane	10	6.9E-05
	Mitochondrial envelope	10	5.2E-04
	Mitochondrial membrane	9	3.0E-03
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	Endoplasmic reticulum lumen	6	2.0E-03
	Cytosol	15	3.8E-02
Nuclear	Intracellular non-membrane-bounded organelle	27	6.3E-08
	Non-membrane-bounded organelle	27	6.3E-08
	Cytosolic ribosome	7	2.1E-05
	Ribosomal subunit	7	3.1E-04
	Ribonucleoprotein complex	11	3.2E-04

v. Cellular components associated with differentially regulated proteins in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using GO analysis through DAVID showing an emphasis on ribosomal, organelle, mitochondial components at both timepoints.

Panther timepoint analysis with fractionation and temperature shift

Biological Process 8hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	Protein metabolic process	10	7.96E-03
Cytoplasmic	Protein metabolic process	7	2.52E-02
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	Translation	15	5.93E-10
	Cellular component organization or biogenesis	17	1.43E-05
	Regulation of translation	6	1.02E-03
	Cellular component organization	14	1.12E-03
	mRNA splicing, via spliceosome	8	4.62E-03

Biological Process 24hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	mRNA splicing, via spliceosome	7	4.33E-03
	RNA splicing	6	7.86E-03
	RNA splicing, via transesterification reactions	6	7.86E-03
	mRNA processing	7	1.30E-02
Cytoplasmic	Metabolic process	35	1.26E-03
	DNA replication	6	2.19E-03
	Primary metabolic process	30	9.44E-03
Nuclear	Generation of precursor metabolites and energy	5	0.000968
	Tricarboxylic acid cycle	2	0.0438

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	Protein metabolic process	26	1.82E-05
	Protein folding	7	5.05E-04
	Primary metabolic process	40	1.30E-03
Nuclear	Translation	9	5.33E-05
	mRNA processing	8	1.03E-03
	mRNA splicing, via spliceosome	7	3.29E-03
	Cellular component organization or biogenesis	11	6.08E-03
	DNA metabolic process	7	9.35E-03

vi. Biological processes differentially regulated in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using PANTHER showing cell component organisation and mRNA differential regulation at 8 hr and RNA splicing, RNA processing, translation, metabolic processes and tricarboxylic acid cycle at 24 hr

Molecular Function 8hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	RNA helicase activity	2	4.05E-02

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	Structural molecule activity	22	3.93E-09
	RNA binding	14	1.91E-06
	Actin binding	8	3.32E-05
	Cytoskeletal protein binding	9	3.36E-05
	Structural constituent of ribosome	8	4.85E-05

Molecular Function 24hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	Structural molecule activity	12	3.52E-03
	mRNA binding	6	1.33E-02
	Oxidoreductase activity	8	1.51E-02
	RNA binding	8	1.59E-02
	Structural constituent of ribosome	5	1.75E-02
Cytoplasmic	Helicase activity	6	1.61E-04
	DNA helicase activity	4	4.08E-03
	RNA binding	8	1.37E-02
	Translation initiation factor activity	4	1.41E-02
	Catalytic activity	25	1.69E-02
Nuclear	Oxidoreductase activity	5	0.0367

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	Isomerase activity	6	2.37E-03
Nuclear	Structural molecule activity	16	9.72E-07
	Binding	31	1.17E-05
	Cytoskeletal protein binding	8	2.26E-05
	Structural constituent of ribosome	7	5.41E-05
	Nucleic acid binding	21	3.60E-04

vii. Molecular functions differentially regulated in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using PANTHER which show structural and RNA binding functions affected at 8 hr and a wider range of functions affected at 24 hr including structural and RNA binding functions but also isomerase, helicase, translational initiation and catalytic activity.

Cellular Component 8hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	NO SIG		
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	Mitochondrion	4	4.12E-04
	Mitochondrial inner membrane	3	8.02E-03
Cytoplasmic	NO SIG		
Nuclear	Cytoskeleton	16	1.11E-07
	Organelle	17	9.62E-07
	Actin cytoskeleton	10	1.83E-05
	Intracellular	17	7.69E-05
	Cell part	17	3.12E-04

Cellular Component 24hr			
Fraction	**UP at 31oC**	**Count**	**Adjusted p-value**
Membrane	Intracellular	12	3.30E-03
	Cell part	12	8.86E-03
	Macromolecular complex	7	2.93E-02
Cytoplasmic	NO SIG		
Nuclear	NO SIG		

Fraction	**DOWN at 31oC**	**Count**	**Adjusted p-value**
Membrane	Cell	12	2.77E-03
	Cell part	12	2.77E-03
	Cellular component	12	3.37E-02
	Membrane	8	4.59E-02
Cytoplasmic	NO SIG		
Nuclear	Cytoskeleton	10	3.03E-04
	Organelle	11	6.16E-04
	Actin cytoskeleton	6	9.03E-03
	Ribonucleoprotein complex	4	9.84E-03
	Intracellular	11	1.01E-02

viii. Cellular components associated with differentially regulated proteins in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using PANTHER showing an emphasis on cytoskeletal, membrane, intracellular and ribonucleoprotien components enriched.

Time Course Temperature shift tables

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
HSPE1	10 kDa heat shock protein, mitochondrial						3.71
EFTUD2	116 kDa U5 small nuclear ribonucleoprotein component	Infinity					
PRS6A	26S protease regulatory subunit 6A				1.25		
PSMC2	26S protease regulatory subunit 7				1.37		
PSMD2	26S proteasome non-ATPase regulatory subunit 2			1.26			
PSMD4	26S proteasome non-ATPase regulatory subunit 4						2.52
PSMD6	26S proteasome non-ATPase regulatory subunit 6			1.41			
HACL1	2-hydroxyacyl-CoA lyase 1		1.47				
RM01	39S ribosomal protein L1, mitochondrial		1.58				
MRPL15	39S ribosomal protein L15, mitochondrial	1.75					
MRPL53	39S ribosomal protein L53, mitochondrial		1.92				
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase	2.81	4.19				
HMGCS2	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial)				1.38		
OXCT1	3-oxoacid CoA transferase 1			1.74			
RPS11	40S ribosomal protein S11			1.26			
RPS12	40S ribosomal protein S12				1.48		
RPS15A	40S ribosomal protein S15a				1.63		
RPS18	40S ribosomal protein S18			1.28			
RPS19	40S ribosomal protein S19				1.36		
RPS2	40S ribosomal protein S2			1.27	1.59		
RPS20	40S ribosomal protein S20				1.62		
RPS21[a]	40S ribosomal protein S21			1.36			
RPS26	40S ribosomal protein S26			1.22			
RPS27	40S ribosomal protein S27			1.64			
RPS29[a]	40S ribosomal protein S29-like				1.44		
RPS3	40S ribosomal protein S3			1.26	1.63		
RPS4	40S ribosomal protein S4				1.59		
RPS5	40S ribosomal protein S5			1.64			
RPS7	40S ribosomal protein S7			1.43			
RPS8	40S ribosomal protein S8				1.31		
RPSA	40S ribosomal protein SA				2.05		
SLC3A2	4F2 cell-surface antigen heavy chain			1.56			
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase				2.06		
HSP60	60 kDa heat shock protein, mitochondrial			1.30			3.74
RLA0	60S acidic ribosomal protein P0				1.72		
RPLP2	60S acidic ribosomal protein P2	2.02			1.64		
RPL10A	60S ribosomal protein L10a			1.22			
RPL11	60S ribosomal protein L11			2.29			
RL12	60S ribosomal protein L12			1.27	1.36		
RPL14	60S ribosomal protein L14		1.82				
RPL17	60S ribosomal protein L17		1.48				
RPL18	60S ribosomal protein L18					2.73	
RPL22	60S ribosomal protein L22		1.81				
RPL23A	60S ribosomal protein L23a			1.79			
RPL3	60S ribosomal protein L3	1.92					
RPL32	60S ribosomal protein L32			1.70			
RPL35	60S ribosomal protein L35		1.68				
RPL36	60S ribosomal protein L36		33.73		1.38		
RPL36A	60S ribosomal protein L36a			1.38	Infinity		
RPL38	60S ribosomal protein L38			1.26			
RPL6	60S ribosomal protein L6			1.63			
RPL7	60S ribosomal protein L7			1.35			
RPL7A	60S ribosomal protein L7a		1.44	1.23			2.32
ACTA2	Actin, aortic smooth muscle		3.32				
ACTB	Actin, cytoplasmic 2	850.80					1.58
ARPC2[a]	Actin-related protein 2/3 complex subunit 2	1.69	1.43				

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
AHCY	Adenosylhomocysteinase						1.81
AK2	Adenylate kinase 2				2801.99		
SLC25A5	ADP/ATP translocase 2			1.21			
AKR1A1	Aldo-keto reductase family 1, member A1 (aldehyde reductase)						2.06
AGPS	Alkylglycerone phosphate synthase		1.90				
AIMP1	Aminoacyl tRNA synthetase complex-interacting multifunctional protein 1				1.80		
MKI67	Antigen KI-67					2.95	1.84
APEX1	APEX nuclease (multifunctional DNA repair enzyme) 1			1.46			
APOO	Apolipoprotein O		1.40				
AIFM2	Apoptosis-inducing factor, mitochondrion-associated			1.29			
RSRC2	Arginine/serine-rich coiled-coil protein 2						1.73
RARS	Arginyl-tRNA synthetase		2.23				
DARS	Aspartyl-tRNA synthetase	1.92					
ATXN2L	Ataxin-2-like protein	3.15					
ACLY	ATP citrate (pro-S)-lyase			1.29			
ATAD3[a]	ATPase family AAA domain-containing protein 3			1.71	1.72		
ABCE1	ATP-binding cassette sub-family E member 1				1.30		1.34
ABCF2[a]	ATP-binding cassette sub-family F member 2				2.03		
DDX3Y	ATP-dependent RNA helicase			1.43			
EPRS	Bifunctional glutamate/proline--tRNA ligase		1.99				
BOD1L1	Biorientation of chromosomes in cell division 1-like					37.64	
BAIP2	Brain-specific angiogenesis inhibitor 1-associated protein 2		2.85		1.90		3.63
BRD2	Bromodomain-containing protein 2						1.64
ATP2A1	Ca2+ transporting ATPase, sarcoplasmic/endoplasmic reticulum			1.34			
CDH3[a]	Cadherin-3					8.01	
ARPP19	CAMP-regulated phosphoprotein 19		1.71				
CSNK2A1	Casein kinase 2, alpha 1 polypeptide						1.84
CSNK2B	Casein kinase 2, beta polypeptide			1.54			
CSK2B[a]	Casein kinase II subunit beta			1.71			
CLPB	Caseinolytic peptidase B protein homolog		1.75				
CTNND1	Catenin delta-1				1.25		
TCPE	Cct5 T-complex protein 1 subunit epsilon					2.19	
CEP250[a]	Centrosome-associated protein CEP250			1.36			
CHD5	CHD5 chromodomain helicase DNA binding protein 5					13.92	
ARCN1	Coatomer subunit delta			1.20			
CCDC43	Coiled-coil domain-containing protein 43			1.40			
CCDC50	Coiled-coil domain-containing protein 50			4.02			
CCDC86	Coiled-coil domain-containing protein 86					25.96	
CHCHD3	Coiled-coil-helix-coiled-coil-helix domain-containing protein 3, mitochondrial			1.72			
CSDE1	Cold shock domain-containing protein E1	2.21					
EMC8	COX4 neighbor				1.66		
CTBP2	C-terminal-binding protein 2						1.95
CTPS	CTP synthase			1.65	1.38		
CUED2	CUE domain-containing protein 2			3.08			
CAND1	Cullin-associated NEDD8-dissociated protein 1		1.69				
CDK11B	Cyclin-dependent kinase 11B				1.32		
CDK6	Cyclin-dependent kinase 6				1.49		
COX6B[a]	Cytochrome c oxidase subunit 6B1-like		2.01				1.90
MT-CO2	Cytochrome c oxidase subunit II		1.94				
COX6B1	Cytochrome c oxidase subunit VIb				3.48		
COX6C	Cytochrome c oxidase subunit Vic		1.75				
COX7A1	Cytochrome c oxidase subunit VIIa		2.69				
CKAP4	Cytoskeleton-associated protein 4			1.29			
DDX19B	DEAD (Asp-Glu-Ala-As) box polypeptide 19B	2.01				2.64	
DDX6	DEAD (Asp-Glu-Ala-Asp) box polypeptide 6	1.90					
DNTTIP2	Deoxynucleotidyltransferase terminal-interacting protein 2						2.47

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
DRG1	Developmentally-regulated GTP-binding protein 1				1.79		
DHODH	Dihydroorotate dehydrogenase	1.43		1.39			
DIS3	DIS3 mitotic control homolog		1.84				
LIG1	DNA ligase 1						3.44
MCM3	DNA replication licensing factor MCM3				1.27		
MCM5	DNA replication licensing factor MCM5				2.14		
MCM6	DNA replication licensing factor MCM6				1.36		
TOP2	DNA topoisomerase II						3.26
YBX1	DNA-binding protein A			2.04			
POLR2E	DNA-directed RNA polymerase II subunit E	1.75					
DNAJA1	DnaJ homolog subfamily A member 1			2.14	1.75		
DNAJC7	DnaJ homolog subfamily C member 7				1.63		
RANBP2	E3 SUMO-protein ligase RanBP2						3.56
EF1B	Elongation factor 1-beta			1.66			
EF1G	Elongation factor 1-gamma	6.18		1.30			
EEF2	Elongation factor EF-2			1.29	2.12		2.63
ENO1	Enolase	346.21					
ESD	ESD esterase D/formylglutathione hydrolase			1.65			
EIF4A2[a]	Eukaryotic initiation factor 4A-II	2.27					
ERF3B[a]	Eukaryotic peptide chain release factor GTP-binding subunit ERF3B						2.28
EEF1A1	Eukaryotic translation elongation factor 1 alpha 1			1.88	2.44		
EEF1D	Eukaryotic translation elongation factor 1 delta	2.59					
EIF1A	Eukaryotic translation initiation factor 1A			1.50	1.47		
EIF2S1[a]	Eukaryotic translation initiation factor 2 subunit 1-like	4.91					
IF2G	Eukaryotic translation initiation factor 2 subunit 3				1.32		
EIF3E	Eukaryotic translation initiation factor 3 subunit E			1.22			
EIF3G	Eukaryotic translation initiation factor 3 subunit G		2.94	1.92			
EIF3I	Eukaryotic translation initiation factor 3 subunit I	4.08		1.26			
EIF3C	Eukaryotic translation initiation factor 3, subunit C			1.33	1.77		
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1	2.39		1.49	1.57		
EIF4G2	Eukaryotic translation initiation factor 4 gamma 2		3.17				
EIF4A1	Eukaryotic translation initiation factor 4A1	2.28	2.36	1.64	1.64		
EIF4A3	Eukaryotic translation initiation factor 4A3				2.56		
EIF4E	Eukaryotic translation initiation factor 4E				1.35		
ELF5A-1	Eukaryotic translation initiation factor 5A-1	1.80					
EIF5A2	Eukaryotic translation initiation factor 5A-2						13.33
EIF5B[a]	Eukaryotic translation initiation factor 5B				4.08		
ECSIT	Evolutionarily conserved signaling intermediate in Toll pathway, mitochondrial			1.33			
ERCC6	Excision repair cross-complementing rodent repair deficiency, complementation group 6			1.32			
EZR	Ezrin		1.69				
CAPZB	F-actin-capping protein subunit beta		1.61				
FUBP1	Far upstream element (FUSE) binding protein 1					11.64	
KHSRP	Far upstream element-binding protein 2						2.84
FASN	Fatty acid synthase, animal type				1.97		
FLNB	Filamin-B	1.97	1.53				
FLNC	Filamin-C	2.60					
FOXK2	Forkhead box protein K2					1.51	
FHL3	Four and a half LIM domains protein 3				1.25		
FXR1	Fragile X mental retardation syndrome-related protein 1	1.77		1.20			
ALDOA	Fructose-bisphosphate aldolase, class I		1.49		1.33		
GSPT2	G1 to S phase transition 2		2.21				
GSPT1	G1 to S phase transition 2	1.64			1.48		
GOT1	Glutamic-oxaloacetic transaminase 1, soluble (aspartate aminotransferase 1)				1.40		
GLRX3	Glutaredoxin-3			1.32			
GSTA3	Glutathione S-transferase alpha 3				1.38		
GSTP2[a]	Glutathione S-transferase P 2				3.59		

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
GSTP1	Glutathione S-transferase pi 1				5.65		
GAPDH[a]	Glyceraldehyde-3-phosphate dehydrogenase			1.29			
GLO1	Glyoxalase I						3.92
GRSF1	G-rich sequence factor 1			1.56			
G3BP1	GTPase activating protein (SH3 domain) binding protein 1			1.37	1.66		
GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1				1.65		
GNL1	Guanine nucleotide-binding protein-like 1			1.23			
HSPA8	Heat shock cognate 71 kDa protein			1.36			1.58
HS105	Heat shock protein 105 kDa			1.38	1.38		3.05
TRAP1	Heat shock protein 75 kDa, mitochondrial	1.34					
HSP90AB1	Heat shock protein HSP 90-beta				1.42		
HMOX1	Heme oxygenase (decycling) 1	4.72	17.08		4.36		
HDGF	Hepatoma-derived growth factor						3.15
HNRNPH2	Heterogeneous nuclear ribonucleoprotein A/B					Infinity	
HNRPK	Heterogeneous nuclear ribonucleoprotein K					Infinity	
HNRPM	Heterogeneous nuclear ribonucleoprotein M					5.08	
SYNCRIP[a]	Heterogeneous nuclear ribonucleoprotein Q isoform 1	4.00					
HK1	Hexokinase 1		3.13				
HTATSF1	HIV Tat-specific factor 1 homolog			1.43			
PRRC2C	HLA-B associated transcript 2-like 2	2.48					
HCFC1	Host cell factor						2.06
CDC37	Hsp90 co-chaperone Cdc37			1.34			
HSD17B10	Hydroxysteroid (17-beta) dehydrogenase 10 1.1.1.178	Infinity					
HYOU1	Hypoxia up-regulated protein 1		Infinity				
IMPDH1	IMP dehydrogenase				1.30		
IMA2	Importin subunit alpha-2				2.78		
IMB1	Importin subunit beta-1		1.60	1.27			
IPO7	Importin-7			1.53			
IMMT	Inner membrane protein, mitochondrial (mitofilin)		5.51	1.46	1.96		
IGF2BP2	Insulin-like growth factor 2 mRNA-binding protein 2	1.82		1.23			
ILK	Integrin-linked kinase					6.90	
AQR	Intron-binding protein aquarius						3.54
DYNC1H1	K10413 dynein heavy chain 1, cytosolic 3.6.4.2					4.33	
KARS	Kars lysyl-tRNA synthetase	1.90	1.82				
KIF15	Kinesin-like protein KIF15			3.52	3.78		
KIF2A[a]	Kinesin-like protein KIF2A						2.93
KIF2C	Kinesin-like protein KIF2C					2.50	3.09
LDHA	Lactate dehydrogenase A						1.95
LMNA	Lamin-A/C		1.95				2.39
TMPO	Lamina-associated polypeptide 2, isoforms alpha/zeta						3.00
LARP4	La-related protein 4	5.19		1.81			
LPPRC	Leucine-rich PPR motif-containing protein, mitochondrial				1.36		
LRCC59	Leucine-rich repeat-containing protein 59			1.40			
LASP1	LIM and SH3 domain protein 1					5.36	
LIN7B	Lin-7 homolog B			1.26			
SSB	Lupus La protein homolog			Infinity			
MED6	Mediator of RNA polymerase II transcription subunit 6						3.75
CLNS1A	Methylosome subunit pICln			1.28			
MAP4	Microtubule-associated protein 4	7.49					
MTCH1	Mitochondrial carrier homolog 1		1.47				
TIMM9[a]	Mitochondrial import inner membrane translocase subunit Tim9-like			1.21			
TOMM34	Mitochondrial import receptor subunit TOM34			1.62			
TAB3	Mitogen-activated protein kinase kinase kinase 7-interacting protein 3				2.03		
BUB3	Mitotic checkpoint protein BUB3				1.46		
MYOF[a]	Myoferlin			1.38			
MYH10	Myosin, heavy chain 10, non-muscle		1.63	1.81			

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
NAA10	N(alpha)-acetyltransferase 10, NatA catalytic subunit				1.52		
NAA50	N(alpha)-acetyltransferase 50, NatE catalytic subunit					28.81	
NDUFA13	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 13		1.57				
NDUFA4	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4, 9kDa		2.29		1.80		
NDUFS1	NADH dehydrogenase (ubiquinone) Fe-S protein 1		1.52				
AHNAK	Neuroblast differentiation-associated protein AHNAK			1.36			
SLC1A5	Neutral amino acid transporter B(0)	1.76					
NSF1C	NSFL1 cofactor p47				1.23		
NASP	Nuclear autoantigenic sperm protein			1.57			
NUFP2	Nuclear fragile X mental retardation-interacting protein 2				1.64		
NUDC	Nuclear migration protein nudC			1.37			
NUP107	Nuclear pore complex protein Nup107					114.86	
NUP153	Nuclear pore complex protein Nup153						4.70
NOG1	Nucleolar GTP-binding protein 1						2.49
NOLC1	Nucleolar phosphoprotein p130				1.78		
NCL	Nucleolin			1.58			
TIAL1	Nucleolysin TIAR			1.80			
TPR	Nucleoprotein TPR		3.60	1.41			
OLA1	Obg-like ATPase			1.26			
CDC73	Parafibromin		2.68				
PPIL2	Peptidylprolyl isomerase (cyclophilin)-like 2					1.58	
PPID	Peptidyl-prolyl isomerase D (cyclophilin D)			1.37			
PRDX1	Peroxiredoxin 1				1.48		3.08
PCYT1A	Phosphate cytidylyltransferase 1, choline, alpha			1.44			
PCYT2	Phosphate cytidylyltransferase 2, ethanolamine			1.27			
PFKL	Phosphofructokinase, liver			1.42			
PGAM1	Phosphoglycerate mutase						2.87
PRPS2	Phosphoribosyl pyrophosphate synthetase 2			1.36			
SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein			1.31			
PLS3	Plastin-3	4.41					
PLEC	Plectin		1.67				
PABP1	Polyadenylate-binding protein 4	1.98					
CCDC72	Predicted gene 4363				Infinity		
VBP1	Prefoldin subunit 3			1.23			
PRPF6	Pre-mRNA-processing factor 6				2.30		2.50
BRR2	Pre-mRNA-splicing helicase BRR2					294.16	
PSAP	Proactivator polypeptide			1.40			
PCNA	Proliferating cell nuclear antigen						2.15
PA2G4	Proliferation-associated protein 2G4	1.58		1.35			
P4HA1	Prolyl 4-hydroxylase			1.57			
PSME3	Proteasome (prosome, macropain) activator subunit 3 (PA28 gamma			1.41			
PRMT1	Protein arginine methyltransferase		2.62		1.49		
CNPY2	Protein canopy homolog 2			1.33			
PPP1CB	Protein phosphatase 1, catalytic subunit	1.39				Infinity	
PPP1CA	Protein phosphatase 1, catalytic subunit, gamma isozyme				1.71		
PPP1R2	Protein phosphatase 1, regulatory (inhibitor) subunit 2			8.23			
SET	Protein SET				1.44		
PCNP	Proteolytic signal-containing nuclear protein			1.42			
RBM15[a]	Putative RNA-binding protein 15						2.28
PDXDC1	Pyridoxal-dependent decarboxylase domain containing	4.29	1.67				
PKLR	Pyruvate kinase						1.97
QDPR	Quinoid dihydropteridine reductase		1.31				
RGAP1	Rac GTPase-activating protein 1						2.32
RAD18	RAD18 homolog (S. cerevisiae)						2.79
RDX	Radixin		1.61	1.47			
RAGP1	Ran GTPase-activating protein 1		1.81				

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
RAN	Ran GTP-binding nuclear protein Ran		1.39				
RANBP1	Ran-specific GTPase-activating protein						2.02
IQGAP1	Ras GTPase-activating-like protein IQGAP1		2.99				
RAB11B	Ras-related protein Rab-11B			1.83			
RAB18[a]	Ras-related protein Rab-18	2.43					
RAB21	Ras-related protein Rab-18	1.66					
RENT1	Regulator of nonsense transcripts 1		2.70				
UPF1	Regulator of nonsense transcripts 1	2.97					
RPA3	Replication protein A 14 kDa subunit		3.89				
RAVER1	Ribonucleoprotein, PTB-binding 1						7.90
RRM2	Ribonucleotide reductase M2			1.59	5.14		
RPN2	Ribophorin II			1.35			
NSA2	Ribosome biogenesis protein NSA2 homolog						1.71
RRBP1	Ribosome-binding protein 1		1.88	1.20			
SEPT11	Septin-11				1.58		
SHMT2	Serine hydroxymethyltransferase 2 (mitochondrial)						1.79
SRRM2	Serine/arginine repetitive matrix protein 2	1.44					2.46
STRAP	Serine-threonine kinase receptor-associated protein	2.89			1.37		
TF[a]	Serotransferrin-like		2.08		3.28		
SRP72	Signal recognition particle 72 kDa protein				1.37		
SRPRB	Signal recognition particle receptor subunit beta			1.48			
SNRPB	Small nuclear ribonucleoprotein-associated proteins B and B'						7.36
LSM8	snRNA-associated Sm-like protein LSm8			1.31			
SPTBN1	Spectrin beta chain, brain 1			1.56			1.23
SF3B2	Splicing factor 3B subunit 2					1.38	
SRSF1	Splicing factor, arginine/serine-rich 1						1.44
SCAF11	Splicing factor, arginine/serine-rich 2, interacting protein						3.18
SRSF6	Splicing factor, arginine/serine-rich 6						1.34
SQLE	Squalene epoxidase			3.80			
SLIRP[a]	SRA stem-loop-interacting RNA-binding protein, mitochondrial-like			1.41			
CTTN	Src substrate cortactin	1.40					
SND1	Staphylococcal nuclease and tudor domain containing 1	2.46			1.30		
SMCHD1	Structural maintenance of chromosomes flexible hinge domain-containing protein 1					2.53	
SMC1A	Structural maintenance of chromosomes protein 1A					1.44	
SMC2	Structural maintenance of chromosomes protein 2					248.52	
SMC4	Structural maintenance of chromosomes protein 4	3.16					
SMRC1	SWI/SNF complex subunit SMARCC1						2.15
TCP1	T-complex protein 1 subunit alpha		1.48	1.38			2.42
CCT2	T-complex protein 1 subunit beta	1.69		1.24	1.54		
CCT4	T-complex protein 1 subunit delta				1.62		
CCT3	T-complex protein 1 subunit gamma				1.23		
CCT8	T-complex protein 1 subunit theta			1.27	1.28		
EMC2	Tetratricopeptide repeat domain 35		1.75				
TXNRD1	Thioredoxin reductase 1			1.53			
TXNL1	Thioredoxin-like protein 1			1.25			
THO3	THO complex subunit 3						4.32
TJP2	Tight junction protein ZO-2			2.29			
SPT6H	Transcription elongation factor SPT6			2.70			
BTF3	Transcription factor BTF3			1.56			
JUNB	Transcription factor jun-B						9.16
GATAD2A	Transcriptional repressor p66 alpha						2.60
TBL2	Transducin beta-like protein 2						2.35
TFRC	Transferrin receptor protein 1			2.02			
TKTL2	Transketolase-like 2				449.72		3.21
EIF3A	Translation initiation factor eIF-3 subunit 10			1.34			
TPT1	Translationally-controlled tumor protein		1.61				

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31°C	37°C	31°C	37°C	31°C	37°C
TOM22	Translocase of outer mitochondrial membrane 22 homolog (yeast)			1.62			
SSRA	Translocon-associated protein subunit alpha				1.35		
SSR1	Translocon-associated protein subunit alpha	1.26					
TMCO1	Transmembrane and coiled-coil domain-containing protein 1		1.30				
TPM3	Tropomyosin alpha-3 chain	1.66					
TUBB4	Tubulin beta-4 chain			2622.72			
TTLL12	Tubulin--tyrosine ligase-like protein 12			1.48			
RRP9	U3 small nucleolar RNA-interacting protein 2				1.61		
PRP31	U4/U6 small nuclear ribonucleoprotein Prp31				1.58		
PRPF4	U4/U6 small nuclear ribonucleoprotein Prp4					7.39	
UBE2C	Ubiquitin-conjugating enzyme E2C				2.23		
UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A9			1.32			
C14orf166	UPF0568 protein C14orf166	2.59					
N/A	UPF0723 protein C11orf83 homolog		1.86		1.29		
UMPS	Uridine monophosphate synthetase			1.32	1.38		
VPS25	Vacuolar protein-sorting-associated protein 25				1.23		
HDLBP	Vigilin	11.35					
VIM	Vimentin			41.31	Infinity		
VDAC2	Voltage-dependent anion-selective channel protein 2			1.24			
ZPR1	Zinc finger protein ZPR1				1.55		
ZHX1	Zinc fingers and homeoboxes protein 1		1.69				

[a]

Protein identification obtained from NCBI nonredudant database which was not present in BBCHO or had a higher confidence score compared to BBCHO database

[b]

Fold change showing increased protein expression in CHO cells grown at 24hr compared to 8hr timepoint at 31°C and 37°C

ix. Down-regulated proteins in CHO-K1 SEAP cells at 24 hr compared to 8 hr time point at 31 and 37 °C in membrane, cytoplasmic and nuclear fractions

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
PSMA4	20S proteasome subunit alpha 4		Infinity				
PSMD4	26S protease regulatory subunit 4			1.31			
PRS6B	26S protease regulatory subunit 6B			1302.12			
MRPS31	28S ribosomal protein S31, mitochondrial			Infinity			
OGDH	2-oxoglutarate dehydrogenase E1 component				1.64		
MRPL12	39S ribosomal protein L12, mitochondrial					1.45	
MRPL39	39S ribosomal protein L39, mitochondrial			1.68			
RPS12	40S ribosomal protein S12						1.39
RPS13	40S ribosomal protein S13					1.36	
RPS15	40S ribosomal protein S15			Infinity	#######		
RPS15A	40S ribosomal protein S15a						1.36
RPS16	40S ribosomal protein S16						1.40
RPS19	40S ribosomal protein S19					789.56	
RPS21	40S ribosomal protein S21					1.26	
RPS24	40S ribosomal protein S24					2.61	
RPS25	40S ribosomal protein S25						1.68
RPS3	40S ribosomal protein S3						2.14
RPS5	40S ribosomal protein S5					Infinity	
RPS9	40S ribosomal protein S9						2.05
HSP60	60 kDa heat shock protein, mitochondrial				4.46		
RPL11[a]	60S ribosomal protein L11						1.25
RPL13	60S ribosomal protein L13						1.62
RPL19	60S ribosomal protein L19	2.71					
RPL21	60S ribosomal protein L21			29.32			
RPL4[a]	60S ribosomal protein L4					2.17	
RPL7	60S ribosomal protein L7				1.26		
GRP78	78 kDa glucose-regulated protein			1.47	2.61		
ACO2	Aconitase 2, mitochondrial	1.56			1.57		
ACTG1	Actin, cytoplasmic 2					9075.00	
ACTR2	Actin-related protein 2					99.60	
ARPC1B	Actin-related protein 2/3 complex subunit 1B			1.28			
ACSF2	Acyl-CoA synthetase family member 2		1.88				
ACAD9	Acyl-Coenzyme A dehydrogenase family, member 9	1.45					
AK3	Adenylate kinase 3				1.56		
FDX1	Adrenodoxin, mitochondrial			1.31	1.32		
AKAP12	A-kinase anchor protein 12			1.77			
AARS	Alanyl-tRNA synthetase			1.20	1.85		
ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)			1.31			
ALDR	Aldehyde reductase			1.20			
AKR1A1	Aldo-keto reductase family 1, member A1 (aldehyde reductase)			1.24			
ACTN1	Alpha-actinin-1				1.26		
SNAA[a]	Alpha-soluble NSF attachment protein		1.30				
ANKFY1	Ankyrin repeat and FYVE domain-containing protein 1	5.03					
ANXA1	Annexin A1	1.87		1.55	2.92		
ANXA2	Annexin A2	1.79					
ANXA4	Annexin A4			1.60			
ANXA5	Annexin A5			1.30	1.37		
ANXA7	Annexin A7				1.88		
ASRGL1	Asparaginase like 1			1.28			
ASNS	Asparagine synthetase				2.54		
PEA15	Astrocytic phosphoprotein PEA-15				1.53		
ATP5H	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit d	1.54					
ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle					1.33	
ABCF1	ATP-binding cassette sub-family F member 1						1.91
DHX9	ATP-dependent RNA helicase A		2.58				
DDX42	ATP-dependent RNA helicase DDX42					1.34	

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
BAG3[a]	BAG family molecular chaperone regulator 3			1.55			
HSPG2	Basement membrane-specific heparan sulfate proteoglycan core protein		1.95				
BOP1	Block of proliferation 1			1.70			
BOLA1[a]	BolA-like protein 1	1.49					
CAMK2D	Calcium/calmodulin-dependent protein kinase (CaM kinase) II delta			1.32			
SCMC1	Calcium-binding mitochondrial carrier protein SCaMC-1				1.60		
CALR[a]	Calreticulin	1.59		1.77	79.15		
PRKAR1A	cAMP-dependent protein kinase regulator			1.28			
CHST11	Carbohydrate (chondroitin 4) sulfotransferase 11					1.44	
CES1	Carboxylesterase 1 (monocyte/macrophage serine esterase 1)	1.74			2.08		
CPT2	Carnitine O-palmitoyltransferase 2	1.37	1.29				
CATB	Cathepsin B		1.76				
CATD	Cathepsin D				2.07		
CATL1	Cathepsin L1				1.42		
CTSZ	Cathepsin Z		1.85	1.45	1.44		
CD109	CD109 antigen	Infinity					
CD59	CD59 glycoprotein	2.71				8.49	
CD63	CD63 antigen						1.89
CD9[a]	CD9 antigen	2.42					
CLIC1	Chloride intracellular channel 1				1.28		
PCYT1A	Choline-phosphate cytidylyltransferase		3.78				
CSPG4	Chondroitin sulfate proteoglycan 4			Infinity			
CS	Citrate synthase	1.76					
MAPK8IP3	C-jun-amino-terminal kinase-interacting protein 3				1.84		
CLTC	Clathrin heavy chain 1	1.32		148.79			
CLTB	Clathrin light chain B					3.17	
COPB1	Coatomer protein complex, subunit beta 1	1.62					
COPB2	Coatomer subunit beta'	2.08					
COPG2	Coatomer subunit gamma-2		1.89				
CIRBP[a]	Cold-inducible RNA-binding protein	2.76		2.74			
NCAPG	Condensin complex subunit 3						2.71
CCS	Copper chaperone for superoxide dismutase				1.36		
COR1B	Coronin-1B					5.09	
CRN7	Coronin-7			1.23			
CDK2	Cyclin-dependent kinase 2					1.46	
NFS1	Cysteine desulfurase			1.50			
CREL2	Cysteine-rich with EGF-like domain protein 2				1.56		
QCR1	Cytochrome b-c1 complex subunit 1, mitochondrial					2.06	
UQCRFS1[a]	Cytochrome b-c1 complex subunit 7					1.74	
CYCS	Cytochrome c, somatic						1.29
CKAP4	Cytoskeleton-associated protein 4						2.22
CTLA2A	Cytotoxic T lymphocyte-associated protein 2 alpha			1.88			
DDX1	DDX1 DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	1.68					
DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17	2.26	2.12				
DDX47	DEAD (Asp-Glu-Ala-Asp) box polypeptide 47		2.92				
DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 50		Infinity				
DDX56	DEAD (Asp-Glu-Ala-Asp) box polypeptide 56	1.91					
DRG1	Developmentally-regulated GTP-binding protein 1						1.88
DAB2	Disabled homolog 2				2.26		
DBN1	DNA damage-binding protein 1			1.42			
MCM3	DNA replication licensing factor MCM3		1.22				
MCM4	DNA replication licensing factor MCM4			183.99			
DNAJB11	DnaJ homolog subfamily B member 11						1.57
RPN1[a]	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1			1.47		2.14	1.94
DBNL	Drebrin-like protein			1.35			
DNM2	Dynamin 2						2.46

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
OPA1	Dynamin-like 120 kDa protein, mitochondrial	1.99					
ELAV1	ELAV-like protein 1		1.90				
ETFDH	Electron-transferring-flavoprotein dehydrogenase	1.71					
EF1G	Elongation factor 1-gamma						168.79
HSP90B1	Endoplasmin			1.26	1.27		
EDF1	Endothelial differentiation-related factor 1					1.74	
HADHA	Enoyl-CoA hydratase / long-chain 3-hydroxyacyl-CoA dehydrogenase			1.42			
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1					1.78	
EIF4A2	Eukaryotic translation initiation factor 4A2					1.53	
EIF4A3	Eukaryotic translation initiation factor 4A3	1.74					
IF4B	Eukaryotic translation initiation factor 4B					1.42	
EZR	Ezrin					1.87	
FAH	FAH K01555 fumarylacetoacetase				1.59		
FUBP1	Far upstream element (FUSE) binding protein 1		2.09				
KHSRP	Far upstream element-binding protein 2		1.66				
FERMT2	Fermitin family homolog 2			1.31			
FLNB	Filamin-B				1.44	1.74	
FKBP10	FK506 binding protein 10	1.70					
FKBP4	FK506 binding protein 4, 59kDa				1.26		
FLNA	FLNA Filamin-B			1.20			
FLOT1	Flotillin-1				1.42		2.94
FHL1	Four and a half LIM domains protein 1				1.77		
ALDOA	Fructose-bisphosphate aldolase, class I	1.44					
FH	Fumarate hydratase, class II	1.58					
LEG3	Galectin-3				1.49		1.79
NAPG	Gamma-soluble NSF attachment protein		1.62				
GTF2I	General transcription factor II-I						168.11
GAA	Glucosidase, alpha				1.69		
GLUD1	Glutamate dehydrogenase 1	1.40		1.32			1.70
GOT2	Glutamic-oxaloacetic transaminase 2, mitochondrial (aspartate aminotransferase 2)	1.33			1.24		
GSTM7	Glutathione S-transferase, mu 7			2769.51			
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase		1.52				
GARS	Glycyl-tRNA synthetase				1.56		
GLOD4	Glyoxalase domain-containing protein 4			1.49			
G3BP1	GTPase activating protein (SH3 domain) binding protein 1						1.60
RAN	GTP-binding nuclear protein Ran					1.30	
GNAI1	Guanine nucleotide-binding protein G(i), alpha-2 subunit	1.43	202.84				
GNA13	Guanine nucleotide-binding protein subunit alpha-13			1.23			
GNB2L1	Guanine nucleotide-binding protein subunit beta-2-like 1						1.69
HSPA1A	Heat shock 70 kDa protein 1A/1B					Infinity	
HSPB1	Heat shock protein beta-1	1.82					
HSP90AA1	Heat shock protein HSP 90-alpha				2.88		
HP1BP3	Heterochromatin protein 1-binding protein 3						4.97
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1		Infinity				
ROA3	Heterogeneous nuclear ribonucleoprotein A3	Infinity					
HNRNPCL1	Heterogeneous nuclear ribonucleoprotein C-like 1	1984.36					
HNRPD	Heterogeneous nuclear ribonucleoprotein D0		1.98				
HNRNPDL	Heterogeneous nuclear ribonucleoprotein D-like	2.58					
HNRNPF	Heterogeneous nuclear ribonucleoprotein F						1.44
HNRH3	Heterogeneous nuclear ribonucleoprotein H3		2.91				
HNRNPK	Heterogeneous nuclear ribonucleoprotein K	1.66					
HNRPL	Heterogeneous nuclear ribonucleoprotein L		3.60	2.56			
HNRNPLL	Heterogeneous nuclear ribonucleoprotein L-like			1.27			
HNRPM	Heterogeneous nuclear ribonucleoprotein M		2.62				
SYNCRIP	Heterogeneous nuclear ribonucleoprotein Q					1.95	
HNRNPR	Heterogeneous nuclear ribonucleoprotein R		3.04	4.44			2.55

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
HNRNPU	Heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)	1.97	2.33	1.36			1.37
HNRNPUL2	Heterogeneous nuclear ribonucleoprotein U-like protein 2		3.19				
HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1		2.32				
HNRNPC	Heterogeneous nuclear ribonucleoproteins C1/C2		5.76				
HIST1H1C	Histone H1.2			5.29	3.88	2.75	
H2AZ	Histone H2A.Z				1.79	2.21	
HIST1H2AG[a]	Histone H3				2.00	2.30	
H33	Histone H3.3				1.96		
H3F3A[a]	Histone H3.3			3.93			
HIST1H4A	Histone H4						2.07
HSD17B4	Hydroxysteroid (17-beta) dehydrogenase 4					2.91	
HSDL2	Hydroxysteroid dehydrogenase-like protein 2	1.59					
HMGB1P1[a]	Hypothetical protein I79_003161				2.88		
N/A[a]	Hypothetical protein I79_008707			2.65			
HYOU1	Hypoxia up-regulated protein 1				1.33		
ILF3	Ilf3 Interleukin enhancer-binding factor 3	2.22	2.77				
ITGA5	Integrin alpha-5		2.27				
ITIH5	Inter-alpha-trypsin inhibitor heavy chain H5	2.39					
ICAM1	Intercellular adhesion molecule 1	Infinity					
ILF2	Interleukin enhancer-binding factor 2		3.65				
ISCA1[a]	Iron-sulfur cluster assembly 1, mitochondrial	1.41					
IDH2	Isocitrate dehydrogenase		1.59				
IARS2	Isoleucyl-tRNA synthetase 2, mitochondrial	1.60					
KHDRBS1	KH domain-containing, RNA-binding, signal transduction-associated protein 1	2.43					
MFGM	Lactadherin		1.47				
LMNA	Lamin-A/C			6.21			
LAMB1	Laminin subunit beta-1						4.78
LAMC1	Laminin subunit gamma-1				1.36		
LAMA3	Laminin, alpha 3	2.56	2.19				
LDHA	LDHA lactate dehydrogenase A	1.84		1.23			
LGMN	Legumain		2.26		3.25		
LAP3	Leucine aminopeptidase 3		1.36				
SERPINB1A	Leukocyte elastase inhibitor A			1.60			
LMO7	LIM domain only protein 7					5.65	
ACADL	Long-chain-acyl-CoA dehydrogenase		1.27				
LPL	LPL lipoprotein lipase						2.57
LAMP2	Lysosome-associated membrane glycoprotein 2		1.61				
LOXL1	Lysyl oxidase-like 1	3.66					
MDH2	Malate dehydrogenase 2, NAD (mitochondrial)			1.25			
MATR3	Matrin-3		3.01				
MANF	Mesencephalic astrocyte-derived neurotrophic factor		1.35				
MESDC2	Mesoderm development candidate 2		1.83		1.47		
MET	Met proto-oncogene (hepatocyte growth factor receptor)				2.29		
MTHFD2	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2			1.22			
MAP1B	Microtubule-associated protein 1B				1.58		
MAP4	Microtubule-associated protein 4			1.56	1.53		
SLC25A20	Mitochondrial carnitine/acylcarnitine carrier protein	2.00					
TOM70	Mitochondrial import receptor subunit TOM70	1.37					
MOSC2	MOSC domain-containing protein 2, mitochondrial						5.42
MYH1	Myosin I					1.76	
MYH9	Myosin-9					1.61	2.96
NDUFB11	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 11, 17.3kDa	1.34					
NDUFV3	NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa					117.40	
NDUFA3[a]	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 3					2.15	
AHNAK	Neuroblast differentiation-associated protein AHNAK	1.21	1.75		1.53		1.66
GANAB[a]	Neutral alpha-glucosidase AB			1.37			

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
NCLN	Nicalin	1.29					
NME1	Non-metastatic cells 1						2.28
CALD1	Non-muscle caldesmon					1.67	
NSF	NSF K06027 vesicle-fusing ATPase	1.65					
NOL10	Nucleolar protein 10		2.32				
NCL	Nucleolin						1.47
PPIL1	Peptidylprolyl isomerase (cyclophilin)-like 1			1.31			
PLIN3	Perilipin 3				1.86		
PLIN4	Perilipin 4				2.35		
PEBP1	Phosphatidylethanolamine-binding protein 1						1.27
PCK1	Phosphoenolpyruvate carboxykinase (GTP)		1.50				
PHGDH	Phosphoglycerate dehydrogenase	2.00	2.03		1.57		
PGK1	Phosphoglycerate kinase 1			1.28			
PSAT1	Phosphoserine aminotransferase 1				1.55		
PLEKHA5	Pleckstrin homology domain-containing family A member 5						1.67
PUF60	Poly(U)-binding-splicing factor PUF60	7.77					
PABP1	Polyadenylate-binding protein 1					1.95	1.55
PTRF	Polymerase I and transcript release factor					3.43	
PTBP1	Polypyrimidine tract-binding protein 1		1.67				
PPIH	PPIH peptidylprolyl isomerase H (cyclophilin H)			1.57			
PRPF6	Pre-mRNA-processing factor 6	5.77					
PSAP	Proactivator polypeptide		1.94				
DDX5[a]	Probable ATP-dependent RNA helicase DDX5		1.99				
EBNA1BP2	Probable rRNA-processing protein EBP2						1.54
PHB	Prohibitin		Infinity				
PHB2	Prohibitin-2					1.87	
AATF	Protein AATF					18.97	
CNPY2	Protein canopy homolog 2				1.31		
CNPY3[a]	Protein canopy-like 3				1.43		
DEK	Protein DEK						1.73
PDIA3	Protein disulfide isomerase family A, member 3	1.44		1.21	1.26		
PDIA4	Protein disulfide isomerase family A, member 4				3.04		
PACSIN2	Protein kinase C and casein kinase substrate in neurons protein 2				1.40		
PRKCDBP[a]	Protein kinase C delta-binding protein						2.06
PRKCSH	Protein kinase C substrate 80K-H				1.57		
MAGOH	Protein mago nashi homolog 1-related		3.16				
NIPSNAP1	Protein NipSnap homolog 1				1.77		
GBAS	Protein NipSnap homolog 2			2.78			
PPP3CA	Protein phosphatase 3, catalytic subunit				1.65		
CRK	Proto-oncogene C-crk				1.59		
PNP	Purine-nucleoside phosphorylase			1.22			
PDHA1	Pyruvate dehydrogenase (lipoamide) alpha 1	1.32					
PKLR	Pyruvate kinase					1628.83	
GDI1[a]	Rab GDP dissociation inhibitor alpha				1.58		
RDX	Radixin						1.35
IQGAP1	Ras GTPase-activating-like protein IQGAP1			1.22			
RPA3	Replication protein A 14 kDa subunit				780.22		
RHG18	Rho GTPase-activating protein 18			1.48			
ARHGAP18[a]	Rho GTPase-activating protein 18, partial			1.79			
RNH1	Ribonuclease inhibitor			1.39			
RSL1D1	Ribosomal L1 domain-containing protein 1	2.24					
RPL37A	Ribosomal protein L37a	Infinity					
RRBP1	Ribosome-binding protein 1						1.86
RBM9	RNA binding motif protein 9	2.85					
RBMX	RNA binding motif protein, X-linked	3.13					
RBM12	RNA-binding protein 12			2.05			

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
RBM14	RNA-binding protein 14		2.16			2.22	
RALY	RNA-binding protein Raly		2.44				2.08
FBL	rRNA 2'-O-methyltransferase fibrillarin	184.05					
RUVBL2	RuvB-like protein 2			1.27			
SUN2	Sad1 and UNC84 domain containing 2					1.68	1.92
AT2A2[a]	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	1.85					
SELENBP1	Selenium-binding protein 1			2.09			
SBP[a]	Selenium-binding protein 1-like			2.09			
SEPT7	Septin-7					1.90	
SHM2	Serine hydroxymethyltransferase 2 (mitochondrial)	1.30				1.33	
SRRM2[a]	Serine/arginine repetitive matrix protein 2			1.26			
STK24	Serine/threonine kinase 24 (STE20 homolog, yeast)			1.36			
SPB6	Serpin B6				1.39		
SHOX2	Short stature homeobox protein 2						1.40
SNRPD2	Small nuclear ribonucleoprotein Sm D2	1.41					
SNRPD3	Small nuclear ribonucleoprotein Sm D3				1.24		
SRI	Sorcin				1.40		
SNX1	Sorting nexin-1				1.31		
SPTAN1	Spectrin alpha					1.72	1.73
SF3A3	Splicing factor 3A subunit 3						1.67
SRSF1	Splicing factor, arginine/serine-rich 1	1.44					
SRSF9	Splicing factor, arginine/serine-rich 1/9						10.67
SRSF11	Splicing factor, arginine/serine-rich 11						206.79
SFPQ	Splicing factor, proline- and glutamine-rich	4.23					
SSR1	SSR1 Translocon-associated protein subunit alpha						1.50
SSR4	Ssr4 Translocon-associated protein subunit delta	1.27					
STMN1	Stathmin				6.34		
SCP2	Sterol carrier protein 2			1.28			
SDHA	Succinate dehydrogenase (ubiquinone) flavoprotein subunit			1.43	1.33		
SUCLG1	Succinate-CoA ligase, alpha subunit			Infinity			
SUCLA2	Succinyl-CoA synthetase beta subunit		1.44				
SBSN	Suprabasin				1.34		
STX4	Syntaxin-4				1.85		
TLN1	Talin-1					1.79	
TAP2	TAP2 Antigen peptide transporter 2	1.71					
TADBP	TAR DNA-binding protein 43		2.86				
TCPE	T-complex protein 1 subunit epsilon			1.46			
TST	Thiosulfate sulfurtransferase	1.26					
THOC2	THO complex 2						2.49
THRAP3	Thyroid hormone receptor-associated protein 3	1.47		1.51			
TSTA3	Tissue specific transplantation antigen P35B			1.89			
I79_018912[a]	TPR repeat-containing protein C10orf93-like	1.34					
TCEB2[a]	Transcription elongation factor B polypeptide 2			1.39			
TCERG1	Transcription elongation regulator 1					1.36	
RHOA	Transforming protein RhoA	1.33					
TPI1	Triosephosphate isomerase 1	2.87			1.43		
TMOD3	Tropomodulin-3						1.49
TUBA1B	Tubulin alpha-1B chain	259.80					
TUBB6	Tubulin beta-6 chain		1235.10				
SNRNP70	U1 small nuclear ribonucleoprotein 70 kDa	1.50	1.61				
SNRPA	U1 small nuclear ribonucleoprotein A						2.04
SNRPA1	U2 small nuclear ribonucleoprotein A'		1.37		2692.86		1.84
UQCRC1	Ubiquinol-cytochrome c reductase core protein I	1.27					
UQCRB	Ubiquinol-cytochrome c reductase core protein I					1.91	
UQCRC2	Ubiquinol-cytochrome c reductase core subunit 2	1.39					
UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A9	1.43					1.75

		8hr v 24hr Fold Change[b]					
		Membrane		Cytoplasmic		Nuclear	
Gene name	Description	31oC	37oC	31oC	37oC	31oC	37oC
UAP1L1	UDP-N-acteylglucosamine pyrophosphorylase 1-like 1				1.53		
UGDH	UGDH UDP-glucose dehydrogenase	2.35					
VARS2[a]	Valyl-tRNA synthetase-like			1.20			
VNN1	Vanin 1	2.57					
YARS	YARS tyrosyl-tRNA synthetase				1.28		
ZC3H15	Zinc finger CCCH domain-containing protein 15					1.97	
ZYX	Zyxin				1.36		

[a]

Protein identification obtained from NCBI nonredudant database which was not present in BBCHO or had a higher confidence score compared to BBCHO database

[b]

Fold change showing decreased protein expression in CHO cells grown at 24hr compared to 8hr timepoint at 31°C and 37°C

x. Up-regulated proteins in CHO-K1 SEAP cells at 24 hr compared to 8 hr time point at 31 and 37 °C in membrane, cytoplasmic and nuclear fractions

DAVID analysis 16hr time course at 31 and 37 °C with subcellular fractionation

Biological Process			
Fraction	**UP regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	RNA processing	16	6.1E-05
	Generation of precursor metabolites and energy	12	3.6E-04
	mRNA processing	12	4.7E-04
	RNA splicing	11	1.3E-03
	mRNA metabolic process	12	1.9E-03
Cytoplasmic	Protein folding	7	9.2E-02
Nuclear	Translation	9	4.2E-03
	Translational elongation	6	8.2E-03
	Cytoskeleton organization	9	3.0E-02

Fraction	**UP regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	RNA splicing	14	6.8E-10
	mRNA processing	14	3.2E-09
	mRNA metabolic process	14	1.9E-08
	RNA processing	15	1.9E-07
	RNA splicing, via transesterification reactions with bu	10	3.1E-07
Cytoplasmic	Maintenance of location	6	6.4E-03
Nuclear	Translational elongation	8	1.8E-05
	Translation	10	5.3E-04
	mRNA metabolic process	9	1.2E-02
	RNA splicing	8	1.8E-02
	RNA processing	10	3.0E-02

Biological Process			
Fraction	**DOWN regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Translation	12	2.3E-06
	Posttranscriptional regulation of gene expression	8	2.1E-03
	Regulation of translation	6	3.5E-02
Cytoplasmic	Translation	29	3.3E-19
	Translational elongation	20	6.0E-19
	Translational initiation	7	7.8E-04
Nuclear	Cell cycle process	7	3.8E-02

Fraction	**DOWN regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Translation	15	5.9E-09
	Translational elongation	6	2.2E-02
Cytoplasmic	Translation	21	1.5E-13
	Translational elongation	14	2.4E-12
Nuclear	NO SIG		

xi. Top five biological processes (Bonferonni $p<0.05$) differentially regulated in CHO cells grown at 31 and 37 °C over 16hr between time points of 8 and 24 hr after temperature shift using GO analysis through DAVID. Similar process can be seen to be enriched for at both temperatures with posttranscriptional regulation of gene expression, regulation of translation, translational initiation and cell cycle process being specifically associated with downregulated proteins at 31 °C.

Molecular Function			
Fraction	**UP regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	RNA binding	19	2.3E-06
	Nucleotide binding	25	3.5E-02
Cytoplasmic	NO SIG		
Nuclear	Cytoskeletal protein binding	11	6.4E-04
	Actin binding	9	1.5E-03
	Structural constituent of ribosome	7	2.6E-03
	Structural molecule activity	11	4.7E-03
	RNA binding	11	1.3E-02

Fraction	**UP regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	RNA binding	21	9.7E-13
	Nucleotide binding	26	1.0E-07
	RNA helicase activity	5	2.3E-04
	RNA-dependent ATPase activity	4	5.0E-03
	Purine NTP-dependent helicase activity	5	2.9E-02
Cytoplasmic	NO SIG		
Nuclear	RNA binding	16	9.0E-07
	Structural constituent of ribosome	7	3.0E-03

Molecular Function			
Fraction	**DOWN regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	RNA binding	17	3.8E-07
	Translation factor activity, nucleic acid binding	7	2.1E-04
	Translation initiation factor activity	5	1.0E-02
Cytoplasmic	RNA binding	31	1.5E-12
	Structural constituent of ribosome	18	2.1E-12
	Structural molecule activity	23	4.4E-07
	mRNA binding	9	9.2E-06
	Translation factor activity, nucleic acid binding	9	2.1E-04
Nuclear	ATP binding	9	4.5E-02
	Adenyl ribonucleotide binding	9	4.9E-02

Fraction	**DOWN regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Actin binding	8	2.6E-02
Cytoplasmic	Structural constituent of ribosome	12	5.6E-07
	Translation factor activity, nucleic acid binding	8	3.0E-04
	Purine ribonucleotide binding	27	1.0E-03
	Ribonucleotide binding	27	1.0E-03
	Purine nucleotide binding	27	2.3E-03

xii. Top five molecular functions (Bonferonni $p<0.05$) differentially regulated in CHO cells grown at 31 and 37 ^{o}C over 16hr between time points of 8 and 24 hr after temperature shift using GO analysis through DAVID. Similar process can be seen to be enriched for at both temperatures with posttranscriptional regulation of gene expression, regulation of translation, translational initiation and cell cycle process being specifically associated with down-regulated proteins at 31 ^{o}C.

Cellular Component			
Fraction	**UP regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Ribonucleoprotein complex	16	1.1E-05
	Mitochondrial part	16	7.3E-05
	Mitochondrion	21	9.6E-05
	Membrane-enclosed lumen	25	2.9E-03
	Intracellular organelle lumen	24	4.7E-03
Cytoplasmic	Intracellular organelle lumen	24	3.2E-04
	Organelle lumen	24	4.8E-04
	Membrane-enclosed lumen	24	6.8E-04
	Melanosome	6	1.1E-02
	Pigment granule	6	1.1E-02
Nuclear	Intracellular non-membrane-bounded organelle	23	4.9E-04
	Non-membrane-bounded organelle	23	4.9E-04
	Ribosomal subunit	7	6.4E-04
	Cytosolic ribosome	6	1.2E-03
	Cytosolic small ribosomal subunit	5	1.6E-03

Fraction	**UP regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Spliceosome	10	7.6E-08
	Ribonucleoprotein complex	15	1.3E-07
	Heterogeneous nuclear ribonucleoprotein complex	6	3.3E-07
	Nucleolus	14	5.3E-05
	Intracellular organelle lumen	20	3.0E-04
Cytoplasmic	Soluble fraction	11	3.8E-04
	Cell fraction	17	6.2E-03
	Endoplasmic reticulum lumen	6	7.2E-03
	Cytosol	18	2.1E-02
Nuclear	Ribonucleoprotein complex	18	1.2E-10
	Cytosolic ribosome	8	2.3E-06
	Ribosome	10	9.5E-06
	Non-membrane-bounded organelle	26	3.9E-05
	Intracellular non-membrane-bounded organelle	26	3.9E-05

Cellular Component			
Fraction	**DOWN regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Ribonucleoprotein complex	14	2.9E-06
Cytoplasmic	Cytosol	49	8.5E-20
	Cytosolic part	20	7.1E-16
	Cytosolic ribosome	16	7.1E-15
	Ribonucleoprotein complex	28	1.7E-13
	Ribosome	20	5.2E-13
Nuclear	Intracellular non-membrane-bounded organelle	13	2.5E-03
	Non-membrane-bounded organelle	13	2.5E-03
	Chromosome	7	2.6E-03
	Condensed chromosome	5	2.6E-03
	Chromosomal part	6	1.5E-02

Fraction	**DOWN regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Non-membrane-bounded organelle	27	4.4E-05
	Intracellular non-membrane-bounded organelle	27	4.4E-05
	Large ribosomal subunit	6	9.9E-04
	Cytosolic large ribosomal subunit	5	2.4E-03
	Ribosome	8	2.9E-03
Cytoplasmic	Cytosol	34	5.5E-14
	Vytosolic part	14	9.0E-11
	Cytosolic ribosome	11	1.3E-09
	Cytosolic small ribosomal subunit	9	3.2E-09
	Ribonucleoprotein complex	19	8.3E-09
Nuclear	Intracellular organelle lumen	18	8.4E-04
	Organelle lumen	18	1.1E-03
	Nuclear lumen	16	1.5E-03
	Membrane-enclosed lumen	18	1.5E-03
	Ribonucleoprotein complex	9	1.8E-02

xiii. Cellular components associated with differentially regulated proteins in CHO cells grown over 16hr at 31 ^{o}C and 37 ^{o}C temperature shift using GO analysis through DAVID.

Panther analysis 16hr time course at 31 and 37 ^{o}C with subcellular fractionation

Biological Process			
Fraction	**UP regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Metabolic process	61	7.71E-10
	Primary metabolic process	52	2.39E-07
	RNA splicing	10	1.11E-05
	RNA splicing, via transesterification reactions	10	1.11E-05
	mRNA processing	12	1.15E-05
Cytoplasmic	Metabolic process	51	3.12E-05
	Protein folding	8	6.63E-05
	Primary metabolic process	45	9.75E-05
	Protein metabolic process	23	5.79E-03
Nuclear	Translation	11	4.14E-07
	Cellular component organization or biogenesis	15	4.72E-06
	Cellular component organization	12	8.23E-04
	Cell cycle	13	2.43E-03
	Metabolic process	35	2.60E-03

Fraction	**UP regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	mRNA splicing, via spliceosome	17	2.17E-14
	mRNA processing	17	3.69E-13
	RNA splicing	12	1.94E-09
	RNA splicing, via transesterification reactions	12	1.94E-09
	Primary metabolic process	42	4.37E-07
Cytoplasmic	Primary metabolic process	47	4.76E-05
	Metabolic process	51	2.22E-04
	Protein folding	7	1.18E-03
	Protein metabolic process	24	3.75E-03
Nuclear	Translation	11	1.81E-06
	mRNA splicing, via spliceosome	8	1.09E-03
	mRNA processing	8	3.83E-03
	RNA splicing	6	1.67E-02
	RNA splicing, via transesterification reactions	6	1.67E-02

Biological Process			
Fraction	**DOWN regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Translation	17	2.35E-14
	Regulation of translation	11	8.07E-12
	Protein metabolic process	25	4.29E-07
	Primary metabolic process	37	1.98E-05
	Metabolic process	39	2.35E-04
Cytoplasmic	Translation	31	1.91E-23
	Protein metabolic process	56	2.58E-16
	Metabolic process	91	4.31E-12
	Primary metabolic process	80	1.31E-10
	Regulation of translation	12	5.59E-09
Nuclear	mRNA processing	6	2.25E-03
	Nuclear transport	4	2.70E-03
	Chromosome segregation	4	1.29E-02
	mRNA splicing, via spliceosome	5	1.41E-02
	Meiosis	3	4.87E-02

Fraction	**DOWN regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Translation	11	5.46E-06
Cytoplasmic	Translation	26	3.52E-21
	Protein metabolic process	45	7.72E-15
	Metabolic process	68	2.00E-09
	Regulation of translation	11	3.23E-09
Nuclear	Unclassified	15	0.00E+00
	mRNA splicing, via spliceosome	11	7.29E-07
	mRNA processing	11	4.30E-06
	RNA splicing	7	1.70E-03
	RNA splicing, via transesterification reactions	7	1.70E-03

xiv. Biological processes differentially regulated in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using PANTHER. Similar processes are seen differentially regulated with protein up regulation such as metabolic process, translation, RNA processing and protein processing. Down-regulated proteins show specific dysregulation of nuclear transport, meiosis and chromosome segregation at 31 ^{o}C compared to 37 ^{o}C.

Molecular Function			
Fraction	**UP regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	RNA binding	19	1.56E-10
	Catalytic activity	48	2.74E-09
	mRNA binding	12	2.27E-07
	Oxidoreductase activity	12	4.36E-04
	Poly(A) RNA binding	4	4.49E-02
Cytoplasmic	Oxidoreductase activity	10	8.25E-03
	Structural constituent of ribosome	6	1.10E-02
	Isomerase activity	5	4.45E-02
Nuclear	Structural molecule activity	17	2.31E-07
	Structural constituent of cytoskeleton	11	4.77E-04
	Binding	29	7.62E-04
	Nucleic acid binding	20	3.19E-03
	Single-stranded DNA binding	4	4.42E-03

Fraction	**UP regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	RNA binding	24	2.11E-19
	mRNA binding	14	1.88E-11
	Poly(A) RNA binding	7	6.94E-07
	Nucleic acid binding	27	4.15E-06
	Catalytic activity	34	2.39E-05
Cytoplasmic	Isomerase activity	6	4.89E-03
Nuclear	Structural constituent of ribosome	10	2.44E-08
	RNA binding	13	1.41E-06
	Structural molecule activity	17	2.21E-06
	Nucleic acid binding	26	8.14E-06
	Binding	35	1.12E-05

Molecular Function			
Fraction	**DOWN regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	RNA binding	19	8.37E-14
	Translation regulator activity	11	4.95E-12
	Translation factor activity, nucleic acid binding	11	7.38E-12
	Translation initiation factor activity	9	7.74E-10
	Nucleic acid binding	28	6.25E-08
Cytoplasmic	Structural constituent of ribosome	19	2.08E-15
	RNA binding	22	2.38E-09
	Translation regulator activity	12	3.38E-09
	Translation factor activity, nucleic acid binding	12	5.16E-09
	Structural molecule activity	29	1.48E-08
Nuclear	NO SIG		

Fraction	**DOWN regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Structural molecule activity	14	1.93E-03
	Translation initiation factor activity	4	3.75E-02
Cytoplasmic	Translation regulator activity	11	1.99E-09
	Translation factor activity, nucleic acid binding	11	2.95E-09
	Translation initiation factor activity	10	3.75E-09
	Structural constituent of ribosome	11	1.88E-07
	RNA binding	16	1.57E-06
Nuclear	RNA binding	12	1.67E-05
	Nucleic acid binding	23	9.24E-04
	mRNA binding	7	3.26E-03
	Binding	29	3.07E-02

xv. Molecular functions differentially regulated in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using PANTHER. Up-regulated proteins show specific dysregulation of oxidoreductase activity, structural constituent of the cytoskeleton and single stranded DNA binding at 31 ^{o}C compared to 37 ^{o}C. Similar processes are seen differentially regulated with protein down regulation such as binding, translational regulation and structural molecule activity.

Cellular Component			
Fraction	**UP regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Macromolecular complex	13	2.24E-05
	Ribonucleoprotein complex	7	2.63E-05
	Vesicle coat	3	2.00E-02
Cytoplasmic	NO SIG		
Nuclear	Organelle	15	3.48E-07
	Intracellular	17	4.24E-07
	Actin cytoskeleton	10	7.51E-07
	Cell part	17	1.97E-06
	Cytoskeleton	12	7.15E-06

Fraction	**UP regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Ribonucleoprotein complex	13	2.44E-15
	Macromolecular complex	16	5.29E-10
Cytoplasmic	NO SIG		
Nuclear	Ribonucleoprotein complex	5	1.24E-03

Cellular Component			
Fraction	**DOWN regulated at 31oC**	**Count**	**Adjusted p-value**
Membrane	Ribonucleoprotein complex	5	9.49E-04
	Actin cytoskeleton	6	1.74E-02
Cytoplasmic	Ribonucleoprotein complex	9	4.13E-06
	Ribosome	3	1.67E-03
	Macromolecular complex	13	3.92E-03
Nuclear	Ribonucleoprotein complex	3	2.39E-02
	Macromolecular complex	5	3.98E-02

Fraction	**DOWN regulated at 37oC**	**Count**	**Adjusted p-value**
Membrane	Intracellular	14	2.14E-03
	Actin cytoskeleton	7	5.75E-03
	Cell part	14	6.54E-03
	Organelle	11	8.62E-03
	Cytoskeleton	9	1.84E-02
Cytoplasmic	Ribosome	2	4.45E-02
Nuclear	NO SIG		

xvi. Cellular components associated with differentially regulated proteins in CHO cells grown at 31 ^{o}C 8 and 24 hr after temperature shift using PANTHER. Up-regulated proteins show specific dysregulation to many structural elements such as vesicle coat, organelle, intracellular, actin cytoskeleton and cytoskeleton terms at 31 ^{o}C compared to 37 ^{o}C. Down-regulated proteins show specific dysregulation to many structural elements such as vesicle ribonucleoprotein complex and the macromolecule complex terms at 31 ^{o}C compared to 37 ^{o}C.

Temperature shift target transfections

Cyclon

	siRNA1		siRNA2		siRNA3	
Target	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**
Total viable cells/ml	7.67E-02	-1.39	**4.19E-02**	**-1.47**	**3.00E-02**	**-1.52**
Cell viability	6.83E-01	-1.02	3.27E-01	-1.05	5.47E-01	-1.03
Cell average diameter	9.89E-01	1.00	1.75E-01	1.03	4.04E-01	1.01
Cell average area	5.52E-02	-1.27	**3.16E-02**	**-1.30**	**8.34E-05**	**-1.43**
Cell average perimeter	6.36E-02	-1.22	**1.85E-02**	**-1.25**	**4.84E-04**	**-1.34**

Ezrin

	siRNA1		siRNA2		siRNA3	
Target	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**
Total viable cells/ml	7.73E-01	-1.10	9.42E-01	1.01	5.86E-02	-1.42
Cell viability	6.01E-01	1.02	4.19E-01	0.95	4.29E-01	-1.01
Cell average diameter	4.58E-02	-1.02	4.89E-01	1.00	1.35E-01	1.00
Cell average area	4.80E-01	-1.10	4.62E-01	-1.10	5.35E-02	-1.32
Cell average perimeter	5.80E-01	-1.06	4.72E-01	-1.08	5.52E-02	-1.25

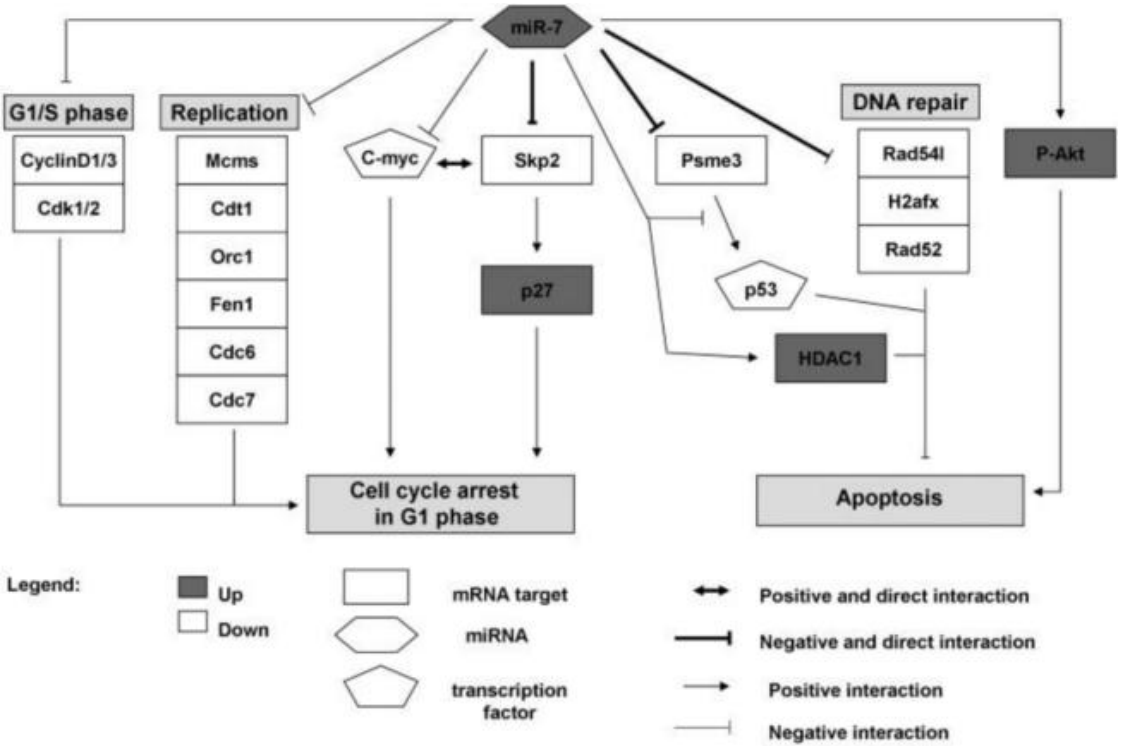
Moesin

	siRNA1		siRNA2		siRNA3	
Target	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**
Total viable cells/ml	6.12E-02	-1.41	**6.12E-02**	**-1.58**	**6.12E-02**	**-1.74**
Cell viability	8.33E-01	1.01	8.33E-01	-1.04	8.33E-01	-1.04
Cell average diameter	6.25E-01	1.00	6.25E-01	-1.01	6.25E-01	-1.02
Cell average area	**4.06E-02**	**-1.30**	**4.06E-02**	**-1.44**	**4.06E-02**	**-1.58**
Cell average perimeter	**4.32E-02**	**-1.24**	**4.32E-02**	**-1.33**	**4.32E-02**	**-1.44**

Lamin A

	siRNA1		siRNA2		siRNA3	
Target	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**	**Mean p-value (n=3)**	**Mean fold change (n=3)**
Total viable cells/ml	**4.95E-02**	**-1.46**	1.50E-01	-1.28	1.08E-01	-1.42
Cell viability	6.63E-01	-1.03	7.87E-01	1.01	9.08E-01	-1.01
Cell average diameter	4.80E-01	1.01	4.24E-01	1.01	7.26E-01	1.00
Cell average area	**8.31E-03**	**-1.19**	1.97E-01	-1.05	7.33E-02	-1.23
Cell average perimeter	5.15E-02	-1.14	1.78E-01	-1.04	8.60E-02	-1.17

xvii. Effect of siRNA mediated knockdown of Cyclon, Ezrin, Moesin or Lamin A in CHO K1-SEAP cells after 72 hr on total viable cell/ml, cell viability, cell average diameter, cell average area and cell average perimeter. Mean p-value and mean fold change are calculated against the siRNA negative control (n=3). Significant values were calculated using a two tailed, two sample unequal variance student's t-test and fold changes denoted with a minus sign represent a fold decrease compared to the negative control.



xviii. **Schematic of the proposed transcriptional effect of miR-7 on CHO cell growth**. Potential direct interactions between miR-7 and targets SKP2 and PSME3 mediate cell cycle arrest and apoptosis resistance respectively. Source - (Sanchez et al. 2013)