UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Cancer Sciences Academic Unit

Characterisation of Skeletal Stem Cell populations by Western Blotting, Aptamer Selection and Mass Spectrometry

by

Nunzia Sposito

Thesis for the degree of Doctor of Philosophy

September 2015

Supervisors: Professor Richard O. C. Oreffo and Professor Paul A. Townsend

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE

Stem cells and regenerative medicine

Thesis for the degree of Doctor of Philosophy

CHARACTERISATION OF SKELETAL STEM CELL POPULATIONS BY WESTERN BLOTTING, APTAMER SELECTION AND MASS SPECTROMETRY

By Nunzia Sposito

With an aging population, the spectrum of treatable conditions broadens, especially within the orthopoedic field. The burden of bone related diseases has been foreseen to be potentially alleviated by skeletal stem cells (SSCs). However, the clinical potential of these cells is hampered by the absence of specific isolation techniques, due to the lack of specific markers. The aim of this work was to identify specific markers for SSC isolation, thereby allowing the therapeutic potential of SCC to be harnessed. Advanced quantitative mass spectrometry was used to analyse SSC proteome of four patients, 2 males aged 75 and 64 and 2 females aged 42 and 67. Cells were sorted for the Stro-1 marker (a cell surface marker widely used for SSCs), and cultured in osteogenic and basal media. Changes in protein expression were investigated. More than 7,000 proteins were identified and quantified, 6 proteins were significantly upregulated in the osteogenic group. Data suggest that protein modulation was highly governed by the chemical stimuli artificially introduced in the culture condition through the addition of dexamethasone and ascorbic acid (osteogenic media). In addition to the mass spectrometry experiments, characterisation of the Stro-1 antigen was undertaken by Western Blotting. The Stro-1 antibody was observed to react with an epitope of 50 kDa. Furthermore characterisation of the surface markers of SSC was performed adopting an aptamer SELEX process. The results presented here are a further step towards the characterisation of the proteomic profile of SSCs, and identification of unique markers for the isolation of a homogeneous SSC population.

A POSSE A ESSE

VERBA VOLANT, SCRIPTA MANENT

SEMPER AD MELIORA

Contents

ABSTRACT iii

Contents vii

List of tables xiii

List of figures xiv

List of Accompanying Materials xix

DECLARATION OF AUTHORSHIP xxi

Acknowledgements xxiii

Definitions and Abbreviations xxv

1. Chapter 1: Introduction 1

1.1 Section I: Bone Structure and composition 3
  1.1.1 The skeletal system 3
  1.1.2 Bone macrostructure 3
  1.1.3 Bone microstructure organisation 5
  1.1.4 Bone remodelling: role of osteoblasts and osteoclasts. 7
    1.1.4.1 Osteoblasts 9
    1.1.4.2 Osteocytes 9
    1.1.4.3 Osteoclasts 9
  1.1.5 Bone histogenesis: intramembranous and endochondral ossification. 10
  1.1.6 Control of osteogenesis 11
    1.1.6.1 Transcription factors 12
    1.1.6.2 Growth factors 13
    1.1.6.3 Proteins 15
1.2 Section II: Introduction to marrow stromal and skeletal stem cells 17
  1.2.1 Overview 17
  1.2.2 History 17
  1.2.3 Stem cells for regenerative medicine and tissue engineering 18
  1.2.4 Stem cell niche 21
  1.2.5 Mesenchymal stem cells, skeletal stem cells and bone marrow stromal cells 22
  1.2.6 SSC markers 23
  1.2.7 *In vivo* identity of SSCs and MSCs 27
  1.2.8 Clinical needs and applications 29
1.3 Section III: Introduction to Mass Spectrometry 31
  1.3.1 Overview 31
  1.3.2 Mass spectrometry 31
  1.3.3 Brief history of mass spectrometry 31

1.3.4 Proteomics 34
1.3.5 The LTQ-Velos Pro Orbitrap-Elite 35
1.3.6 Isobaric tags for the relative and absolute quantification (iTRAQ) 37
1.4 Section IV: Aptamers technology for biomarker discovery 38
1.4.1 Aptamers: History 38
1.4.2 SELEX 39
1.4.3 Aptamer libraries 41
1.4.4 Target molecules 41
1.4.5 Aptamers versus antibodies 42
1.5 Aim and Hypothesis 45
2. Chapter 2: Materials and Methods 47
2.1 Cell culture and osteoinduction of human bone marrow cells for mass spectrometry analysis 49
2.2 Magnetic Activated Cell Sorting (MACS) isolation 49
2.3 Alkaline Phosphatase staining 50
2.4 Molecular analysis 51
2.4.1 RNA extraction 51
2.4.2 cDNA synthesis 51
2.4.3 Primer design 51
2.4.4 Real-time PCR (RT-PCR) 51
2.4.5 Statistics 52
2.5 Stro-1 antibody production 52
2.6 Fluorescent Immunostaining 53
2.7 Cell lysis for Western Blotting (WB) 53
2.8 Protein quantification (BCA assay) 54
2.9 Western Blotting 55
2.10 Aptamer selection process (SELEX) .... Error! Bookmark not defined.
2.11 Liquid chromatography/Mass spectrometry 57
2.12 Aptamer selection process (SELEX) 57
3. Chapter 3: Characterisation of Human Skeletal Stem cells 59
3.1 Introduction 61
3.2 Aims and hypothesis 62
3.3 Material and Methods 63
3.3.1 Characterisation of bone marrow cells 63
3.3.1.1 Cell culture, MACS and osteoinduction 63
3.3.1.2 Alkaline phosphatase staining 63
3.3.1.3 Molecular analysis 63
3.3.2 Stro-1 characterisation experiments 63
3.3.2.1 Cell culture 63
3.3.2.2 Cell lysis for Stro-1 WB 65
3.3.2.3 Fluorescent immunostaining 66

3.3.2.4 Immunoblotting: Western Blotting ... 66
3.3.2.5 Statistics ... 66
3.3.2.6 Immunoprecipitation ... 66
3.3.3 Aptamer selection ... 68
3.3.3.1 Cell culture ... 68
3.3.3.2 Reagents for the aptamer selection ... 68
3.3.3.3 Aptamer library ... 70
3.3.3.3.1 DNA random library ... 70
3.3.3.4 Purification of the Aptamer library ... 70
3.3.3.4.1 DNA extraction from acrylamide gels ... 70
3.3.3.4.2 Ethanol precipitation ... 71
3.3.3.5 Overview of the selection process ... 71
3.3.3.6 Aptamer incubation with cells and cell lysis ... 72
3.3.3.7 Automated reverse transcription, cycle course PCR ... 74
3.3.3.8 Gel electrophoresis: amplified DNA ... 74
3.3.3.9 Manual PCR: DNA amplification ... 74
3.3.3.10 Overnight automated transcription ... 74
3.4 Results ... 75
3.4.1 Characterisation of bone marrow extracted cells ... 75
3.4.1.1 Alkaline phosphatase expression increases in the osteogenic phenotype ... 75
3.4.1.2 Molecular analysis demonstrated osteogenesis ... 76
3.4.2 Stro-1 characterisation experiments ... 80
3.4.2.1 Stro-1 positive and negative cell lines ... 80
3.4.2.2 Western Blotting analysis of the Stro-1 antigen ... 87
3.4.2.3 Characterisation of the Stro-1 antibody derived by hybridoma supernatant ... 89
3.4.2.4 WB band analysis ... 91
3.4.3 Aptamer selection ... 93
3.4.3.1 Representative results ... 93
3.4.3.2 Summary of results ... 94
3.4.3.3 Trouble-shooting approaches ... 96
3.5 Discussion ... 97
3.5.1 Characterisation of SSCs ... 97
3.5.2 Stro-1 antibody and antigen analysis ... 97
3.5.3 WB band analysis ... 98
3.5.4 Aptamers technology ... 99
3.6 Conclusion ... 101
4. Chapter 4: Mass Spectrometry 1 ... 103
4.1 Introduction ... 104
4.2 Aims and hypothesis: ... 105
4.3 Materials and Methods ... 107

4.3.1 Sample preparation 107
4.3.2 Magnetic Activated Cell Sorting (MACS) isolation 107
4.3.3 Cell culture and harvesting 107
4.3.4 Reagents and Chemicals 109
4.3.5 Sample preparation for liquid chromatography 111
4.3.6 LC-MS Analysis 113
4.3.7 MS data processing 115
4.3.7.1 Downstream data analysis 115
4.3.8 Hierarchical clustering 117
4.3.9 Protein annotation and interactions: FunRich and STRING. 117
4.4 Results 119
4.4.1 Off-line alkaline reverse phase chromatography 119
4.4.2 On-line LC-MS Analysis 120
4.4.3 Hierarchical clustering 122
4.4.4 FunRich and STRING 124
4.5 Discussion 129
4.5.1 Overview 129
4.5.2 Limitation of the study: Proteome coverage 136
4.5.3 Data analysis limitation 138
4.6 Conclusion 138
5. Chapter 5: Mass spectrometry 2 141
5.1 Introduction 143
5.2 Aims and hypothesis 144
5.3 Materials and methods 146
5.3.1 Reagents and chemicals 146
5.3.2 Differences with mass spectrometry study 1 146
5.3.3 Sample preparation 146
5.3.4 MS Data processing 147
5.3.5 Downstream data analysis 148
5.3.5.1 Hierarchical clustering 148
5.3.5.2 Upregulated and downregulated proteins 148
5.3.5.3 Protein annotation and interactions: FunRich and STRING 149
5.3.5.4 MetaCore analysis 149
5.3.6 Western blotting validation 150
5.4 Results 152
5.4.1 Offline alkaline reverse phase chromatography 152
5.4.2 Online LC-MS Analysis 152
5.4.3 Hierarchical clustering 155
5.4.4 Upregulated proteins 159
5.4.5 Downregulated proteins 160
5.4.6 FunRich 160

5.4.7 MetaCore analysis 167
5.4.8 Western blotting-based validation of selected differentially expressed proteins 168
5.4.9 *In silico* analysis of 50 kDa proteins for correlation and association to Stro-1 based on immunostaining 171
5.4.10 Method 171
5.4.11 Results 171
5.5 Discussion 175
5.5.1 Overview 175
5.5.2 Hierarchical clustering provides for a comprehensive overview of the protein expression highlighting patient and gender differences 176
5.5.3 Upregulation and downregulation analysis by 1.5 and 2 fold highlighted that protein modulation was driven by osteogenesis 177
5.5.3.1 Peptidyl-prolyl cis-trans isomerase 177
5.5.3.2 Carboxypeptidase M 177
5.5.3.3 Amine oxidase [flavin-containing] A 178
5.5.3.4 Fibrillin-2 178
5.5.4 Classification of the proteins highlighted a high catalytic activity 179
5.5.5 Pathway analysis using FunRich and MetaCore attributed the protein modulation to osteoinduction by downregulation of the GCR-alpha cycle 179
5.5.6 Validation of the mass spectrometry results by Western Blotting 180
5.5.7 Further observations 181
5.5.8 Considerations 182
5.5.8.1 Gender differences 182
5.5.8.2 Differences with MS1 study 183
5.5.9 Discussion of *in silico* analysis of 50 kDa proteins for correlation and association to Stro-1 based on immunostaining 183
5.6 Conclusion 185
5.6.1 Hypotheses iteration 185
5.6.2 Limitation of the MS study 186
6. Chapter 6: General discussion 187
6.1 Setting the scene 189
6.2 Results in context 193
6.3 Limitations 194
6.4 Future perspective 195
7. Chapter 7: Conclusion 197
7.1 Conclusion 199
8. List of References 201
9. Appendix 1 227
9.1 ALP staining of 3 patients SSC Stro-1 sorted and unsorted cells. 227
9.2 Flow cytometry 228

9.2.1 Sample preparation ... 228
9.2.2 Flow cytometry measurements ... 228
9.2.3 Results ... 228
9.3 Cell fractionation ... 236
9.4 MACS IP ... 238
9.5 Blue native page (BN-PAGE) ... 240
9.6 WB Band Analysis (documentation given by Dr. Omar Jallow, St. Georges University of London) ... 240
9.7 Solubility test with different lysis buffers ... 241
9.8 Preliminary experiments for the aptamer process ... 243
10. Appendix 2 ... 245
10.1 Protein concentration correlates with cell number ... 245
10.2 Correlation between lysis buffers ... 245
10.2.1 Guava cell count versus haemocytometer count ... 246
10.2.2 Protein quantification according to cell number ... 246
10.2.3 Correlation between lysis buffers ... 247
10.2.4 Correlation between lysis buffers ... 249
10.3 Separation of the Stro-1 bright fraction by flow cytometry ... 250

List of tables

Table 1: Positive SSC selection markers. 24

Table 2: Negative SSC selection markers. 25

Table 3: Forward and reverse primers used for RT-PCR analysis in this work.. 52

Table 4: BCA assay total protein calculation. 55

Table 5: Antibody specifications used in this study. 58

Table 6: Reagents employed during the aptamer selection. 68

Table 7: Details of each selection round. 73

Table 8: Mass spectrometry reagents information. 109

Table 9: iTRAQ Reagent-8plex Multiplex Kit. 109

Table 10: Off-line HPLC phases details. 110

Table 11: On-line HPLC phases details. 110

Table 12: iTRAQ labelling. 112

Table 13: Nodes identified by FunRich protein-protein interaction map. 124

Table 14: 7 common elements in "F85", "MALDI MS" and "ESI MS". 135

Table 15: 7 common elements in "F85" and "ESI MS". 135

Table 16: Method variation between MS1 and MS2. 146

Table 17: Antibodies used for the Western Blotting validation of mass spectrometry studies 1 and 2 150

Table 18: Clinical studies using MSC in the world. 189

Table 19: Subcellular proteome extraction buffers. 236

Table 20: Lysis buffers. 242

Table 21: Cell density preliminary study. 244

Table 22: Description of adopted lysis buffers for protein determination..... 248

List of figures

Figure 1: The principal components of the femur. 4

Figure 2: Bone internal structure 6

Figure 3: High magnification of cortical bone structure. 7

Figure 4: Bone remodelling process, consisting of a cycle of resorption, reversal, formation and resting. 8

Figure 5: Rat osteoclast and resorption pit 10

Figure 6: The stages of endochondral ossification. 11

Figure 7: Canonical Wnt/β-catenin signaling pathway. 14

Figure 8: Scientific Publication on MSCs from 1980 to 2014 in PubMed. 18

Figure 9: SSCs differentiation potential. 20

Figure 10: SSC *in vivo* location within the bone marrow. 29

Figure 11: Bone diseases. 30

Figure 12: Thomson's instrument. 32

Figure 13: Inside the LTQ-Velos Pro Orbitrap-Elite. 36

Figure 14: iTRAQ labels chemistry. 37

Figure 15: SELEX process. 40

Figure 16: BCA assay standard curve example. 54

Figure 17: WCL, SOL and INS fractions. 65

Figure 18: Aptamer process overview. 72

Figure 19: Human bone marrow extracted stromal cells in culture. 75

Figure 20: Alkaline phosphatase expression based on image-based analysis. 76

Figure 21: ALP gene expression. 77

Figure 22: OCN gene expression. 77

Figure 23: SOX9 gene expression. 78

Figure 24: Alkaline phosphatase staining. 79

Figure 25: HK cells Stro-1 immunostaining. 81

Figure 26: MG63 cells Stro-1 immunostaining. 82

Figure 27: BXPC3 cells Stro-1 immunostaining. 83

Figure 28: C28I2 cells Stro-1 immunostaining. 84

Figure 29: Raji cells immunostaining for Stro-1. 85

Figure 30: 721.221 cells immunostaining for Stro-1. 86

Figure 31: Isolation of the soluble and unsoluble fractions within cell lysate.. 87

Figure 32: Soluble and insoluble fractions of HK cells probed with Stro-1 supernatant antibody. 88

Figure 33: Stro-1 hybridoma supernatant western blotting characterisation... 90

Figure 34: Mass spectrometry analysis of the 50 kDa gel band. 92

Figure 35: Selection round 4 results. 93

Figure 36: Round 7 of the aptamer selection process. 94

Figure 37: DNA from previous rounds. 95

Figure 38: Round 7 and 9. 95

Figure 39: Trouble-shooting. 96

Figure 40: Skeletal profiling methodology. 106

Figure 41: F85 extraction and cell culture. 108

Figure 42: Offline HPLC elution gradient profile. 113

Figure 43: Online chromatography elution gradient profile. 114

Figure 44: iTraq ratios rationale. 116

Figure 45: High pH reverse phase chromatography, MS1. 119

Figure 46: Reproducible proteins overlapping (Venn diagram). 121

Figure 47: Clustering of the 644 differentially regulated proteins in F85...... 122

Figure 48: Clustering of the 26 highly differentially regulated proteins in F85. 123

Figure 49: Biological pathway classification. 125

Figure 50: Cellular component classification... 125

Figure 51: Biological process classification. 126

Figure 52: Molecular function classification. 126

Figure 53: Network of interaction between the discovered proteins. 127

Figure 54: Heatmap of the 26 proteins found in F85 when present in other tissue.. 128

Figure 55: F85 BM population enrichment. 129

Figure 56: Current study compared to Mindaye, Ra et al. 134

Figure 57: Process flow diagram of mass spectrometry analysis.................. 145

Figure 58: iTRAQ labelling.. 147

Figure 59: iTRAQ ratios rationale.. 148

Figure 60: High pH reverse phase chromatography (C_8 column), MS2.......... 152

Figure 61: Common proteins between genders... 153

Figure 62: LC/MS representative results of carboxypeptidase M peptide SLTPDDDVFQYLAHTYASR. 154

Figure 63: Hierarchical clustering of proteins found in M75 and M64. 2877 proteins... 156

Figure 64: Most differentially expressed proteins in M75 and M64. 157

Figure 65: Hierarchical clustering of proteins found in F42 and F67............. 158

Figure 66: Upregulated proteins by 2 and 1.5 fold...................................... 159

Figure 67: Downregulated proteins by 1.5 fold. 160

Figure 68: Biological pathway of M75, M64, F42 and F67. 162

Figure 69: Biological process of M75, M64, F42 and F67. 163

Figure 70: Cellular component of M75, M64, F42 and F67. 164

Figure 71: Molecular function of M75, M64, F42 and F67. 165

Figure 72: Interaction of the protein processes found in M75, M64, F42 and F67. 166

Figure 73: MetaCore Pathway analysis. 167

Figure 74: Carboxypeptidase M validation. 169

Figure 75: Fibrillin 2 validation. 169

Figure 76: Decorin and Alkaline phosphatase validation. 170

Figure 77: PANTHER protein classification based on cell compartments. 172

Figure 78: PANTHER protein subgroup of membrane proteins. 173

Figure 79: Protein expression of 50 kDa membrane proteins (left) and Stro-1 antibody staining (right). 174

Figure 80: The three pillars of the Humane Proteome Project (HPP). 194

Figure 81: ALP staining panel of 3 patients Stro-1 sorted and unsorted SSC. 227

Figure 82: Stro-1 expression in Raji cells analysed by flow cytometry. 229

Figure 83: Stro-1 expression in 721.221 cells analysed by flow cytometry 230

Figure 84: Stro-1 expression in DU145 cells analysed by flow cytometry 231

Figure 85: Stro-1 expression in PC3 cells analysed by flow cytometry. 232

Figure 86: Stro-1 expression in C28I2 cells analysed by flow cytometry. 233

Figure 87: Raji cells immunostaining for Stro-1. 234

Figure 88: 721.221 cells immunostaining for Stro-1. 235

Figure 89: HK cells fractionation by solvent extraction................................ 237

Figure 90: MACS IP and WB. 239

Figure 91: HK cells solubility under different lysis buffers........................... 242

Figure 92: WB with different lysis buffers. 242

Figure 93: Flow cytometry cell counts against manual cell count with a haemocytometer. 246

Figure 94: C2C12 protein concentration at different cell densities. 247

Figure 95: Comparisons of proteins extracted from hBM cells using two different lysis buffers. 249

Figure 96: Stro-1 bright cells separation. 250

List of Accompanying Materials

Appendix 1....................227

Appendix 2....................245

DECLARATION OF AUTHORSHIP

I, Nunzia Sposito

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Mass spectrometry-based proteomics for the characterisation of skeletal cell populations.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. [Delete as appropriate] None of this work has been published before submission [or] Parts of this work have been published as: [please list references below]:

Signed:...

Date: ....

Acknowledgements

I wish to thank my supervisors Prof Richard OC Oreffo and Prof Paul A Townsend for their continuous support during these long 4 years.

I wish to thank Dr. Spiro Garbis and the proteomics group, mainly Harvey Johnston for support during the last months; Dr. Christopher Woelk and the bioinformatics group of the University of Southampton.

I would also like to thank the entire Bone and Joint research group at the University of Southampton and in particular May De Andrés González for the many late evenings spent with me running PCR analysis.

I wish to thank you the PAT group and including Tom Jackson and Sam Larkin, and especially Charlie Cotter for her support during very difficult times.

I wish to thank you Otto Muskens, Stef Inglis, David Gibbs, David Gothard and Jon Dawson for the thorough corrections that helped this thesis to be in the shape it is now.

I want to thank you my PhD friend Agnieszka Janeczek, the one I met 4 years ago at the bus stop on my first PhD day. She helped me getting through these 4 very hard years with helpful life advice. She has also been the latest editor of this thesis.

A special thank you to my friend Giusi Manfredi for her daily support during this last year.

I would like to thank the people at BTG, especially Sade Sobande and Salma Yunis for their moral support, Ben Minihane, David Hedley and Harshini Patel for correcting my English, Sean Willis and Koorosh Ashrafi for spending two long hours preparing me for the VIVA and Stephen Jones for helping me with last minute formatting.

Finally I would like to thank my family for being the original source from which all this work has derived.

Definitions and Abbreviations

ABS	Absolute number
ALP	Alkaline phosphatase
BCA	Bicinchoninic acid assay
BM	Bone marrow
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
BSP	Bone sialoprotein
CD	Cluster of differentiation
CFU-F	Colony forming unit fibroblast
CID	Collision induced dissociation
CPM	Carboxypeptidase M
CSK	Tyrosine-protein kinase
DC	Direct Current
DCN	Decorin
DPSC	Dental pulp stem cells
EGF	Epidermal growth factor
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FACS	Fluorescent activated cell sorting
FBN2	Fibrillin-2
FCS	Foetal calf serum
FGF	Fibroblast growth factor
FKBP5	Peptidyl-prolyl cis-trans isomerase
FT	Fourier Transformation
FTH1	Ferritin heavy chain
GC	Gas Chromatography
GCR	Glucocorticoid receptor
h	Hours
hBMSCs	Human bone marrow stromal cells
HCD	High collision dissociation
HPLC	High pressure liquid chromatography
IGF	Insulin like growth factors
IgG	Immunoglobulin G
IgM	Immunoglobulin M

INS	Insoluble
IP	Immunoprecipitation
ISCT	International Society for Cellular Therapy
iTRAQ	Isobaric tags for relative and absolute quantitation
kDa	kilo Dalton
LC	Liquid chromatography
LTQ	Linear trap quadrupole
m/z	Mass to charge ratio
m/z	Mass-to-charge ratio
MACS	Magnetic activated cell sorting
MALDI	Matrix-assisted laser desorption ionisation
MAOA	Amine oxidase [flavin-containing] A
METTL7A	Methyltransferase-like protein 7A
min	Minutes
MS	Mass spectrometry
MS1	Mass spectrometry (study) 1
MS2	Mass spectrometry (study) 2
MSC(s)	Mesenchymal stem cell(s)
MudPIT	Multidimensional Protein Identification Technology
nanoRP	Nano reverse phase
nESI	Nano Electrospray ionisation
NICE	The National Institute for Health and Care Excellence
OCN	Osteocalcin
OSP	Osteopontin
OSX	Osterix
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PDGF	Platelet derived growth factor
PSEN2	Presenilin 2
PTX3	Pentraxin-related protein 3
RF	Radio Frequency
RP	Reverse phase
rpm	Revolution per min
RT	Room temperature
RT-PCR	Real time polymerase chain reaction

RUNX2	Runt related transcription factor 2
s	Seconds
SB	Sorted basal
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SELEX	Systematic evolution of ligands by exponential enrichment
SLC16A3	Monocarboxylate transporter 4
SO	Sorted osteogenic
SOL	Soluble
SOX9	Sex determining region Y-box 9
SSC(s)	Skeletal stem cell(s)
ssRNA	Single stranded RNA
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TEAB	Triethylammonium bicarbonate
TGF-β	Transforming growth factor beta
TOF	Time of flight
TRAF6	TNF receptor-associated factor 6
TRIP6	Thyroid receptor-interacting protein 6
UB	Unsorted basal
UHPLC	Ultra-high pressure liquid chromatography
UO	Unsorted osteogenic
VEGF	Vascular endothelial growth factor
vs	Versus
WB	Western blotting
WCL	Whole cell lysate
α	Anti
α-MEM	Alpha-minimum essential medium

Publications

"MicroRNA-146a Regulates Human Foetal Femur Derived Skeletal Stem Cell Differentiation by Down-Regulating SMAD2 and SMAD3".
Cheung KSC, Sposito N, Stumpf PS, Wilson DI, Sanchez-Elsner T, et al. (2014).
PLoS ONE 9(6): e98063

Abstracts (non-published)

- Faculty of Medicine Annual Conference, University of Southampton, Southampton (2012).

- 4th International Satellite Symposium AICC–GISM, Brescia, Italy (2013).

……A mio padre Antonio, a mia madre Rosanna, alle mie sorelle Teresa e Viviana e a mia nonna Teresa……

1. Chapter 1: Introduction

1.1 Section I: Bone Structure and composition

1.1.1 The skeletal system

The skeletal system is composed of bone and cartilage, both of which are classified as connective tissues. Cartilage acts as a shock absorber, and provides a low friction surface to facilitate joint movement. Bone serves several important functions, including mechanical, metabolic and a haematopoietic function. The mechanical functions of bone include acting as a site for attachment of muscles and tendons, permitting locomotion, and affording protection to the organs. Bone acts as a reservoir for calcium and controls levels of calcium and phosphate, critical components of metabolism. Haematopoiesis, the production of erythrocytes, leukocytes and platelets also takes place within bone (Ramachandran and Bates, 2007).

Bone is composed of 10% cells and 90% matrix (Baron, 1993). The macroscopic and microscopic features are critical to the function of bone, and are discussed in the following section.

1.1.2 Bone macrostructure

Bone can be classified according to its macrostructure as: *long, short, flat and irregular*.

- Long bones include the clavicles, humeri, radii, ulnas, femora, and tibiae. Long bone structure consists of two extremities, the epiphysis and an elongated cylindrical shaped tube, the diaphysis or midshaft that connects the two epiphyses. The region of bone that connects epiphysis to diaphysis is termed the metaphysis (Figure 1). Developing bones have a cartilage interface (epiphyseal cartilage) between the metaphysis and the epiphysis. This epiphyseal cartilage consists of cartilage matrix and proliferative cells (chondrocytes), which are responsible for the longitudinal bone growth. The cartilage layer is constantly remodelled and mineralises at skeletal maturity.

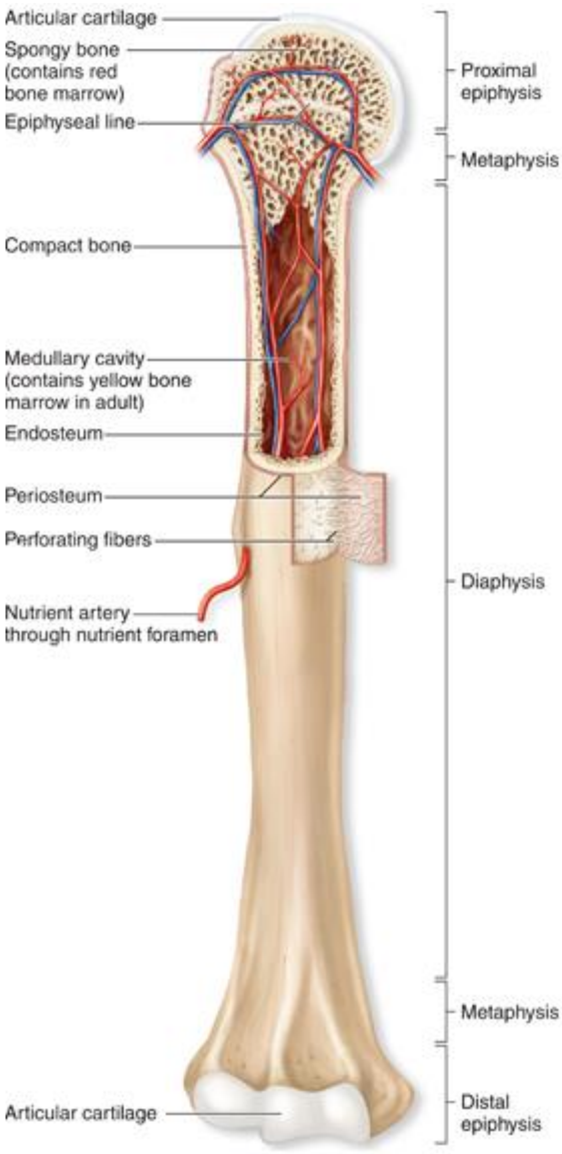

Articular cartilage
Spongy bone (contains red bone marrow)
Epiphyseal line
Compact bone
Medullary cavity (contains yellow bone marrow in adult)
Endosteum
Periosteum
Perforating fibers
Nutrient artery through nutrient foramen
Articular cartilage
Proximal epiphysis
Metaphysis
Diaphysis
Metaphysis
Distal epiphysis

Figure 1: The principal components of the femur.

The long bone is divided into three portions, the two epiphysis which are connected by the diaphysis of tubular shape. Long bone is covered with periosteum and articular cartilage at the extremities. The interior of the bone consists of spongy bone (trabecular bone). From: http://hexanatomy.weebly.com/skeletal-system.html (22/08/2014)

- *Short bones* include the sesamoids, carpal and tarsal bones.
- *Flat bones* include the skull, mandible, scapulae, sternum and ribs.
- *Irregular* bones include the vertebrae, sacrum, and coccyx bone.

The exterior surface of bone, cortical bone, is solid, compact, and dense while the interior is comprised of trabecular bone, characterised by an open, porous structure. Cortical bone is compact and rigid while trabecular bone is spongy or cancellous; their different morphology provides each structure with distinct biomechanical characteristics. The turnover rate of cancellous bone is eight times greater than the cortical bone due to the larger surface area of cancellous bone. The fully developed human skeleton consists of 80% cortical and 20% trabecular bone (Clarke, 2008). It is noteworthy that these percentages are not evenly distributed throughout all bones, and thus, the specific cortical to trabecular ratio varies. For example, vertebrae have a 25:75 ratio, femoral head 50:50, and 95:5 at the femoral diaphysis. Trabecular and cortical bone are formed by the same cell type, with varied organisation giving rise to structural and functional differences. Cortical bone is 80-90% mineralised while only the 15-25% of trabecular bone is mineralised. Variation in the level of mineralisation contributes to the different levels of mechanical strength. Cortical bone serves predominantly a mechanical and protective role while trabecular bone has a predominantly metabolic function.

The external surface of bone is covered by a fibrous layer, the periosteum; while the inner cortical surface is named the endosteum. From this nomenclature the layer of cells on these surfaces is defined respectively as *periosteal* and *endosteal* (Figure 2). The periosteum of the epiphysis is covered with a layer of cartilage that, as indicated above, facilitates movement and absorbs mechanical forces.

1.1.3 Bone microstructure organisation

While the bone macrostructure defines the mechanical functionality, the microstructure is critical to providing support, access to nutrients and growth. Different bone types contain different microstructures. Both cortical and trabecular bone present as a lamellar structure. Lamellae are formed by sheets of high density collagen fibers in alternating orientation. Lamellar structures are organised in concentric sheets either laid in parallel or in a concentric manner around a blood vessel (Haversian system). Haversian systems are only present in cortical bone. In cases of rapid bone turnover such as Paget's disease, (Delmas, 1999), fractures or tumours, the alternating collagen fiber pattern is absent. In woven bone, collagen is laid down in a random disoriented manner (Su, Sun *et al.*, 2003). Figure 2 shows a schematic of the internal structures of bone, including the Haversian canals.

`

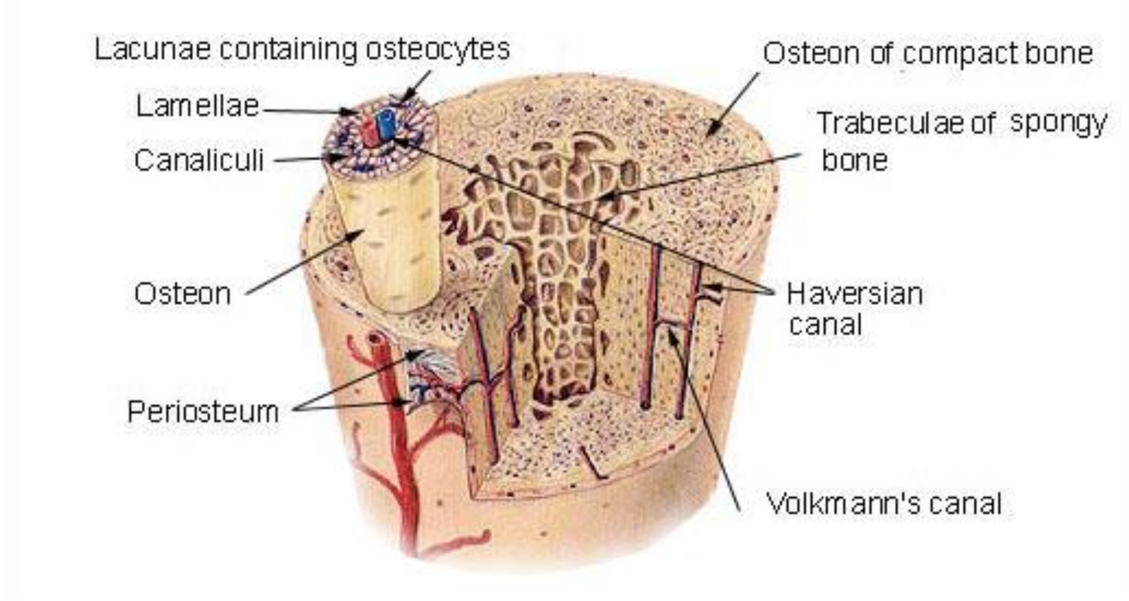

Lacunae containing osteocytes
Lamellae
Canaliculi
Osteon
Periosteum
Osteon of compact bone
Trabeculae of spongy bone
Haversian canal
Volkmann's canal

Figure 2: Bone internal structure.

The outer layer of the bone in contact with other tissues is called periosteum. From the periosteum going towards the centre of the bone, cortical bone comprised of parallel lamellae. After a group of parallel lamellae, five to seven concentrically distributed lamellae can be found that surround blood vessel in a complex called osteons or Haversian systems (50 μm). Between each lamellae bone cells are found (osteocytes). Towards the centre of the bone the cortical lamellar bone structure is replaced by a honeycomb-like structure termed trabecular bone. Bone marrow stroma is contained within this structure. From: http://training.seer.cancer.gov/index.html (03/12/2015).

Within the collagen fibers, hydroxyapatite crystals form. These crystals become fixed to the collagen fibers resulting in bone mineralisation. This process is facilitated by glycoproteins and proteoglycans. Between each lamellae reside osteocytes (trapped osteoblasts), present within the lacunae (small cavities found in bone tissue) within each lamella, as illustrated by the image shown in Figure 3.

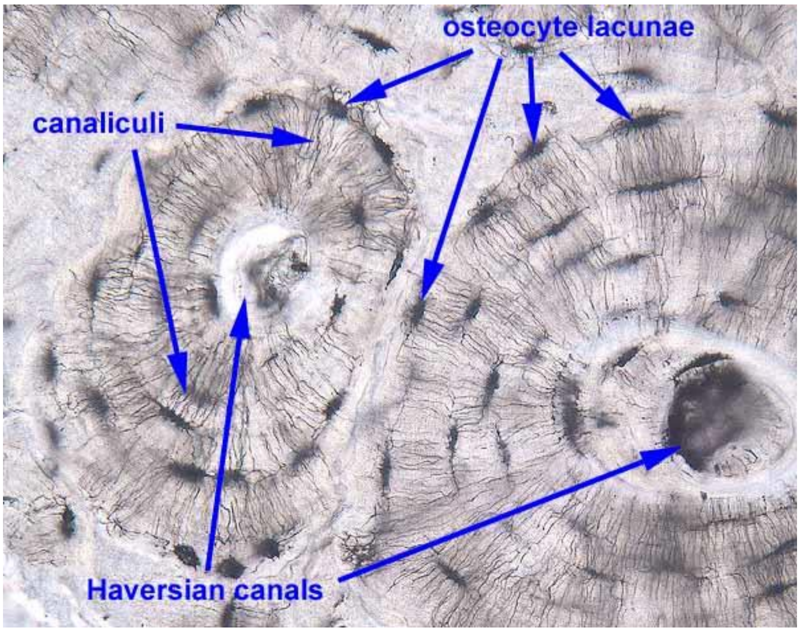


Figure 3: High magnification of cortical bone structure.

From the picture the location of osteocytes is apparent, Haversian canals and canaliculi. From: http://www.siumed.edu/~dking2/ssb/skeleton.htm (18/07/2014).

1.1.4 Bone remodelling: role of osteoblasts and osteoclasts.

Bone remodelling is a complex mechanism whereby old bone is replaced with new bone. Bone remodelling consists of two coupled reactions: bone deposition by osteoblasts and bone resorption by osteoclasts (**Error! eference source not found.**). During *resorption* osteoclasts create a cavity by resorbing bone until they receive a signal to stop, after which in a mechanism termed *reversal*, osteoblast precursors are attracted to the site vacated by osteoclasts, differentiate into osteoblasts and begin to deposit bone. With bone formation, the resorption pit is filled with new bone by osteoblasts. Ultimately, during *resting*, the osteoblast becomes encased within the bone (Coxon, Oades *et al.*, 2004). These activities are influenced by mechanical and physiological stimuli (Clarke, 2008). Both osteoblast and osteoclast activity is subject to quantitative regulation within time and space to generate the healthy bone. When the balance between bone formation and resorption is altered or lost, as in the case of aged individuals (Oreffo, Bord *et al.*, 1998), diseased bone status can occur (osteoporosis) (Rucci, 2008). The osteoclast bone resorption cycle takes between 2 to 4 weeks to complete. In contrast, bone deposition involving osteoblasts takes approximately 4 to 6 months (Sims and Martin, 2014).

The balance between bone formation and resorption is not uniformly distributed. Bone formation predominates at periosteal sites, while bone resorption is greater in endosteal and trabecular bone. This feature of the remodelling process is the main reason for age related trabecular bone thinning (Gabet and Bab, 2011). Osteoblasts and osteoclasts are derived from separate and distinct stem cell progenitors; in fact osteoblasts together with chondrocytes differentiate from mesenchymal progenitors *in situ*, while osteoclasts have a haematopoietic origin, and reach the bone tissue through blood vessels (Day, 2006). The properties of the important cells in bone remodelling, osteoblasts, osteocytes and osteoclasts, are discussed below.

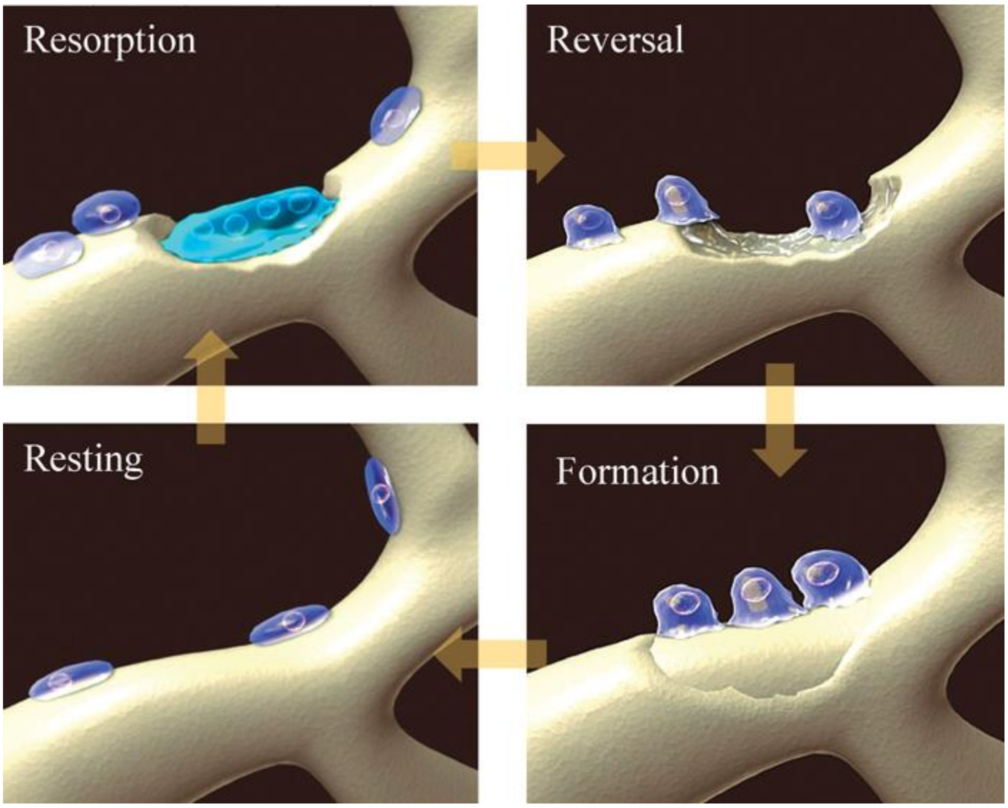


Figure 4: Bone remodelling process, consisting of a cycle of resorption, reversal, formation and resting.

Resorption: Osteoclast create an erosion cavity by resorbing bone until they receive signal to stop. *Reversal:* Osteoblasts precursors are attracted to the site vacated by osteoclasts, differentiate into osteoblasts and begin to lay down bone. *Formation:* Resorption pit is filled up with new bone by osteoblasts. *Resting:* Part of the osteoblast become bone lining cells by covering the new bone surface (Coxon, Oades *et al.*, 2004). Reprinted by permission of the nature publishing group.

1.1.4.1 Osteoblasts

Osteoblasts are bone-forming cells which produce bone matrix termed osteoid. Osteoid is composed predominantly of type I collagen. Osteoblasts are responsible for the mineralisation of osteoid. Osteoblasts have a high alkaline phosphatase activity and express a variety of proteins including bone morphogenetic proteins (BMPs), critical to osteogenesis (see section 1.1.6.2). Osteoblasts may become inactive bone-lining cells, with the central role to prevent osteoclast activity or become entrapped in matrix (osteocytes) or die by apoptosis. Bone lining cells are metabolically inactive but can undergo reactivation if bone formation is required.

1.1.4.2 Osteocytes

As stated earlier, osteocytes are osteoblasts entrapped within bone matrix, located within the lacunae at the interface between each lamella. Osteocytes play a critical role for calcium and phosphate metabolism (Figure 3) (Bonewald, 2011).

1.1.4.3 Osteoclasts

Osteoclasts are the cells responsible for bone resorption (see SEM image in Figure 5). Osteoclasts are large cells containing multiple nuclei (from 3 up to 20), which are located on the periosteal and endosteal surface, within cavity sites named Howship's lacunae. Resorption activity is driven by a complex system of lysosomal enzymes and carbonic acid to convert matrix and calcium minerals from an insoluble to a soluble form. Together with the osteoblasts, osteoclasts contribute to the maintenance of healthy bone (Crockett, Rogers *et al.*, 2011).

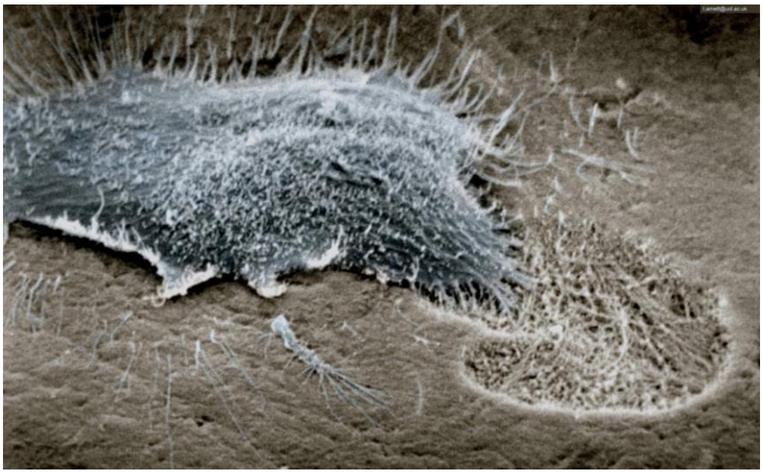


Figure 5: Rat osteoclast and resorption pit.

Scanning electron microscopy (SEM) image of a rat osteoclast in action: removing bone matrix and creating a resorption pit. From: http://www.ucl.ac.uk/cdb/research/arnett/gallery_bone1/Osteoclast___pit_-_Mar_2010.jpg?hires (18/08/2014).

1.1.5 Bone histogenesis: intramembranous and endochondral ossification

Bone histogenesis is initiated by undifferentiated progenitor cells. For flat bone formation the process is referred to as intramembranous ossification, while long bones are formed by endochondral ossification.

Intramembranous ossification occurs when undifferentiated mesenchymal stem cells differentiate into osteoblasts forming an ossification centre. A proportion of osteoblasts then deposit osteoid, which is subsequently mineralised and entraps other osteoblasts, resulting in an osteocyte population. Calcification is random at this stage and this irregular bone formation gives rise to woven bone. Ensuing blood vessels subsequently invade between the woven bone and marrow. Subsequent remodelling transforms woven bone into lamellar bone (Shapiro, 2008).

Endochondral ossification is the process by which all other types of bones are formed (Figure 6). Endochondral ossification begins during the second month of embryonic development when the embryonic cartilage is replaced by bone and, critically is the process through which long bones extend longitudinally until adult height is achieved. The process originates from undifferentiated mesenchymal stem cells, which differentiate into chondroblasts. The chondrocytes secrete cartilaginous matrix which is broken down prior to migration of osteoblasts to the midshaft (bone collar). The bone collar is then invaded by blood vessels, bringing osteoclasts, bone marrow (BM) and preosteoblasts to create the primary ossification centre.

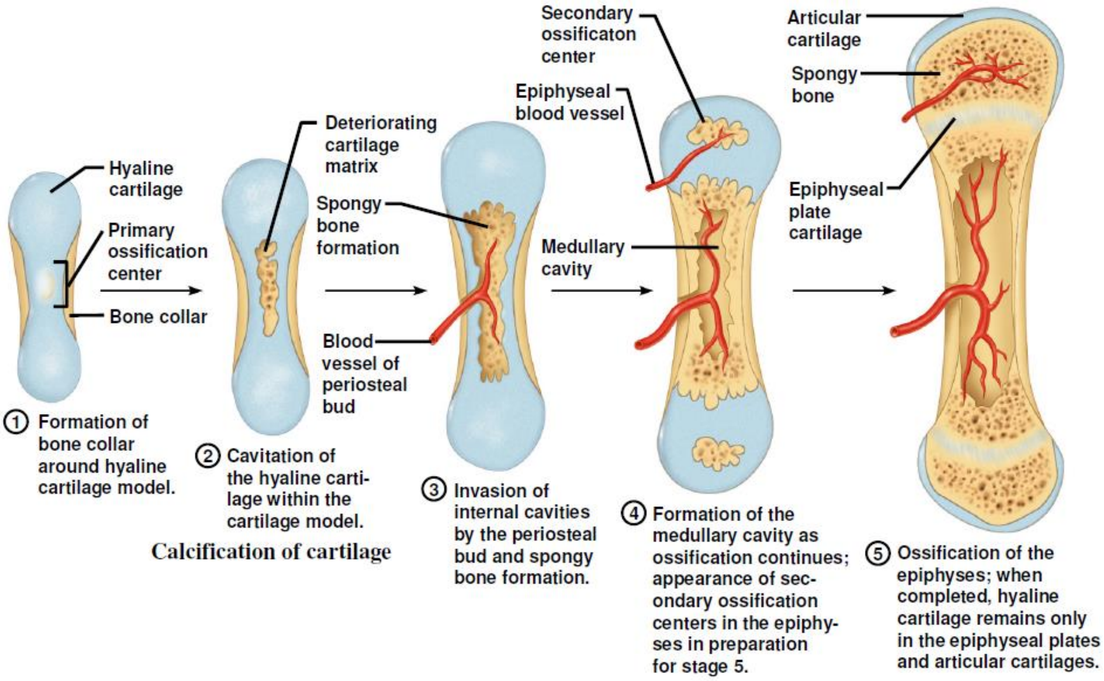

Hyaline cartilage
Primary ossification center
Bone collar
① Formation of bone collar around hyaline cartilage model.
Deteriorating cartilage matrix
Spongy bone formation
② Cavitation of the hyaline cartilage within the cartilage model.
Calcification of cartilage
Blood vessel of periosteal bud
③ Invasion of internal cavities by the periosteal bud and spongy bone formation.
Secondary ossificaton center
Epiphyseal blood vessel
Medullary cavity
④ Formation of the medullary cavity as ossification continues; appearance of secondary ossification centers in the epiphyses in preparation for stage 5.
Articular cartilage
Spongy bone
Epiphyseal plate cartilage
⑤ Ossification of the epiphyses; when completed, hyaline cartilage remains only in the epiphyseal plates and articular cartilages.

Figure 6: The stages of endochondral ossification.

The image depicts the *stadia* of ossification from formation to ossification. From: http://apbrwww5.apsu.edu/thompsonj/Anatomy%20&%20Physiology/2010/2010%20Exam%20Reviews/Exam%202%20Review/Ch%206%20Modes%20of%20Ossification.htm (18/7/2014).

From the primary ossification centre, osteoblasts migrate towards the epiphysis laying down more bone. Secondary ossification centres arise at the two epiphyses. Between the middle ossification centre and the epiphyseal ossification centres a layer of cartilage is left, the growth plate (physeal) that allows longitudinal growth and remodelling. Long bone cartilage is fully replaced by bone at the end of adult growth. The only cartilage remaining following skeletal maturity in long bones is articular cartilage. Endochondral ossification is highly regulated by circulating hormones, growth factors and extracellular matrix secreted proteins (Mackie, Tatarczuch *et al.*, 2011).

1.1.6 Control of osteogenesis

Osteogenesis, the process of bone formation, is a complex process regulated by growth factors which exert different effects at different intensities and time points. The main growth factors and signalling pathways are described in the following paragraphs.

1.1.6.1 Transcription factors

Runt related transcription factor 2 (RUNX2) is also known as core binding factor alpha1 (Cbfa1), PEBP2αA, AML-3 and osteoblast specific transcription factor 2 (Osf2). The gene that encodes this factor is the central regulatory gene within the osteogenic process. RUNX2 is believed to regulate all subsequent genes involved in bone formation. The importance of RUNX2 is demonstrated by the fact that RUNX2 in knock out or expression mutation studies, can have fatal outcomes, such as skeletal cancer development or skeletal abnormalities (Kirkham and Cartmell, 2007). RUNX2 is also the key regulator that directs mesenchymal stem cells to become skeletal cells. RUNX2 is pivotal for the commitment into the osteoblast lineage as inhibits adipocyte differentiation (Komori, 2005). A number of transcription factors and regulators interact with RUNX2. SMAD1 and SMAD6 for example enhance transcriptional osteogenic activity, while BMP2 in concert with RUNX2 and SMAD5 has been shown to induce osteoblast differentiation (Lee, Kim *et al.*, 2000).

Sex determining region Y-box 9 (SOX9) is an important transcription factor for early embryonic development and for osteoprogenitor stem cell differentiation fate. SOX9 is a transcriptional activator of chondrogenic genes, including type 2 collagen for cartilage formation. It has been shown that SOX9 activity can directly affect RUNX2. Experiments have shown that SOX9 acts as an antagonist of RUNX2 by diverting stem cell fate from osteoblasts into chondrocytes. Therefore, SOX9 represses RUNX2 and both are important for directing osteoprogenitor fate (Zhou, Zheng *et al.*, 2006).

Osterix (OSX) is a zinc-finger-containing transcription factor that promotes osteogenic differentiation and maturation of osteoblasts at the expense of chondrogenic differentiation. OSX is important for skeletal development, with OSX knockout mice demonstrating inadequate bone formation and mineralisation (Nakashima, Zhou *et al.*, 2002). The precise interaction of OSX with RUNX2 and other key transcription factors in the osteogenic process remains unclear. However; retroviral transduction of OSX in BM cells, while not altering RUNX2 expression, did appear to increase the expression of other osteogenic markers such as alkaline phosphatase, bone sialoprotein, osteocalcin and osteopontin (Tu, Valverde *et al.*, 2006).

1.1.6.2 Growth factors

The *Transforming growth factor beta (TGF-β) superfamily* includes the BMP family of growth factors. BMP plays an important role in the post-natal bone formation by inducing the formation of cartilage and bone. BMPs have been found to be involved in heart, neural and cartilage development (Chen, Zhao *et al.*, 2004). The function of BMPs was first postulated by Urist in 1965; subsequently BMPs 2A, 3 and 4 were identified (Wozney, Rosen *et al.*, 1988). Currently, over 20 isoforms of BMPs have been identified. Relevant BMPs for the treatment of bone diseases and osteoporosis include BMP2, which has been shown to have significant ability to induce bone formation by controlling alkaline phosphatase expression (Rawadi, Vayssiere *et al.*, 2003). Other key BMP in bone metabolism is BMP7. Currently BMP2 and BMP7 are licensed and used in clinical practice. BMP2 is used in spinal arthrodesis (Samartzis, Khanna *et al.*, 2005) and BMP7 in used in the treatment of tibial nonunions (fractures) (Axelrad and Einhorn, 2009). Recently there have been some concerns over the use of BMPs in clinical practise after perioperative complications in spinal fusion including radiculopathy (mechanical compression of the nerve) due to excess bone growth promoted by BMP (Owens, Glassman *et al.*, 2011). Although some studies demonstrate there is not proven risk in developing postoperative complications (Savage, Kelly *et al.*, 2015), others argue that the use of BMP in clinical practise does not come without potential adverse events.

Wnt signaling pathways are important for the developing foetus as each pathway contribute to skeletal formation and organ development. Different *Wnt* signaling molecules have been discovered and found to have multiple pathways. Mutations in *Wnt* signaling molecules have been implicated in amelia (absence of all four limbs) (Niemann, Zhao *et al.*, 2004). *Wnt* signaling drives osteogenesis at the expense of chondrocyte formation (Day, Guo *et al.*, 2005). *Wnt* action is strictly dependent on β-catenin protein inhibition by phosphorylation. β-catenin protein activation results in bone formation rather than cartilage (Figure 7). The effect of *Wnt* signaling on osteoblast formation is mediated by the activation of transcription factor *RUNX2* which stimulates osteoblastic differentiation (Gaur, Lengner *et al.*, 2005).

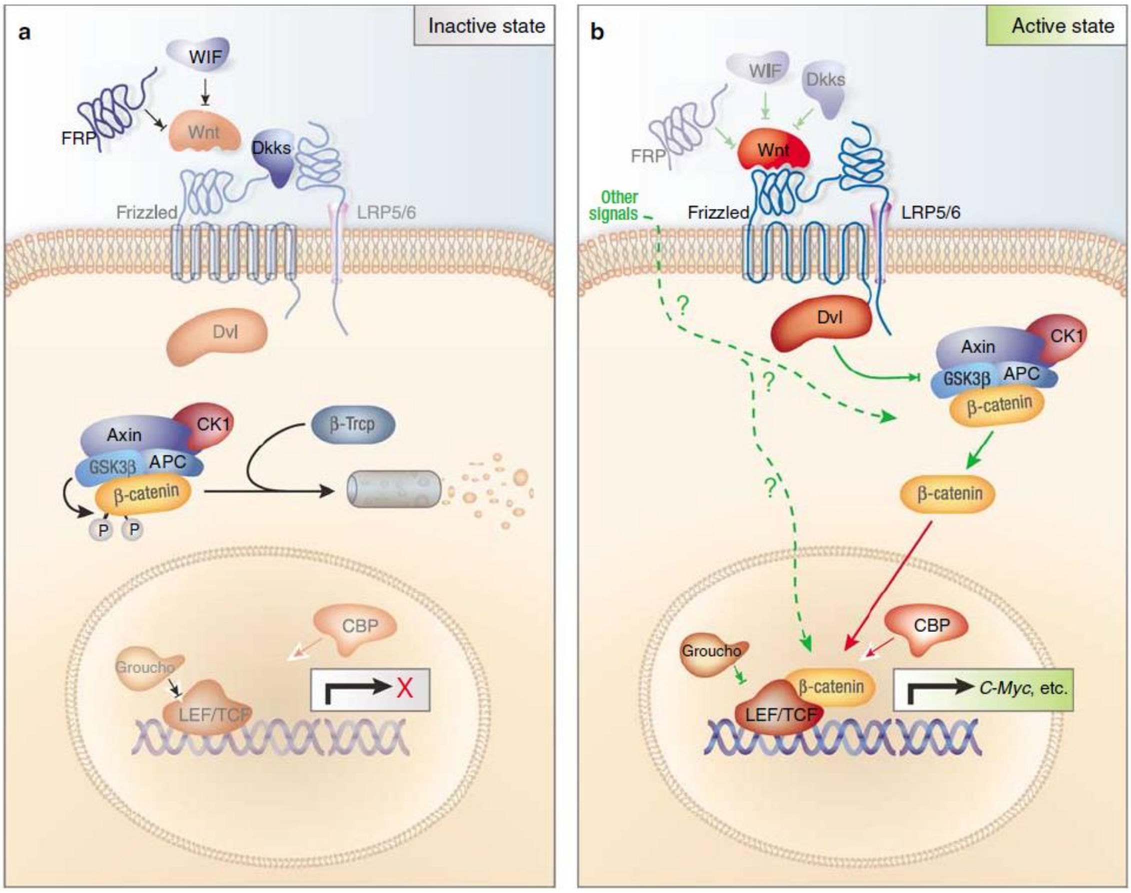

a
Inactive state
WIF
FRP
Wnt
Dkks
Frizzled
LRP5/6
Dvl
Axin
CK1
GSK3β
APC
β-catenin
P
P
β-Trcp
CBP
Groucho
LEF/TCF
X
b
Active state
WIF
Dkks
FRP
Wnt
Other signals
Frizzled
LRP5/6
?
Dvl
Axin
CK1
GSK3β
APC
β-catenin
?
?
β-catenin
CBP
Groucho
β-catenin
LEF/TCF
C-Myc, etc.

Figure 7: Canonical Wnt/β-catenin signaling pathway.

a) When Wnt is inactive β-catenin is recruited into the APC/Axin/GSK3b complex, and phosphorylated. The phosphorylated β-catenin is degraded upon binding to β-Trcp. No β-catenin enters the nucleus. (b) When Wnt is in its active form, binds to its frizzled receptors and LRP5/6 co-receptor and Dvl get activated. β-catenin phosphorylation is inhibited and β-catenin can enter the nucleus and form a transcriptional complex with LEF/Tcf and activates downstream targets such as c-Myc (Luo, Chen *et al.*, 2007). Reprinted by permission of the nature publishing group.

The *Fibroblast growth factor (FGF)* family is responsible for both osteogenesis and chondrogenesis (Mansukhani, Bellosta *et al.*, 2000). Malfunctioning of FGF receptors 1 and 2 has been attributed to cranial abnormalities (Mansukhani, Bellosta *et al.*, 2000); while FGF receptor 3 has been found to limit osteogenesis by negative regulation of the endochondral bone formation (Deng, WynshawBoris *et al.*, 1996). FGF receptors appear to have a variety of effects on osteoblasts, including induction of apoptosis, and stimulation of differentiation and migration (Mansukhani, Bellosta *et al.*, 2000).

Platelet derived growth factor (PDGF) is a mitogen stimulating proliferation and migration of mesenchymal stem cells (MSCs) (Yarden, Escobedo *et al.*, 1986). PDGF is involved in cytoskeletal arrangement, actin reorganisation and plasma membrane raffling (Hammacher, Mellström *et al.*, 1989). Under the vascular endothelial growth factor –A (VEGF-A) stimulation, PDGF promotes recruitment of MSCs at the site of vascularisation, therefore contributing to angiogenesis (Ball, Shuttleworth *et al.*, 2007).

Vascular endothelial growth factor (VEGF) is pivotal in angiogenesis (Ferrara, Carver-Moore *et al.*, 1996). It has been shown that in VEGF knockout mice, the invasion of physeal cartilage by blood vessels during endochondral ossification was completely inhibited (Ferrara, Carver-Moore *et al.*, 1996). Ensuing activation of VEGF allowed vascularisation of the developing bone during endochondral ossification, resulting in bone growth, substitution of cartilage with bone and restoration of the growth plate architecture (Gerber, Vu *et al.*, 1999).

Insulin like growth factors (IGF) mediates osteoblast proliferation through interaction with RUNX2 (Celil and Campbell, 2005). The IGF-1 mechanism of enhancing MSC proliferation has been investigated and was found to increase the MSC migratory response by enhancing expression levels of the chemokine receptor CXCR4 which is PI3/Akt dependent (Li, Yu *et al.*, 2007)

Epidermal growth factor (EGF) promotes migration and proliferation of MSCs (Tamama, Fan *et al.*, 2006). EGF has been shown to enhance osteogenic differentiation of adipose tissues through interaction with BMP2, BMP6 and BMP receptors (Elabd, Chiellini *et al.*, 2007).

1.1.6.3 Proteins

Alkaline phosphatase (ALP) is an enzyme. It is found in bone and in other tissues such as the liver. In tissue engineering, ALP is widely used as an indicator of osteogenic activity. ALP is a marker of early osteogenesis and is the first gene expressed in the mineralisation process (Golub and Boesze-Battaglia, 2007).

Osteopontin (OSP) is a phosphoprotein found in bone and body fluids such as urine, bile, plasma and milk. Although not bone specific, OSP plays an important role in bone remodelling. OSP knockout mice develop similar to wild types with no apparent difference in skeletal development, although osteoclast formation has been found to be elevated in OSP knockout mice (Rittling, Matsumoto *et al.*, 1998). Contrary to these findings a study has suggested that OSP promotes osteoclast differentiation and activity (Standal, Borset *et al.*, 2004).

Osteocalcin (OCN) is the second most abundant protein in bone after collagen, it is specific to bone and is a late osteogenesis marker. OCN is pivotal in osteoblast differentiation, and recently a study has implicated OCN in vascular calcification by regulating osteochondrogenic differentiation of diseased vascular smooth muscle cells (Kapustin and Shanahan, 2011).

Bone sialoprotein (BSP) is postulated to be able to nucleate hydroxyapatite crystals, and therefore regarded as an early osteogenic marker involved in the early mineralisation of osteoblasts. It has been demonstrated to mediate fibroblast attachment on tissue culture dishes and to be involved in cell signalling and collagen binding (Ganss, Kim *et al.*, 1999).

1.2 Section II: Introduction to marrow stromal and skeletal stem cells

1.2.1 Overview

This section firstly discusses the importance of skeletal stem cell (SSC)/mesenchymal stem cell (MSC) research from the point of regenerative medicine. Secondly, an overview of the discovery of MSCs, MSC *in vivo* identity and therapeutic applications is given. Thirdly, the study of MSCs through the use of a proteomic-based approach is introduced. In this work the nomenclature SSCs, unless stated otherwise, will be used to refer BM stem cells that differentiate into bone, cartilage, and fat. These are often (erroneously) referred to in the literature as MSCs. The rationale for the two different terminologies will be discussed.

1.2.2 History

SSCs are a rare adult stem cell population that are plastic adherent, and found mainly in the non-haematopoietic fraction of BM. These cells have the capacity to differentiate into various types of mesenchymal tissue. The presence of non-haematopoietic stem cells in BM was first hypothesised 130 years ago by a German pathologist 'Cohnheim', cited by Prockop, whilst observing wound healing (Prockop, 1997). Much later, the first extensive works on SSCs were executed by Friedenstein and colleagues in 1966 (Friedenstein, Petrakova *et al.*, 1968; Friedenstein, Piatetzky-Shapiro *et al.*, 1966; Owen and Friedenstein, 1988). Friedenstein and colleagues investigated the isolation of SSCs based on their fibroblast-like colony forming (CFU-F) ability (Friedenstein, Chailakhjan *et al.*, 1970) and their plastic adherence (Luria, Panasyuk *et al.*, 1971). Friedenstein noticed that a subset of cells from BM were plastic adherent and able to form colonies (agglomeration of at least 50 cells (Krampera, Franchini *et al.*, 2007)) from a single cell. Friedenstein also observed bone formation originating from a subset of cells which were plastic adherent and of non-haematopoietic origin following transplantation of whole mouse BM within a diffusion chamber (Friedenstein, Piatetzky-Shapiro *et al.*, 1966). Other groups investigating longer periods *in vitro* observed that these cells were able to differentiate into multiple lineages including osteoblasts, adipocytes and chondrocytes (Pittenger, Mackay *et al.*, 1999). These cells were popularised as Mesenchymal Stem Cells (MSCs) by Caplan (Caplan, 1991). As a consequence of this work, interest in MSCs has grown rapidly as manifested by an increase in publications pertaining to MSCs, particularly in the last 10 years (Figure 8). This tremendous growth in studies of MSCs reflects the belief that MSCs are vital to the clinical translation of regenerative medicine therapies.

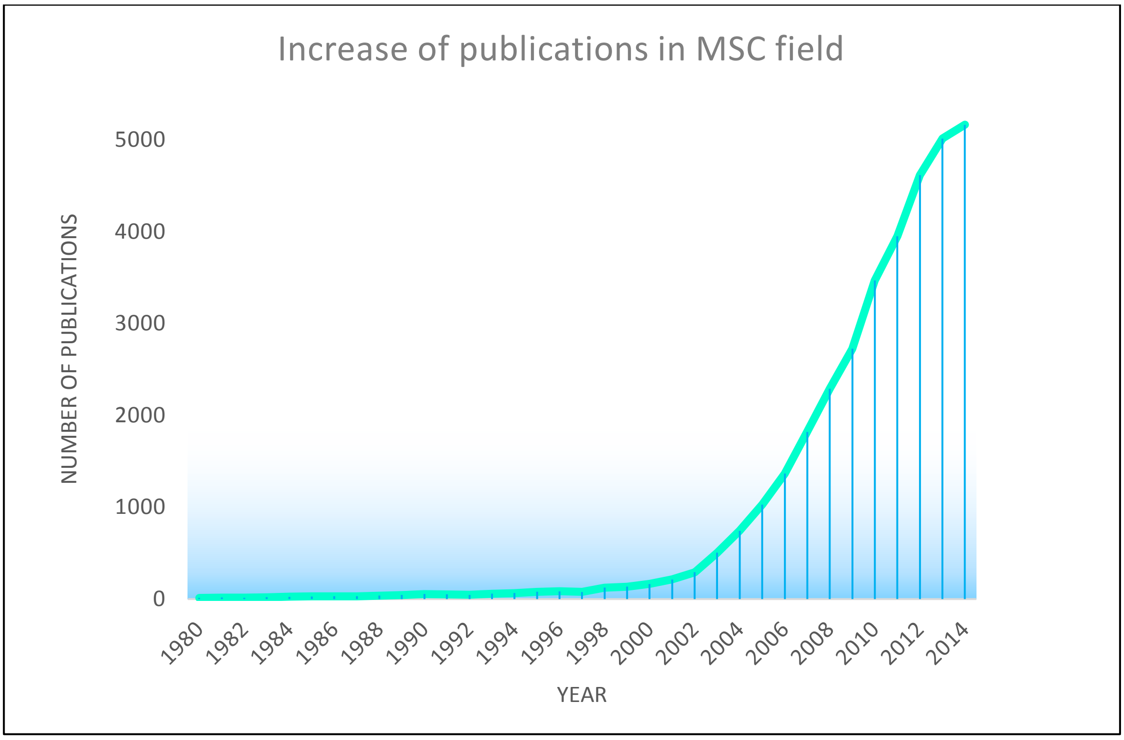


Figure 8: Scientific Publication on MSCs from 1980 to 2014 in PubMed.

The graph shows the increasing interest in MSC research in the last 34 years. Search criteria: “mesenchymal stem cell” was input into PubMed and total results arranged by year into an excel spreadsheet.

1.2.3 Stem cells for regenerative medicine and tissue engineering

The world is faced with an increasing healthcare burden, mainly due to an ageing population (Cracknell, 2010). Average life expectancy has increased, however with longer life comes the increased risk of tissue and organ failure (Kassem, 2006). To remedy this potential health catastrophe significant attention has been given to regenerative medicine. Regenerative medicine is a branch of medicine which is devoted to “replacing or regenerating human cells, tissues or organs to restore or establish normal function” (Mason and Dunnil, 2008). Regenerative medicine aims to treat patients presenting diseased, damaged or degenerated tissues due to accident, illness or natural causes. The potential of regenerative medicine to repair and regenerate tissues with avant-garde techniques may only be achieved with the use of stem cells, which have exquisite and unusual characteristics. The concept of the existence of stem cells was introduced in the middle of the 20th century by Leroy Stevens while observing embryonic teratoma cells (http://www.the-scientist.com/?articles.view/articleNo/12717/title/A-Stem-Cell-Legacy--Leroy-Stevens). Stem cells are distinctive from other cells, being capable of self-renewal and of multilineage differentiation (Kolf, Cho *et al.*, 2007). According to the differentiation potential, stem cells can be totipotent, pluripotent, multipotent (Krampera, Franchini *et al.*, 2007) or unipotent. Totipotent stem cells are early embryonic stem cells which can give rise to all lineages. Pluripotent stem cells are able to give rise to the three embryonic germ layers: endoderm (cells forming the gastrointestinal tract and internal organs such as the liver), mesoderm (blood, blood vessels, bone and connective tissue) and ectoderm (skin and nervous system). Multipotent stem cells can give rise to only one embryonic germ layer (Sell, 2004). Unipotent stem cells can give rise to only one lineage (Visvader and Lindeman, 2011). One type of multipotent stem cells considered to be the future for bone tissue engineering are SSCs, referred to as MSCs by Caplan (Caplan, 2007). SSCs have the capacity for prolonged or unlimited self-renewal under specific conditions, and have the potential to differentiate into multiple lineages such as bone, cartilage and fat (Figure 9). The self-renewal ability of stem cells enables their expansion *in vitro* for a target tissue, a key feature for generating sufficient cells to meet the demand for tissue replacement (Cohen and Chen, 2008). One important feature of a stem cell is stemness (self-renewal and multipotency), which at a molecular level indicates an absence of fatal telomere shortening. Researchers have proposed high telomerase activity as a prerequisite of any stem cell and therefore of SSCs (Weissman, 2000) as telomere shortening leads to eventual cell death; not an attribute of a true stem cell. On this note, some scientists (Simonsen, Rosada *et al.*, 2002) have shown how SSCs (SSCs) transduced with a retroviral vector containing the gene telomerase reverse transcriptase SSCs-TERT had increased telomerase length, which led the cells to undergo up to 160 population doublings. These cells maintained the production of osteoblasts and formed more bone than control cells following subcutaneous transplantation within immunodeficient mice. Interestingly, no tumour growth was observed in these mice. These findings support the association of cellular senescence with telomere shortening (Simonsen, Rosada *et al.*, 2002).

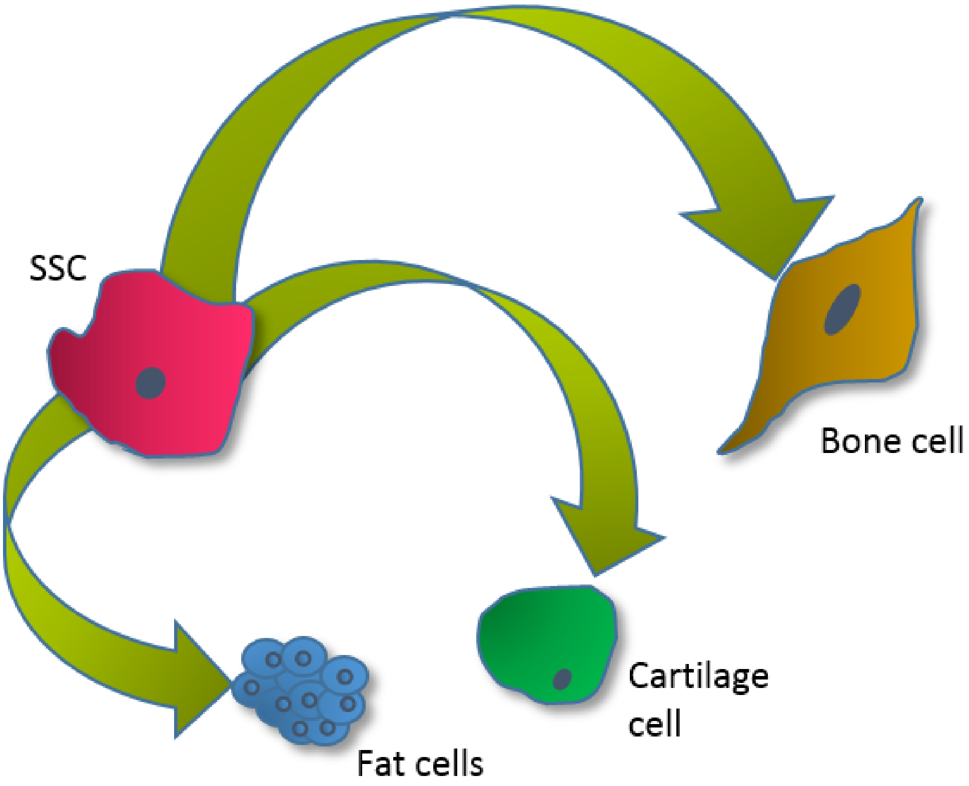


Figure 9: SSCs differentiation potential.

SSCs can differentiate into multiple lineages: bone, cartilage and fat.

SSCs are considered to display “homing capacity” and to direct other cells to specific tissues (Sohni and Verfaillie, 2013). The mechanisms by which SSCs migrate to some tissues, particularly following injury or under pathological conditions, have been studied but remains poorly understood. However, it has been shown that chemokines, chemokine receptors and adhesion molecules are involved (Chamberlain, Fox *et al.*, 2007).

A limitation in the use of SSCs for regenerative medicine remains an absence of protocols to obtain homogeneous populations (Lu, Neff *et al.*, 2011). Heterogeneity can be observed under a light microscope depicted by variation in colony size, growth rate and cell morphology (Bianco, Riminucci *et al.*, 2001). It is assumed that a homogeneous cell phenotype will strictly correlate to a specific function (Di Girolamo, Stokes *et al.*, 1999). In contrast, unpredictable functions can arise from a heterogeneous cell phenotype. Therefore it is hypothesised that heterogeneity impacts on the outcome of the experiment and there is no current methodology to predict and standardise such results. From a clinical point of view it is important to have absolute control over the starting material, *i.e.* cells. It is believed that more homogeneous cell populations will yield more consistent and reproducible clinical outcomes. To date, significant heterogeneity exists within the same SSC population (intra-population heterogeneity) and between donors (inter-donor heterogeneity) (Phinney, 2012). Based on these considerations researchers have raised several questions while studying SSC heterogeneity; specifically the unique epitope profile of SSCs, how to study these cells, and whether it is possible to predict a specific function (Phinney, 2012). Studies published to date suggest that in spite of a similar phenotype, function can vary (Kuznetsov, Krebsbach *et al.*, 1997). However, identification inaccuracy suggests that the true SSC phenotype and its ability to predict function remain unknown.

The quest for SSC identification and proteomic profile characterisation remains a key topic in the field of regenerative medicine and will be discussed in this work.

1.2.4 Stem cell niche

The concept of the stem cell niche was introduced by Schofield in 1978 (cited in (Bianco, 2011)) when describing haematopoietic stem cells (HSCs). Schofield stated that stem cells are seen in association with other cells that can influence their behaviour; Schofield gave the name “niche” to the micro-environment these cells occupy *in vivo* . The modern concept of niche was developed from *Drosophila* studies and now it is identified as a microstructure of cells that work in concert to maintain stem cell activity and aid tissue homeostasis. There is a complex signalling network between stem cells and their niche which safeguards the high production of stem cells avoiding tumourigenesis. Ultimately it has been stated that the niche controls the balance between cell quiescence (state of inactivity) and activity (Moore and Lemischka, 2006). Recently much attention has been given to this context after it was hypothesised that the niche itself can influence stem cell function, and potentially induce a stem cell state, via a complex signalling network which allows the exchange of information between the cells. Therefore niche manipulation may enable manipulation of niche resident stem cells. In the BM, the stem cell niche is predominantly composed of haematopoietic cells, although the identity of the cells forming the niche was unclear until it was suggested that SSCs expressing nestin are actually responsible for maintaining the haematopoietic environment (Mendez-Ferrer, Michurina *et al.*, 2010). In 2007 seminal studies from Bianco's group showed that after transplantation of a single human BM cell, there is the possibility to support the haematopoietic environment (Bianco, Kuznetsov *et al.*, 2006). CD146$^+$ cells were used and it was proven that both before and after transplantation, the newly formed cell population *in situ* expressed CD146. This study demonstrated the coexistence of SSC and HSCs in the BM (Sacchetti, Funari *et al.*, 2007).

1.2.5 Mesenchymal stem cells, skeletal stem cells and bone marrow stromal cells

Recent years have seen a significant debate on the correct nomenclature and description of MSCs. MSCs were first termed "precursors of mechanocytes" where mechanocytes indicated connective tissue cells, and the name was indicative of mechanical properties (Friedenstein, 1976 cited in (Bianco, Kuznetsov *et al.*, 2006)). Later the term "stromal cells" was coined and adopted to describe marrow-derived osteogenic precursors (Owen and Friedenstein, 1988). Subsequently, the name "mesenchymal" was suggested (Caplan, 1991), to encompass a set of cells considered pluripotent progenitor cells at the embryonic level, that after many divisions commit to specific lineages such as cartilage, bone, tendon, ligament, marrow stroma, and connective tissues. Caplan explained that the name mesenchymal is derived from the Greek and signifies 'middle infusion', referring to the ability of mesenchymal cells to spread during embryonic development to other germ layers such as ectoderm and endoderm. Although these cells were described by Caplan as stem cells, later other investigators highlighted the existing uncertainties about their claimed multipotency and self-renewal coupled with a growing tendency to doubt the very use of the term "mesenchymal stem cells" for unselected BM cells (Bianco, Robey *et al.*, 2008). Self-renewal ability indicates the reconstitution *in vivo* of the stem cell environment with identical phenotype and properties to the starting cell population; while multipotency describes the ability of a single cell to differentiate into other cell types. While the latter affirmation has been proven, the former remains to be shown. Furthermore, Caplan postulated that the same MSC progenitors were able to differentiate into skeletal and non-skeletal lineages, however the non-skeletal ability remains unproven, moreover it has been shown that bone and skeletal muscle originate from different progenitors (Majka, Jackson *et al.*, 2003). Thus, only a subset of cells from the BM that are plastic adherent are able to form colonies from a single cell, and display clonogenic potential *in vitro*; nonetheless *in vivo* assays demonstrating the ability of the progeny derived from a single clonogenic cell to give rise to skeletal tissues are required (Bianco, Robey *et al.*, 2008). It has been shown that CD146$^+$ clonogenic cells have the ability to self-renew *in vivo* but few other studies have shown this at a larger scale or using other markers (Sacchetti, Funari *et al.*, 2007). A number of studies suggest that MSCs can give rise to only skeletal cell types such as osteoblasts, chondrocytes, adipocytes, fibroblasts, and adventitial reticular cells; and less rigorous studies suggest the ability to differentiate into tissues of mesodermal origin (skeletal muscle, smooth muscle, cardiac muscle, endothelial cells) (Bianco, Riminucci *et al.*, 2001). However it is common-place in the literature to claim that MSCs can also give rise to cells of mesodermal origin and not just skeletal tissues. It is for this reason that Bianco and colleagues have felt the need to be more specific and use the name “skeletal stem cell” for the BM multipotent stromal cells capable of generating skeletal cell types *in vivo* (Bianco, Cao *et al.*, 2013; Bianco, Robey *et al.*, 2008). A key problem remains that the term MSCs is a recognised and used worldwide acronym, thus, proposing to change it for another term that more closely resembles the biology of bone marrow stem cells, seems futile.

Based on this argument, the nomenclature Skeletal Stem Cells (SSCs) will be used in this thesis as the most appropriate term to refer to this clonogenic, plastic adherent cell population from the human bone marrow stromal fraction, which can give rise to skeletal lineages, unless stated otherwise.

1.2.6 SSC markers

As stated previously there are no specific markers for SSCs. Deciphering the unique proteomic profile of SSCs is hindered by cell heterogeneity. This problem has spurred research into surface markers to aid purification of SSCs from heterogeneous populations. In spite of the vast number of combinations of markers exploited thus far for the identification and characterisation of

SSCs, different studies have adopted different techniques and therefore it is difficult to fully compare studies and determine which markers are optimal for SSCs isolation. In addition it is important to consider that the ensemble of surface markers is strictly modulated by cell culture conditions (Bianco, Robey *et al.*, 2008). Some of the most common markers used across different studies as positive and negative markers for SSC isolation are discussed (Table 1 and Table 2).

Table 1: Positive SSC selection markers.

CD	Full name	References
9	Tetraspan	(Maurer, 2011)
10		(Harichandan and Buhring, 2011); (Lecourt, Marolleau et al., 2010)
13		(Harichandan and Buhring, 2011)
15		(Lecourt, Marolleau et al., 2010); (Maurer, 2011)
29		(Lecourt, Marolleau et al., 2010);(Baddoo, Hill et al., 2003)
44		(Maurer, 2011); (Lecourt, Marolleau et al., 2010); (Baddoo, Hill et al., 2003; Phinney and Prockop, 2007)
47		(Maurer, 2011); (Lecourt, Marolleau et al., 2010); (Phinney and Prockop, 2007)
49		(Lecourt, Marolleau et al., 2010)
51		(Lecourt, Marolleau et al., 2010); (Phinney and Prockop, 2007)
56		(Phinney and Prockop, 2007)
62		(Battula, Treml et al., 2009)
63	Hop26	(Lecourt, Marolleau et al., 2010)
73	Membrane-bound ecto-5'-nucleotisidase	(Dominici, Le Blanc et al., 2006); (Maurer, 2011);(Harichandan and Buhring, 2011); (Lecourt, Marolleau et al., 2010); (Phinney and Prockop, 2007); (Chamberlain, Fox et al., 2007); (Docheva, Haasters et al., 2008)
79		(Troyer and Weiss, 2008)
81		(Baddoo, Hill et al., 2003)
90	Thy-1	(Dominici, Le Blanc et al., 2006); (Maurer, 2011); (Lecourt, Marolleau et al., 2010)
105	Endoglin	(Dominici, Le Blanc et al., 2006); (Maurer, 2011); (Harichandan and Buhring, 2011); (Lecourt, Marolleau et al., 2010); (Barry, Boynton et al., 1999; Chamberlain, Fox et al., 2007; Phinney and Prockop, 2007); (Baddoo, Hill et al., 2003)

106	VCAM	(Harichandan and Buhring, 2011; Phinney and Prockop, 2007)
146	MCAM	(Koyama, Okubo et al., 2011; Lecourt, Marolleau et al., 2010)
166	ALCAM	(Maurer, 2011; Phinney and Prockop, 2007)
271	low affinity nerve growth factor receptor LNGFR	(Battula, Treml et al., 2009; Phinney and Prockop, 2007)
273		(Bárcia, Santos et al., 2015)
	Stro-1	(Simmons and Torok-Storb, 1991)
	SSEA-4	(Gang, Bosnakovski et al., 2007)

Table 2: Negative SSC selection markers.

CD	Full name	References
11b	Integrin α M (ITGAM)	(Phinney and Prockop, 2007; Troyer and Weiss, 2008)
14		(Chamberlain, Fox et al., 2007; Dominici, Le Blanc et al., 2006; Phinney and Prockop, 2007; Troyer and Weiss, 2008)
19	B-lymphocyte antigen	(Dominici, Le Blanc et al., 2006)
31	PECAM	(Harichandan and Buhring, 2011)
34		(Battula, Treml et al., 2009; Chamberlain, Fox et al., 2007; Maurer, 2011)
45	Tyrosine phosphatase	(Battula, Treml et al., 2009; Chamberlain, Fox et al., 2007; Harichandan and Buhring, 2011; Maurer, 2011; Phinney and Prockop, 2007)
79a	Immunoglobulin-associated α	(Dominici, Le Blanc et al., 2006)
80		(Harichandan and Buhring, 2011)
117		(Maurer, 2011)
133		(Battula, Treml et al., 2009)
HLA class1		(Maurer, 2011)
HLA-DR		(Maurer, 2011)

Stro-1 is one of the accepted SSC markers, used worldwide (Kolf, Cho *et al.*, 2007; Ning, Lin *et al.*, 2011) as a potent isolation tool for the BM non-haematopoietic fraction of cells that is capable of osteogenic differentiation (Stewart, Walsh *et al.*, 1999). Studies demonstrated Stro-1 to be an early osteogenic precursor being detected in the outer cell walls of BM. As a recognition tool for SSC, Stro-1 has also been employed to detect dental pulp stem cells (DPSCs) (Shi and Gronthos, 2003).

The antibody α-Stro-1, that detects the so like-called Stro-1 antigen, was created in 1991 (Simmons and Torok-Storb, 1991) by fusing murine myeloma cells NS1-Ag4-1 (BM cancer cells) and BALB/c spleen of an animal immunised with a population of BM cells positive for the CD34 marker (CD34$^+$). Stro-1 is of an IgM type and later studies show that the Stro-1 positive (Stro-1$^+$) fraction of cells within human BM contains the cells including the osteogenic precursors (Gronthos, Graves *et al.*, 1994) as this subpopulation has the capacity to reform *in vitro* the haematopoietic environment and is capable of differentiation into chondrocytes, adipocytes and osteoblast (Psaltis, Paton *et al.*, 2010). Studies show that the majority of the CFU-F cells are contained within the Stro-1$^+$ cell fraction. However, there remains controversy regarding the use of Stro-1 for the isolation of the SSCs as it is not specific to bone. Stro-1$^+$ cells have also been found in the endothelial tissue of blood vessels, muscles and in a population of CD34 cells (marker of endothelial and haematopoietic cells), casting doubt on the use of Stro-1 as a specific marker of SSCs (Lin, Liu *et al.*, 2011). Furthermore, Stro-1 can also bind to glycophorin-A-positive nucleated erythrocytes and to a subset of B-lymphocytes (Gronthos, Zannettino *et al.*, 2003). Since the discovery of Stro-1, the identity of the cell surface antigen to which it reacts remains an enigma. Few research groups have attempted a Western Blotting (WB) approach with the Stro-1 antibody and initial results identified a detected band within the 27-18 kDa range and a minor band at 50 kDa (Castrechini, Murthi *et al.*, 2010). Ning and co-workers in 2011 published an article stating that Stro-1 was an endothelial antigen of 75 kDa, by showing a WB band on the gel at that correspondent molecular weight (Ning, Lin *et al.*, 2011). Since these two investigations no further studies have corroborated these findings.

In the absence of specific SSC markers, each research group adopts a marker or a specific marker set they favour. However despite multiple efforts to find a unique marker, it remains elusive.

Due to the variable results for SSC marker studies, in 2006 the need to establish criteria to denominate a cell as a “SSC” led to the formation of the Mesenchymal and Tissue Stem Cell Committee of the International Society for

Cellular Therapy (ISCT) (Dominici, Le Blanc *et al.*, 2006). The committee agreed on ‘minimal criteria to define human SSCs’. These include: (i) plastic adherence under standard culture conditions; (ii) expression of CD105 (endoglin), CD73 (membrane-bound ecto-5'-nucleotisidase) and CD90 (Thy-1); (iii) lack of expression of the haematopoietic markers CD45 (tyrosine phosphatase), CD34, CD14 or CD11b (Integrin alpha M (ITGAM)), CD79a (immunoglobulin-associated alpha) or CD19 (B-lymphocyte antigen) and HLA-DR surface molecules. Finally the cells must differentiate into osteoblasts, adipocytes and chondroblasts *in vitro* (Dominici, Le Blanc *et al.*, 2006). These criteria are just an enforcement of what has been found to be the recurrent phenotype of MSCs in multiple studies; and it should be noted that in spite of the application of these criteria, the BM stem cell population remains heterogeneous. In fact, a controversial point of view is held by those who consider these criteria not stringent enough, meaning that it is virtually always possible to culture MSCs from every type of tissue, as these criteria are just shared characteristics among cells deriving from any connective tissue (Bianco, Cao *et al.*, 2013).

Moreover, it should be considered that the set of markers expressed in culture not only change according to culture conditions, but may also differ from the set of markers expressed *in vivo* . Further variables to account for are differences between patients. These issues limit the ability to harness the potential of BM cells for clinical therapies.

It is apparent that there is a great need to specifically characterise SSCs. This aim constitutes the basis of the work presented in this thesis. Since transcriptome and genomic analysis have yet to elucidate the specific function or the specific surface marker set of SSCs, proteomic approaches have emerged as the most promising method to address these issues (Maurer, 2011; Mindaye, Ra *et al.*, 2013; Park, Shin *et al.*, 2007; Roubelakis, Pappa *et al.*, 2007). These approaches will be discussed further in Section 1.3.4.

1.2.7 *In vivo* identity of SSCs and MSCs

Conflicting views remain on the location of MSCs *in vivo* . The general concept of MSCs identified by minimal criteria (CFU-F positive for CD105, CD37, CD90, and able to differentiate into osteoblast, adipocytes and chondrocytes) is accepted and it is thought that MSCs can be extracted from a variety of tissues.

Based on the loose criteria as a consequence of a lack of specific markers, MSCs have been extracted from a variety of locations, such as intestinal epithelial niches called crypts in the mucosal surface of the small intestine and colon (Day, 2006), synovial fluid (Sekiya, Ojima *et al.*, 2012), from the ocular surface (Dua and Azuara-Blanco, 2000), deciduous teeth (Barry and Murphy, 2004), dental pulp (Shi and Gronthos, 2003), in the trachea (Popova, Bozyk *et al.*, 2010), umbilical cord blood (Lee, Kuo *et al.*, 2004), amniotic fluid (Roubelakis, Pappa *et al.*, 2007), and most other tissues within the human body (Docheva, Haasters *et al.*, 2008) (Da Silva Meirelles, Chagastelles *et al.*, 2006). Observations in the literature also suggests that Whartons jelly-derived cells (Wharton's jelly is a cushioning matrix within the umbilical cord (Walker, 2008)) are primitive forms of MSCs (Troyer and Weiss, 2008) as they present a mesenchymal morphology, display mesenchymal cell surface markers and produce cytokines similar to MSCs (Walker, 2011). MSCs have also been found in the stromal fraction of lipo-aspirates of adipose tissue, and have been shown to exhibit similar plasticity to BM-derived MSCs (Lee and Kemp, 2006). Given the variety of tissues in which MSCs can be found, tissue localisation cannot be used as a tool for MSC characterisation (Maurer, 2011).

Given the lack of specific biomarkers, there has been some debate over a hypothesised perivascular origin of MSCs, arising from adventitial reticular cells. Pericytes are perivascular cells that form a venule wall between capillary and post-capillary (Shepro and Morel, 1993). Studies have shown that long-term cultured perivascular cells display osteogenic, chondrogenic and adipogenic potential (Farrington-Rock, Crofts *et al.*, 2004), and expressed MSC markers, the same for native uncultured perivascular cells; this suggests that pericytes could be the progenitors of MSCs (Crisan, Yap *et al.*, 2008). Crisan, Yap et al. and Caplan speculate that MSCs indeed “are pericytes” (Crisan, Yap *et al.*, 2008) (Caplan, 2008).

Criteria identifying the location of SSCs *in vivo* are more stringent than for MSCs. SSCs are considered to reside at the outer surface of sinusoids in the BM at a mural subendothelial position (Figure 10) (Bianco, Cao *et al.*, 2013).

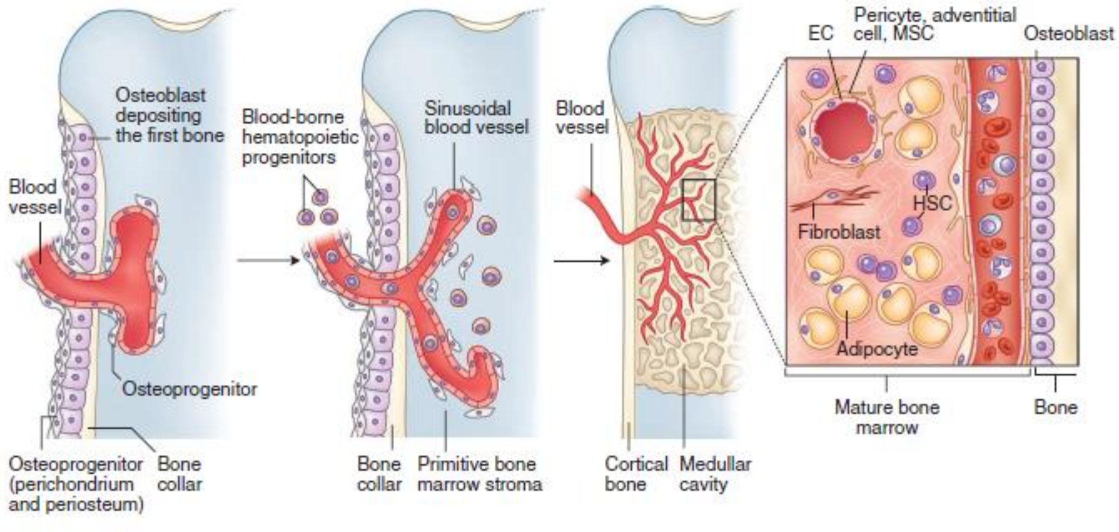

Osteoblast depositing the first bone
Blood vessel
Osteoprogenitor
Osteoprogenitor (perichondrium and periosteum)
Bone collar
Blood-borne hematopoietic progenitors
Sinusoidal blood vessel
Bone collar
Primitive bone marrow stroma
Blood vessel
Cortical bone
Medullar cavity
EC
Pericyte, adventitial cell, MSC
Osteoblast
HSC
Fibroblast
Adipocyte
Mature bone marrow
Bone

Figure 10: SSC *in vivo* location within the bone marrow.

Recruitment of osteoprogenitor cells occurs in postnatal bone marrow by recruitment of osteoprogenitors cells from outside the marrow cavity to the interior wall of the blood vessels of the bone marrow (Bianco, Cao *et al.*, 2013). Reprinted by permission of the nature publishing group.

1.2.8 Clinical needs and applications

Despite the natural capacity of bone tissue to regenerate, there are situations in which bone does not heal. This includes non-or delayed union of fractures (Harwood, Newman *et al.*, 2010), cases of substantial bone loss resulting from trauma, disease or tumour, or simply the inability to regenerate bone due to diseases or old age (Kraus and Kirker-Head, 2006) (Figure 11). In contrast to bone, cartilage does not regenerate spontaneously. Chondral injury or disease may result in bone abnormalities. Improving bone repair and self-regeneration is the driving force for the field of tissue engineering and SSCs are the protagonists embodying a plausible solution. The basis of tissue engineering is to mediate the capacity of patient expanded SSCs to repair, with the aid of appropriate scaffolds and growth factors. SSCs have been exploited clinically for their differentiation potential and paracrine effect (Meirelles Lda, Fontes *et al.*, 2009). Regarding their paracrine effect, SSCs have been reported to have a trophic, nurturing effect, immunomodulatory, anti-scarring and chemoattractant (Meirelles Lda, Fontes *et al.*, 2009). Despite some evidence of clinical efficacy, opinions on the state and efficacy of SSC therapy are incongruent and differences in the methodologies employed remain an obstacle to fully compare studies and understand efficacy. Moreover the heterogeneity of the cell populations used and the lack of specific stem cell identifier molecules constitute a further problem to the realisation of tissue engineering strategies based on SSCs.

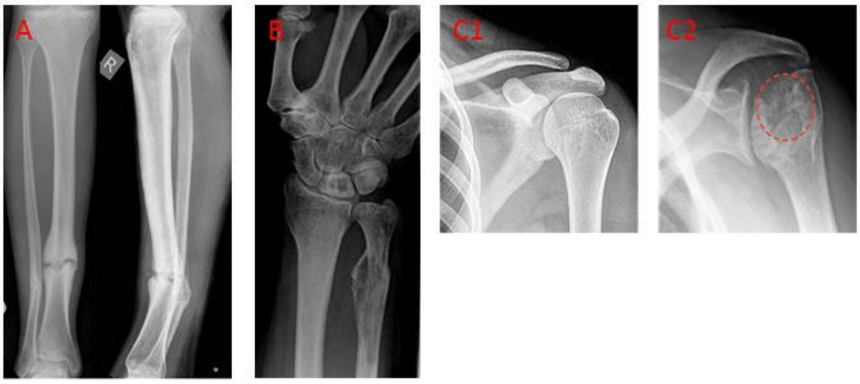


Figure 11: Bone diseases.

A) Tibial nonunions (Harwood, Newman *et al.*, 2010); B) Bone tumour from: (http://www.imageinterpretation.co.uk/images/tumour/Osteosarcoma.jpg, 08/08/2013); C1) Healthy shoulder, C2) Bone deterioration caused by osteonecrosis. C1 and C2 from: http://www.arthrosurface.com/wp-content/uploads/2013/03/ON.png (08/08/2014).

1.3 Section III: Introduction to Mass Spectrometry

1.3.1 Overview

In the previous sections the need for new tools to access homogeneity within cell heterogeneity was introduced to address the sample complexity. To approach the issue, a Multidimensional Protein Identification Technology (MudPIT) proteomic approach is used. Following those lines, state of the art mass spectrometry and its utility for SSC phenotypic analysis are discussed.

1.3.2 Mass spectrometry

Mass spectrometry (MS) is a gas phase technique that determines the molecular weight of chemical compounds (by elements) and biological compounds (organic molecules) in their ionised form as a mass-to-charge ratio (*m/z*). The unitless *m/z* ratio is defined as the mass of the ion on the atomic scale divided by its charge state (Watson and Sparkman, 2008). A compound becomes ionised when it either loses or gains an electron. Types of ionisation include protonation (when a proton is added to a molecule (1+)), deprotonation (a proton is removed from a molecule (1-)) and cationisation (non-covalent addition of a cation to a neutral molecule). Ionised molecules are introduced into the mass spectrometer where they are subjected to electrostatic or electromagnetic fields and become separated according to their *m/z* (Watson and Sparkman, 2008). Detection follows this separation as the strength of the fields is varied so that only a specific fraction of molecules with given *m/z* ratio arrives on the detector for a given field strength. The detection of the various molecular components as *m/z* gives rise to the spectrum, which is then recorded and eventually processed with computational *in silico* or manual *de novo* methodology.

1.3.3 Brief history of mass spectrometry

The basis of mass spectrometry is to understand the *m/z* ratio of ions. Over 100 years ago Sir Joseph John Thomson, a British physicist working at the Cavendish laboratory, University of Cambridge, approached for the first time the question of understanding the mass to charge ratio of ions and invented the first mass spectrometer analyser. His first mass spectrometer (Figure 12) consisted of a tube in which the ions were injected across an electromagnetic field which deviated the ion parabolic trajectory. He observed and studied the diversion of the ions under electromagnetic forces. Ion detection occurred when an ion collapsed upon the fluorescent screen or photographic plate on in the tube. Thomson won the Nobel Prize in Physics in 1906 for the discovery of the electron and for his studies on the conduction of electricity in the gas phase (http://www.nobelprize.org/nobel_prizes/physics/laureates/1906/thomson-facts.html).

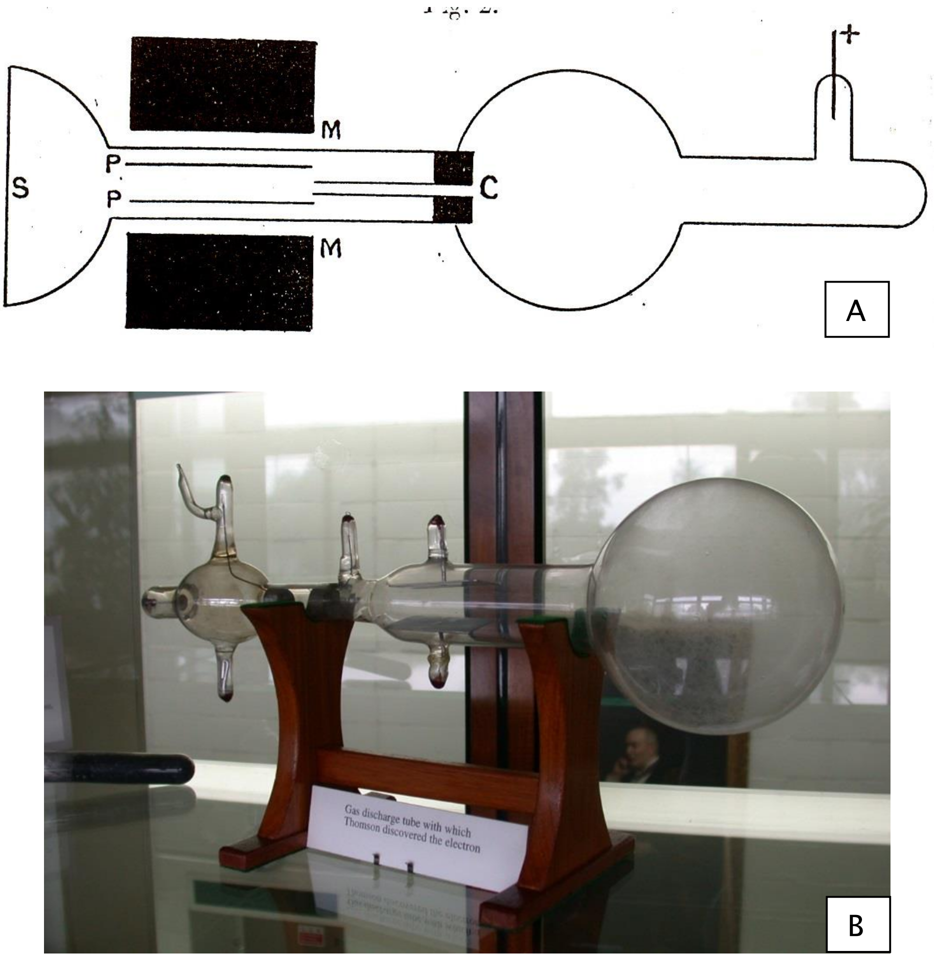

+
M
S
P
P
C
M
A
Gas discharge tube with which Thomson discovered the electron
B

Figure 12: Thomson's instrument.

(A) Thomson's drawing of his own instrument. S= screen where the ions fall, P= parallel plates, M= magnetic fields, C= perforated cathode where the ions enter the unit. From: http://masspec.scripps.edu/mshistory/timeline/time_pdf/1911_ThomsonJJ.pdf 25/05/2014. (B) A picture of Thomson's instrument. From: http://www.outreach.phy.cam.ac.uk/camphy/museum/area2/images/cabinet3_2.jpg (25/05.2014).

Later Francis W. Aston from the same laboratory built a mass spectrometer with enhanced resolution. This was further improved by A. J. Dempster from the University of Chicago who improved the resolution by providing accurate ion abundances as opposite to accurate mass measurements (Watson and Sparkman, 2008). Thus, Thomson, Aston and Dempster are considered the fathers of mass spectrometry.

Since its invention, mass spectrometry has been widely used in a range of applications, including even the fabrication of the first nuclear bomb during World War II; Professor Alfred O.C. Nier used a mass spectrometry device to separate uranium 235 from uranium 238 to make the explosive device.

A different type of mass spectrometer was proposed by William E. Stephens in 1946 called the time-of-flight (TOF) mass spectrometer. This instrument consisted in measuring the time spent by flying electrons to reach the detector. This type of device is still widely used today.

In 1950 the quadrupole mass filter was proposed by Wolfgang Paul at the University of Bonn. In this device the quadrupole is connected to a radio frequency (RF) generator and a direct current (DC) potential. The fundamental concept of this separation relies in the theory that at a specific RF and DC potential, ions with a specific *m/z* ratio will be separated (Kelly, Tolmachev *et al.*, 2010).

In 1960, there was a significant improvement in the field of mass spectrometry with the introduction of gas chromatography (GC). GC separated volatile compounds into individually purified components directly to the mass spectrometer analyser in the gas phase. However this advance had a major drawback. The amount of analyte separated was small and diluted in the gas mobile phase. To overcome this issue, Einar Stenhage, Ragnar Ryhage, J. Trock Watson, Duane Littlejohn and Peter Llewellyn developed a filter system to allow GC/MS to become the instrument that could produce more information about various analytes from less initial sample that any other technique (Watson and Sparkman, 2008).

More recently, electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI) have enabled the application of MS to biology. The former technique (ESI) consists of producing multiple charged ions by protonating a peptide with multiple basic amino acids residues. The ions are kept in solution phase, which allows ESI to be installed on-line with the high pressure liquid chromatography (HPLC) device. Highly charged ions are sprayed from a capillary to an electric field and later into the MS inlet. In comparison, MALDI consists of a laser that desorbs the sample from a solid or liquid UV-absorbing matrix; the ions are then separated. Both these techniques are still adopted.

With time, further advances and improvements have been made in the field until the present day with the “Orbitrap Elite” (Thermo Scientific) that is used in this study and that can be considered one of the “state of the art” of mass spectrometry devices at the current time.

1.3.4 Proteomics

Before describing the use of proteomics for the characterisation of SSCs, it is important to introduce the idea of proteomics and genomics based research. Both words were coined in early 1990s (Liebler, 2002), and are defined as follows:

Genomics is the study of the entire set of genes of an organism, whilst proteomics is the study of the total ‘protein complement of a genome’ (terminology coined by (Wilkins, Pasquali *et al.*, 1996)) and its post-translational modifications (Haga and Wu, 2014).

There are two schools of thought regarding utilising gene expression (genomics) or protein expression (proteomics) for cell characterisation. Both routes present limitations.

Limitations in genomics studies are predominantly: (i) failure to consider protein post-translational modifications and (ii) failure to account for gene splicing and alternative promoters which can code distinct proteins from the same gene (Gygi, Rist *et al.*, 1999).

Limitation in the use of proteomics in research include: (i) the limited capacity of the current protein databases (Duncan, Aebersold *et al.*, 2010); (ii) the finite effectiveness of the current ionization sources in ionizing the full spectrum of proteins and peptides possible; (iii) the physico-chemical complexity and sheer number proteins comprising a typical biological specimen; (iv) protein identification success may vary with the sensitivity of the mass spectrometer; v) the presence of isobaric tags (Garbis, Lubec *et al.*, 2005). Continuous advancements made to liquid chromatography and mass spectrometry designs, however, show promise in addressing these limitations.

1.3.5 The LTQ-Velos Pro Orbitrap-Elite

The LTQ-Velos Pro Orbitrap-Elite is the mass spectrometry instrument used in this thesis. The Orbitrap Elite is a hybrid instrument where two types of mass analysers, the linear ion trap or linear trap quadrupole (LTQ) and the orbitrap are combined (Figure 13). It is a platform where ions are stored using electrostatic fields and are detected by the Fourier transformation law of oscillating image current as explained in more detail below. The fragmentation modes offered are: CID (collision induced dissociation), ETD (electron transfer dissociation) and HCD (high collision dissociation). The linear ion trap has a low resolution ion detection power while the orbitrap has a high resolution power, and high mass accuracy ion detection (Kalli, Smith *et al.*, 2013). The LTQ is a 3-D quadruple cell where ions are subjected to a range of RF and DC potentials allowing them to get stored, fragmented or ejected. These ions can then be analysed either by the photomultipliers adjacent to the ion trap or by the FT (Fourier Transform)-orbitrap detectors.

The orbitrap is essentially an advanced electrostatic design that incorporates a spindle shaped inner electrode on the central axis. Ions rotate around this spindle at harmonic oscillation frequencies directly dependent on specific *m/z* ratio. These oscillation frequencies are recorded within the Orbitrap component and converted to exact *m/z* values by means of the Fast Fourier Transformation algorithm thus constructing the high-resolution mass spectrum at the precursor and product ion level (Thermo Fisher Scientific Orbitrap elite software manual, Revision A-1288170).

In this work protein quantification will be carried out by using isobaric tags for the relative and absolute quantification (iTRAQ).

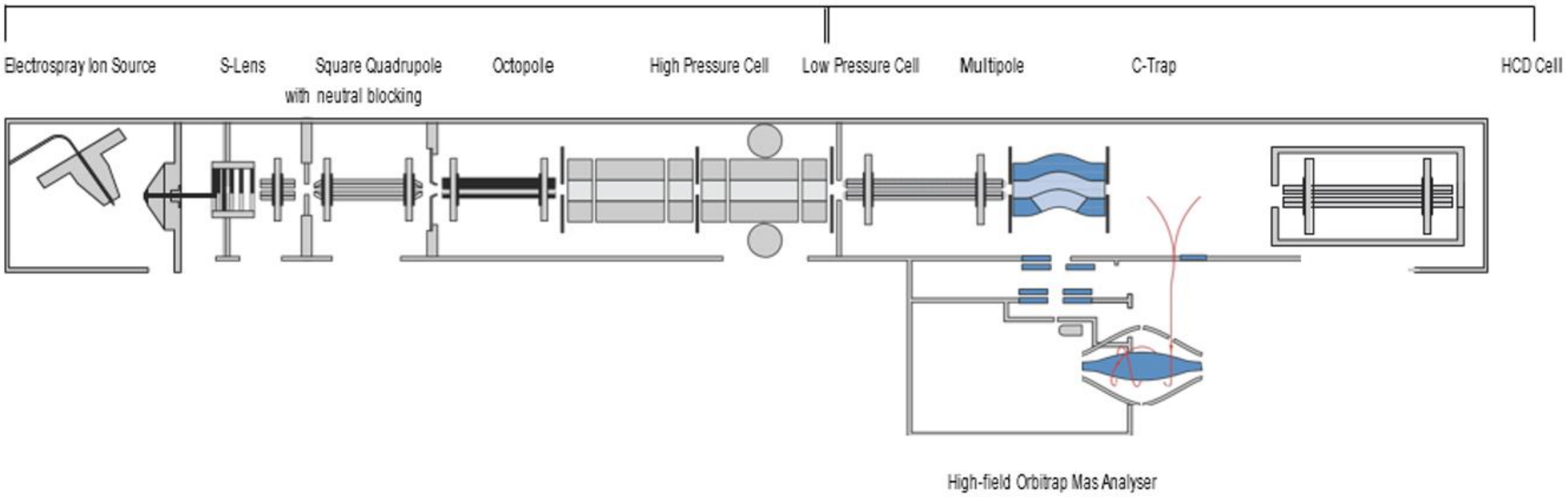
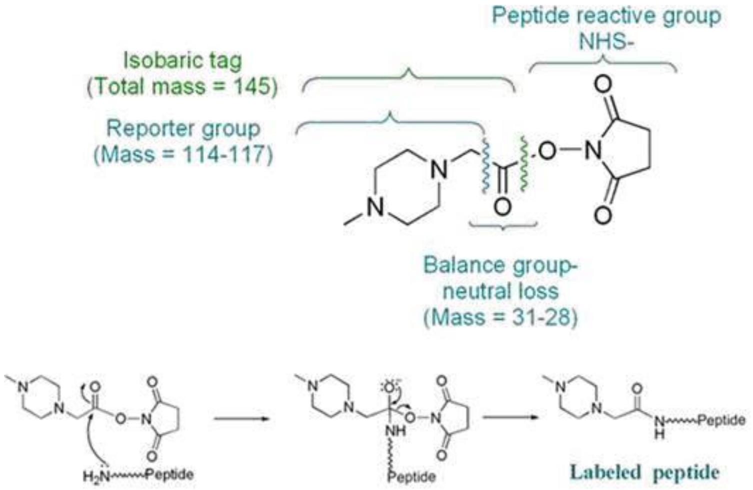

Electrospray Ion Source
S-Lens
Square Quadrupole with neutral blocking
Octopole
High Pressure Cell
Low Pressure Cell
Multipole
C-Trap
HCD Cell
High-field Orbitrap Mas Analyser

Figure 13: Inside the LTQ-Velos Pro Orbitrap-Elite.

A beam of ion is cationised and electrosprayed through the electrospray ion source. Ion beam crosses the S-Lens, square quadrupole with near blocking and the octopole where the mass of the precursor ions is analysed according to RF/DC potential. Subsequently ions migrate to the C-Trap where are diverted to the HCD cell where each precursor ion collide with nitrogen particles (tandem MS) dividing in smaller ionised molecules. Ionised molecules are trapped by the C-Trap which directs the ion beam to the Orbitrap analyser. In the Orbitrap ions start to oscillate according to their m/z at a specific RF/DC potential. Fast Fourier Transformation algorithm transforms oscillations into *m/z* value. From: http://planetorbitrap.com/orbitrap-elite (25/05/2014).

1.3.6 Isobaric tags for the relative and absolute quantification (iTRAQ)

The use of isobaric stable isotope labelling constitutes a recent development as an analytical tool that allows the multiplex quantification of proteins. Such is the case for the isobaric Tags for Relative and Absolute Quantitation (iTRAQ) developed by Darryl Pappin (Ross, Huang *et al.*, 2004). The iTRAQ reagent principally covalently modifies the N-terminal amine moiety of proteolytic peptides. Each isobaric (same weight) tag has a total mass of 145 kDa and is constituted by a reporter group and a balance group. The reporter group varies in mass and it is traceable to the proteolytic peptides derived from the sample initially labelled with the respective iTRAQ tag. The reporter group is released from the precursor ion during MS/MS analysis. The balance group serves to equilibrate the mass of the entire tag (Figure 14).


Peptide reactive group
NHS-
Isobaric tag
(Total mass = 145)
Reporter group
(Mass = 114-117)
Balance group-
neutral loss
(Mass = 31-28)
H_2N~Peptide
Peptide
N~Peptide
H
Labeled peptide

iTraq tag mass	isobaric tag mass	balance group mass
113	145	32
114	145	31
115	145	30
116	145	29
117	145	28
118	145	27
119	145	26
121	145	24

Figure 14: iTRAQ labels chemistry.

The tag binds covalently to the amine group of a peptide. The label recognition part is the reporter ion that is released during MS/MS analysis. From: http://www.med.wayne.edu/physiology/facultyprofile/chen/Figure%201.jpg (18/08/2014).

1.4 Section IV: Aptamers technology for biomarker discovery

1.4.1 Aptamers: History

Aptamers (from the latin *aptus*- fit, and Greek *meros*– part) are high affinity short single stranded nucleic acid sequences that bind specifically to a target molecule. Aptamers can be described as the analogues of antibodies but constituted by oligonucleotides; therefore like antibodies, aptamers can be employed as tools for target recognition and biomarker discovery for sensing as well as diagnostic applications (Kong and Byun, 2013). The main benefit of aptamers as tools for biomarker discovery is the ability to bind virtually to any given molecule(s), thanks to a versatile three dimensional structure characterised by stems, loops and hairpins which facilitate the anchorage and attachment to molecules (Hermann and Patel, 2000).

Historically, inadequate emphasis was given to nucleic acids as biomarkers, nucleic acids being predominantly associated to the storage of the genetic code and considered less complex than proteins. Interest in nucleic acids has steadily increased with the discovery of non-coding ribonucleic acids exhibiting binding (due to structural motifs) and catalytic properties similarly proteins (Weigand and Suess, 2009). In 1990 Ellington and Szostak and Tuerk and Gold (Ellington and Szostak, 1990; Tuerk and Gold, 1990) reported the identification of given target molecules through the amplification and selection of specific nucleic acids. The “Ellington and Szostak” laboratory called the process “in vitro selection” to select RNA ligands against organic dyes, and coined the name “aptamers”, whilst the “Tuerk and Gold” laboratory coined the term SELEX (Systematic evolution of ligands by exponential enrichment) for selecting RNA ligands against T4 DNA polymerase.

Since the achievement of the first SELEX process, relatively little commercial research has been undertaken in the field of nucleic acids as molecular probes due to a patent protecting the intellectual property of the SELEX process (Gold and Tuerk, 1997). This patent expired in 2012 and since then multiple companies have acted as ‘fast followers’ adopting and examining applied aptamer technology.

1.4.2 SELEX

The SELEX process employs a combinatorial library of 10^{13} to 10^{15} molecules (Figure 15). These molecules of nucleic acids contain a random base pair sequence constituted by 30 to 60 nucleotides, flanked by primers for amplification purposes. The SELEX process includes: i) binding, ii) selection and iii) amplification (Ozer, Pagano *et al.*, 2014). These steps are repeated for as many cycles as desired until a population of nucleic acid highly specific for the desired targeted molecule is obtained and no further improvements in binding are observed. Given the huge variety in the sequences and shapes contained in the library, it is hypothesised that at least one sequence will bind to the sought target. In the case of biomarker discovery on cell surface, a typical SELEX process will consist of an ss-DNA (single strand) or RNA aptamer library incubation with cells. Some sequences will bind, other will not and will be removed with the washes. Bound sequences are then amplified by PCR in case of DNA libraries or RT-PCR in case of RNA libraries. The resulting PCR product containing enriched aptamer sequences for the target used in the previous round, is employed for the next round of selection.

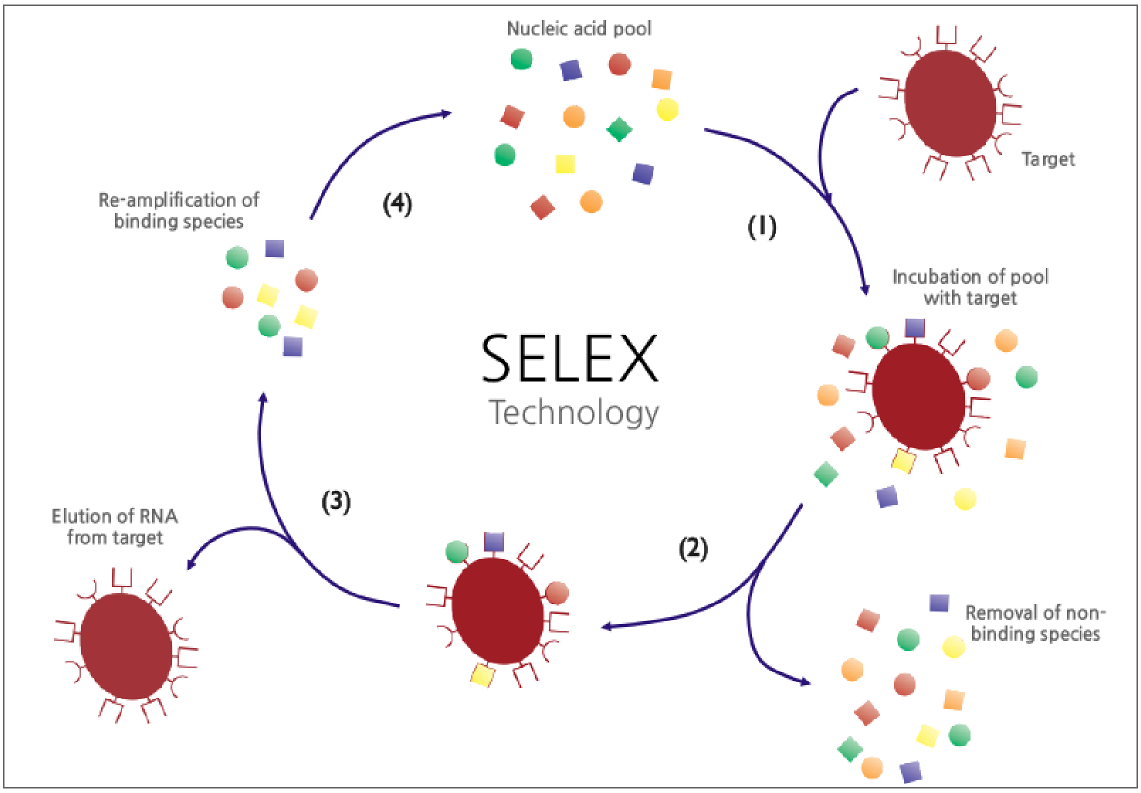

Nucleic acid pool
Target
(1)
Incubation of pool with target
SELEX
Technology
(2)
Removal of non-binding species
(3)
Elution of RNA from target
(4)
Re-amplification of binding species

Figure 15: SELEX process.

The target cell is allowed to mix with the aptamer library. The cell bound sequences are amplified by PCR (positive selection) whilst the negative are washed away. The cycle is repeated multiple times until the aptamer pool is populated by only highly specific sequences. http://www.biois.co.kr/?page_id=106&lang=en (03/12/2015).

1.4.3 Aptamer libraries

Aptamer libraries contain multiple oligonucleotide sequences. These sequences are characterised by a central region of random 30 to 60 nucleotides flanked with primer sequences of around 20 nucleotides. In a classical SELEX experiment DNA or RNA libraries can be employed. The choice between RNA or DNA libraries is not obvious as both libraries present advantages; generally, DNA is favoured as considered more stable than RNA, however it has been shown that RNA has a higher affinity to molecules and, furthermore, RNA aptamers can undertake more conformational motifs than DNA libraries. Although some reviews claim that there is no evident difference between RNA and DNA aptamers, the debate on the benefit of DNA over RNA libraries and *vice versa* remains open. Aptamer libraries are often modified with the aim to increase stability. The modifications are introduced at the 2' position of the pyrimidines bases using a 2-aminopyrimidine, 2-fluoropyrimidine, 2-O-methyl nucleotides, position 5' also, and also at position 4' using UTP and CTP (Eaton, Gold *et al.*, 1995; Kimoto, Yamashige *et al.*, 2013; Kusser, 2000).

1.4.4 Target molecules

Aptamers can be generated against virtually any given target molecule. The prerequisite for binding is that the target molecule should be sufficiently abundant or display high purity. Target molecules include nucleotides (Sassanfar and Szostak, 1993), amino acids (Famulok, 1993), carbohydrates (Srisawat, Goldstein *et al.*, 2001), natural products such as theophylline, theobromine and caffeine (Johnson, Kumar *et al.*, 2003), fluorescent dyes (Werstuck and Green, 1998; Wilson and Szostak, 1998), antibiotics (Niazi, Lee *et al.*, 2008), proteins, peptides, synthetic polymers, polysaccharides, glycoproteins, hormones, receptors, and small molecules such as drugs, metabolites, cofactors, transition state analogues and toxins (Ellington and Szostak, 1990) . Aptamers bind to their targets due to electrostatic and van der Walls interactions, hydrogen bonding or a combination thereof. SELEX has been most commonly used for the generation of aptamers against protein targets (Gopinath, 2007).

1.4.5 Aptamers versus antibodies

Aptamers are alternatives to antibodies in that they are totally different molecules; yet the application of aptamers correspond closely to antibody applications (Chen and Yang, 2015). While the aptamer field is still in its infancy, the potential of aptamers for therapeutics, diagnostics, and biomarker discovery is considered to be significantly greater than their antibody counterpart.

The following advantages of aptamers over antibodies have been suggested:

- Aptamers are more versatile than antibodies. Aptamers can fold and refold into different three dimensional shapes without losing specificity (Porciani, Signore *et al.*, 2014).
- Aptamer stability, ease of immobilisation, and labelling is higher than antibodies (Song, Lee *et al.*, 2012).
- Aptamers, having a smaller size (8-15 kDa) compared to antibodies (150 kDa), are better candidates for therapeutics than antibodies as small molecules are most suited for rapid tissue penetration and blood clearance (Hicke, Marion *et al.*, 2001).
- Being synthesised in vitro, aptamers do not require production in animals. Animal loss, work load and batch variation that characterise antibodies production are eliminated with aptamers technology (Xiang, Shigdar *et al.*, 2015).
- Molecules such as biotin, fluorescein can be attached to aptamers without loss of binding affinity (Davis, Abrams *et al.*, 1998).
- Antibodies are highly susceptible to high temperature, denaturing conditions, harsh storage conditions and pH changes, whilst aptamers can denature and reconstitute within min (Yao, Zhu *et al.*, 2010).
- With aptamers the problem of cross reactivity and un-specificity that is common with antibodies is overcome by their higher affinity to only one target (Kong and Byun, 2013).
- Considering antibody production needs animals, molecules such as toxins are note well tolerated by animals; aptamers instead can be selected against toxins too (Lauridsen and Veedu, 2012).

- Aptamers can be selected against not only any proteins but also against different proteins isoforms with high affinity and specificity (Zimbres, Tárnok *et al.*, 2013).
- The cost associated with the chemical synthesis of aptamers is significantly lower compared to antibody production costs (Sun, Zhu *et al.*, 2014).
- Aptamers are non-immunogenic and have been approved for human use (the FDA approved Pegaptanib sodium (Macugen) in 2004, for the treatment of all types of neovascular age-related macular degeneration, (Ng, Shima *et al.*, 2006)).

Various drawbacks do, however, attend the use of aptamers against antibody applications (McKeague and Derosa, 2012):

- Aptamer tertiary structure is dependent on solution temperature and pH.
- Aptamers can be easily degraded in blood.
- The field of aptamers technology is still young and there are not many aptamers available as "off the shelf' tools such as many antibodies.

The future prospective envisages an expansion of the aptamer technology, with fading disadvantages that hinder the field.

1.5 Aim and Hypothesis

The aim of this thesis is to further the understanding of SSC identification from a mixed BM stromal cell population. As a means to normalise this intrinsic heterogeneity of SSCs, I aim to identify novel protein markers by applying cell sorting with subsequent multiplex proteomic analysis.

In addition this thesis will be examining different human cell lines for their Stro-1 expression with the aim to characterise the Stro-1 epitope by WB, immunoprecipitation and cell fractionation techniques.

The hypotheses of the thesis were:

I. The Stro-1 subpopulation of cells represents the *bona fide* SSCs and as such exhibits high differentiation potential. *Objective:* To assess the osteogenic differentiation potential of the Stro-1 fraction of cells by molecular and bimolecular techniques.

II. The Stro-1 antigen size can be identified by WB analysis. *Objective:* to develop a WB-based method that allows detection of the Stro-1 antigen.

III. Within the Stro-1 fraction resides a subset of cells with a unique proteomic profile that can be revealed by mass spectrometry. *Objective:* to focus on the differences in mass spectrometry profiles between the Stro-1 sorted fraction and the unsorted fraction.

IV. Protein modulation is a result of sorting and not osteoinduction. *Objective:* to assess the source of protein modulation in the mass spectrometry results.

V. The aptamer SELEX process identifies potential ligands for the isolation of SSCs. *Objective:* to run a SELEX experiment on MG63 cells mimicking the osteogenic phenotype.

By testing these hypotheses experimentally, this thesis could identify potential candidates for SSC selection.

2. Chapter 2: Materials and Methods

2.1 Cell culture and osteoinduction of human bone marrow cells for mass spectrometry analysis

Human bone marrow (BM) stromal cells (BMSC) were extracted from human BM collected from patients (2 males aged 75 and 64 and 2 females aged 42 and 67) undergoing a routine total hip replacement surgery at Southampton General Hospital, under informed consent and compliant to the Southampton and South West Hampshire Local Research Ethics Committee (LREC194-199). BM was washed in α-MEM (Minimal essential Medium alpha, phenol free) (GIBCO Life Technologies), and passed through a cell strainer (70 μm cut off membrane) to remove any bone chips, fat and tissue. The resuspended cells were centrifuged at 1,300 rpm for 5 min at 18 °C. Cells were separated from red blood cells by density gradient centrifugation (2,200 rpm, 40 min at 18 °C) using Lymphocyte Separation Medium LSM1077 (GE Healthcare). Following centrifugation, BM cells were found at the interphase in a buffy layer. Cells were collected and an aliquot of cells was subjected to the magnetic separation protocol (2.2) to obtain Stro-1 positive (Stro-1^+) cells. The bone marrow samples of two patients were processed contemporaneously and the cell seeding density was standardised according to the sample with a minor cell count. This allowed the same number of sorted and unsorted cells between the 2 patients to be the same. Unsorted and Stro-1^+ cells were resuspended in basal media, α-MEM phenol free, containing 10% foetal calf serum (FCS, Sigma) and 1% penicillin/streptomycin (PAA The cell culture company), and cultured in flasks for 14 days. Subsequently cells were trypsinised, re-seeded and cultured for 21 days in basal or osteogenic media (α-MEM phenol free, L-Ascorbic acid 2-phosphate sesquimagnesium salt hydrate (100 μM, Sigma-Aldrich), dexamethasone (10 nM, Sigma-Aldrich), 10% FCS, and 1% penicillin/streptomycin. Cells were incubated at 37 °C, and 5% CO_2 in a humidified atmosphere.

2.2 Magnetic Activated Cell Sorting (MACS) isolation

Following lymphoprep separation, cells recovered from the buffy layer were resuspended in blocking buffer (17ml α-MEM phenol free, 2 mL AB human serum (Sigma), 0.2g Bovine Serum Albumin (BSA), 1ml FCS) for 30 min, washed in MACS buffer (500ml PBS, 2.5g BSA, 0.37 EDTA) 3 times and incubated with the Stro-1 antibody hybridoma supernatant for 30 min at 4 °C. After three washes with MACS buffer, cells were resuspended in rat α-mouse IgM microbeads (Miltenyi Biotec) (20 μL per 10^7) diluted in MACS buffer (1:5) and incubated for 30 min at 4 °C. Labelled cells were then added to a MACS column pre-conditioned with MACS buffer. The first eluent was collected as the Stro-1 negative fraction (Stro-1^-) and discarded. Subsequently the column was detached from the magnet and a plunge was used to elute the Stro-1^+ cell fraction. Three cell counts were performed using a disposable haemocytometer chamber (FastRead Counting Slides; ISL, Immune Systems Ltd.) and Axiovert 40 C (Zeiss) light microscope. The positive fraction was transferred to tissue culture flasks with α-MEM phenol free (Modified Eagle's Medium) (GIBCO life Technologies), containing 10% foetal calf serum (FCS) (Sigma), and 1% penicillin/streptomycin (PAA The cell culture company) and incubated at 37 °C, in 5% CO_2. Media changes were performed at day 7 after the isolation, then twice a week thereafter.

2.3 Alkaline Phosphatase staining

Human BM cells were cultured in a monolayer on tissue culture plastic, fixed in 95% ethanol for 10 min and washed in phosphate buffer saline (PBS). Cells were fixed at day 7, 14 and 21 after the addition of osteogenic media. ALP stain was prepared by adding 0.003g of Fast Violet (Sigma) into 10 mL Naphtol solution (Sigma). ALP stain was added to the fixed cell monolayer and incubated for 40 min at 37 °C. ALP stain was removed and cells were washed twice with PBS. Images were captured with a light microscope Axiovert 200 equipped with an Axiocamera (Zeiss). An image-based approach to quantify the ALP stain was executed by the free software CellProfiler: image analysis software for identifying and quantifying cell phenotypes (Broad institute), version 2.1.1. Intensity of ALP stain was calculated per cell based on nuclei count.

2.4 Molecular analysis

2.4.1 RNA extraction

RNA extraction from human BM samples was performed using an RNeasy Mini Kit (Qiagen), according to manufacturer's instructions. Cells in monolayer culture were scraped with 200 µL of RLT buffer and homogenised with the same volume of 100% ethanol on ice. Samples were transferred to spin columns and washed once with RW1 buffer, and then twice with RPE buffer. A final wash with RNase-free water was performed to elute the RNA. RNA was quantified using a NanoDrop-1000 spectrophotometer (Thermo Scientific).

2.4.2 cDNA synthesis

cDNA synthesis was performed using the SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen). 300 ng of RNA was diluted in ultra-pure water to a volume of 7 µL, to which 2 µL and 1 µL of VILO Reaction Mix and VILO 10x SuperScript were added respectively. Samples were plated in a thermocycler with the following conditions: 25 ˚C for 10 min, 42 ˚C for 120 min, 85 ˚C for 5 min and then maintained at 4 ˚C until retrieved and stored at -20 ˚C.

2.4.3 Primer design

Primers specific to bone genes were designed by Dr. Rahul Tare (Table 3).

2.4.4 Real-time PCR (RT-PCR)

RT-PCR was performed by using SYBR-Green master mix (2x concentrated). ALP, OCN, SOX9 and GAPDH forward and reverse primers were used (Table 3). SYBR-Green, primers and cDNA were added in triplicate in a 96-well PCR plate. The plate was loaded into the Applied Biosystem (Life Technologies, USA), 7500 RT-PCR system. Thermocycler conditions were 50 ˚C for 2 min, 95 ˚C for 10 min, followed by 40 cycles at 95 ˚C for 15 s and 60 ˚C for 1 min. Melting curves were generated at the end of the run and checked to evaluate specificity of the product. Data were analysed using the Applied Biosystems 7500 Fast System Software, version 1.3.1. ALP, OCN and SOX9 expression level was calculated by subtracting the GAPDH dCT value from the ALP dCT value. Mean and standard deviation was calculated using the three replicates values of each condition.

2.4.5 Statistics

The analysis of the ALP staining of three patients and the molecular analysis of ALP, OCN and SOX9 of the same three patients were statistically analysed with GraphPad Prism (version 6.00 for Windows, GraphPad Software, San Diego California USA) using two-way ANOVA for multiple comparison, (* p<0.05, ** p<0.01, *** p<0.001, n=3). Statistical comparisons were run between the UB vs UO and SB vs SO within the same time point. For OCN, statistical comparisons were run between time points.

Table 3: Forward and reverse primers used for RT-PCR analysis in this work.

Primers were designed by Dr. Rahul Tare (Bone and Joint group, University of Southampton) and manufactured by Sigma.

Gene	Primer
Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	F: 5’ GACAGTCAGCCGCATCTTCTT 3’ R: 5’ TCCGTTGACTCCGACCTTCA 3’
Alkaline phosphatase (ALP)	F: 5’ GGAACTCCTGACCCTTGACC 3’ R: 5’ TCCTGTTCAGCTCGTACTGC 3’
Osteocalcin (OCN)	F: 5’ CTGACCTCACAGATGCCAAG 3’ R: 5’ GTAGCGCCGGAGTCTGTTC 3’
Sex determining region Y-box 9 (SOX9)	F: 5’ CCCCAACAGATCGCCTACAG 3’ R: 5’ GAGTTCTGGTCGGTGTAGTC 3’

2.5 Stro-1 antibody production

Stro-1 hybridoma was re-derived from an original donation provided by Dr. Jon N. Beresford, University of Bath (Stewart, Walsh *et al.*, 1999). Cells were cultured in D-MEM (Dulbecco’s modified eagle’s medium, Lonza), 20% FCS (Sigma) and incubated at 37 ˚C in 5% CO_2. FCS concentration was reduced to 5% once cells started proliferating. Cells were transferred into a cell bioreactor (CELLine, Integra). Cells were harvested every week and the supernatant was collected upon centrifugation at 1,400 rpm ant stored at -80 ˚C. The collected supernatant constituted the Stro-1 antibody, in-house produced and employed in this thesis.

2.6 Fluorescent Immunostaining

Cells on a monolayer were washed twice in PBS and fixed in 4% paraformaldehyde (PFA) in PBS for 20 min and then re-washed twice in PBS. Cells were permeabilised and blocked in blocking buffer (0.3% triton, 5% goat serum, 1% BSA in PBS) for 5 min. Blocking buffer was removed and wells were washed three times in PBS/tween (PBS with 0.05 % (v/v) tween 20 (Sigma)). For primary antibody incubation, the antibodies used were:

- Stro-1 hybridoma supernatant undiluted (in-house produced);
- Stro-1 hybridoma supernatant diluted 1:100 (in-house produced);
- Stro-1 diluted 1:100 (R&D Systems);
- IgM Lambda isotype (Sigma) diluted 1:100;
- No antibody control

Cells were incubated at 4 ˚C overnight with primary antibodies. Cells were washed 3 times in PBS/tween (PBS with 0.05 % (v/v) tween 20 (Sigma)) and incubated with fluorescent secondary antibody specific for IgM (Alexafluor 488 goat a-mouse IgM, (Invitrogen)) (1:200). Wells were washed twice in PBS/tween and counterstained with DAPI, (dilactate, nucleic acid stain, H_2O soluble (4´, 6-diamidino-2-phenylindole, (Invitrogen))) (1:200); re-washed twice in PBS/tween. Pictures were taken with Axiovert 200 with Axiovert camera (Zeiss).

2.7 Cell lysis for Western Blotting (WB)

Cell pellets were lysed employing different lysis buffers according to the application (Table 20). The cell lysis step was performed in the presence of a cocktail of protease inhibitors (1:100, Cell Signaling Technology) and incubated on ice for 30 min. Cell lysates were sonicated 3 times at an amplification of 20% for 1 min each time. Protein content was analysed by bicinchoninic acid (BCA) assay and WB according to the protocols in sections 2.8 and 2.9 respectively.

2.8 Protein quantification (BCA assay)

Proteins were quantified using the Pierce BCA assay according to manufacturer's instructions. A set of standards were first prepared containing a range of bovine serum albumin (BSA) concentrations from 2,000 µg/µL to 25 µg/µL in ultra-pure water (range: 2,000 µg/µL, 1,500 µg/µL, 1,000 µg/µL, 750 µg/µL, 500 µg/µL, 250 µg/µL, 125 µg/µL, 25 µg/µL, and 0 µg/µL). Following cell lysis, proteins were diluted (1:10 or 1:20) in ultra-pure water and 10 µL of diluted proteins were plated twice on a 96 well plate. 10 µL of each standard were also added to the plate in duplicates upon vortexing. Reaction was started by the addition of 200 µL of Pierce working reagents (solution A and solution B, 50:1 dilution respectively) and each plate was incubated for 30 min at 37 ˚C. Absorbance was read with a BIO-RAD model 680 microplate reader (BIO-RAD) at a wavelength of 570 nm and calculations were performed to find the protein concentration in µg/µL in the sample under investigation (example of calculations is reported in Table 4 and a standard curve example is reported in Figure 16). The developed protocol was applied to all subsequent studies where a BCA assay was employed.

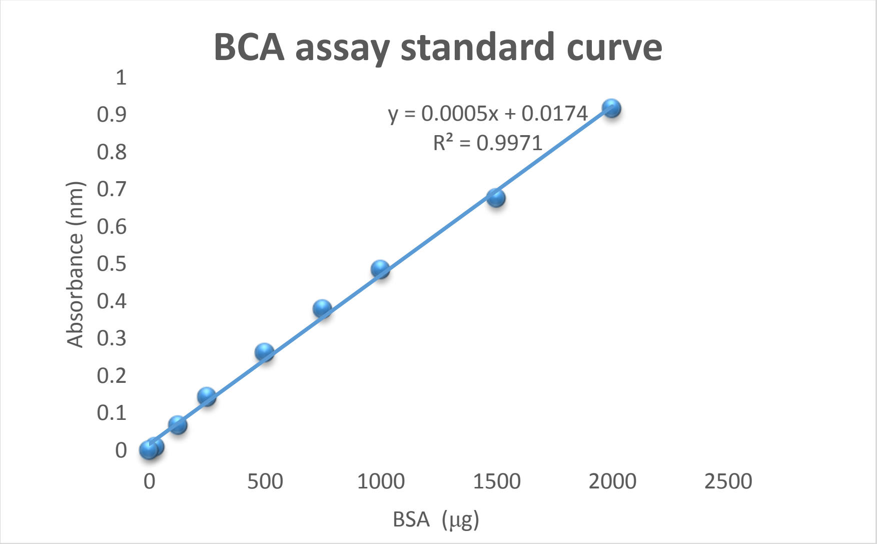


Figure 16: BCA assay standard curve example.

Standard curve of BSA absorbance at 570 nm against BSA concentrations. The equation is employed to determine protein concentration.

Table 4: BCA assay total protein calculation.

Total protein amount in the sample was calculated by measuring the absorbance of the sample emitted at 570nm and comparing against the absorbance given by a set of standards of known total protein concentration. The equation employed refers to the standard curve in Figure 16. Values of the sample absorbance and blank considered in the table are hypothetical but realistic.

Sample ID	Absorbance @570nm	Absorbance minus blank	Transformation of absorbance reading (µg/mL)	Dilution factor (10) (µg/mL)	µg/µL unit transformation
SMP1	0.345	(0.345-0.055) =0.290	(0.290-0.0174)/0.0005 =545.2	545.2*10 =5452	5452/1000=5.45

2.9 Western Blotting

Cell lysates were prepared according to the protocol described in section 2.7 and protein concentration was calculated as per protocol in section 2.8. Cell lysates were mixed with 3x SDS (sodium dodecyl sulfate) loading buffer (Biolabs) with the addition of 30x dithiothreitol (DTT, Cell Signalling). Samples and a molecular weight marker (Precision Plus Protein All Blue prestained standard, BioRad) were loaded onto precast gels (10 well combo 30 µL any kD Mini-PROTEAN TGX, precast gels, BioRad). Gel electrophoresis was run for 40 min at 150 V with a Bio-Rad power supply. Samples were transferred onto polyvinylidene fluoride (PVDF) immobilon-FL transfer membrane (Millipore) for 1.5 hours at 75 V. Non-specific sites were blocked in a 5% (w/v) non-fat milk powder in PBS/tween (PBS with 0.05% (v/v) tween 20 (Sigma)) for 1 h at room temperature (RT). Blots were washed for 5 min 3 times in PBS/tween. Blots were subsequently incubated with primary antibody overnight at 4 °C. Blots were washed for 5 min 3 times in PBS/tween and incubated in secondary antibody for 1 h at RT. Blots were washed 3 times for 5 min in PBS/tween and once for 5 min in PBS only. Antibody dilution specifications, buffer employed for dilution, and incubation time for each antibody are indicated in Table 5.Liquid chromatography/Mass spectrometry

2.10 The method and the reagents employed for the two mass spectrometry studies are details in Chapter 4 and 5 respectively.

Aptamer selection process (SELEX)

The method followed for the aptamer selection process (SELEX) is described in 3.3.3.

Table 5

2.11 Liquid chromatography/Mass spectrometry

The method and the reagents employed for the two mass spectrometry studies are details in Chapter 4 and 5 respectively.

2.12 Aptamer selection process (SELEX)

The method followed for the aptamer selection process (SELEX) is described in 3.3.3.

Table 5: Antibody specifications used in this study.

Antibody type	**Antibody/manufacturer**	**Antibody dilution used**	**Diluent**	**Incubation time**
Primary antibody	Stro-1 Hybridoma Supernatant/produced in house (IgM)	Used neat	N/A	Overnight at 4 ˚C
Primary antibody	Monoclonal α-β-Actin−Peroxidase antibody produced in mouse/Sigma (IgM)	1:1,0000	5% semi-skimmed milk in PBS/tween	1 h at RT
Secondary antibody	α-mouse IgM (μ chain specific) horseradish peroxidase (HRP) conjugate/Sigma	1:40,000	5% semi-skimmed milk in PBS/tween	1 h at RT
Secondary antibody	IRDye 680 Goat α-mouse IgM (μ chain specific)/ LI-COR, Biosciences	1:30,000	5% semi-skimmed milk in PBS/tween + 0.01%SDS	1 h at RT
Secondary antibody	IRDye 800CW Rabbit α-HRP Properties / LI-COR, Biosciences	1:5,000	5% semi-skimmed milk in PBS/tween + 0.01%SDS	1 h at RT

3. Chapter 3: Characterisation of Human Skeletal Stem cells

I am grateful to Mrs Christine Penfold (Tenovus building, Cancer Science division, University of Southampton) for running a blue native gel of the Stro-1 IgM antibody.

I am grateful to the Aptamer group (York) for hosting me during the aptamer selection studies.

The 50 kDa band analysis by mass spectrometry was outsourced to Dr. Omar Jallow (St. Georges University of London).

3.1 Introduction

As previously described in section 1.2.3, SSCs found in bone marrow (BM) offer a renewable cell source for skeletal reparative medicine to regenerate and reconstruct damaged tissues (Wakitani, Goto *et al.*, 1994). In this chapter SSC will be characterised based on analysis of osteogenic differentiation potential, the Stro-1 antigen will be investigated by WB and ultimately a new approach using synthetic molecule for biomarker identification will be explored. This thesis was prevalently focused on the osteogenic potential, thus the adipogenic and the chondrogenic potentials were not considered.

Although SSC represent only a small proportion of the BM (1/100,000 nucleated cells (Connolly, Guse *et al.*, 1989) sufficient cell numbers can be obtained by *in vitro* expansion of plastic adherent cells to generate desired numbers (minimum 500,000 cells for molecular characterisation, minimum 100 cells for staining). Furthermore, direct differentiation of expanded SSC along the adipogenic, chondrogenic and osteogenic pathway can be prompted through addition of specific inductive chemicals to tissue culture media or biological stimuli by applying chemistries to the growth surface material (Curran, Pu *et al.*, 2011).

Expanded SSC populations present morphologically different subpopulations which exhibit, hypothetically, diverse differentiation. This apparent heterogeneity is believed to affect reproducible population analysis; hence recognition and isolation of SSC with high differentiation potential is an urgent clinical imperative. A variety of methods have been employed to distinguish and subsequently isolate homogeneous SSC populations. Classic methods comprise surface marker profiling and lineage differentiation assays. Diverse panels of antibodies are, and have been, proposed for the identification and isolation of SSCs (Table 1 and Table 2), however a consensus remains elusive.

One interesting, yet highly debated marker is Stro-1 (Simmons and Torok-Storb, 1991). Stro-1 is used to isolate the non-haematopoietic fraction of cells from human BM capable of osteogenic differentiation (Stewart, Walsh *et al.*, 1999). The Stro-1$^+$ subpopulation has been shown to be capable of differentiation into chondrocytes, adipocytes and osteoblasts (Psaltis, Paton *et al.*, 2010). Indeed studies showed that the Stro-1$^+$ fraction exhibits a higher proportion of cells with CFU-F capacity (Pittenger, Mackay *et al.*, 1999).

However, despite the potential utility of Stro-1, the identity of the cell surface antigen for Stro-1 remains unknown. In 2011 Ning, Lin et al. (2011) indicated that the Stro-1 antigen is an endothelial marker of 75 kDa based on results obtained with the WB technique. Previously, it was published that Stro-1 reacted with an epitope at 50 kDa and at 18 kDa (Castrechini, Murthi *et al.*, 2010). Currently there is still no consensus as to the size of the Stro-1 antigen and, moreover, on the identity of Stro-1. Interestingly, little discussion was prompted following these previous suggestive publications. Clearly, identification of the Stro-1 antigen would allow classification of the relative selectivity and specificity of the Stro-1$^+$ population and define the function of the antigen following expression and functional regulation.

Stro-1, although the most selective SSC isolation marker currently available, is not specific to SSC, hence alternative methods to identify markers have been explored in this chapter. Described herein is a technique of employing nucleic acids aptamers with the aim of selecting specifics recognition tools to isolate SSC. The work performed with aptamer molecules served the purpose of a pilot, proof of concept, feasibility study.

3.2 Aims and hypothesis

To summarise this chapter aimed:

- To further characterise and highlight the osteogenic differentiation potential of SSC following selection by the expression of the Stro-1 antigen.
- To characterise the Stro-1 antibody and antigen by WB.
- To select aptamer molecules that specifically and uniquely bind to the surface of MG63 (positive target) cells with a counter-selection against Raji cells (negative target) to identify novel surface markers.

Hypothesising:

I. Stro-1 selected cells exhibit higher osteodifferentiation potential in comparison to unselected cells.
II. The Stro-1 antibody reacts to an epitope at 50 kDa.
III. Aptamers can be selected specifically against the surface markers of MG63 cells using SELEX technology.

3.3 Material and Methods

3.3.1 Characterisation of bone marrow cells

3.3.1.1 Cell culture, MACS and osteoinduction

Bone marrow stromal cells (BMSC) were extracted from BM collected from three patients (M57, M21 and M54) undergoing routine total hip replacement surgery at Southampton General Hospital under informed consent and compliant to the Southampton and South West Hampshire Local Research Ethics Committee (LREC194-199). The specific methodology was described in section 2. Following extraction, cells were magnetically separated on the basis of exhibiting the Stro-1 epitope. The protocol for the Magnetic Activated Cell Sorting (MACS) isolation was described in section 2.2. After culturing cells for 14 days, cells underwent osteoinduction by the addition of L-Ascorbic acid 2-phosphate sesquimagnesium salt hydrate (100 µM) and dexamethasone (10 nM) to the basal culture media constituted by α-MEM phenol free, 10% FCS, and 1% penicillin/streptomycin. Cells were incubated at 37 °C, and 5% CO_2 in a humidified atmosphere.

3.3.1.2 Alkaline phosphatase staining

BM cells from from three patients (M57, M21 and M54) were cultured in a monolayer, fixed in 95% ethanol and stained for alkaline phosphatase (ALP). The detailed protocol can be found in section 2.3.

3.3.1.3 Molecular analysis

The method for RNA extraction, cDNA synthesis, and RT-PCR was described in section 2.4. RNA was extracted from three patients (M57, M21 and M54). The list of primers employed in this study is detailed in Table 3.

3.3.2 Stro-1 characterisation experiments

3.3.2.1 Cell culture

HK cells (human follicular dendritic cell line) (Kim, Zhang *et al.*, 1994) were cultured in IMDM (Iscove's Modified Dulbecco's Media, Lonza), 10% FCS, 1%

sodium pyruvate (Life Technologies), 1% non-essential amino acids solution (NEAA, Life Technologies), 1% penicillin/streptomycin and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere until cultures were sufficiently expanded for further investigation and testing.

MG63 cells (human osteosarcoma) were expanded in D-MEM (Dulbecco's modified eagle's medium, Lonza), 10% FCS and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere.

BXPC3 cells (pancreatic adenocarcinoma) were kindly donated by Dr. Veronika Jenei, Cancer Sciences Research Unit, University of Southampton. BXPC3 cells were cultured as monolayers in RPMI-1640 (Roswell Park Memorial Institute, PAA) media supplemented with 10% FCS and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere.

C28I2 cells (immortalised human chondrocytes) were expanded in DMEM/F12 supplemented with 5% FCS, 1% insulin-transferrin-selenium (Sigma-Aldrich,), 1% of penicillin/streptomycin, and 100µg/mL of ascorbic acid (Sigma-Aldrich) and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere.

Raji cells (B cells) were kindly donated by the Dr. Ali Roghanian, Cancer Sciences Research Unit, University of Southampton. Raji cells were cultured in suspension in RPMI-1640 media supplemented with 10% FCS and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere.

721.221 cells (B cells) were kindly donated by PhD student Dannielle Wellington the Cancer Sciences Research Unit, University of Southampton. 721.221 were expanded in suspension in RPMI-1640 media supplemented with 10% FCS and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere.

PC3 cells (prostate cancer cells) were kindly donated by Dr. Samantha Larkin Cancer Sciences Research Unit, University of Southampton. PC3 cells were cultured in suspension in RPMI-1640 media supplemented with 10% FCS and incubated at 37 ˚C and 5% CO_2 in a humidified atmosphere.

DU145 cells (prostate cancer cells derived from brain cancer metastasis) were kindly donated by Dr. Samantha Larkin, Cancer Sciences Research Unit, University of Southampton. DU145 cells were grown in suspension in RPMI-

1640 media supplemented with 10% FCS and incubated at 37 °C and 5% CO_2 in a humidified atmosphere.

3.3.2.2 Cell lysis for Stro-1 WB

HK cell pellets were lysed in Pierce immunoprecipitation (IP) (Pierce Biotechnology Inc.) buffer constituted by 0.025 M Tris, 0.15 M NaCl, 0.001M EDTA, and nonidet P40 1% (v/v), 5% glycerol, pH 7.4. The cell lysis step was performed in the presence of a cocktail of protease inhibitors (1:100, Cell Signaling Technology) and incubated on ice for 30 min. Cell lysates were sonicated 3 times at an amplification of 20% for 1 min each time. An aliquot of the lysate was removed and named whole cell lysate (WCL). The remaining portion of the lysate was centrifuged at 13,000 rpm for 20 min in an Eppendorf centrifuge (Eppendorf centrifuge 5415 R). The supernatant (soluble fraction (SOL)) was transferred to another clean Eppendorf tube. The remaining pellet (insoluble fraction (INS)) was re-suspended in 30 µL of lysis buffer and sonicated 3 times at an amplification of 20% for 1 min each time. Protein content within the three cell fractions: WCL, SOL and INS was further analysed by bicinchoninic acid assay (BCA) and WB according to the protocols in sections 2.8 and 2.9 respectively (Figure 17). Different lysis buffers were tried and results are detailed in Appendix 1, section 9.7.

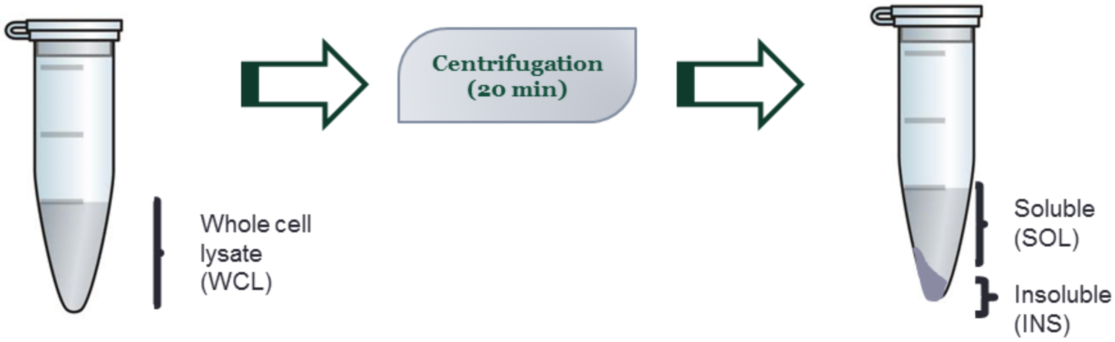


Figure 17: WCL, SOL and INS fractions.

HK cells were lysed with the addition of lysis buffers (Table 20). Upon centrifugation 2 distinct cell fractions were formed: the supernatant or soluble (SOL), and the pellet or insoluble (INS).

3.3.2.3 *Fluorescent immunostaining*

The protocol followed was described in section 2.6. HK, MG63, BXPC3 and C28I2 cells are adherent cells and therefore immunostaining was performed on a monolayer; whilst Raji and 721.221 cells are non-plastic adherent and immunostaining was performed in suspension.

3.3.2.4 *Immunoblotting: Western Blotting*

50 µg of extracted proteins were analysed by gel electrophoresis combined with Dithiothreitol (DTT, Cell Signalling) and 3x red loading buffer (Cell Signalling). Precast gels '10 well combo 30 µL any kD precast gels' (Mini-PROTEAN TGX; BioRad) were used. Gel electrophoresis was run for 40 min at 150 V with a Bio-Rad power supply. The gels were transferred onto polyvinylidene fluoride (PVDF) immobilon-FL transfer membranes (Millipore) for 1.5 h at 75 V. The membrane was incubated at RT for 1 h in blocking buffer (5% (w/v) non-fat milk powder in PBS/tween. The membrane was subsequently incubated with primary antibody overnight at 4 °C, and secondary antibody added for 1 h at RT. Antibody dilution, diluent and incubation time for each antibody followed description as per Table 5. Following each incubation with primary and secondary antibody solutions, samples were washed with PBS/tween 3 times for 5 min. A further wash in PBS only was performed prior to imaging. Membranes were imaged with Odyssey fluorescent instrument (LI-COR, Biosciences) or with a bioimager (BD systems). Densitometry analysis were run using Image Studio™ 4 software (Li-Cor).

3.3.2.5 *Statistics*

Statistical analysis between the WCL, SOL and INS densitometry data were performed using ANOVA in the statistics software GraphPad Prism version 6.07. The WB experiments were performed using 4 separate populations of cells. Values of $p<0.05$ were considered significant. The experiment was run four times.

3.3.2.6 *Immunoprecipitation*

HK and Raji cells were cultured as indicated in section 3.3.1.1. Cells were centrifuged at 13,000 rpm, washed 4 times in ice cold PBS, re-suspended in 1 mL of Ripa buffer (Table 20) supplemented with a cocktail of protease inhibitors (1:100) and disrupted through a 21 gauge needle (BD Microlance ™ 3). Cell lysates were pre-cleared with protein-L beads (Santacruz) for 2 h at 4 ˚C and then incubated with Stro-1 probed protein-L beads overnight. Probed protein-L beads were obtained by incubating the Stro-1 antibody with the beads for 2 h at 4 ˚C. Samples were recovered and washed 5 times in Ripa buffer for 5 min each, and centrifuged at 10,000 rpm at 4 ˚C. Three different elution buffers were employed including i) glycine pH 2.8, ii) IP Pierce elution buffer, iii) 3x red loading buffer. Samples were incubated with the elution buffers for 15 min at RT. Glycine pH 2.8 and IP Pierce elution buffers were neutralised with 1 M Tris pH 8.0 and 9.7 respectively. For elution with 3x red loading buffer no neutralisation was necessary, the sample was boiled at 95 ˚C for 5 min. Samples were analysed by gel electrophoresis and WB.

Immunoprecipitation was also attempted employing the MACS protocol for cell separation (section 2.2) on HK cells WCL (Appendix 1, section 9.4).

3.3.3 Aptamer selection

3.3.3.1 Cell culture

MG63 cells (human osteosarcoma) were expanded in D-MEM (Dulbecco's modified eagle's medium, Lonza) with 10% foetal calf serum (FCS) (Sigma) and incubated at 37 °C in 5% CO_2 to expand the population required for subsequent testing. Cells were seeded at a density of 20,000 cells/cm^2 24h prior to the aptamer selection. Cells were seeded in different tissue culture plastic areas according to the stage of the selection process (Table 7).

Raji cells (B cells) were kindly donated by the Dr. Ali Rogan, Cancer Sciences Research Unit, University of Southampton Raji cells were cultured in suspension in RPMI-1640 (Roswell Park Memorial Institute, PAA) media supplemented with 10% FCS (Sigma) and incubated at 37 °C in 5% CO_2. The required cell number for the selection (Table 7) was collected from T25 tissue culture flasks.

3.3.3.2 Reagents for the aptamer selection

Table 6: Reagents employed during the aptamer selection.

Solution name	Solution content
TBE buffer	55 g boric acid (Sigma), 7.5 g Ethylenediaminetetraacetic acid disodium salt concentrate pH 8 (Sigma) in a litre of PBS.
Loading buffer	20% Ficoll 400 (GE Healthcare), 0.1 M Ethylenediaminetetraacetic acid disodium salt concentrate pH 8 (Sigma), 1% sodium dodecyl sulphate (Sigma).
Binding buffer	DPBS 500 mL (Lonza), 2.25 g D-(+)-glucose (Sigma), 25 mg tRNA, 0.5 g BSA (Lonza), 2.5 mL $MgCL_2$ (1M) 5 mM final concentration (Sigma), 0.5 mL $CaCl_2$ (1M) 1 mM final concentration (Sigma).
Wash buffer	DPBS 500 mL (Lonza), 2.25 g D-(+)-glucose (Sigma), 2.5 mL $MgCL_2$ (1M) 5 mM final concentration (Sigma), 0.5 mL $CaCl_2$ (1M) 1 mM final concentration (Sigma).

2x Folding buffer	DPBS 500 mL (Lonza), 5 mL Mg Cl_2 1 M 10 mM final concentration (Sigma), 1 mL $CaCl_2$ 1 M 2 mM final concentration (Sigma).
Transcription mix (50 μL reaction)	5 μL 10x transcription buffer (NEB), 2 μL 1 M DTT (Melford), 16.25 μL Milli-Q water (Millipore), 2.5 μL 100 mM 2"fluoroNTP (TriLink), 1.25 μL DMSO (ApplyChem), 2.5 μL T7 polymerease enzyme (Epicentre biotechnology), 0.5 μL YIP (Yeast inorganic pyrophosphatase) (NEB), 20 μL template.
DNase mix	6 μL 10x DNase buffer (NEB), 3 μL water Milli-Q water (Millipore), 1 μL DNase (NEB) (10 μL of DNase mix is added to transcription mix after the incubation is complete to degrade remaning DNA).
Reverse-Transcription mix	10 μL 5x Reverse transcription buffer (NEB), 4 μL 10 mM DNTPs (KAPA), 2.5 μL of 50 mM reverse primer (IDT), Milli-Q water (Millipore), 1 μL of reverse transcriptase enzyme (NEB), 30 μL of cells lysed in quick extract (CAMbio).
PCR mix for the robot (for the cycle course)	16 μL 5x KAPAZG buffer B 1.5 mM Mg at 1x (Kapa Biosystems), 16 μL 5x KAPA enhancer 1 KA5010 (Kapa Biosystems), 1 μL 10 mM KAPA dNTPMix (Kapa Biosystems) 0.6 μL T7 forward selection primer (TriLink) (5' TTC AGG TAA TAC GAC TCA CTA TAG GGA AGA GAA GGA CAT ATG AT 3'), 0.6 μL reverse primer (TriLink) (5' TCA AGT GGT CAT GTA CTA GTC AA 3'), 0.25 μL KAPA2G Robust DNA Polymerase 5U/μL 250 U (Kapa Biosystems).
RNA purification	108 μL Axygen ® AxyPrep Magnetic Bead (FisherBiotech), 27 μL 2-Propanol, (isopropanol) (Sigma). 70% ethanol is used during RNA purification step subsequent to RNA production during transcription overnight.
Manual PCR (DNA amplification)	12 μL buffer B (KAPA), 12 μL enhancer (KAPA), 1.2 μL DNTPs (KAPA), 0.6 μL forward T7 primer (IDT), 0.6 μL reverse primer (IDT), 0.24 μL enzyme (KAPA), Milli-Q water

	(Millipore), 30 µL template (robot archived sample from transcription).
Native gel mix	10 mL 10% polyacrylamide (Sigma), 100 µL 10% Ammonium persulfate (Sigma), 10 µL TEMED (Sigma).
RNA gel mix	10 mL 10 % polyacrylamide Urea gel (Sigma), 100 µL 10% Ammonium persulfate (Sigma), 10 µL TEMED (Sigma).
Further reagents for the automated system	40 µL of Vapour lock liquid vapour barrier (Qiagen) was added to the PCR mix together with the samples to prevent evaporation during the PCR cycle course.

3.3.3.3 Aptamer library

3.3.3.3.1 DNA random library

A 40 bp random region DNA library was acquired from TriLink (5' TAG GGA AGA GAA GGA CAT ATG ATN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN TTG ACT AGT ACA TGA CCA CTT GA 3') (TriLink Biotechnologies). The library was amplified in a PCR reaction to produce multiple copies of each molecule.

3.3.3.4 Purification of the Aptamer library

3.3.3.4.1 DNA extraction from acrylamide gels

The TriLink DNA library (TriLink DNA template (N40) TD-DC01A (100 µM)) was loaded onto a 10% acrylamide gel and run at 300 V for 35 min, stained with ethidium bromide solution (10mg/mL, BioRad) and visualised in a UV light box. DNA bands were sectioned and transferred into a cylindrical valve for the liquid extraction via electro dialysis. Each valve contained a DNA band and a membrane (SnakeSkin Dialysis Tubing, 10K MWCO, Pierce). The system was run for 30 min at 200 V. Extracted DNA was collected in eppendorf tubes and precipitated with 100% ethanol (3.3.3.4.2).

3.3.3.4.2 Ethanol precipitation

Sodium acetate was added to the eluted DNA (3.3.3.4.1) to a final concentration of 10%. $MgCl_2$ was added to a final concentration of 0.01 M. Two volumes of 100% ethanol were added, vortexed and incubated on ice for 30 min. Sample was centrifuged for 45 min at 15,000 rpm in a Heraeus Fresco 21 centrifuge (Thermo scientific). Ensuing centrifugation, supernatant was discarded and the DNA pellet was washed in 70% ethanol twice and centrifuged at 13,000 rpm for 2 min. Supernatant was discarded and the sample was left to dry for 1 min on a heating block at 55 °C. Subsequently the DNA was resuspended in TBE buffer (Table 6).

3.3.3.5 Overview of the selection process

Each round of selection lasted approximately 15 h. The selection process initially consisted in incubating the MG63 cells with the aptamer sequences (3.3.3.4.2), reverse transcribing the RNA aptamers into complementary DNA aptamers with a robot, amplifying the DNA and verifying the results by gel electrophoresis (Figure 18). The amplified DNA was transcribed into RNA and purified overnight by an automated process utilising a liquid handling robot (Biomek 3000 Beckam Coulter equipped with T robot (Biometra). The subsequent day, the RNA obtained was verified on a 10% denaturing urea-acrylamide gel. The RNA aptamers were employed for the following selection round. For the first two rounds of selection only the positive target (MG63) was used whilst for the consequent rounds the introduction of Raji cells for negative selection was adopted prior to the MG63 incubation. In other words, in every step the aptamer pool was filtered negatively for one aliquot (or more) of Raji cells, then the remaining aptamers were incubated with MG63 cells. The scope of the selection process was to enrich the aptamer library with high affinity specific sequences (against MG63) by increasing the stringency of the process at every selection round. Increasing the stringency of the process consisted of increasing the number of negative cells, increasing the incubation time with the negative target (Raji cells), decreasing the incubation time with the positive target (MG63), and gradually increasing washing steps and duration. The SELEX process normally consists of 12 selection rounds but it can vary depending on applications. The details of each selection round are found in Table 7.

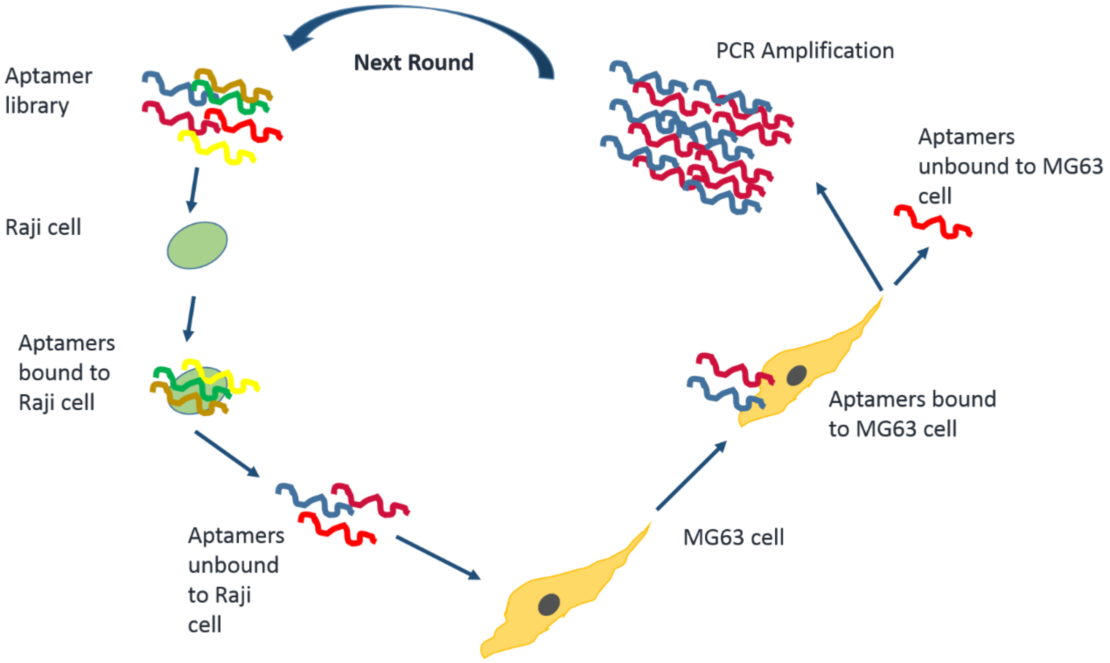

Aptamer library
Next Round
PCR Amplification
Aptamers unbound to MG63 cell
Raji cell
Aptamers bound to Raji cell
Aptamers bound to MG63 cell
MG63 cell
Aptamers unbound to Raji cell

Figure 18: Aptamer process overview.

A library of aptamer molecules was incubated with Raji cells (negative target) and washed. The unbound fraction of aptamer molecules was incubated with MG63 (positive target). The aptamer sequences that attached to MG63 were amplified by PCR. The process was started again.

3.3.3.6 Aptamer incubation with cells and cell lysis

The entire selection process was performed at 4 ˚C. Raji cells were collected and 1 mL of cell suspension was counted; calculations were performed to provide aliquots of Raji cells similar to MG63. Raji cells were conditioned on ice for 5 min, then washed three times in PBS (5,000 rpm for 5 min at 4˚C) and blocked in binding buffer for 5 min on ice. Raji cells were resuspended with aptamers to a final volume of 500 µL and incubated for the required amount for that round of section (Table 7). Raji cells with aptamers were centrifuged at 5,000 rpm for 5 min at 4 ˚C and the supernatant containing the unbound aptamers was collected and incubated with MG63 cells previously washed and blocked. Incubation time varied at each round of selection and details are provided in Table 7. Both cell lines were washed, pelleted and lysed in 30 µL of quick extract (QE) (Epicentre biotechnologies) and vortexed for 1 min.

Table 7: Details of each selection round.

Round number	MG63 growth area used	Raji cells equivalent in MG63 growth area	Incubation time		wash MG63	wash Raji
			MG63	Raji		
1	T75		1 h		5x 1 min	
1.1	T25		1 h		5x 1 min	
2	T25		45 min		5x 1 min	
3	2x 6 well	1x 6 well	30 min	5 min	7x 1 min	7x 1 min
4	1x 6 well	2x 6 well	15 min	5 min	7x 1 min	7x 1 min
5	1x 12 well	2x 6 well	10 min	5 min	7x 1 min	7x 1 min
5.1	1x 12 well		10 min	5 min	7x 1 min	7x 1 min
5.1	1x 12 well	1x 12 well	10 min	5 min	7x 1 min	7x 1 min
5.1	1x 12 well	1x 6 well	10 min	5 min	7x 1 min	7x 1 min
5.2	1x 12 well		5 min	5 min	7x 1 min	7x 1 min
5.2	1x 12 well	1x 12 well	5 min	5 min	7x 1 min	7x 1 min
6	1x 12 well	1x 12 well	5 min	5 min	7x 1 min	7x 1 min
6	1x 12 well	2x 12 well	5 min	5 min	7x 1 min	7x 1 min
7	1x 12 well	2x 12 well	5 min	5 min	7x 1 min	7x 1 min
7	1x 12 well	3x 12 well	5 min	5 min	7x 1 min	7x 1 min
5 and 7	1x 12 well	1x 12 well	5 min	5 min	7x 1 min	7x 1 min
5 and 7	1x 12 well	1x 12 well	5 min	5 min	7x 1 min	7x 1 min
6 and 8	1x 12 well	1x 12 well	5 min	5 min	7x 1 min	7x 1 min
6 and 8	1x 12 well	1x 12 well	5 min	5 min	7x 1 min	7x 1 min

3.3.3.7 Automated reverse transcription, cycle course PCR

Cell lysate was loaded onto a robot (Biomek 3000 Beckam Coulter) equipped with T robot (Biometra). The machine was programmed to automatically reverse transcribe the RNA aptamers into DNA aptamers and amplify the DNA for 2, 4, 6 and 8 cycles course. Samples for each cycle course were checked on a 10% polyacrylamide gel. The reagents for the reverse transcription are detailed in Table 6.

3.3.3.8 Gel electrophoresis: amplified DNA

Amplified DNA of Raji and MG63 bound aptamers was loaded with loading on a 10% acrylamide gel. 10 bp DNA ladder (Life Technologies) was run in one well. Gel was run for 1 h at 200 V. Gel was then stained in ethidium bromide solution (10mg/mL, BioRad) for 5 min and then imaged with a UV detector.

3.3.3.9 Manual PCR: DNA amplification

The DNA was amplified using a manual PCR reaction (Table 6). The PCR thermocycler conditions were the followings: 95 ˚C for 270 s (hot start), 95 ˚C for 30 s (denature), 65 ˚C for 30 s (anneal), 72 ˚C for 30 s (extend) and 72 ˚C for 240 s (final extension). The number of cycles to amplify the DNA was dictated by the results of the automated cycle course PCR.

3.3.3.10 Overnight automated transcription

Following the DNA amplification and visualisation on a native gel, the DNA was fed to a liquid handling automated system (Biomek 3000 Beckam Coulter equipped with T robot (Biometra)) for overnight transcription and RNA purification with the Axygen® AxyPrep Magnetic Bead Purification Kit following manufacturer's instructions.

3.4 Results

3.4.1 Characterisation of bone marrow extracted cells

SSC extracted from the BM show a spindle shape fibroblastic-like morphology when grown on tissue culture plastic (Figure 19).

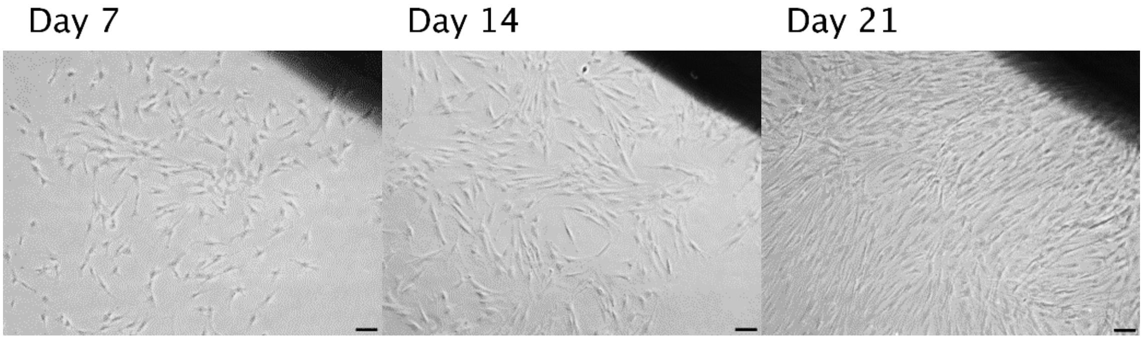


Figure 19: Human bone marrow extracted stromal cells in culture.

Images were taken at the same point on the flask at days 7-14 and 21 following isolation. The black mark delineates the area where the picture was taken. Scale bar: 100 µm.

3.4.1.1 Alkaline phosphatase expression increases in the osteogenic phenotype

The bone marker ALP was examined over 21 days by imaged-based analysis using CellProfiler (Figure 20). ALP specific activity was significantly increased following osteogenic culture in comparison to the basal control. Histological analysis confirmed biochemical data shown in Figure 24.

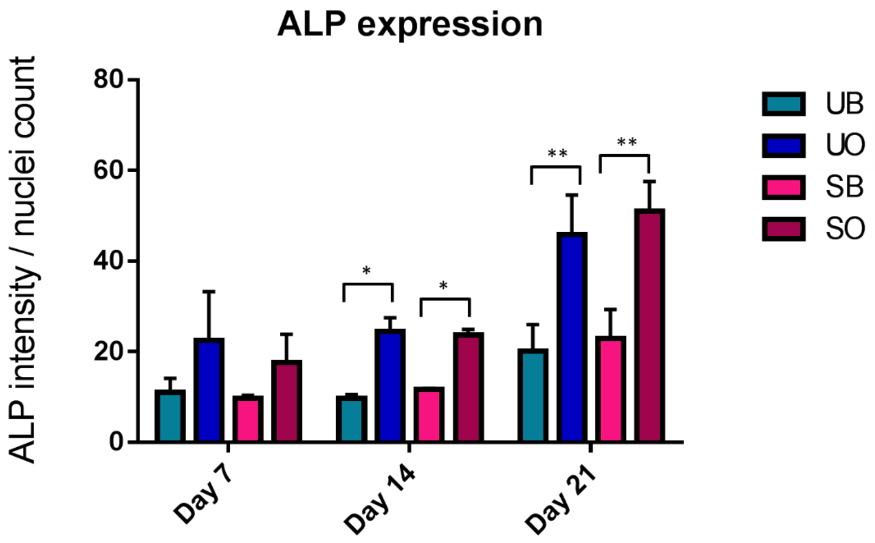


Figure 20: Alkaline phosphatase expression based on image-based analysis.

Graphical representation of mean and standard deviation of alkaline phosphatase expression of Stro-1 sorted and unsorted cells cultured under basal and osteogenic conditions, of three patients (M57, M21, M54). The analysis was imaged-based using CellProfiler. Images were taken at day 7, 14 and 21 from the addition of osteogenic media. Results between basal and osteogenic groups were analysed with two-way ANOVA comparison. (* p<0.05, ** p<0.01, n=3). UB: unsorted basal; UO: unsorted osteogenic; SB: sorted basal; SO: sorted osteogenic.

3.4.1.2 Molecular analysis demonstrated osteogenesis

ALP was used in this study as an indicator of bone formation and mineralisation. ALP expression was observed to increase over the 21 days culture with a significant increase in the osteogenic induced cultures in comparison to basal cultures. OCN gene expression was measured to compare samples in basal and osteogenic conditions over a period of 21 days. OCN is considered a late bone marker. OCN expression was found significantly upregulated in the osteogenic samples on day 21 with a distinctive increase in the sorted sample versus the unsorted (Figure 22). SOX9 gene expression was measured as a negative control for osteogenesis. SOX9, in fact is a chondrogenic marker, as such, expression significantly decreased in osteogenic samples by day 21 (Figure 23).

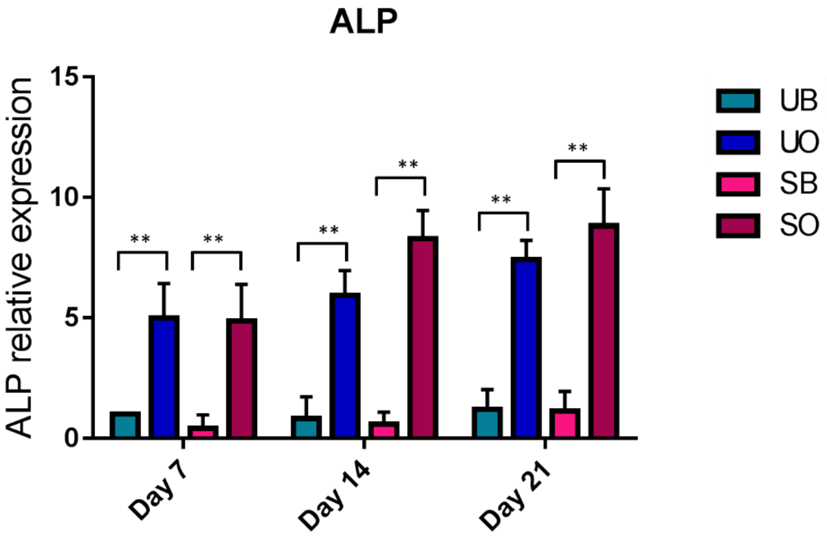


Figure 21: ALP gene expression.

ALP gene expression was measured in cells extracted from human bone marrow of three patients (M57, M21, M54), sorted for the Stro-1 marker and cultured under basal and osteogenic conditions. Gene expression was compared between basal and osteogenic treated samples at days 7, 14 and 21. ALP expression was normalised to GAPDH expression. Data presented as mean ± standard deviation; n=3. Results between basal and osteogenic groups were analysed with two-way ANOVA comparison. (* p<0.05, ** p<0.01, n=3). UB: unsorted basal; UO: unsorted osteogenic; SB: sorted basal; SO: sorted osteogenic.

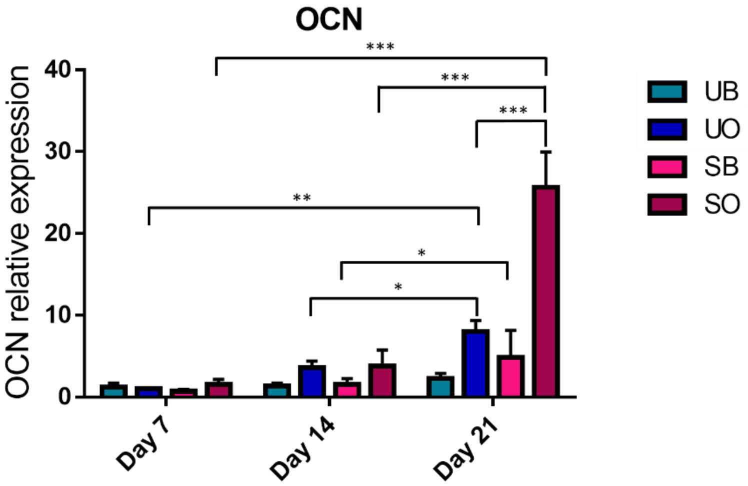


Figure 22: OCN gene expression.

OCN gene expression was measured in cells extracted from human bone marrow of three patients (M57, M21, M54), sorted for the Stro-1 marker and cultured under basal and osteogenic conditions. Gene expression was compared between basal and osteogenic treated samples at days 7, 14 and 21. OCN expression was normalised to GAPDH expression. Data presented as mean ± standard deviation; n=3. Results between basal and osteogenic groups were analysed with two-way ANOVA comparison. (* p<0.05, ** p<0.01, *** p<0.001, n=3). UB: unsorted basal; UO: unsorted osteogenic; SB: sorted basal; SO: sorted osteogenic.

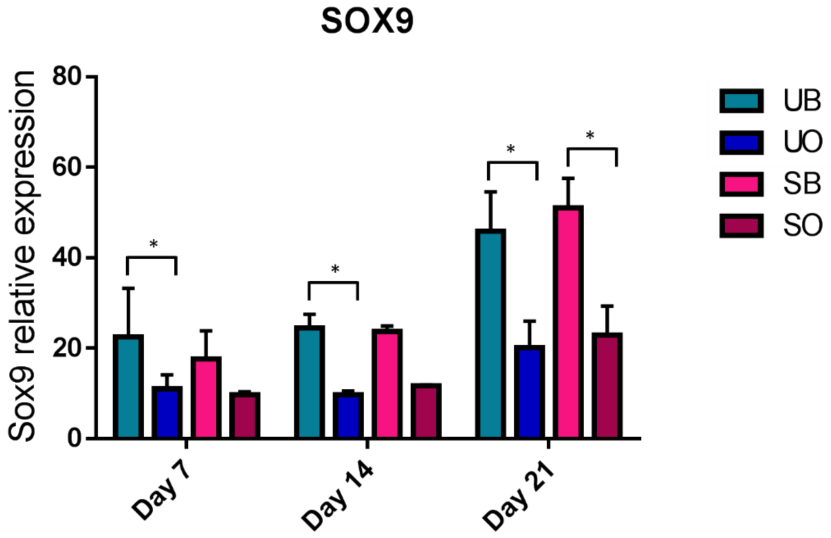


Figure 23: SOX9 gene expression.

SOX9 gene expression was measured in cells extracted from human bone marrow of three patients (M57, M21, M54), sorted for the Stro-1 marker and cultured under basal and osteogenic conditions. Gene expression was compared between basal and osteogenic treated samples at days 7, 14 and 21. SOX9 expression was normalised to GAPDH expression. Data presented as mean ± standard deviation; n=3. Results between basal and osteogenic groups were analysed with two-way ANOVA comparison. (* $p<0.05$, ** $p<0.01$, n=3). UB: unsorted basal; UO: unsorted osteogenic; SB: sorted basal; SO: sorted osteogenic.

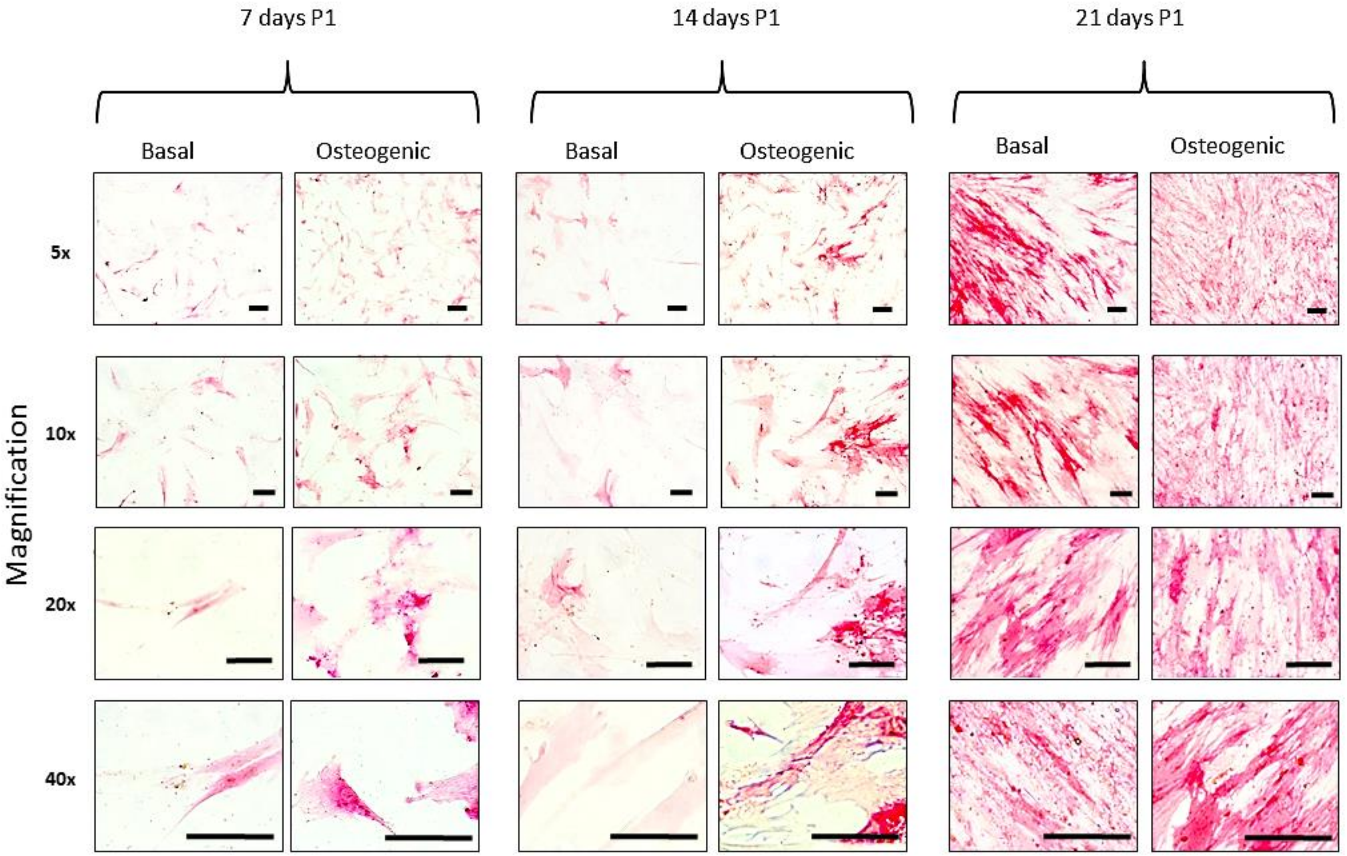

7 days P1
14 days P1
21 days P1
Basal
Osteogenic
Basal
Osteogenic
Basal
Osteogenic
Magnification
5x
10x
20x
40x

Figure 24: Alkaline phosphatase staining.

Human bone marrow stromal cells after 7, 14, and 21 days osteogenic induction, fixed in 90% ethanol and stained for alkaline phosphatase. Scale bar: 100 µm. P: passage.

3.4.2 Stro-1 characterisation experiments

3.4.2.1 Stro-1 positive and negative cell lines

Elucidation of Stro-1 antigen characteristics by WB required the use of positive and negative cells lines for the Stro-1 marker. Consequentially 6 cells lines: i) HK, ii) MG63, iii) BXPC3, iv) C28I2, v) Raji and vi) 721.221) were screened by fluorescence immunostaining to gain a wider insight of the Stro-1 expression. High level of Stro-1 expression was detected in HK, MG63, BXPC3 and C28I2 cells (Figure 25, Figure 26, Figure 27, and Figure 28 respectively). In contrast, no reactivity was detected in Raji and 721.221 (Figure 29 and Figure 30, respectively). Cell lines were also screened by flow cytometry (Appendix 1, section 9.2).

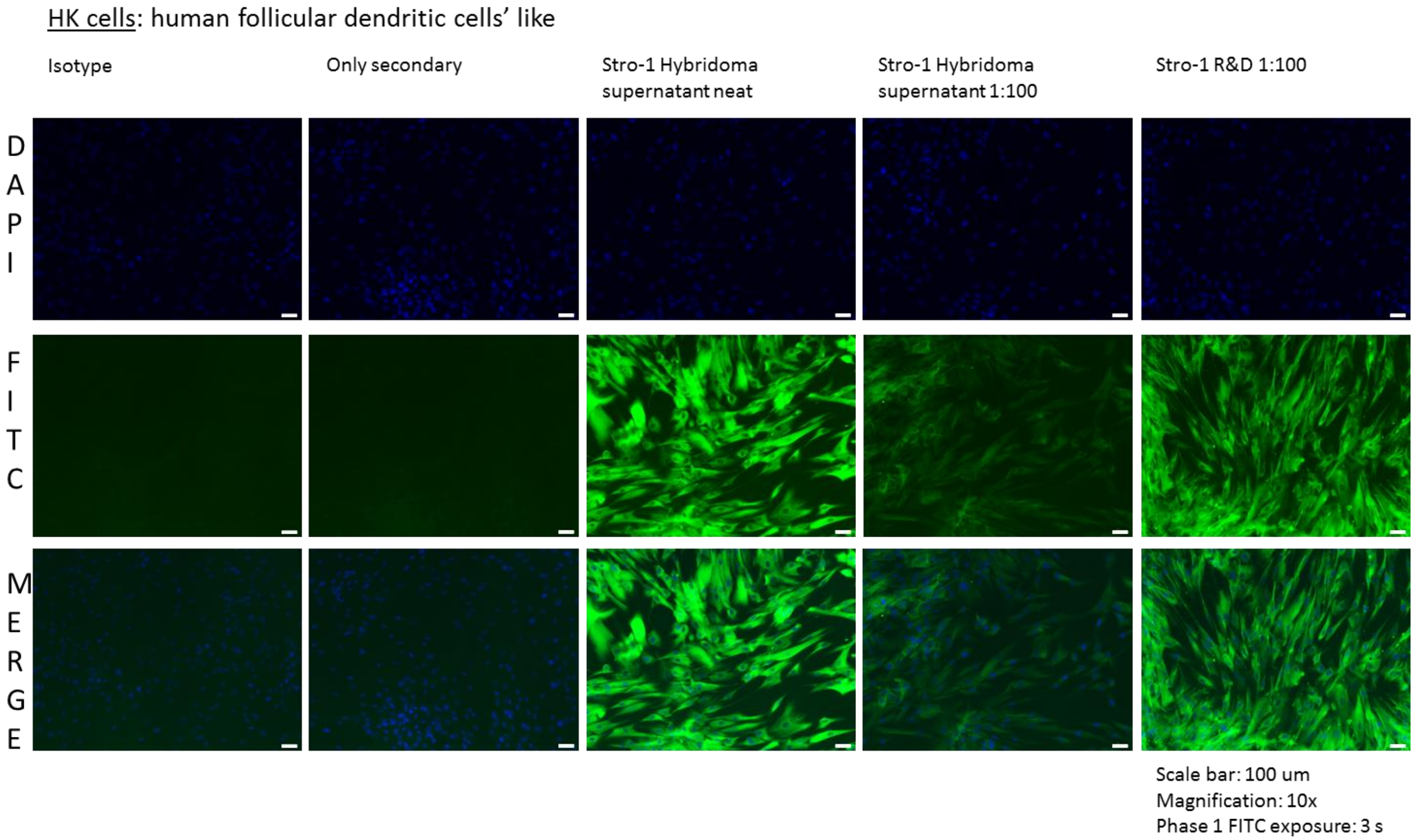

HK cells: human follicular dendritic cells' like
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
DAPI
FITC
MERGE
Scale bar: 100 um
Magnification: 10x
Phase 1 FITC exposure: 3 s

Figure 25: HK cells Stro-1 immunostaining.

Stro-1 expression (green) and cell nuclei (blue) in cells grown in monolayer. Scale bar: 100 µm.

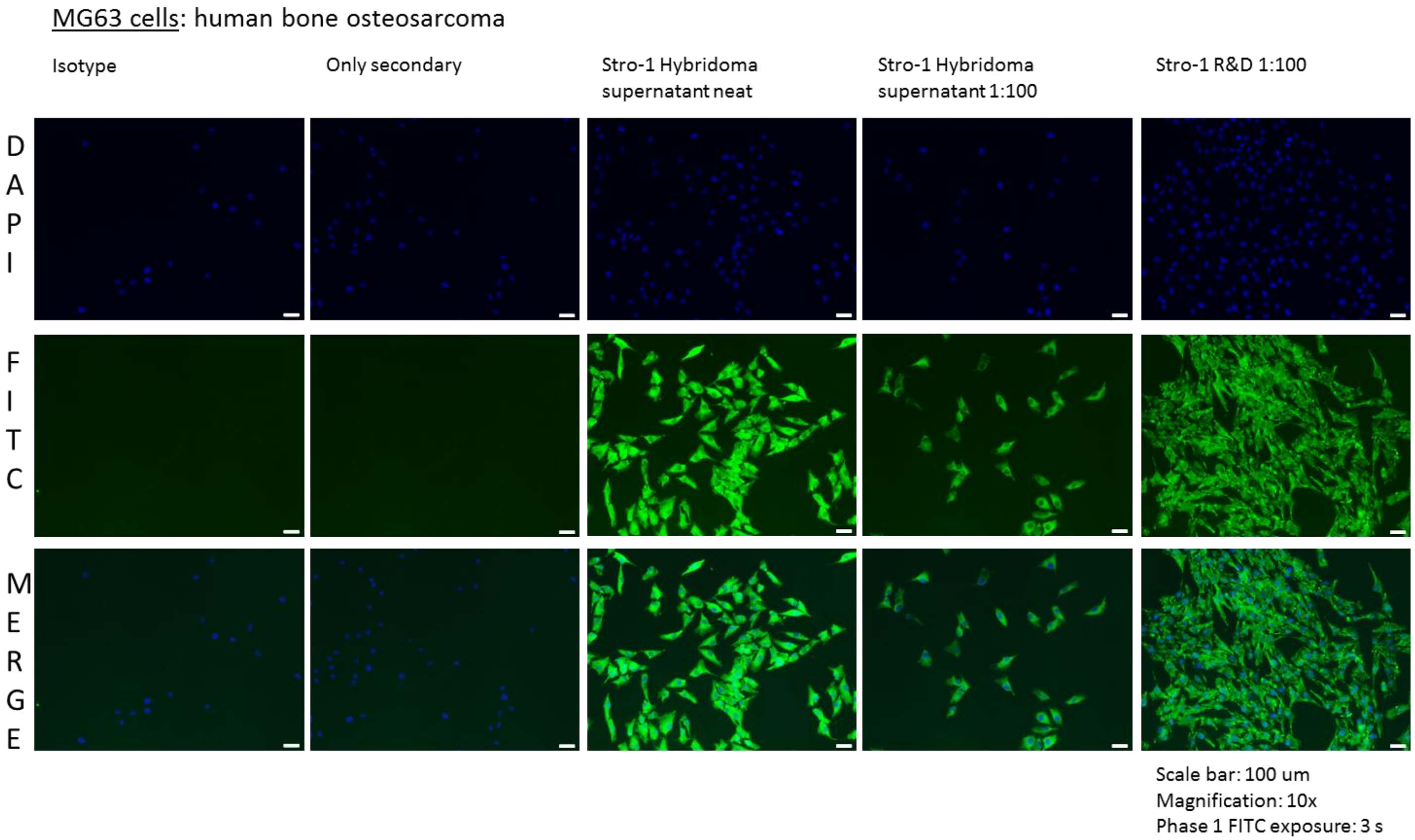

MG63 cells: human bone osteosarcoma
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
DAPI
FITC
MERGE
Scale bar: 100 um
Magnification: 10x
Phase 1 FITC exposure: 3 s

Figure 26: MG63 cells Stro-1 immunostaining.

Stro-1 expression (green) and cell nuclei (blue) in cells grown on a monolayer. Scale bar: 100 μm.

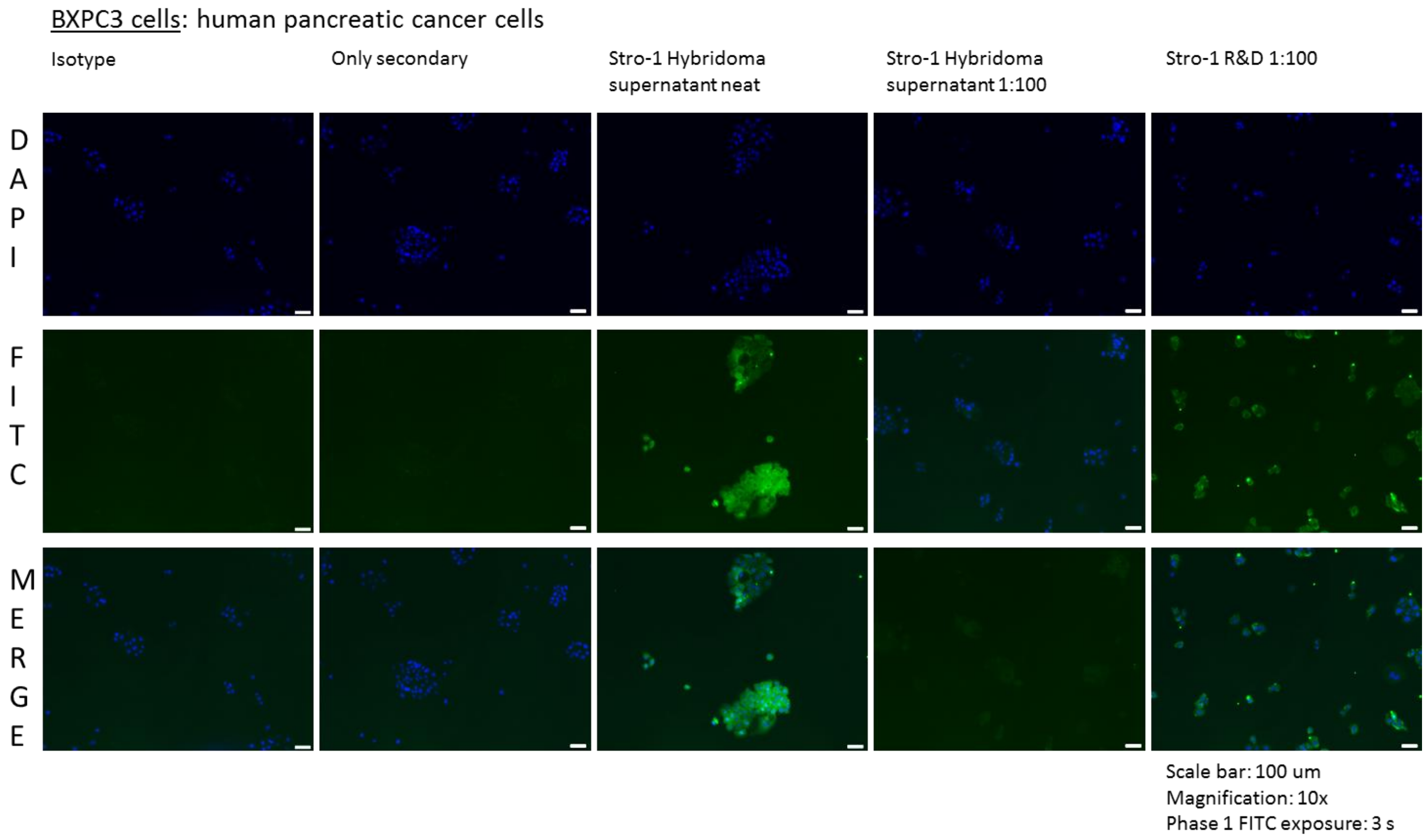

BXPC3 cells: human pancreatic cancer cells
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
DAPI
FITC
MERGE
Scale bar: 100 um
Magnification: 10x
Phase 1 FITC exposure: 3 s

Figure 27: BXPC3 cells Stro-1 immunostaining.

Stro-1 expression (green) and cell nuclei (blue) in cells grown on a monolayer. Scale bar: 100 μm.

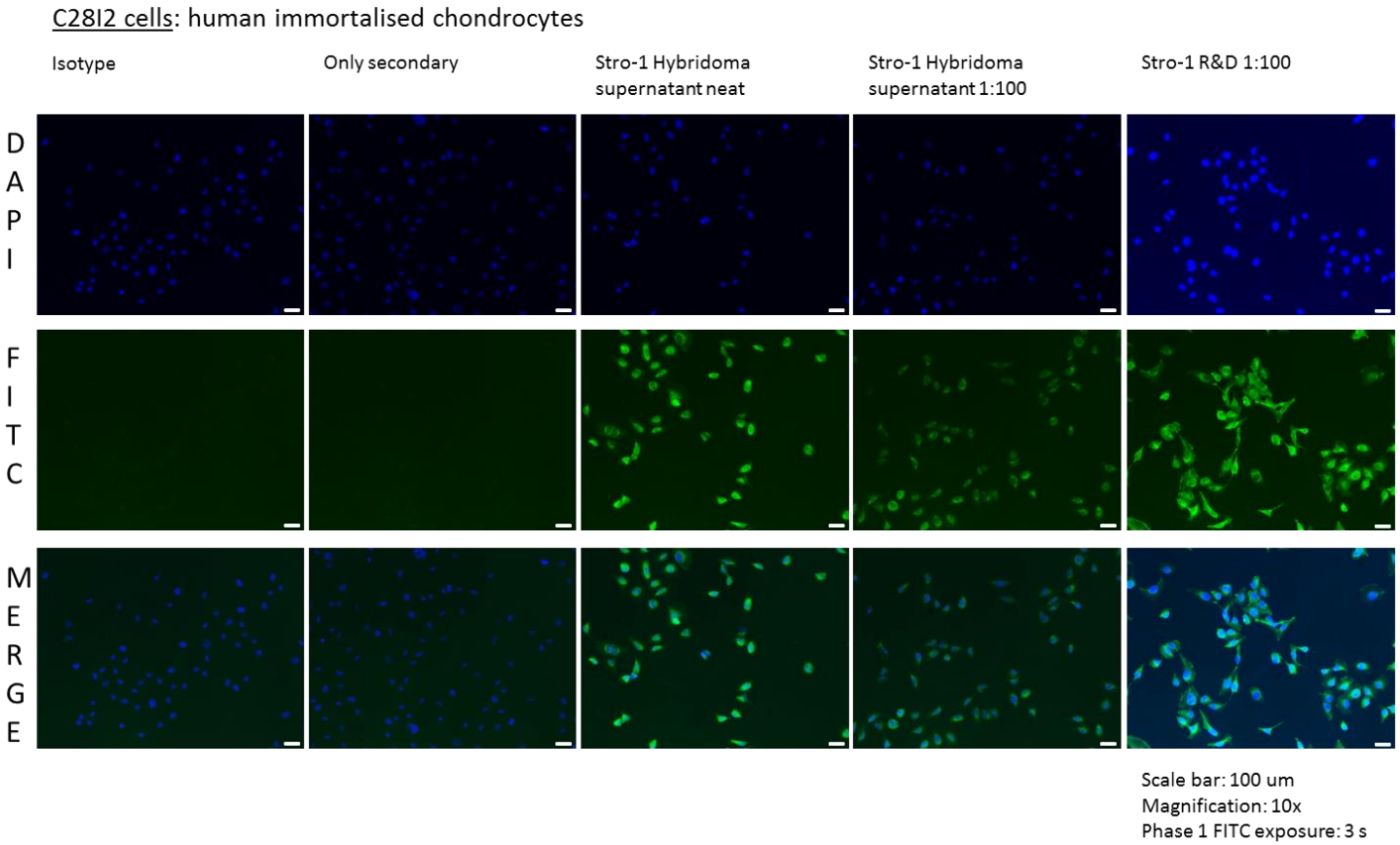

C28I2 cells: human immortalised chondrocytes
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
DAPI
FITC
MERGE
Scale bar: 100 um
Magnification: 10x
Phase 1 FITC exposure: 3 s

Figure 28: C28I2 cells Stro-1 immunostaining.

Stro-1 expression (green) and cell nuclei (blue) in cells grown on a monolayer. Scale bar: 100 µm.

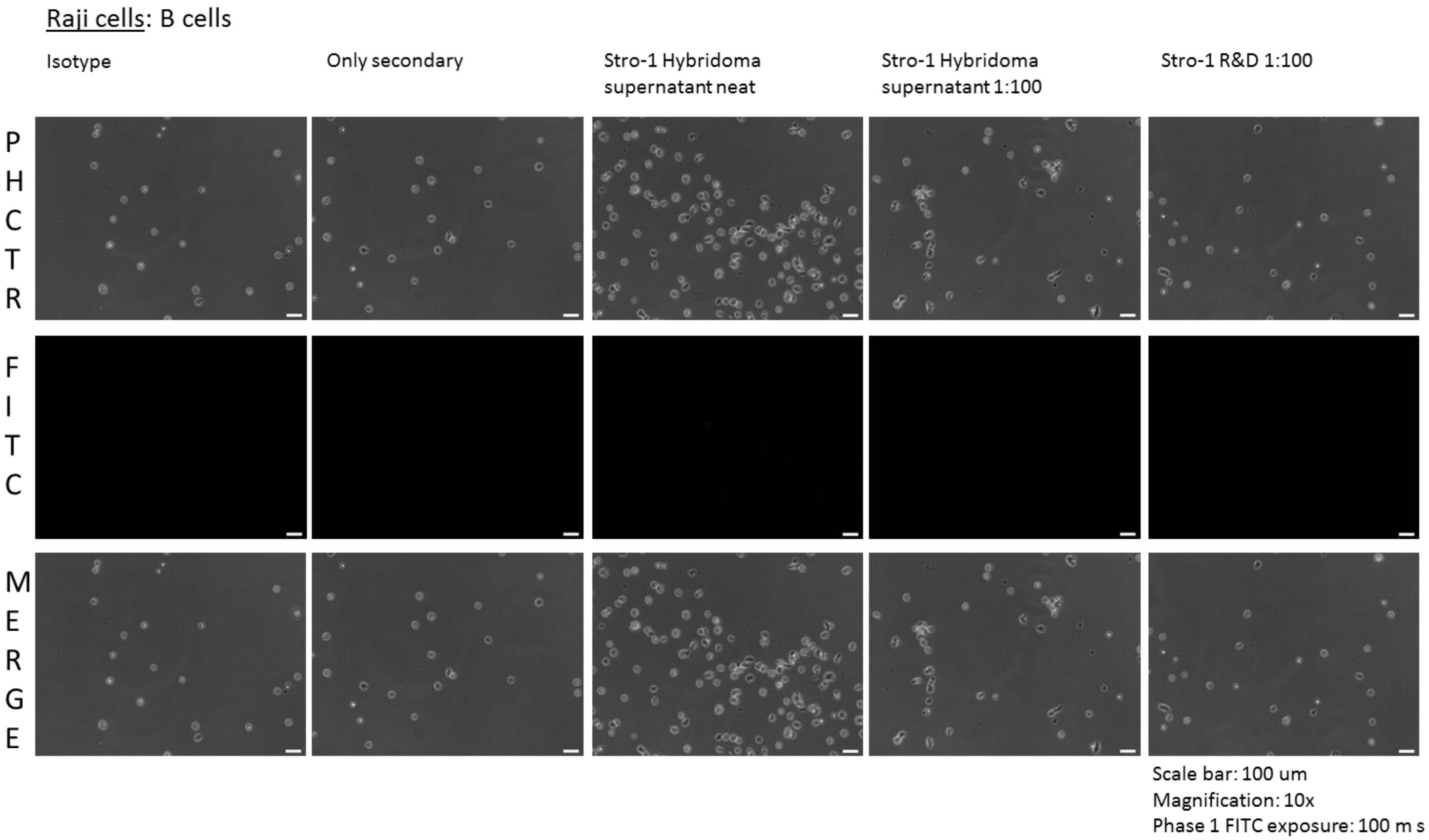

Raji cells: B cells
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
PHCTR
FITC
MERGE
Scale bar: 100 um
Magnification: 10x
Phase 1 FITC exposure: 100 m s

Figure 29: Raji cells immunostaining for Stro-1.

Stro-1 expression (green) in suspension cells previously analysed with flow cytometry. An IgM isotype and a test with just secondary antibody was run as a negative control. Stro-1 hybridoma supernatant (produced in-house) was used undiluted and diluted 1 in 100. Stro-1 antibody sourced from R&D systems was used diluted 1 in 100. Scale bar: 100 µm. Magnification 10 x

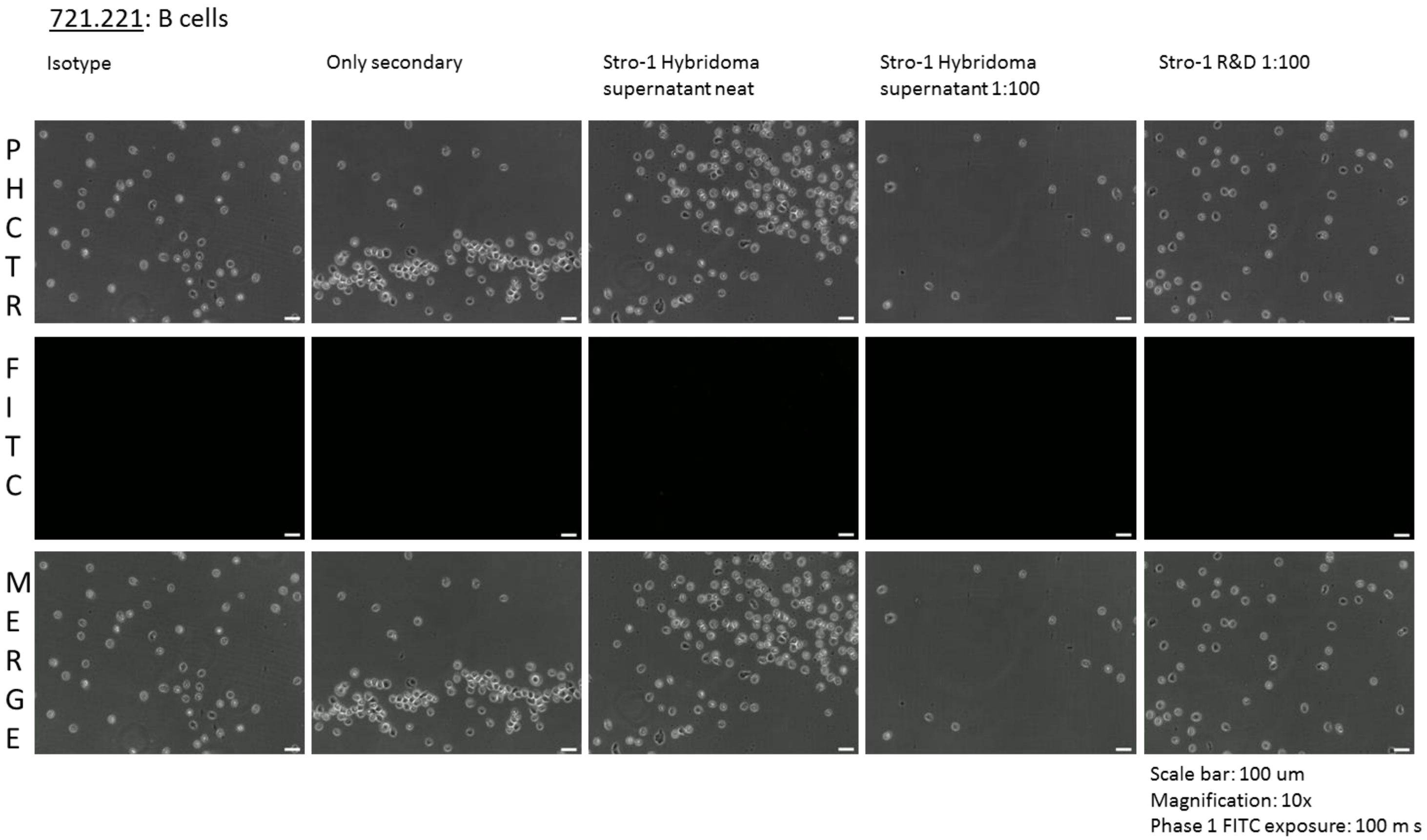

721.221: B cells
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
PHCTR
FITC
MERGE
Scale bar: 100 um
Magnification: 10x
Phase 1 FITC exposure: 100 m s

Figure 30: 721.221 cells immunostaining for Stro-1.

Stro-1 expression (green) in suspension cells previously analysed with flow cytometry. An IgM isotype and a test with just secondary antibody was run as a negative control. Stro-1 hybridoma supernatant (produced in-house) was used undiluted and diluted 1 in 100. Stro-1 antibody sourced from R&D systems was used diluted 1 in 100. Scale bar: 100 µm.

3.4.2.2 *Western Blotting analysis of the Stro-1 antigen*

WB technique was employed to detect and determine the Stro-1 antigen size. Results indicated that the Stro-1 antibody reacted to an epitope at 50 kDa. Expression was enriched in the INS fraction and diminished in the SOL fraction of the HK cell lysate (Figure 31). Similar results were obtained using the MG63 cells. In contrast, no binding was detected when the WCL of Raji and 721.221 was analysed (Figure 32).

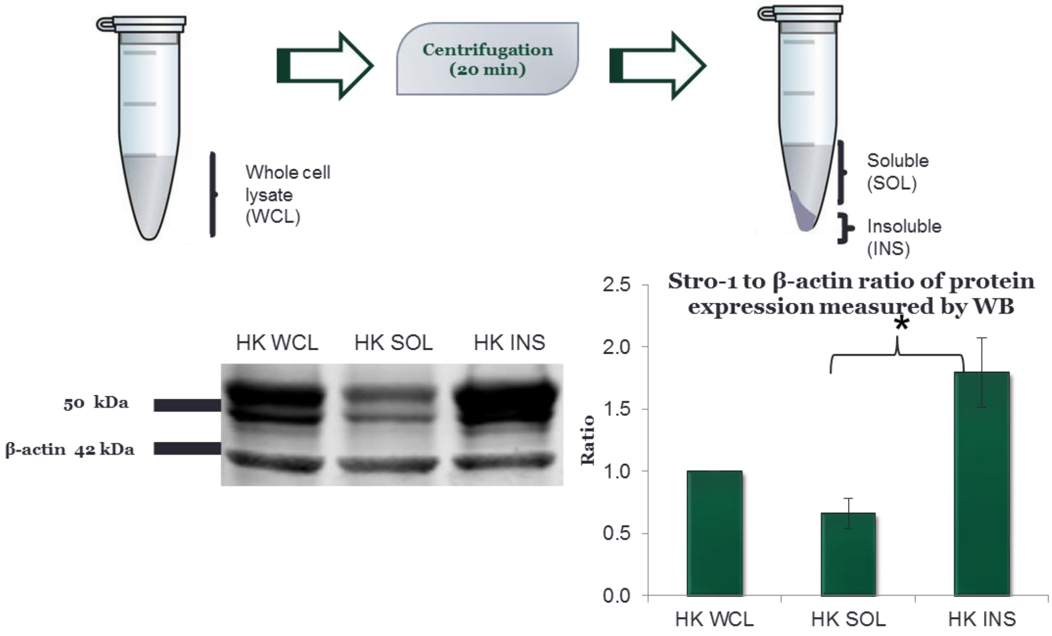


Figure 31: Isolation of the soluble and unsoluble fractions within cell lysate.

Following the lysis and centrifugation of HK cells a supernatant and a pellet were visible. The supernatant fraction was labelled soluble (SOL) fraction. The pellet fraction was resuspended in 30 µL of lysis buffer and sonicated; this was labelled insoluble (INS) fraction. Protein content was analysed by BCA assay. 50 µg of protein was loaded onto a gel and probed with the Stro-1 antibody overnight and then re-probed with IRDye 680 Goat α-mouse IgM (µ chain specific) secondary antibody (1:30,000) for 1 h. The blot was imaged using an Odyssey fluorescent imaging unit. Densitometry analysis was run with Image Studio software (Li-Cor). Stro-1 expression was compared to the loading control β-actin. Data presented as mean ± standard deviation (SD); n=4 (4 different blots). Significant results (* $p<0.05$) of expression differences were found between the SOL and INS fractions with two-way ANOVA comparison.

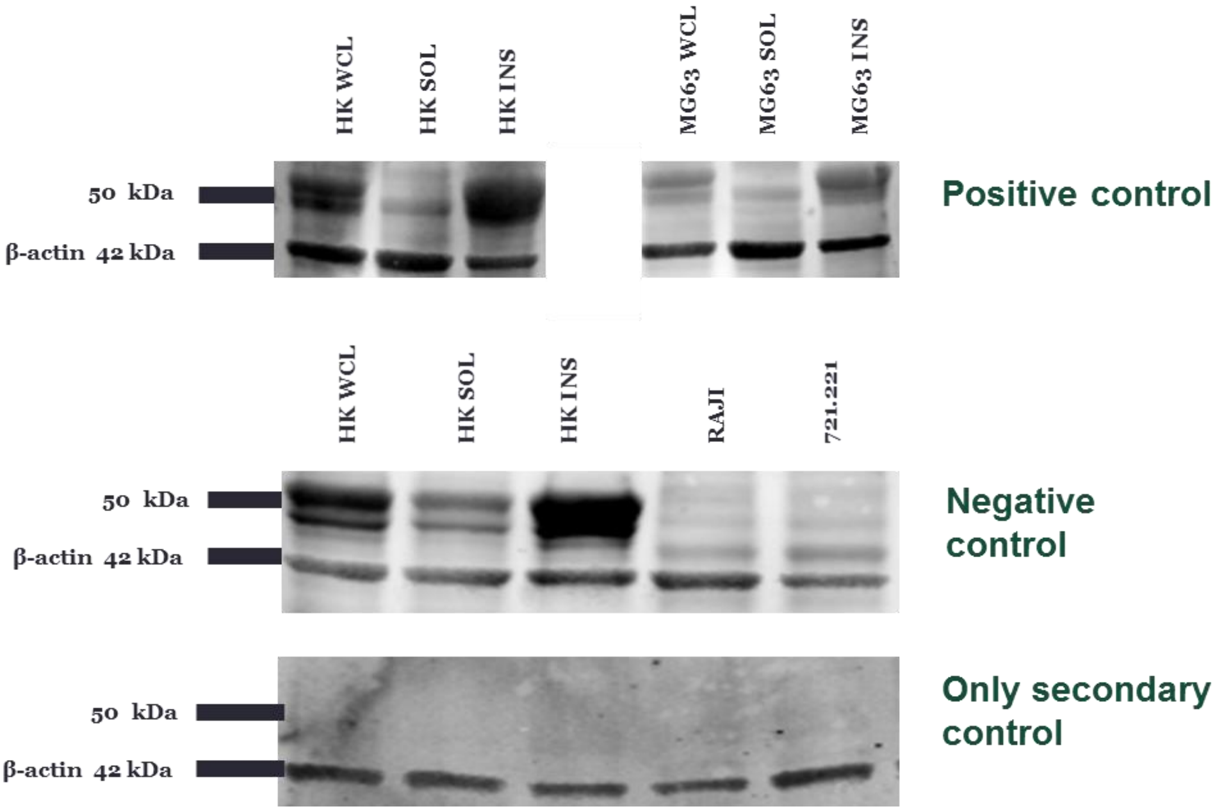

HK WCL
HK SOL
HK INS
MG63 WCL
MG63 SOL
MG63 INS
50 kDa
β-actin 42 kDa
Positive control
HK WCL
HK SOL
HK INS
RAJI
721.221
50 kDa
β-actin 42 kDa
Negative control
50 kDa
β-actin 42 kDa
Only secondary control

Figure 32: Soluble and insoluble fractions of HK cells probed with Stro-1 supernatant antibody.

Following cell lysis, the protein content was analysed by BCA assay. 50 µg of protein were loaded onto a gel and probed with Stro-1 antibody overnight and then re-probed with IRDye 680 Goat α-mouse IgM (µ chain specific) secondary antibody (1:30,000). β-actin was used as a loading control. The blot was imaged with an Odyssey fluorescent imaging unit. WCL: whole cell lysate, SOL: soluble, INS: insoluble.

3.4.2.3 Characterisation of the Stro-1 antibody derived by hybridoma supernatant

It is known that batch variations between each cycle of Stro-1 antibody production can occur. To eliminate variation and achieve consistency of results studies were undertaken to examine the Stro-1 antibody concentration (Figure 33). First the presence of IgM in the culture supernatant was verified (Figure 33, A). Secondly, the Stro-1 antibody concentration was compared to an IgM isotype. Correlation between the Stro-1 antibody concentration and the isotype control was not possible as bands were more intense in the isotype (although the same thickness) than the Stro-1 antibody's band (Figure 33, B and C).

The Stro-1 supernatant to protein-L beads titration results showed that the protein-L beads bind to the Stro-1 antibody and to the IgM isotype. Although considering the intensity of the bands when the antibody alone was analysed (Figure 33, B), it showed that the beads did not bind to all the IgM in the antibody or alternatively it could indicate that the beads purified the antibody from FCS which could have contributed to the stronger bands in Figure 33, D.

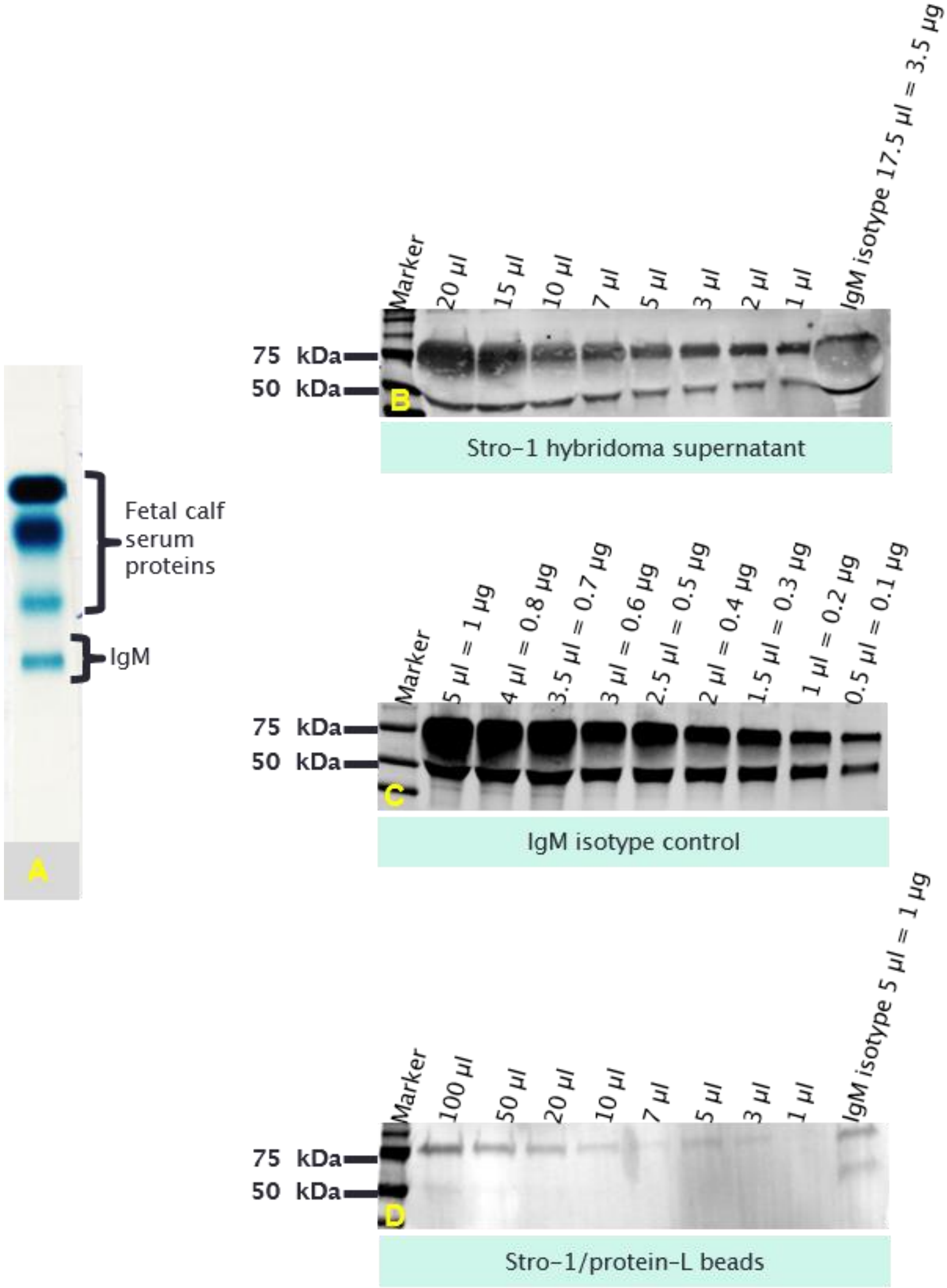
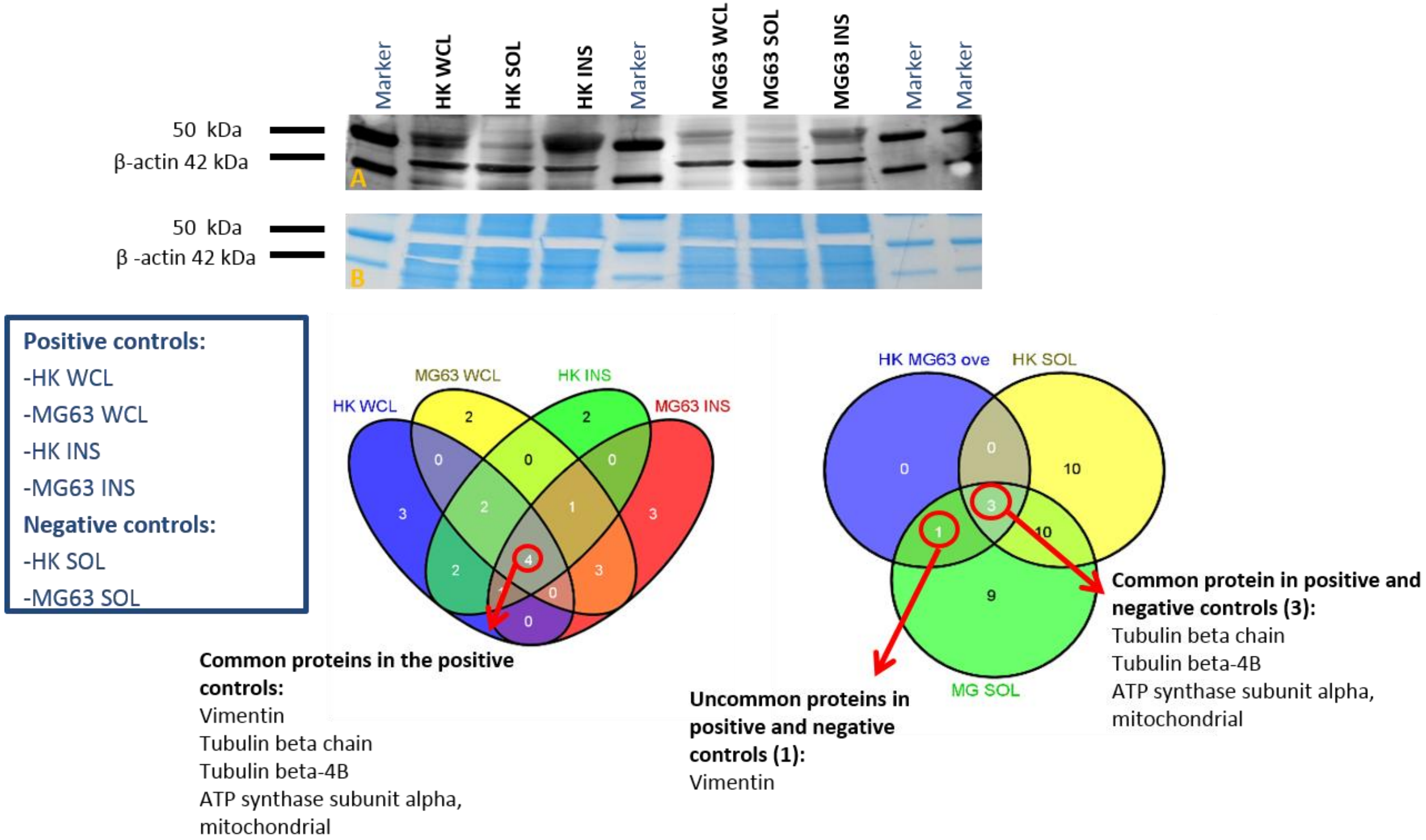

Fetal calf serum proteins
IgM
A
Marker
20 µl
15 µl
10 µl
7 µl
5 µl
3 µl
2 µl
1 µl
IgM isotype 17.5 µl = 3.5 µg
75 kDa
50 kDa
B
Stro-1 hybridoma supernatant
Marker
5 µl = 1 µg
4 µl = 0.8 µg
3.5 µl = 0.7 µg
3 µl = 0.6 µg
2.5 µl = 0.5 µg
2 µl = 0.4 µg
1.5 µl = 0.3 µg
1 µl = 0.2 µg
0.5 µl = 0.1 µg
75 kDa
50 kDa
C
IgM isotype control
Marker
100 µl
50 µl
20 µl
10 µl
7 µl
5 µl
3 µl
1 µl
IgM isotype 5 µl = 1 µg
75 kDa
50 kDa
D
Stro-1/protein-L beads

Figure 33: Stro-1 hybridoma supernatant western blotting characterisation.

A) Blue native page of Stro-1 hybridoma supernatant (Appendix 1, 9.5). B) Stro-1 hybridoma supernatant titration on Western Blotting (WB) (µL of Stro-1 hybridoma supernatant). C) IgM isotype titration on WB (isotype concentration 0.2 µg/µL). D). Stro-1 hybridoma supernatant bound to protein-L beads. B-C and D were probed with probed with IRDye 680 Goat α-mouse IgM secondary antibody.

3.4.2.4 WB band analysis

To investigate further the identity of the proteins likely to be present at around 50 kDa with the cell lysates, two gels were run with HK WCL, SOL and INS. Raji was along run as negative control. One gel was used for protein transfer onto a membrane and probed for Stro-1 (to confirm the sample used reacted to the Stro-1 antibody at 50 kDa (Figure 34, A) and a second gel was employed for band excision (Figure 34, B). Mass spectrometry (MS) analysis of the excised bands were outsourced to St. Georges University of London and executed by Dr. Omar Jallow (method employed in Appendix 1, 9.5). Mass spectrometry results, comprising in a list of identified proteins, were overlapped in a Venn diagram. The list of proteins in the negative sample, was used as a negative filter for the proteins contained in the positive samples. The band identified in both HK and MG63 WCL and INS fraction was characterised by vimentin. Vimentin was not identified in the HK SOL and was weakly identified in the MG63 SOL.


Marker
HK WCL
HK SOL
HK INS
Marker
MG63 WCL
MG63 SOL
MG63 INS
Marker
Marker
50 kDa
β-actin 42 kDa
A
50 kDa
β -actin 42 kDa
B
HK WCL
MG63 WCL
HK INS
MG63 INS
2
2
0
0
0
3
2
1
3
2
4
3
0
0
HK MG63 ove
HK SOL
0
0
10
3
1
10
9
MG SOL

Positive controls:

-HK WCL

-MG63 WCL

-HK INS

-MG63 INS

Negative controls:

-HK SOL

-MG63 SOL

Common proteins in the positive controls:

Vimentin

Tubulin beta chain

Tubulin beta-4B

ATP synthase subunit alpha, mitochondrial

Uncommon proteins in positive and negative controls (1):

Vimentin

Common protein in positive and negative controls (3):

Tubulin beta chain

Tubulin beta-4B

ATP synthase subunit alpha, mitochondrial

Figure 34: Mass spectrometry analysis of the 50 kDa gel band.

The previously identified band detected with the Stro-1 antibody at 50 kDa was analysed by mass spectrometry (gel A). Bands were excised from the gel B. The Venn diagrams indicate the common and uncommon proteins identified.

3.4.3 Aptamer selection

During the aptamer selection, a contamination was encountered that affected the outcome of the process. On this basis the best representation of the results was chosen to be highlighted in the following paragraphs in conjunction with troubleshooting approaches to identify the artefact.

3.4.3.1 Representative results

For each selection round the RNA transcribed overnight by the automated liquid handling system was verified on a gel (Figure 35, A). Subsequently, the selection process was performed and an aliquot of the recovered aptamer sequences stock was amplified up to 10 cycles in a cycle course PCR collecting an aliquot every 2 cycles and analysed by gel electrophoresis (Figure 35, B). On the basis of the cycle course results, the highest number of cycles absent from unspecific amplifications was employed to amplify the original RNA aptamer stock (Figure 35, C). Round 4 was selected as the best representation of the selection process step (Figure 35). Cell numbers used and incubation time are listed in Table 7 for each of the selection rounds.

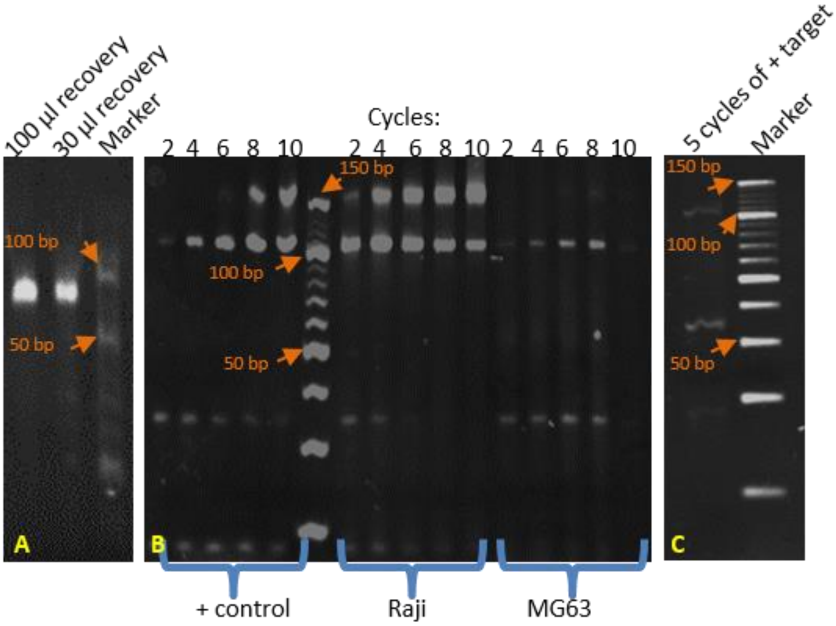

100 µl recovery
30 µl recovery
Marker
100 bp
50 bp
A
Cycles:
2 4 6 8 10
2 4 6 8 10
2 4 6 8 10
150 bp
100 bp
50 bp
B
+ control
Raji
MG63
5 cycles of + target
Marker
150 bp
100 bp
50 bp
C

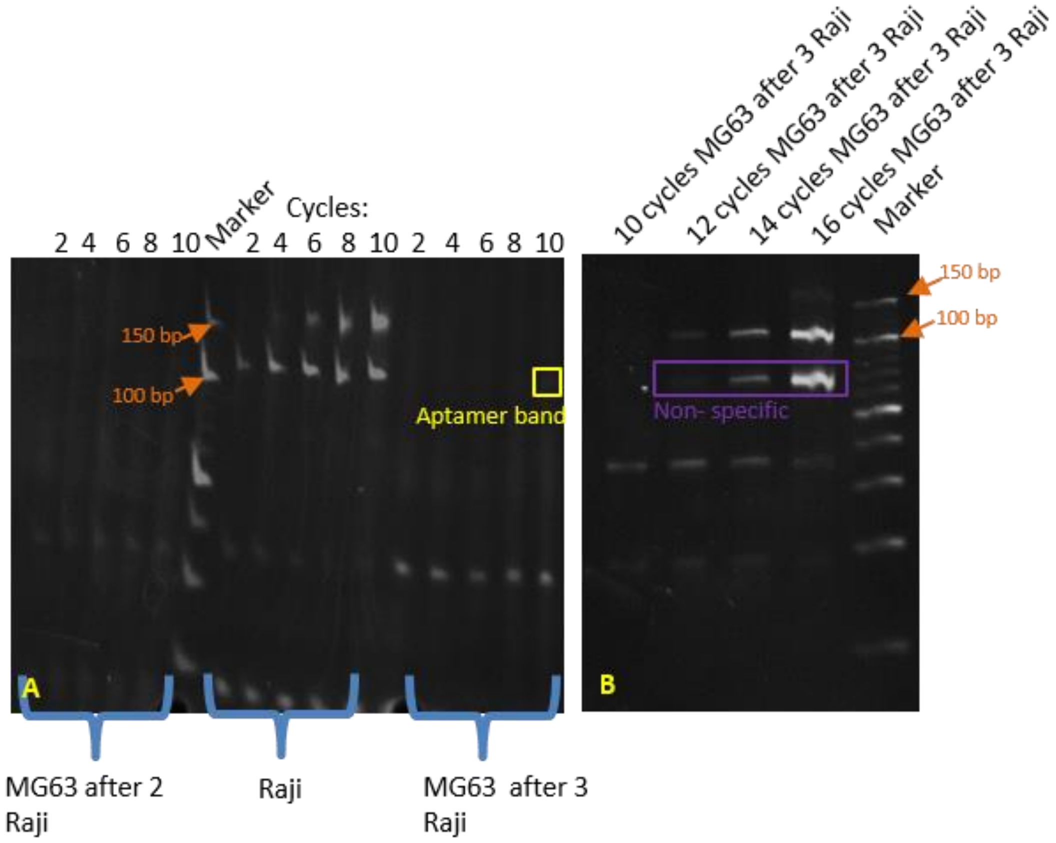
Figure 35: Selection round 4 results.

A) RNA gel of overnight automatically transcribed RNA; B) cycle course PCR of an aliquot of positive RNA aptamers following the selection process with positive and negative controls; C) Stock RNA manually amplified by 5 cycles. Aptamer band at 100 bp. P1: forward primer. P2: reverse primer.

3.4.3.2 Summary of results

At each selection round technical challenges were presented that were addressed iteratively. Selection was continued up to round 7 when a non-specific band was found to have entered the amplification cycle (Figure 36, B). In order to continue the selection, DNA samples from previous rounds were defrosted and verified on a gel (Figure 37). DNA from round 4 and 6 were chosen to take the selection forward. A further two rounds of selection were accomplished before a further contamination was subsequently detected (Figure 38). The contamination consisted in a band of the same size as the aptamer band (100 bp), in a band below of 50 bp and in band above of 150 bp (Figure 38).

In this section only salient results are reported. Results from all the selection rounds are not shown.



Cycles:
2 4 6 8 10 Marker 2 4 6 8 10 2 4 6 8 10
150 bp
100 bp
Aptamer band
A
MG63 after 2 Raji
Raji
MG63 after 3 Raji
10 cycles MG63 after 3 Raji
12 cycles MG63 after 3 Raji
14 cycles MG63 after 3 Raji
16 cycles MG63 after 3 Raji
Marker
150 bp
100 bp
Non- specific
B

Figure 36: Round 7 of the aptamer selection process.

A) Cycle course PCR, an aptamer band at 100 bp is visible in the MG63 sample filtered with Raji cells three times by cycle 10. B) Stock RNA manually amplified by 10, 12, 14 and 16 cycles with a manual PCR. Unspecific bands were found to contaminate the sample.

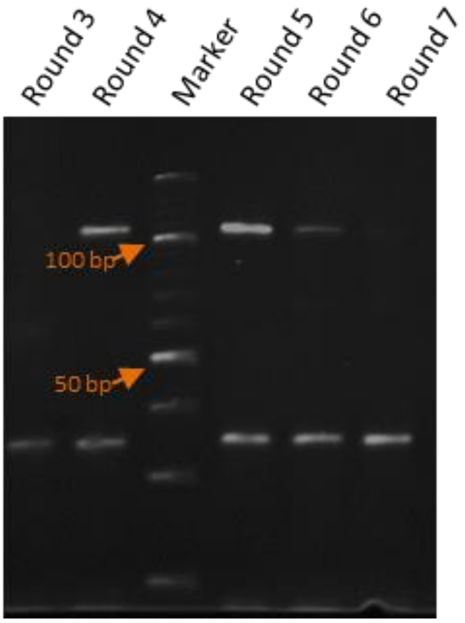

Round 3
Round 4
Marker
Round 5
Round 6
Round 7
100 bp
50 bp

Figure 37: DNA from previous rounds.

Frozen aliquots of previous DNA from round 3,4,5,6,7 were defrosted and verified on a gel for integrity. DNA from rounds 4 and 6 was chosen as template to re start the selection process.

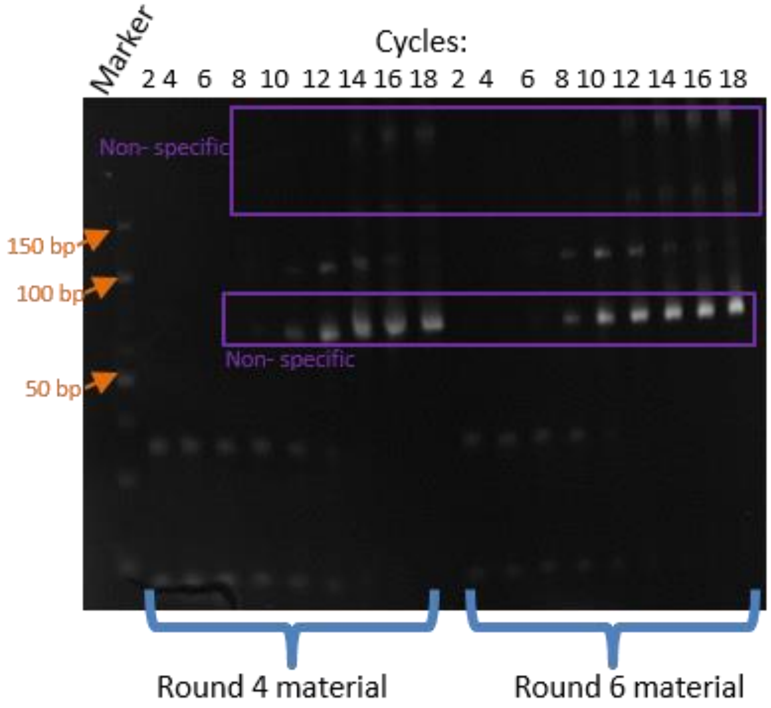

Marker
Cycles:
2 4 6 8 10 12 14 16 18 2 4 6 8 10 12 14 16 18
Non- specific
150 bp
100 bp
50 bp
Non- specific
Round 4 material
Round 6 material

Figure 38: Round 7 and 9.

The results showed amplified product at 100 bp, at 150 and at 50 bp without any aptamers sequences being present. These bands were unspecific considering no target was present during the cycle course PCR. An unspecific amplified band, not present previously, appeared in round 6 material. These unspecific bands implicated the entrance of unspecific sequences into the process.

3.4.3.3 *Trouble-shooting approaches*

Considering the non-specific sequences found in the selection process, the regent employed for cell lysis and the TriLink library were tested for contamination. It was identified that amplification of primers, water and QE (with no DNA) resulted in a band of same size as the aptamer band, after 10 cycle courses (Figure 39, A) and after 4 cycle courses (Figure 39, B).

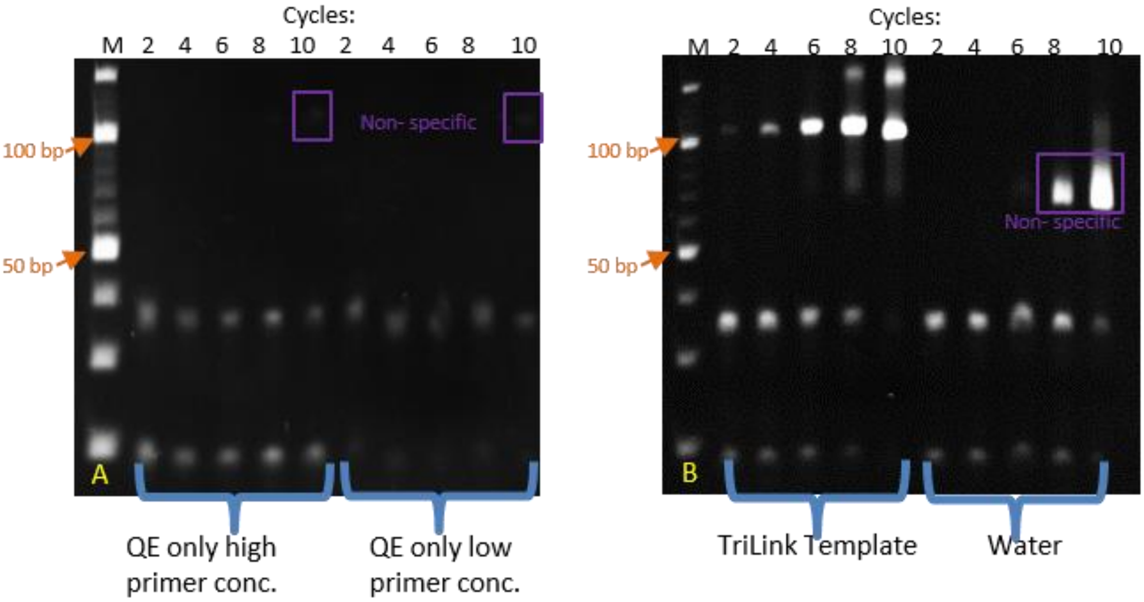

Cycles:
M 2 4 6 8 10 2 4 6 8 10
100 bp
50 bp
Non- specific
A
QE only high primer conc.
QE only low primer conc.
Cycles:
M 2 4 6 8 10 2 4 6 8 10
100 bp
50 bp
Non- specific
B
TriLink Template
Water

Figure 39: Trouble-shooting.

A) QE was alone amplified with high primer concentration and low primer concentration. After 4 cycle course a fainted non-specific band appeared at the same size (bp) as the aptamer band. B) The TriLink library template was amplified alone with a negative control of water and after 4 cycle courses a non-specific sequence was highly amplified by cycle 8 and 10.

3.5 Discussion

The aim of this chapter was to examine molecular differences between Stro-1 selected and unselected cells in basal and osteogenic media; characterise the Stro-1 antigen by WB and, ultimately, identify aptamer molecules as biomarkers for SSCs.

3.5.1 Characterisation of SSCs

The ability of SSCs to respond to osteogenic induction was demonstrated by analysis of alkaline phosphatase (ALP) expression using histological and molecular approaches. The aim of this chapter section was to verify the higher osteoblast differentiation potential in Stro-1 cells compared to unsorted cells (Stewart, Walsh *et al.*, 1999). It was demonstrated that cells undergoing osteogenesis by the addition of dexamethasone and ascorbic acid presented an OCN profile by day 21 as OCN is known to be a late bone marker. As a proof of cells undergoing differentiation towards the osteogenic pathway, the chondrogenic marker SOX9 was shown to decrease with time in osteogenic culture.

3.5.2 Stro-1 antibody and antigen analysis

The Stro-1 antigen was examined using WB analysis. Contrasting studies have shown the Stro-1 antibody to react with an antigen at 75, 50 and 18 kDa (Castrechini, Murthi *et al.*, 2010; Ning, Lin *et al.*, 2011). The data in this chapter indicated that the antibody reacted at 50 kDa. An improvement over previous studies consisted in the difference in expression of the 50 kDa band according to the cell fraction analysed. An enrichment in the INS fraction of the cells was obtained which is consistent with membrane proteins being poorly soluble. These findings are consistent to a certain extent to P. Bianco who in one of his latest papers defines (although showing no evidence) the Stro-1 antigen as having a glycolipid nature, hence hydrophobic (Bianco, 2014).

Once the WB with the Stro-1 antibody was optimised, to pursue the intent of identifying the Stro-1 antigen an IP protocol was attempted. In order to perform IP, antibody concentration and affinity to protein-L beads was investigated. The Stro-1 antibody was titrated against an IgM standard (isotype) (Figure 33, B and C). The results revealed that the Stro-1 band intensity was notably minor although presenting the same thickness. Possibly, the secondary antibody was targeting light chains of the isotype control which differ from the chains of the primary antibody (ultimately, the light chains type of the Stro-1 antibody deriving from the hybridoma supernatant is unknown; the isotype control should be chosen on the basis of having equal light chains to the primary antibody under test). Chromatography experiments would have played a key role in verifying the Stro-1 hybridoma supernatant antibody concentration. Affinity to protein-L beads investigation (Figure 33, D) showed that Stro-1 antibody binds to protein L-beads. It was also shown that the antibody signature on a gel presents a thick band at 75 and at 50 kDa when probed with an α-IgM. The 50 kDa band is the same molecular weight as the band found in the WCL, SOL, INS fractions. These findings represent a concern when attempting IP followed by WB verification as it is impossible to distinguish between the antibody and the antigen bands as both appear at the same molecular weight. MS could have been adopted as a technique to circumvent these obstacles, although issues would have been raised when using a cell pellet that was not completely lysed, (as the INS fraction), when separating proteins by nanoRP-chromatography prior to MS. The cell lysate sample for chromatography analysis should be clear and although it was shown that Laemmli buffer solubilised the majority of the cell pellet (Appendix 1, section 9.7), it contained SDS, not compatible with chromatography analysis. Restriction of time and MS resources were also main reasons for not investigating the lysate further. The IgM nature of the antibody restricted the field of available techniques for IP and antibody labelling.

3.5.3 WB band analysis

In order to understand the proteins contained within the WB band at 50 kDa, the gel sections at 50 kDa were cut and sent for mass spectrometry analysis to St. Georges University of London (Appendix 2, 9.5). The gel section results were characterised by abundant cytoskeletal and housekeeping proteins such as tubulin and vimentin. It is impossible to say if the band *i.e.* the Stro-1 antigen, is vimentin, or if vimentin was the most abundant protein in the band which obscured the less abundant proteins contained within the sample. Furthermore, although vimentin is a 50 kDa protein the immunostaining of cells in the literature is markedly different from the Stro-1 staining reported here. For these reasons the results were not considered further; a search for commercial companies were undertaken to repeat the experiment, however due to time and cost restrains the work was not performed. It must be noted that immunostaining differences on a monolayer on tissue culture plastic might also not be realistic of the true protein expression; more evidence is arising from importance of the mechanotransduction when culturing cells on topographically patterned surfaces (Dalby, Gadegaard *et al.*, 2014) .

3.5.4 Aptamers technology

A novel approach, to characterise the surface marker profile of SSC, based on aptamer technology, was explored. To develop aptamers that discriminate between SSCs and other cells in the BM stroma, a SELEX process was adapted to adherent and non-adherent cell lines (MG63 and Raji respectively). Cell lines were preferred over primary cells in order to standardise cell conditions and diminish variables in this pilot/feasibility study. The osteosarcoma cell line MG63 shares with SSCs common features such as morphology and Stro-1 expression. Hence a panel of nucleic acids aptamers was screened against the MG63 cell line filtering for the Raji cell line as negative control for Stro-1 with the hope of identifying novel markers on the cells surface.

The selection process started following a standard protocol developed by the Aptamer Group based in York. However, technical challenges prevented the accomplishment of the process limiting the interpretation of the results as a non-specific sequence contaminating the PCR was detected. Although attempts to avoid the observed contamination were taken, such as a re-start from pre-contamination selection rounds, the non-specific sequence was re-amplified in the amplification process. The process was stopped after the realisation by gel electrophoresis, that a sequence of 100 bp (same size as the aptamer sequence) was present even when the aptamer sequences were not present in the sample. Unfortunately, time constraints prevented further work, hence the initial hypotheses could not be tested.

Despite the encountered contamination, the knowledge acquired during the troubleshooting will be useful for subsequent rounds of selection. Troubleshooting key findings included:

1. The optimisation of the cell lysis process utilising a higher volume of quick extract buffer when dealing with a higher number of cells.

2. Decrease the stringency of the process by dropping gradually the cell number to avoid primers self-amplification due to lack of sequences.

3. Improved reagent quality, *e.g.* gel purified selection of aptamers.

In a repeat process it would be advisable to run in parallel a conservative and a stringent selection. The use of RT-PCR would also result in further advantages, with not only process acceleration but also enhanced sensitivity. It would be possible to capture contamination in real time, thanks to dissociation curves, hence intervene earlier.

3.6 Conclusion

The results in the context of the stated hypothesis in section 3.2 were examined:

I. Stro-1 selected cells exhibit higher osteogenic differentiation potential in comparison to unselected cells. This hypothesis is **true**.

The molecular analysis OCN expression demonstrates a higher expression of OCN in the Stro-1 sorted cells than in the unsorted cells.

II. The Stro-1 antibody reacts to an epitope at 50 kDa. This hypothesis is **true**.

WB analysis of the Stro-1 antibody illustrated that the antibody reacts to an epitope at 50 kDa; furthermore it was demonstrated that there is a shift in antibody reactivity at 50 kDa between the SOL and INS fraction of cell.

III. Aptamers can be selected specifically against the surface markers of MG63 cells using SELEX technology. This hypothesis was **not verified** due to interruption of the protocol.

A protocol for planning the cell culture in accordance to the aptamer selection method was validated and could be retained feasible for further studies.

In summary, an optimised protocol for WB analysis of the Stro-1 antigen was developed. These results generated further knowledge regarding the use of the Stro-1 antibody for proteomic studies. Future work must be undertaken to find the identity of the Stro-1 antigen. Techniques such as IP and MS should be optimised. Considering the difficulty of working with an IgM antibody, the Stro-1 antibody should be cloned into an IgG antibody to ease the research; IgM antibodies display low affinity for the target and their solubility is less than the IgGs counterpart, furthermore the majority of commercial available secondary antibodies and proteomic kits are validated for the use with IgG antibodies. In spite of the fact that Stro-1 is not the sole marker for SSCs, there is an increased interest in the Stro-1 antibody reflected by increased commercial availability. Regarding the work based around the aptamer technology, optimisation should include kinetic experiments to model the random chance of an aptamer sequence to be in contact with its binding site. Reagents need to be ultra-pure, and multiple attempts are needed to establish optimal working conditions such as cell number, primer concentration and gel purification strategies.

4. Chapter 4: Mass Spectrometry 1

I want to thank Dr. Spiro Garbis, Antigoni Manousopoulou and Harvey Johnston for their help and input with the running of the mass spectrometry.

4.1 Introduction

Harnessing the differentiation potential of SSC to provide replacement or renewal material for damaged tissue has been hampered by the lack of specific markers to effectively recognise SSCs in the presence of a mixed bone marrow (BM) cell population (Dominici, Le Blanc *et al.*, 2006). Such specific markers, are generally proteins located on the surface of cells which can be targeted using antibodies. The intrinsic variations in these anchorage molecules and their structural complexity form a substantial obstacle in the discovery of a unique molecule that target just one cell phenotype.

Classic approaches for targeted antibody-based biomarker discovery have been developed but no successful results have been reported to date. A salient limitation include the inefficacy of formulating an antibody against a unique unknown target (O'Hurley, Sjöstedt *et al.*, 2014). In fact criteria used to isolate SSC are defined as "loose" as stated in Bianco, Cao et al., 2013 as the markers employed are "non-specific" to the surface of SSCs (*i.e.* markers can select other types of cells) (Bianco, Cao *et al.*, 2013).

Recently, thanks to the latest progress made in various bio-analytical techniques, instrumentation platforms and software tools, the field of proteomics has been envisaged to offer a comprehensive approach for biomarker discovery. Proteomics, in contrast to genomics, presents distinct advantages as it accounts for variability between the production of mRNA and the production of proteins (Kulasingam and Diamandis, 2008). Changes in the mRNA can result in significant changes in the corresponding proteins. To address the issue of target identification from a novel angle, rather than testing one marker at a time (as per classical antibody techniques), a comprehensive non-targeted, global mass spectrometry based-proteomics approach was undertaken in this thesis.

Proteomics has been previously employed to compare the protein profile of SSC isolated from the bone marrow of healthy and leukemic patients, this study identified 73 differentially regulated proteins involved in the haematopoiesis (Seshi, 2006). In another study designed to shed light on the osteogenic key regulators, collagen IV, fibronectin, vitronectin, and thrombospondin were identified as osteogenic hallmarks in SSC (Lo, Tsai *et al.*, 2012).

The proteomic method used in this thesis combined the techniques of orthogonal two dimensional liquid chromatography and ultra-high resolution mass spectrometry. This method will increase the effectiveness in profiling novel candidates for cell surface identification.

The method applied constituted a pilot study to establish new knowledge prior to a further comprehensive study (see Chapter mass spectrometry 2). SSCs (female individual aged 85) were collected, selected for Stro-1 and grown under basal and osteogenic conditions. Research aims, methodology and results are pictographically shown in Figure 40, and described and discussed in the following paragraphs.

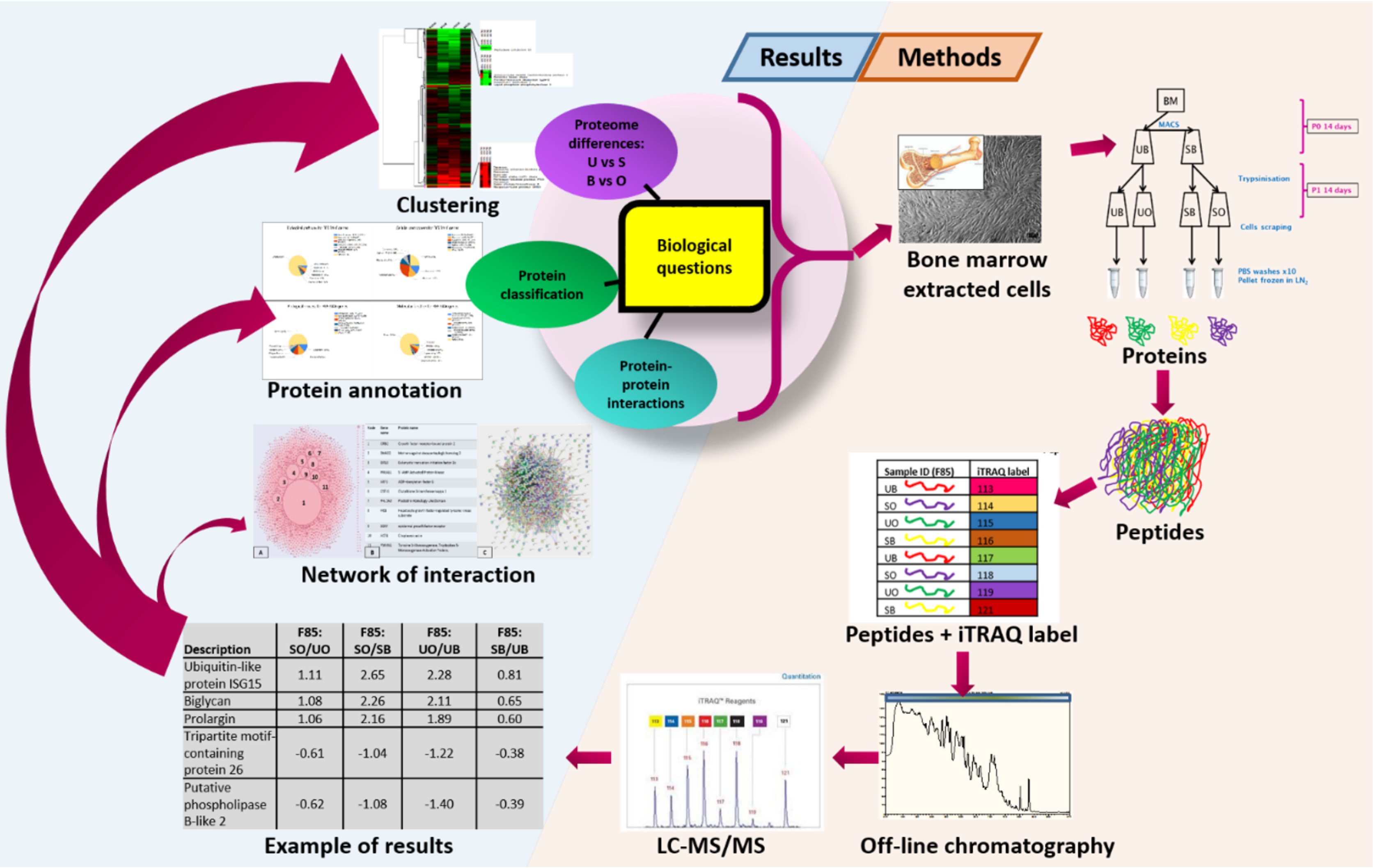
4.2 Aims and hypothesis:

To summarise this chapter aimed:

- To analyse the proteome of enriched human SSCs.
- To compare the proteome of Stro-1 sorted SSCs with unselected SSC, following culture in basal and osteogenic conditions.
- Identify a set of epitopes capable of selecting for the most optimal cellular source in the context of regenerative medicine.

Hypothesising:

I. The hyphenated use of Liquid Chromatography-Mass spectrometry is a sufficiently sensitive approach to compare differences in protein expression between cells treated in basal and osteogenic conditions and sorted by Stro-.1
II. Stro-1 selected cells exhibit a different set of modulated proteins in comparison to unselected cells.


Results
Methods
Clustering
Proteome differences: U vs S B vs O
Biological questions
Protein classification
Protein-protein interactions
Protein annotation
Network of interaction
A
B
C
Bone marrow extracted cells
BM
MACS
UB
SB
UB
UO
SB
SO
P0 14 days
Trypsinisation
P1 14 days
Cells scraping
PBS washes x10
Pellet frozen in LN_2
Proteins
Peptides
Peptides + iTRAQ label
Off-line chromatography
Quantitation
iTRAQ™ Reagents
LC-MS/MS
Example of results

Sample ID (F85)	iTRAQ label
UB	113
SO	114
UO	115
SB	116
UB	117
SO	118
UO	119
SB	121

Description	F85: SO/UO	F85: SO/SB	F85: UO/UB	F85: SB/UB
Ubiquitin-like protein ISG15	1.11	2.65	2.28	0.81
Biglycan	1.08	2.26	2.11	0.65
Prolargin	1.06	2.16	1.89	0.60
Tripartite motif-containing protein 26	-0.61	-1.04	-1.22	-0.38
Putative phospholipase B-like 2	-0.62	-1.08	-1.40	-0.39

Figure 40: Skeletal profiling methodology.
The diagram represents a schematic outline of the study aims, experimental design and results. The proteins of bone marrow extracted cells are labelled with iTRAQ probes and processed for LC-MS/MS. Mass spectra are translated into ratios of protein abundance. Protein data is then analysed to create a cluster gene map, protein annotation and a network of interactions map.

4.3 Materials and Methods

4.3.1 Sample preparation

Human BM stromal cells were extracted and cultured as detailed in section 1.1 of the Materials and Methods chapter.

4.3.2 Magnetic Activated Cell Sorting (MACS) isolation

Cells positive for the Stro-1 antibody were magnetically labelled and separated as detailed in section 2.2 of the Material and Methods chapter.

4.3.3 Cell culture and harvesting

A schematic of the extraction and culture procedure is shown in Figure 41. 14 days after isolation, cells were trypsinised and re-seeded in the same flasks. After a further 14 days, cells were trypsinised and divided into four T175 flasks, of which 2 flasks were supplemented with osteogenic media (consisting of α-MEM phenol free (Modified Eagle's Medium, GIBCO life Technologies), L-Ascorbic acid 2-phosphate sesquimagnesium salt hydrate (100 μM, Sigma-Aldrich), and dexamethasone (10 nM, Sigma-Aldrich), containing 10% foetal calf serum (FCS, Sigma), and 1% Penicillin Streptomycin (PAA The cell culture company). The remaining flasks were kept in basal media conditions. All flasks were incubated at 37 ˚C, supplemented with 5% CO_2. Media change was performed twice a week. After 14 days in osteogenic and basal media conditions, cells were harvested with a cell scraper and decanted into 50 mL falcons respectively. Falcons were centrifuged at 1,500 rpm for 5 min in a Heraeus Multifuge 3 S-R. After centrifugation the supernatant was discarded, the pellet was transferred into an eppendorf and washed 10 times with 1 mL of PBS. Between each wash, samples were centrifuged in a Heraeus centrifuge 5415 R at 13,000 rpm at 4 ˚C for 2 min. After the final wash, the supernatant was completely removed to obtain a dry pellet. The pellet was subsequently frozen in liquid nitrogen for 10 min and stored at -80 ˚C.

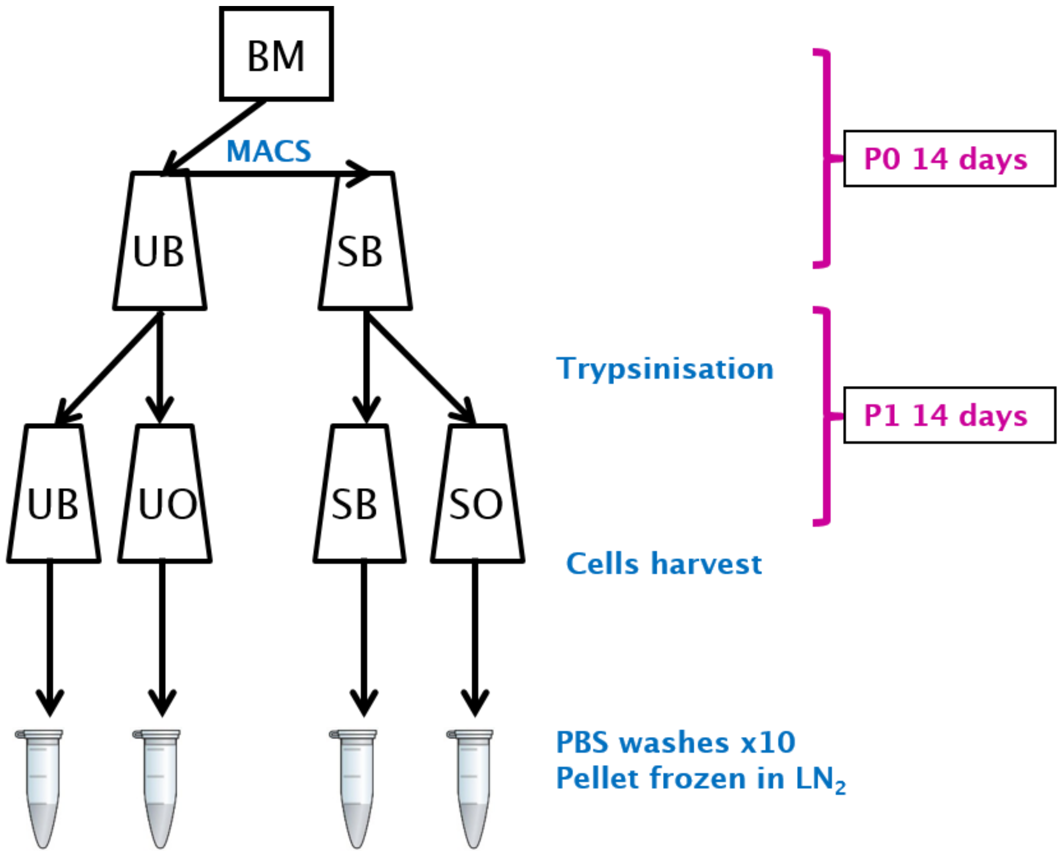

BM
MACS
UB
SB
P0 14 days
Trypsinisation
P1 14 days
UB
UO
SB
SO
Cells harvest
PBS washes x10
Pellet frozen in LN2

Figure 41: F85 extraction and cell culture.

Unsorted cells extracted from the marrow of a female patient aged 85 were sorted using the Stro-1 marker and expanded for 14 days under basal conditions. Cells were harvested by scraping, thoroughly washed and frozen in liquid nitrogen. BM: bone marrow; UB: unsorted basal; UO: unsorted osteogenic; SB: sorted basal; SO: sorted osteogenic.

4.3.4 Reagents and Chemicals

Table 8: Mass spectrometry reagents information.

Reagent	Grade of purity	Company
Acetone	*ACS reagent, ≥99.5%	Sigma-Aldrich
Ammonium hydroxide solution (NH_4OH)	*ACS reagent, 28.0-30.0% NH3 basis	Sigma-Aldrich
Acetonitrile (ACN)	**Anhydrous, 99.8%	Sigma-Aldrich
Ethyl alcohol, pure (ethanol)	*ACS reagent, ≥99.5%	Sigma-Aldrich
2-Propanol, (isopropanol)	**Anhydrous, 99.5%	Sigma-Aldrich
Methanol	**Anhydrous, 99.8%	Sigma-Aldrich
Triethylammonium bicarbonate (TEAB)	For HPLC 1.0 M, pH 8.5±0.1	Sigma-Aldrich
Formic acid (FA)	LC-MS Ultra, eluent additive for UHPLC-MS	Sigma-Aldrich
Trypsin	Proteomics sequencing grade	Roche Diagnostics
Milli-Q water	18 mM-Ohm resistivity; < 3ppb total organic carbon	Millipore

* ACS reagent: Meets or exceeds American Chemical Society specifications. For general procedures requiring stringent quality specifications such as analytical testing (Sigma-Aldrich definition).

** Anhydrous: Ideal use in organic synthesis, organometallics, oligonucleotide synthesis, combinatorial chemistry and other applications requiring extremely low levels of water. Water content ranges from typically 10 to 30ppm and do not require additional drying in your lab due to our low residue levels (Sigma-Aldrich definition).

Table 9: iTRAQ Reagent-8plex Multiplex Kit.

Reagent	Description and content	Company
iTRAQ Reagent-8plex 113	Amine-modifying labelling reagent. Labels up to 100 µg protein.	AB SCIEX
iTRAQ Reagent-8plex 114		AB SCIEX
iTRAQ Reagent-8plex 115		AB SCIEX

iTRAQ Reagent-8plex 116		AB SCIEX
iTRAQ Reagent-8plex 117		AB SCIEX
iTRAQ Reagent-8plex 118		AB SCIEX
iTRAQ Reagent-8plex 119		AB SCIEX
iTRAQ Reagent-8plex 121		AB SCIEX
Reducing agent	Reduce the disulphides bonds of the proteins. 50 mM tris-(-2-carboxyethyl) phosphine (TCEP)	AB SCIEX
Cysteine-blocking Reagent	Reversibly blocks the cysteine group. 200mM methyl methanethiosulfonate (MMTS) in isopropanol.	AB SCIEX

Table 10: Off-line HPLC phases details.

HPLC phase	**Details**
mobile phase A	2% ACN, 0.05% NH_4OH in Milli-Q water
mobile phase B	100% ACN, 0.05% NH_4OH
stationary phase	Modified silica column C_{18} (18 atoms of carbon) Xbridge (Waters) of size 150 mm in length with an internal diameter of 3mm and a particle size of 3.5µm. (150 × 4.6 mm, 3.5 µm particle)

Table 11: On-line HPLC phases details.

HPLC phase	**Details**
mobile phase A	2% ACN and 0.1% FA in Milli-Q water
mobile phase B	100% ACN and 0.1% FA
stationary phase	Acclaim PepMap 100, 100 µm×2cm C18,5 µm particle trapping column with the ul PickUp injection mode

4.3.5 Sample preparation for liquid chromatography

Sample lysis: UO, UB, SO and SB sample pellets were solubilised in 100 µL of TEAB buffer (Table 8), vortexed and sonicated using a sonicator (Processor XL, Misonik incorporated).

Bradford assay: A Bradford assay was performed using the BioRad protein assay kit to quantify total protein content of the samples against a set of BSA standards. Proteins were vortexed to mix and were incubated at RT for 5 min. Absorbance was recorded at a wavelength of 570 nm using an ELx800 absorbance microplate reader (BioTek). Once the protein concentration was established, 100 µg of each protein sample was transferred into a new eppendorf with the addition of TEAB buffer up to a final volume of 20 µL. Each sample was run in duplicates for a total of 8 samples.

Protein reduction and cysteine blocking: To break the disulfide bonds between the proteins, 2 µL of reducing agent (Table 9), were added to the samples containing 100 µg of proteins and incubated on a heating block (Wealtec, HB-2) at 60°C for 1 h. Subsequently 1 µL of Cysteine-blocking Reagent (Table 9) was added to the samples to block (alkylate) cysteine residues and prevent re-folding (oxidation). Samples were mixed by vortexing, and incubated at RT for 15 min.

Protein digestion with trypsin: Samples were subjected to proteolysis through the addition of 6 µL of trypsin solution (500 ng/µL and 14 µL of ultra-pure water); the samples were mixed by vortexing, centrifuged and incubated overnight at RT in the dark.

iTRAQ reagent preparation and labelling: 8plex iTRAQ reagents 113 to 121 (Table 9) were prepared with the addition of 50 µL of isopropanol, mixed with the respective sample and incubated for 2 h at RT. The content of each sample tube was then combined and concentrated under vacuum (Eppendorf concentrator 5301).

Table 12: iTRAQ labelling.

Each iTRAQ label is associated with one unique sample or technical replicate thereof. Samples were run in duplicates using the 8plex iTRAQ kit and to verify technical reproducibility. UB: unsorted basal, UO: unsorted osteogenic SB: sorted basal SO: sorted osteogenic, F85: female patient aged 85.

Sample ID (F85)	iTRAQ label
UB	113
SO	114
UO	115
SB	116
UB	117
SO	118
UO	119
SB	121

High pH reverse Phase C_{18} Chromatography (RP-C_{18}): The HPLC system used was a Dionex P680 pump equipped with PDA-100 photodiode array detector using the Waters, XBridge C_{18} column as previously described (Papachristou, Roumeliotis *et al.*, 2013). The phases used are described in Table 10. Previously concentrated samples were reconstituted in 100 µL of mobile phase A, vortexed, dissolved by bath sonication and centrifuged for 5 min at 13,000 rpm. 100 µL of sample supernatant was loaded onto the HPLC system syringe and initially injected into a loop positioned before the starting valve of the system. Following system equilibration, the starting valve was brought in line with the analytical column. The pressure in the system was 60-80 bar resulting in a flow rate of 200 µL/min. The elution separation program was set as follows: for 10 min isocratic 5% (B), for 55 min gradient up to 60% (B), for 10 min gradient up to 70% (B), for 10 min up to 95% (B) (Figure 42). The column temperature was maintained at 30 ˚C.

Eluting peptides were monitored by absorbance at a wavelength of 215 nm. A total of 28 peptides fractions were collected in a peak-dependent manner over a 90 min chromatographic window. These peptides fractions were dried to residue overnight with a speed-vacuum concentrator (Eppendorf concentrator 5301) at RT and stored at –20 ˚C.

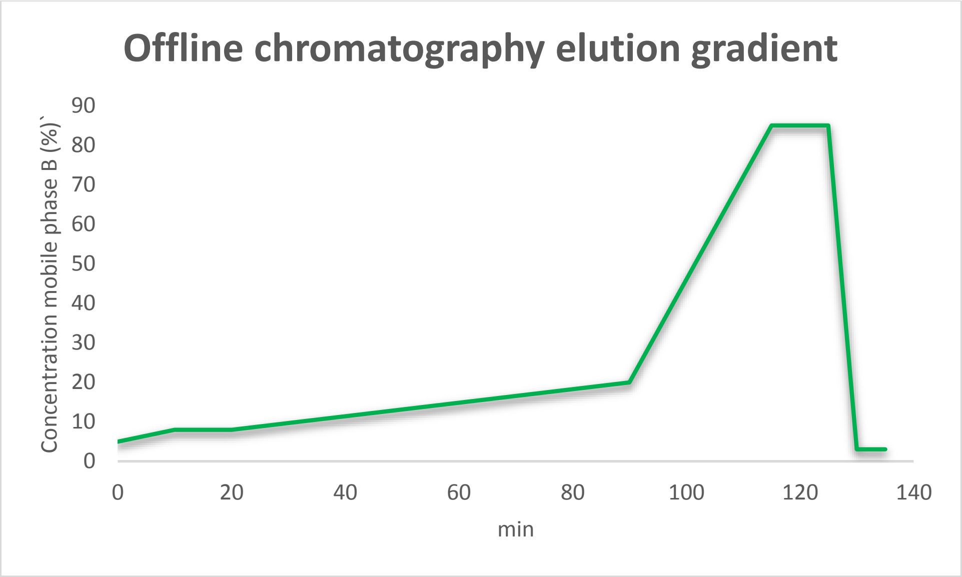


Figure 42: Offline HPLC elution gradient profile.

Elution gradient of peptides during offline chromatography. Mobile phase B: 100% ACN, 0.05% NH_4OH.

4.3.6 LC-MS Analysis

The online peptide fraction analysis was performed with ultra high performance nano-capillary chromatography hyphenated with nanospray ionization–ultra high resolution mass spectrometry (LC –nESI-MS) with a Dionex Ultimate 3000 UHPLC/LTQ-Velos Pro Orbitrap Elite system (Thermo Scientific) as previously reported (Papachristou, Roumeliotis *et al.*, 2013).

The specific conditions for the analysis were as follow:

Ultra high performance nano-capillary chromatography: peptides fractions (see section 4.3.5) were reconstituted in 30 μL of mobile phase A. De-salting was performed with the loading pump by introducing 3 μL of sample onto a trapping column (Acclaim PepMap 100, 100 μm × 2 cm C18, 5 μm particle) at a flow rate of 5 μL/min for 5 min under aqueous mobile phase conditions (98% Milli-Q water with 0.01% FA). For the analytical stage, peptides were further separated with ultra-performance reverse phase chromatography (Acclaim PepMap C_{18}, 75 μm × 25 cm, 2 μm particle, 100 Å, Thermo Scientific) under acidic conditions at a flow rate of 0.3 μL/min at 40 ˚C.

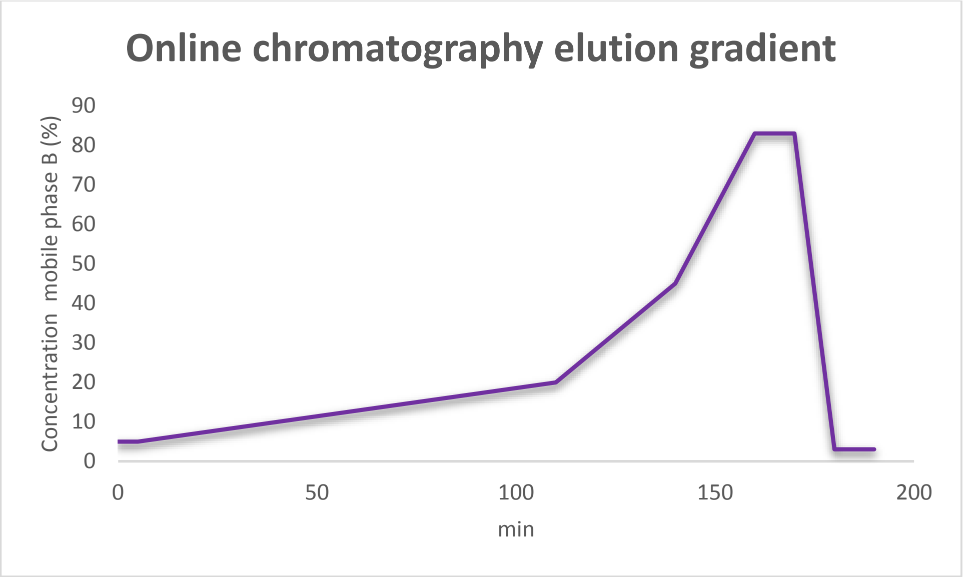


Figure 43: Online chromatography elution gradient profile.

Elution gradient of peptides during online chromatography. Mobile phase B: 100% ACN, 0.05% NH_4OH.

The gradient elution method in this system was set up at a flow rate of 0.3 µL/min and consisted of the following gradient steps: start at 5%, for 80 min gradient up to 40 % (B), for 5 min gradient up to 8 % (B), for 5 min isocratic 85% (B), for 2 min down to 3% (B), for 8 min isocratic equilibration 3% (B) at 35 ˚C (Figure 43).

Nanospray ionization–ultra high-resolution mass spectrometry: Nanospray ionization (nESI) was performed with a PicoTip emitter (FS360-20-10-D-20-C7; New Objective, Massachusetts, USA) at a voltage potential of 2.5 kV. The precursor tryptic peptide ions were monitored at 350-1,900 m/z at 120,000 mass resolution. For each MS scan event, the top 10 most abundant multiple charged ions with a minimum of a +2 charge state were subjected to product ion fragmentation with both N_2 high-collision dissociation (HCD) and He collision induced dissociation (CID). The HCD fragmentation was performed at 40 keV normalised collision energy, a 1.2 Da isolation window and monitored at 15,000 mass resolution within the mass range of 100–1,900 of *m/z*. The CID fragmentation was performed at a 2 Da isolation window, and 35 keV normalised collision energy. A maximum of 200 ms ion injection time was allowed. An ion exclusion rule was applied for the already measured precursors for 30 s.

4.3.7 MS data processing

Peptide identification and protein alignment was performed by processing the raw data generated by the MS with Proteome Discoverer version 1.3 (Thermo Scientific). The SEQUEST algorithm was used to estimate the target decoy search for tryptic peptides, allowing one missed cleavage, a precursor tolerance of 10 ppm, a minimum peptide length of 5, and a maximum of 2 variable (1 equal) modifications of oxidation (M), deamidation (N, Q), or phosphorylation (S, T). Methylthio (C) and iTRAQ (K, Y, and N-terminus) were set as fixed modifications. Only peptides identified with a false discovery rate (FDR) of less than 0.05 were included. Spectra were searched against the UniProtKB SwissProt human proteome.

iTRAQ labels offered a relative quantification, hence one condition was compared to the opposite in a ratio equation. To determine the conditions (SB, SO, UB, UO) to compare, the rationale of analysing protein modulation due to sorting consisted in considering the sorted sample as the nominator of the equation and the unsorted sample as the denominator, (S/U). To study protein modulation due to osteogenesis, the osteogenic sample was considered as the nominator of the equation and the basal sample as the denominator, (O/B). Samples ratios obtained are summarised in Figure 44.

4.3.7.1 Downstream data analysis

The discovered proteins were filtered to obtain three protein lists, each specific to a different analysis:

1. 3446 discovered proteins
2. 644 reproducible proteins
3. 26 reproducible proteins outside the range of ±1 SD

1) The first protein list was used for protein classification, interaction and pathway analysis. These analyses do not require protein quantification data but only the gene name associated with each protein discovered. Details on protein classification and interaction are presented in section 4.3.9.

2) The second list was used for clustering the proteins based on their gene similarities and protein expression, in order to create a heat map (4.4.3). For

this analysis it was necessary to include the protein quantification data, hence the discovered 3446 proteins were filtered in a conservative manner to obtain only the proteins in which duplicates were statistically consistent *i.e.* reproducible.

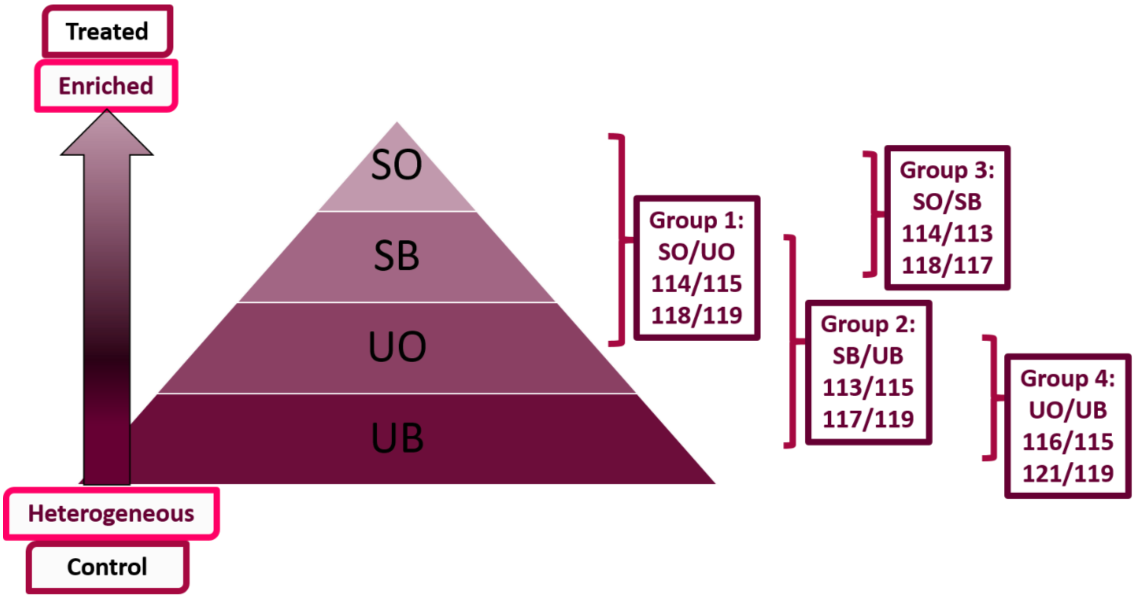

Treated
Enriched
SO
SB
UO
UB
Heterogeneous
Control
Group 1: SO/UO 114/115 118/119
Group 2: SB/UB 113/115 117/119
Group 3: SO/SB 114/113 118/117
Group 4: UO/UB 116/115 121/119

Figure 44: iTraq ratios rationale.

The pyramid represent the bone differentiation hierarchy based on the conditions employed to treat the samples in this study. The bottom of the hierarchy is represented by the unsorted cells in basal media condition (UB). This UB cell population is heterogeneous and it is employed as a baseline control for the iTRAQ ratios. The top of the pyramid is represented by the Stro-1 sorted cells in osteogenic media (SO). This SO population is enriched in respect to the UB, hence is considered as treated. The unsorted osteogenic (UO) and the sorted basal population (SB) are considered as intermediates of this differentiation process. The scope of the ratios is to compare the sorted sample with the unsorted sample under basal and osteogenic conditions. 2 iTRAQ ratios are available, as samples were run in duplicates.

To perform this operation and identify only the reproducible proteins, each ratio was converted in $\log_2$, the converted ratios of the technical replicates were averaged and the absolute value (ABS) of each average was calculated. The standard deviation (SD) between the technical replicates was also determined. Data were all filtered to satisfy the formula ABS-SD>0 so that all technical replicates with a SD greater than the mean were discarded. The filtering process was applied to 4 groups at once (644 filtered), and to each group (*i.e.* SO/UO, SO/SB, UO/UB, SB/UB) individually. When the filter was applied to each group, the result yielded a different number of reproducible proteins per group. The list of reproducible proteins of each group was uploaded on http://bioinfogp.cnb.csic.es/tools/venny/index.html andhttp://www.pangloss.com/seidel/Protocols/venn4.cgi to visualise the overlap using a Venn diagram (Figure 46).

3) A third list was created with the aim to enhance the visualisation of the protein expression and select only the proteins that were highly regulated within the reproducible protein list. For this list the 644 reproducible protein list was employed as a starting point including the expression levels. This list was created by selecting the values falling out of the range ± 1 SD. This analysis yielded only 26 proteins which correspond to the most highly or low expressed proteins across the 4 groups. These proteins were clustered and displayed in a heat map. Clustering was performed as per section 4.3.8, Figure 48.

4.3.8 Hierarchical clustering

Hierarchical clustering was performed using Cluster 3.0 and TreeView 6.0. Node distances were calculated using the Euclidean based metric and then clustered using the complete linkage method. Data were transformed in $\log_2$ prior to analysis.

4.3.9 Protein annotation and interactions: FunRich and STRING.

Protein list 1 (section 4.3.7.1) was uploaded on the free software FunRich (Functional Enrichment Analysis Tool) (Pathan, Keerthikumar *et al.*, 2015) to create a map of protein interaction, and classify the data according to biological pathway, biological processes, molecular function and cellular components (Figure 49, Figure 50, Figure 51 and Figure 52). A map of protein interaction was created and the packed view selected to display the results (Figure 53). The packed view allowed for the visualisation of specifically enriched nodes where a higher degree of distribution corresponded to a higher radius of nodes. Data were emphasized in a circular fashion and placed closely with the other groups having similar degree of distribution. Protein list 3 (4.3.7.1) was used in FunRich to create a heatmap (Figure 54). This type of heatmap uses the Human proteome map quantitative dataset as background, where the values of protein expression across a set of tissues are displayed according to the protein query. Different colours codes display the higher or lower expression. Only 27 proteins were used as the data input to prevent loss of significance.

STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) was used to show the complexity of protein-protein interaction in a protein network. The protein network represents the topological summary of all published proteins interactions detected in the UniProt database (Szklarczyk, Franceschini *et al.*, 2015). The genes associated to each of the 644 proteins were used as the input for the creation of network of proteins interaction map. The 3446 proteins gene list was also used as an input but the list was not compatible with the capacity of the system.

4.4 Results

4.4.1 Off-line alkaline reverse phase chromatography

Prior to mass spectrometry analysis, the proteins surrogate tryptic peptides were fractionated by RP-C_{18} chromatography under alkaline mobile phase conditions (Figure 45). Each peak monitored at 218 nm corresponds to a fraction containing partially resolved peptides with similar hydrophilic or hydrophobic retention time behaviour. From left to right, the scale indicates the hydrophobic and hydrophilic index of the tryptic peptides found in the original pooled sample. A total of 28 fractions were collected.

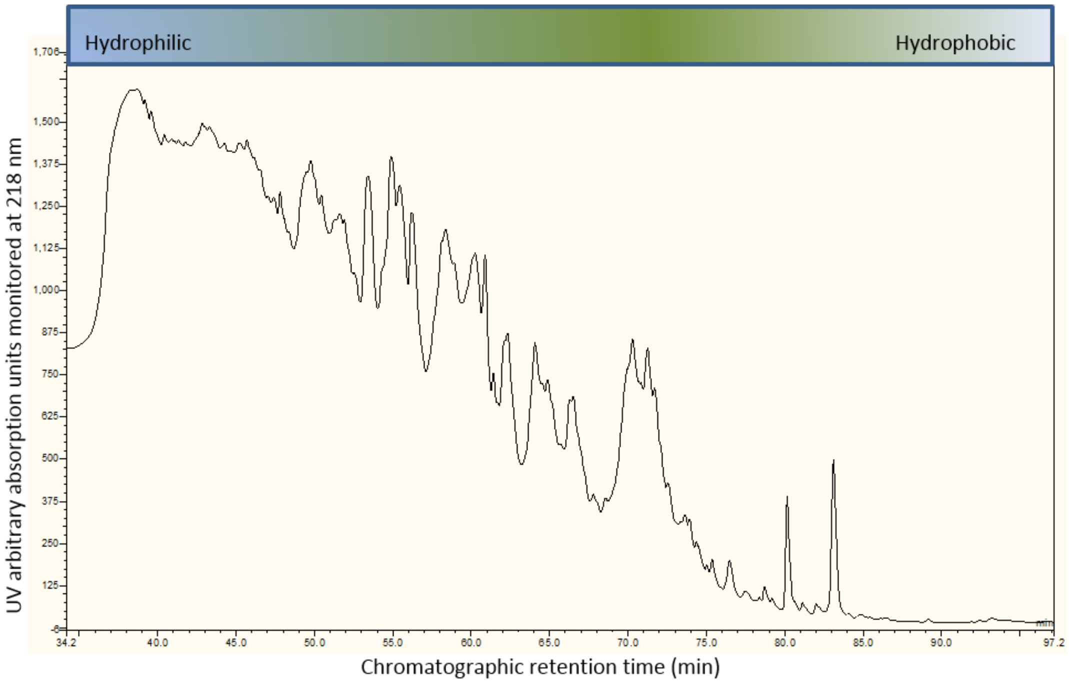
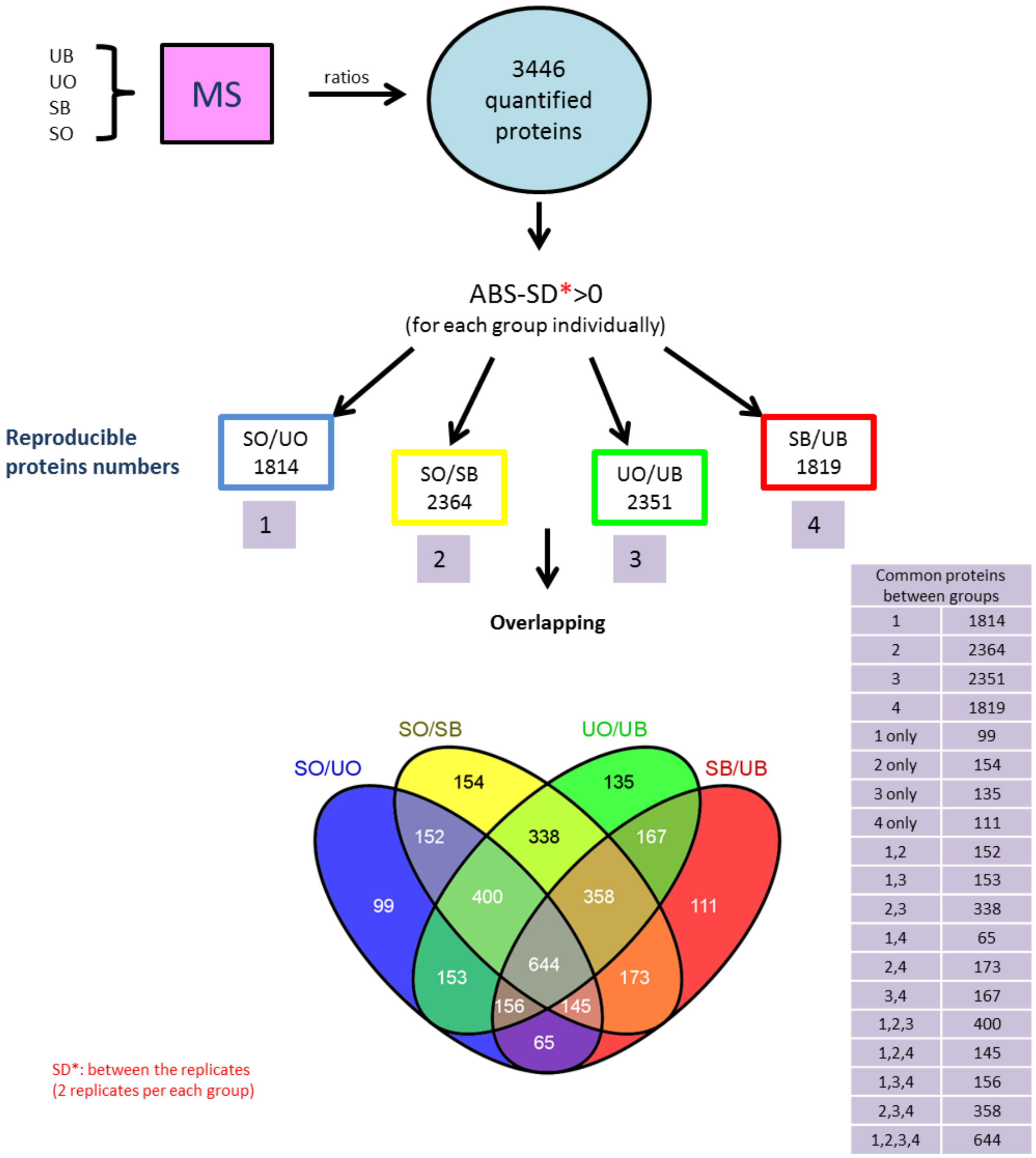


Figure 45: High pH reverse phase chromatography, MS1.

The protein surrogate tryptic peptides were separated in accordance to their differences in hydrophobic properties. A total of 28 peptide fractions were collected prior to the online LC-MS/MS analysis. The y-axis corresponds to UV arbitrary absorption units monitored at 218 nm. The x-axis corresponds the chromatographic retention time in min.

4.4.2 On-line LC-MS Analysis

A total of 3446 proteins were discovered at ≥ 95% confidence and at a false discovery rate (FDR) <0.05%. Proteins were identified and quantified by nanoRP-ESI-LC-MS/MS-HCD-CID. Of the 3446 proteins, a total of 644 were found to be reproducible across the 4 groups by satisfying the equation ABS-SD>0 (4.3.7.1). The equation ABS-SD>0 excluded the technical replicates where the SD was greater than the mean. When each group was filtered independently by the others using the equation ABS-SD>0, the following numbers of proteins were found to be reproducible: SO/UO=1814; SO/SB=2364; UO/UB=2351 and SB/UB=1819 (Figure 46). Among the 644 proteins biglycan displayed the greatest upregulation (2.9 fold) and peptidase inhibitor 16 was the most downregulated (3.8 fold) (respectively in the SO/SB and SB/UB across the 4 groups). Proteins that were more than 2 fold upregulated included ubiquitin-like protein ISG15, adipocyte enhancer-binding protein 1, periostin and tenascin. Protein upregulation was observed according to the osteogenic culture condition and furthermore increased in the Stro-1 fraction.


UB
UO
SB
SO
MS
ratios
3446 quantified proteins
ABS-SD*>0
(for each group individually)
Reproducible proteins numbers
SO/UO 1814
1
SO/SB 2364
2
UO/UB 2351
3
SB/UB 1819
4
Overlapping
SO/UO
SO/SB
UO/UB
SB/UB
99
154
135
111
152
338
167
400
358
153
644
173
156
145
65
SD*: between the replicates
(2 replicates per each group)

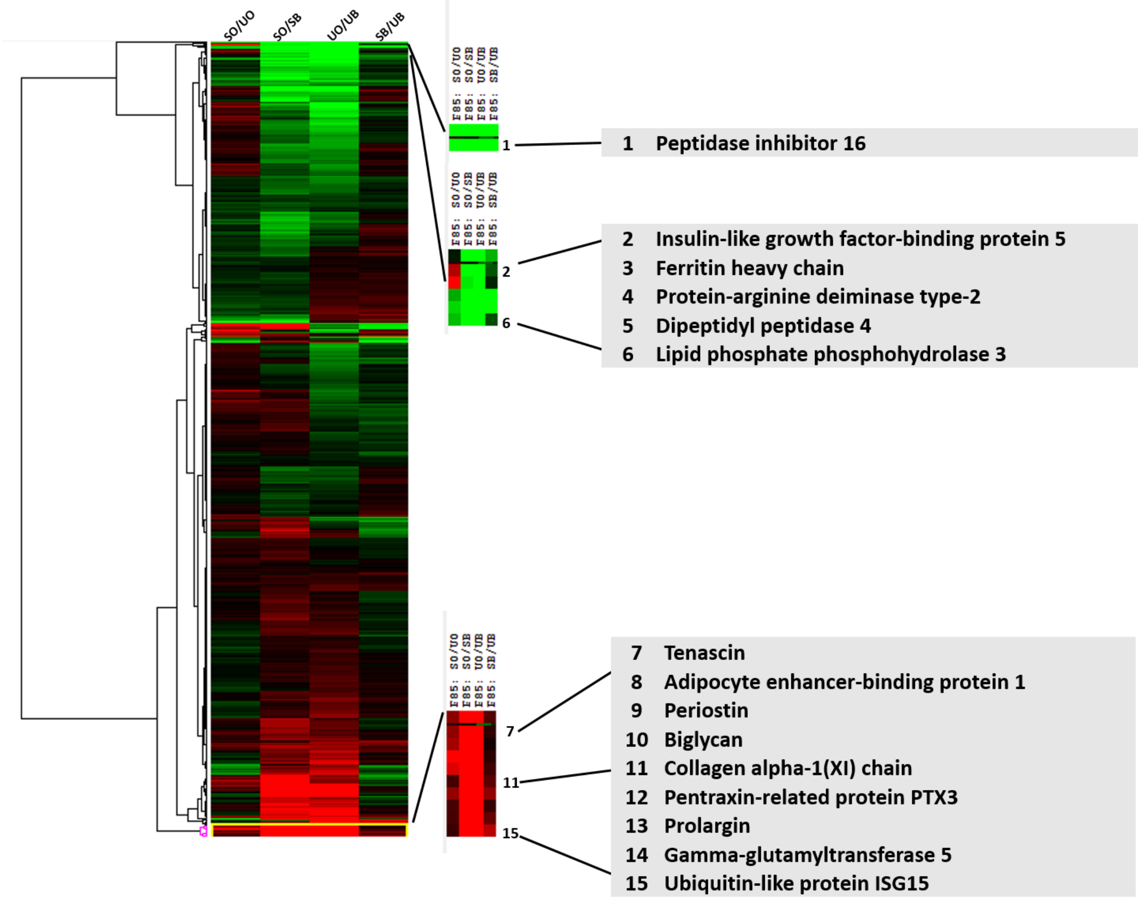
Common proteins between groups	
1	1814
2	2364
3	2351
4	1819
1 only	99
2 only	154
3 only	135
4 only	111
1,2	152
1,3	153
2,3	338
1,4	65
2,4	173
3,4	167
1,2,3	400
1,2,4	145
1,3,4	156
2,3,4	358
1,2,3,4	644

Figure 46: Reproducible proteins overlapping (Venn diagram).

From the MS analysis 3446 proteins were quantified. These proteins were present in all 4 samples. Relative abundances between samples were filtered for the equation ABS-SD>0 to eliminate replicates when the SD was greater than the mean. Different proteins passed the filter and the 4 protein lists were entered in http://bioinfogp.cnb.csic.es/tools/venny/index.html for the Venn diagram to show the overlapping of the same proteins; and on http://www.pangloss.com/seidel/Protocols/venn4.cgi to visualise the overlap. Only 644 proteins out of the 3446 initially quantified, were found to be reproducible across the 4 groups.

4.4.3 Hierarchical clustering

Obtaining a comprehensive visualisation of the protein modulation is helpful to understand the treatment that may have influenced protein expression. On this basis, data were clustered by Euclidian distance, genes that were related on the bases on the genome projects were linked together in a three branch. Hierarchical clustering of the total 644 reproducible proteins highlighted the biological reasons for protein modulation. The groups where the protein modulation is more enhanced where SO/SB and UO/UB demonstrating that in spite of the sorting, the cells responded highly to the osteo-induction (Figure 47). The clustering in Figure 48 represents the 26 proteins highly modulated across the 4 groups. Overall the most upregulated protein was biglycan and the most downregulated was peptidase inhibitors 16.



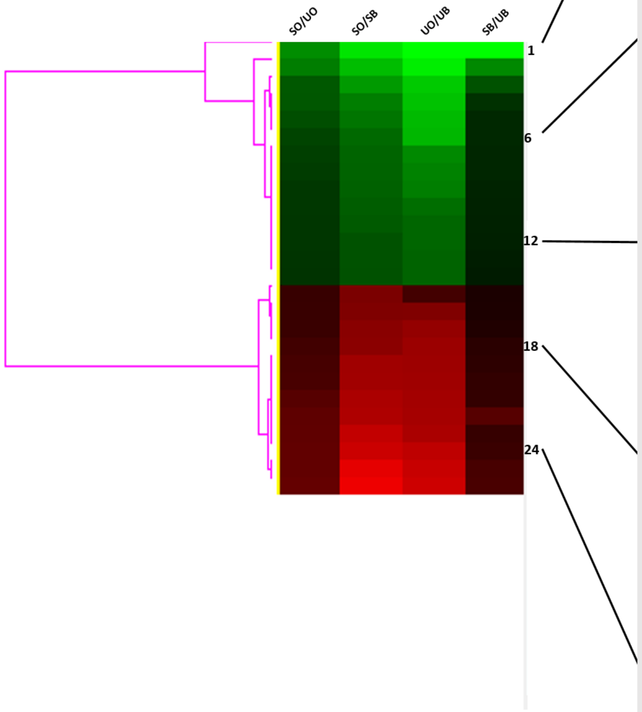
1 Peptidase inhibitor 16

2 Insulin-like growth factor-binding protein 5
3 Ferritin heavy chain
4 Protein-arginine deiminase type-2
5 Dipeptidyl peptidase 4
6 Lipid phosphate phosphohydrolase 3

7 Tenascin
8 Adipocyte enhancer-binding protein 1
9 Periostin
10 Biglycan
11 Collagen alpha-1(XI) chain
12 Pentraxin-related protein PTX3
13 Prolargin
14 Gamma-glutamyltransferase 5
15 Ubiquitin-like protein ISG15

Figure 47: Clustering of the 644 differentially regulated proteins in F85.

644 differentially expressed proteins clustered by Euclidian distance using a complete linkage method (left). On the right, closely clustered groups of proteins with similar expression. The red indicates the upregulation and the green the downregulation.



1 Peptidase inhibitor 16
2 Insulin-like growth factor-binding protein 5
3 Ferritin heavy chain
4 Dipeptidyl peptidase 4
5 Protein-arginine deiminase type-2
6 Lipid phosphate phosphohydrolase 3
7 A-kinase anchor protein 12
8 Vascular cell adhesion protein 1
9 Putative phospholipase B-like 2
10 Tripartite motif-containing protein 26
11 Coiled-coil-helix-coiled-coil-helix domain-containing protein 1
12 Myosin-10
13 Collagen alpha-1(III) chain
14 Bromodomain-containing protein 9
15 Solute carrier family 2, facilitated glucose transporter member 1
16 Neprilysin
17 Tenascin
18 DNA topoisomerase 2-alpha
19 Periostin
20 Collagen alpha-1(XI) chain
21 Pentraxin-related protein PTX3
22 Adipocyte enhancer-binding protein 1
23 Prolargin
24 Biglycan
25 Gamma-glutamyltransferase 5
26 Ubiquitin-like protein ISG15

Figure 48: Clustering of the 26 highly differentially regulated proteins in F85.

Top 26 differentially expressed proteins clustered by Euclidian distance using a complete linkage method (left). On the right, the full list. The red indicates the upregulation and the green the downregulation.

4.4.4 FunRich and STRING

From ontology classification of the proteins using FunRich it was found that the main activated biological process and pathway was metabolism (Figure 51). The majority of the proteins (26.4%) derived from the cytoplasm of the cells. The main molecular function identified was transporter activity (Figure 52). When the 3446 proteins were interrogated for their interactions, a highly populated network of interaction map was found. 11 were the nodes where proteins aligned the most (Table 13). STRING was used to show protein-protein interaction map, resulting in a dense interconnection map between the 644 proteins differetially regulated. No main nodes were identified with STRING, the results were used to show the vast set of protein interactions.

In silico analysis of the tissue expression profile of the most regulated proteins (upregulated or downregulated) (list n3 section 4.3.7.1) in other types of tissues was analysed using a heatmap tool in FunRich. It was discovered that the expression varied greatly across all tissues. The results might indicate a stronger expression in the adult tissue compared to the foetal, and showed that Myosin 10 was highly expressed across the set of tissues analysed in FunRich. These heatmaps were built according to the data collected from the Human genome project. The results were integrated manually with the mass spectrometry results to show similarities, based on the averaged protein expression between conditions. No correlation could be found (Figure 54).

Table 13: Nodes identified by FunRich protein-protein interaction map.

Node	Gene name	Protein name
1	GRB2	Growth factor receptor-bound protein 2
2	SMAD2	Mothers against decapentaplegic homolog 2
3	EIF1B	Eukaryotic translation initiation factor 1b
4	PRKAB1	5'-AMP-Activated Protein Kinase
5	ARF6	ADP-ribosylation factor 6
6	GSTK1	Glutathione S-transferase kappa 1
7	PHLDA3	Pleckstrin Homology-Like Domain
8	HGS	Hepatocyte growth factor-regulated tyrosine kinase substrate

9	EGRF	epidermal growth factor receptor
10	ACTB	Citoplasmic actin
11	YWHAG	Tyrosine 3-Monooxygenase, Tryptophan 5-Monooxygenase Activation Protein,

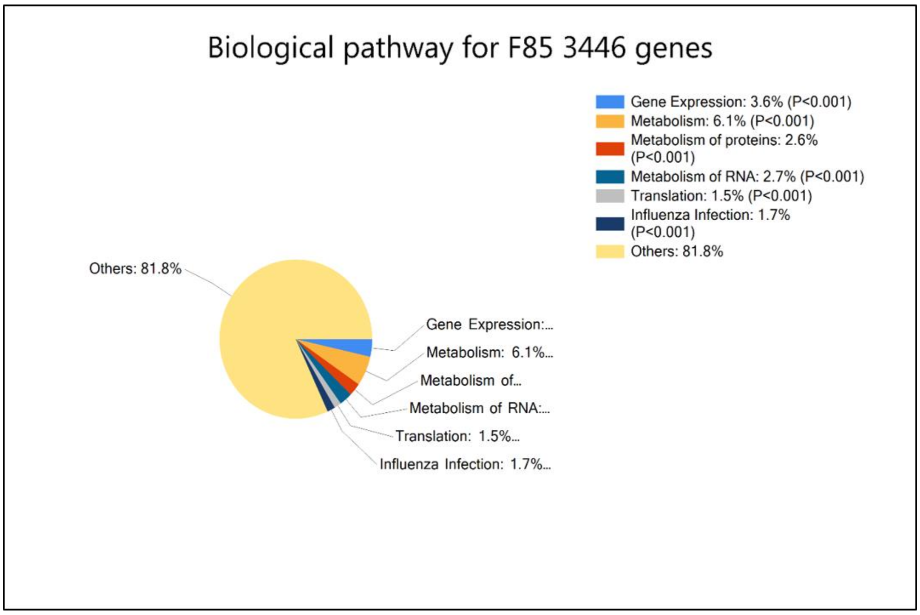


Figure 49: Biological pathway classification.

Proteins were classified according to biological pathway using FunRich.

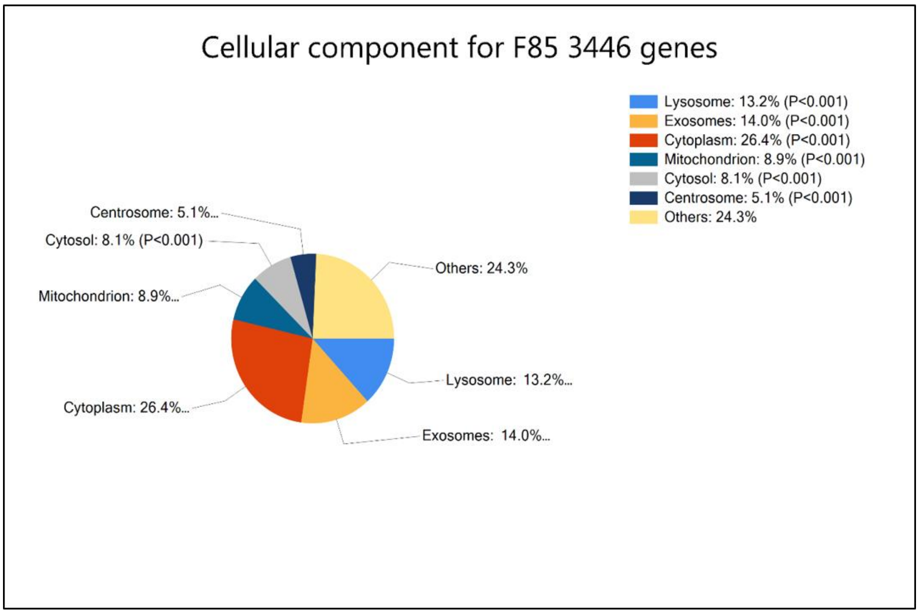


Figure 50: Cellular component classification.

Proteins were classified according to cellular component using FunRich.

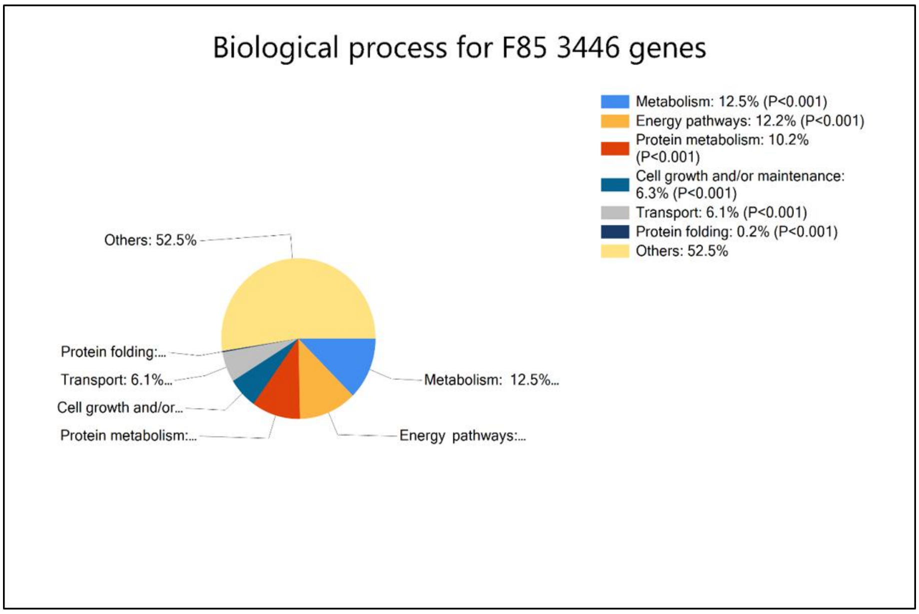


Figure 51: Biological process classification.

Proteins were classified according to biological processes using FunRich.

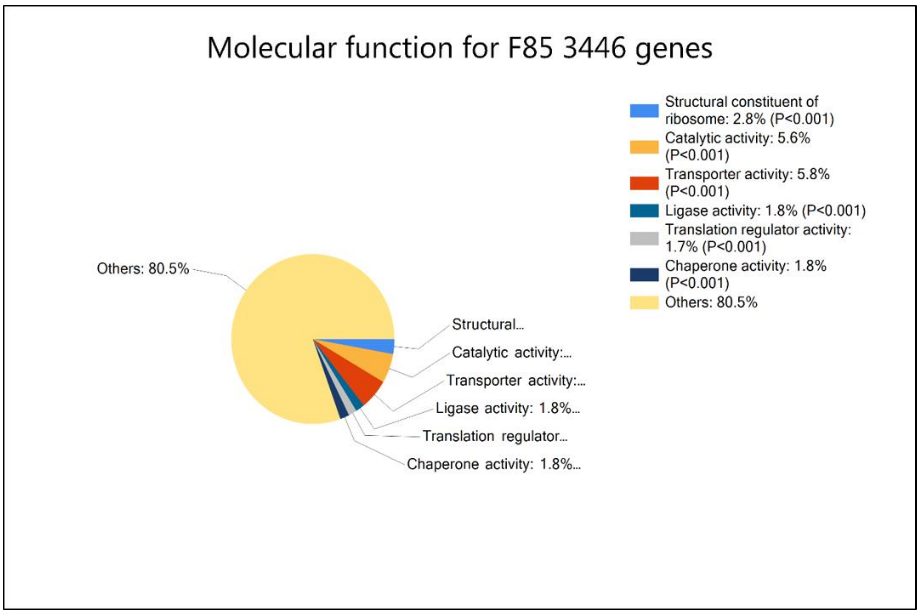
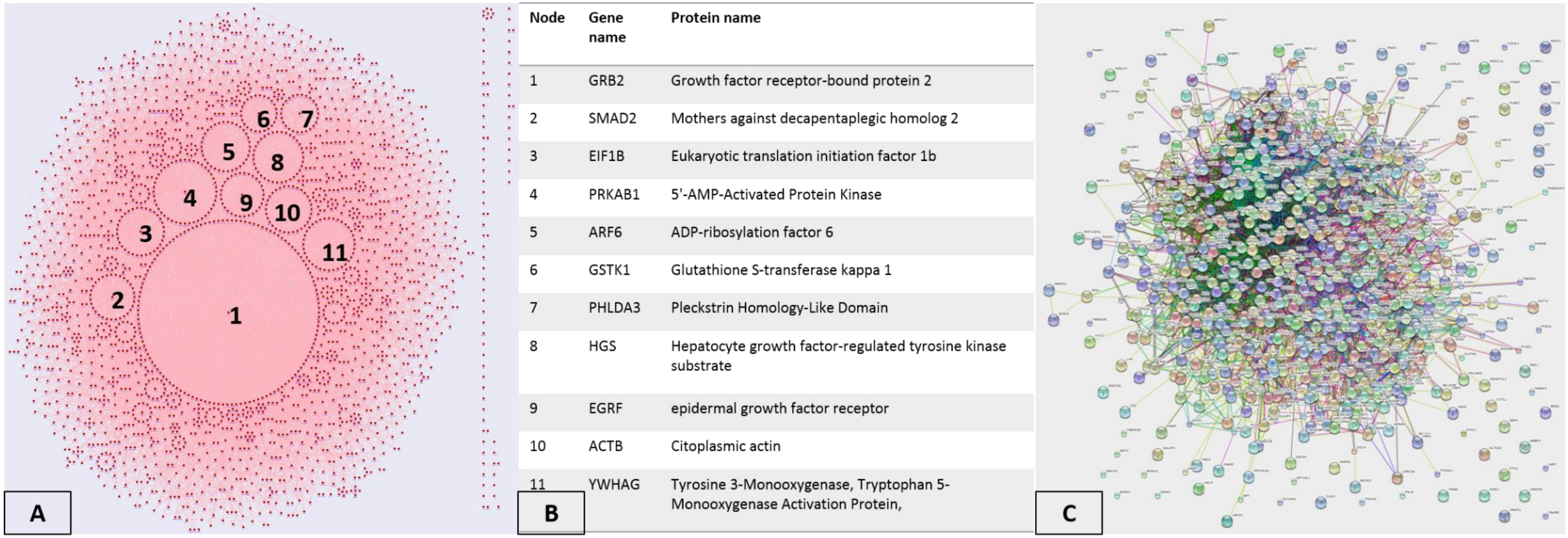


Figure 52: Molecular function classification.

Proteins were classified according to molecular function using FunRich.



Node	Gene name	Protein name
1	GRB2	Growth factor receptor-bound protein 2
2	SMAD2	Mothers against decapentaplegic homolog 2
3	EIF1B	Eukaryotic translation initiation factor 1b
4	PRKAB1	5'-AMP-Activated Protein Kinase
5	ARF6	ADP-ribosylation factor 6
6	GSTK1	Glutathione S-transferase kappa 1
7	PHLDA3	Pleckstrin Homology-Like Domain
8	HGS	Hepatocyte growth factor-regulated tyrosine kinase substrate
9	EGRF	epidermal growth factor receptor
10	ACTB	Citoplasmic actin
11	YWHAG	Tyrosine 3-Monooxygenase, Tryptophan 5-Monooxygenase Activation Protein,

Figure 53: Network of interaction between the discovered proteins.

A: Interactions between 3446 proteins using FunRich. The circles represent protein nodes. B: List of the proteins that form the nodes in image A. C: Interaction between 644 proteins, shown in the confidence view produced by STRING. The different line colours indicate the different types of protein interaction (stable physical associations, transient binding, substrate chaining, information relay and others). In STRING it was only possible to use the 644 protein list and not the 3446 due to system capacity failure.

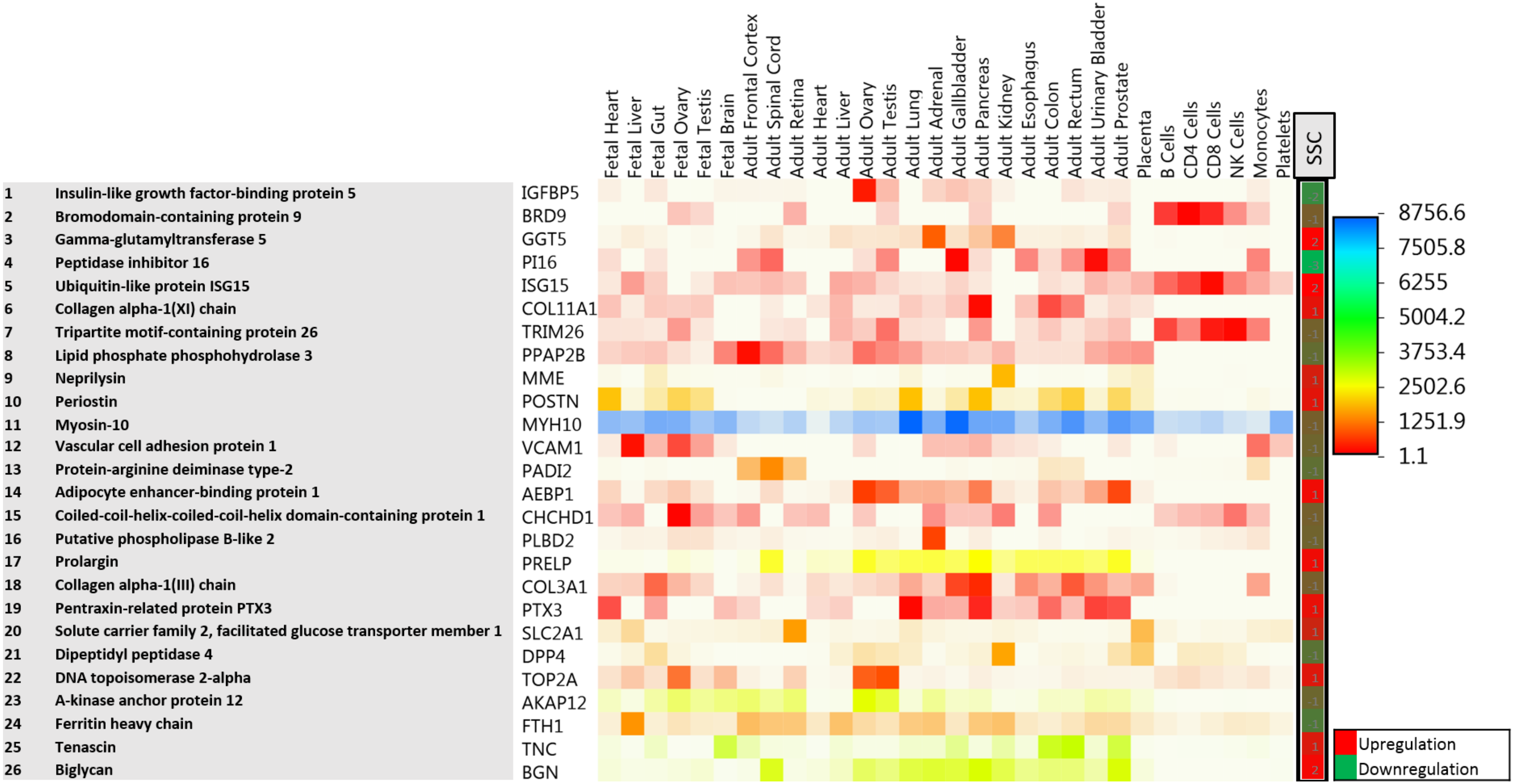


Figure 54: Heatmap of the 26 proteins found in F85 when present in other tissue.

Tissues expression of the top 26 differentially expressed proteins in skeletal stem cells. The heatmap was based on the genome project data and was created using FunRich. Red colour indicates increase of protein abundance and blue colour indicates depletion.

4.5 Discussion

4.5.1 Overview

In this study, high precision mass spectrometry-based proteomics was employed to assess global protein expression differences with an aim to identify specific proteins of potential clinical utility for SSCs. Within this context, the study examined protein expression differences between basal, osteogenic, sorted and unsorted cell populations. Considering the paucity of proteomic studies on SSCs, this study, exploratory in nature, applied an established method encompassing state-of-the-art technologies (offline RP-C_{18} chromatography, and ultra performance nano-capillary chromatography, nanospray ionization, FT-Orbitrap based ultra-high resolution mass spectrometry) to augment knowledge of SSCs proteomics. The overall aim was to evaluate possible marker candidates and set up learning outcomes to consider in subsequent proteomic studies. For this study, an enrichment of the heterogeneous BM stromal population was achieved by sorting cells for the Stro-1 marker (Figure 55).

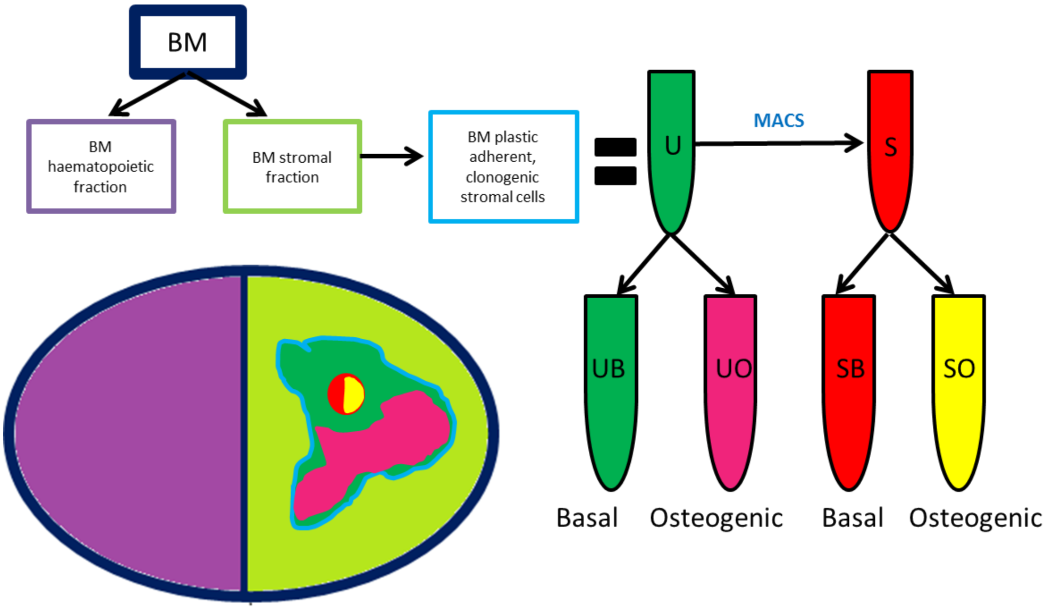

BM
BM haematopoietic fraction
BM stromal fraction
BM plastic adherent, clonogenic stromal cells
U
MACS
S
UB
UO
SB
SO
Basal
Osteogenic
Basal
Osteogenic

Figure 55: F85 BM population enrichment.

BM stromal fraction cells were sorted as expressing the Stro-1 epitope. Both sorted and unsorted populations were osteogenically induced for 14 days. The image represents the cell population distribution within the bone marrow. BM, bone marrow; U, unsorted; UB, unsorted basal; UO, unsorted osteogenic; S, sorted; SB, sorted basal; SO, sorted osteogenic; MACS, magnetic activated cell sorting.

In this study, biglycan was found to be the most upregulated protein and the expression of biglycan was increased following culture in osteo-inductive media. Biglycan was highly upregulated in the SO/SB group followed by UO/UB, SO/UO and SB/UB. Furthermore, biglycan upregulation is not only associated with osteo-induction but also to the Stro-1 isolation. It is known that biglycan in conjunction with decorin is one of the two small chondroitin sulphate proteoglycans present in the developing bone (Young, Kerr *et al.*, 1992). In skeletal tissue, biglycan is specifically localised in cartilaginous and bone components of the developing bone while there is a wide distribution in non-skeletal tissues including kidney, heart, lung, adrenal gland, aorta, skin, tendon, skeletal muscle, eye, peripheral nerve and placenta (Bianco, Fisher *et al.*, 1990). Biglycan is associated with osteoblast differentiation (Parisuthiman, Mochida *et al.*, 2005) and with the induction of cytoskeletal changes in lung fibroblast which increase cell migration (Tufvesson and Westergren-Thorsson, 2003). Recently biglycan has been associated with tumour endothelial cells (TEC) and has been considered a good candidate as a specific marker for TEC (Yamamoto, Ohga *et al.*, 2012). Overall it is known that biglycan plays an important role during osteogenic induction although this protein has not been associated with the Stro-1 phenotype as observed in this work. The phenotype of the sorted cells was observed to modulate the over-expression of biglycan, although considering that no further investigation has been completed at this stage, and recognising these results could be patient specific, this observation cannot be confirmed.

In this study, downregulation of peptidase inhibitors 16 was observed, in particular peptidase inhibitor 16 was found highly downregulated in SB/UB, > UO/UB > SO/SB > SO/UO. A peptidase inhibitors protein is a proteolytic enzyme that attenuates the activity of peptidases (Laskar, Mandal *et al.*, 2010). These proteins are important in medicine and biotechnology and have been studied extensively, however peptidase inhibitors nomenclature is still in its infancy (Rawlings, Tolle *et al.*, 2004). The peptidase inhibitors protein 16 has, as yet, not been extensively studied and the mechanism underlying its function, is poorly understood. It is known that this protein is abundant within semen and its mRNA presence was observed in prostate, testis, ovary and intestine. Furthermore, peptidase inhibitors 16 downregulation have been shown in a murine model to be associated with prostate cancer (Reeves, Xuan *et al.*, 2005). There are no studies that report the presence of this protein in BM cells.

Overall, SO/SB displayed the greatest upregulation, indicating that, in combination, both the Stro-1 phenotype and the osteogenic induction contributed to cellular mechanisms driving the upregulation of a number of proteins.

Proteins that were more than 2 fold upregulated included ubiquitin-like protein ISG15, adipocyte enhancer-binding protein 1, periostin and tenascin. Ubiquitin-like protein ISG15 functions are largely unknown, studies have shown that it is indispensable for STAT1 and interferon signalling (Osiak, Utermohlen *et al.*, 2005), however no connection to the bone phenotype, or to MSCs, SSCs, or BM cells has been published.

Adipocyte enhancer-binding protein 1 is known to play a role in the skeletal system development, especially by enhancing adipocytes proliferation and reducing adipocytes differentiation. This protein also appears to be involved in the differentiation of osteoprogenitors cells into osteoblast (Ohno, Hashimoto *et al.*, 1996). In this study adipocyte enhancer-binding protein 1 was upregulated in the osteogenic conditions which agrees with previous findings.

Periostin and tenascin have been widely associated with the bone phenotype in various studies (Burk, Gittel *et al.*, 2014; Heo, Shin *et al.*, 2015). Periostin is mainly involved in cell adhesion and is predominantly found in epithelial cells, it is also found in many carcinomas. One particular study has localised periostin within bone nodules and has suggested a role in extracellular matrix mineralisation (Coutu, Wu *et al.*, 2008). In this work periostin was upregulated in the SO/SB group, which suggests that both osteogenic induction and sorted phenotype may play an important role in periostin regulation.

Tenascin is involved with osteoblast adhesion and differentiation, it has been demonstrated that its function is regulated by osteogenic induction (Morgan, Wong *et al.*, 2011). This is congruent with the results in this work that demonstrated an upregulation in the SO/SB group which indicated that not only the osteogenic induction regulates protein expression but also the Stro-1 phenotype enhances tenascin regulation. However, the mechanisms underlying these processes remain to be elucidated.

Data was investigated for differences in protein expression and also for heterogeneity of expression. The SD of each group was used to determine the degree of heterogeneity. Results indicate that there are more differences in the samples that were induced osteogenically and sorted by the Stro-1 marker than in the samples where the cells were cultured under basal condition and no antibody selection was applied. The heterogeneity was also visualised by hierarchical clustering (Figure 47 and Figure 48).

With the help of annotation software (FunRich), it was highlighted that the majority of the proteins found were involved in the metabolism of the cell.

Multiple interactions between the proteins were highlighted by the pathway analysis map (Figure 53). The map showed the highly intrinsic set of interactions between the proteins. In Image A of Figure 53, 11 main protein nodes were highlighted (Figure 53, B).

The GRB2 gene resulted to be the centre of the most populated node in the network map. The GRB2 gene is indirectly involved into the formation of mesoderm and ectoderm being the adapter protein that participates in the action of the fibroblast growth factor (FGF) (Thisse and Thisse, 2005). In fact GRB2 binds to the FRS2 protein (a receptor tyrosine kinase), linking it to the Ras/MAPK signalling pathway which is essential for cell proliferation and differentiation (Kouhara, Hadari *et al.*).

It was interesting to see that from the Human genome project data, from the 27 highly modulated proteins found in this study, only myosin 10 presented a different expression in other tissues, being highly upregulated in adult lungs and gallbladder (Figure 54). In relation to osteogenesis myosin 10 has been shown to promote early osteogenesis in cells grown on topographical models (Ozdemir, Xu *et al.*, 2013).

The results in this work were compared to the results of previous proteomic studies of hBMSC. The list of proteins identified in a recently published study (Mindaye, Ra *et al.*, 2013) was correlated to the list of proteins found in this work in a Venn diagram (Figure 56). The study found 7753 proteins in 6 patients against the 3446 proteins found in this work in one patient. Out of the 3446, 1671 proteins were in common with this study and the MALDI MS method; and 7 with the ESI MS method and other 7 proteins with both methods (Table 14, Table 15).

It is a challenge to compare the two studies as they were performed employing different methodologies, and most importantly the cell population selected was different. In the current study, cells isolated from the BM at passage two were used, while in Mindaye, Ra et al. study cells were purchased at passage 1 and expanded to passage 3, 5 and 7. Furthermore, the age cohort is markedly different, F85 (this study) compared to F22, F23, F24, F31, M24, M23 (Mindaye, Ra et al.). In the study by Mindaye, Ra et al., cells were isolated from BM aspirates, while in the current study BM was derived from hip replacement surgery. It has been reported previously that cell populations from BM aspirates present different characteristics when compared to cell populations extracted from BM removed from femoral intramedullary cavity using a curette (Cox, Boxall *et al.*, 2012).

Another major difference is the observation that Mindaye, Ra et al. did not attempt to quantify the protein expression. Mindaye's group concentrated on protein identification while in this work a major focus was given to the quantification of the protein expression employing iTRAQ labelling system.

Mindaye, Ra et al. group employed less stringent analytical criteria, the FDR was 3% against the 0.05% in this study, the precursor tolerance of 20 ppm against 10 in this study. Furthermore, 3 missed cleavage were allowed instead of 1 as per this study. This might have contributed to the discrepancy in proteins identified.

In summary the presented results implicated osteogenesis and Stro-1 differentiation as a major cause for protein modulation, however it is important to state the need for further studies to confirm these findings.

Until now, the lack of specific biomarkers for the identification of SSCs have hampered the development of bioassay for exploiting SSC regeneration potential within clinical settings (Kassem, 2006). The discovery of unique biomarkers for SSCs would constitute a sought after aid to build a technology to allow specific isolation tools for SSCs. There is however still a limited understanding of the surface proteomic profile of SSCs.

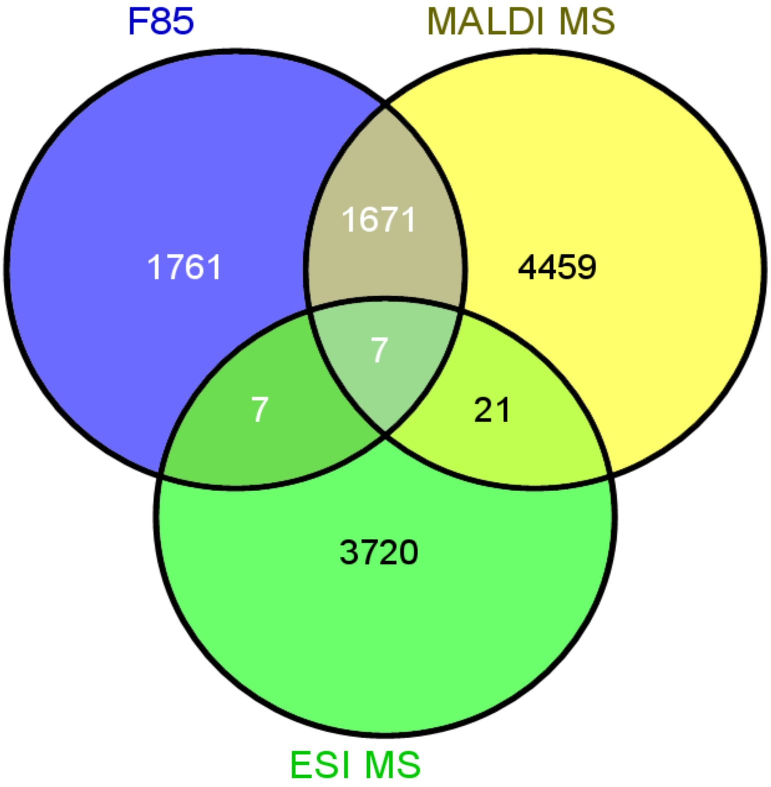


Figure 56: Current study compared to Mindaye, Ra et al.

This diagram shows common proteins between this work and a work recently published (Mindaye, Ra *et al.*, 2013). F85 indicated the patient analysed in this thesis and MALDI MS and ESI MS indicated the results analysed in Mindaye, Ra et al study. Mindaye's study identified 7753 proteins in 6 patients against the 3446 proteins detailed in this work from one patient. Out of the 3446 it was found that 1671 proteins were in common with this thesis work. 7 proteins discovered by MALDI MS method and 7 with the ESI MS were in common with this thesis study (Table 14,Table 15).

Table 14: 7 common elements in "F85", "MALDI MS" and "ESI MS".

#	Protein name
1	Calcium/calmodulin-dependent protein kinase type II subunit delta
2	6-phosphogluconolactonase
3	Ras suppressor protein 1
4	Catenin delta-1
5	B-cell receptor-associated protein 31
6	UDP-glucose 6-dehydrogenase
7	Fibronectin

Table 15: 7 common elements in "F85" and "ESI MS".

#	Protein name
1	S-phase kinase-associated protein 1
2	ADP-ribosylation factor 6
3	ADP-ribosylation factor 5
4	Obg-like ATPase 1
5	ADP-ribosylation factor 3
6	3-ketodihydrosphingosine reductase
7	ADP-ribosylation factor 4

4.5.2 Limitation of the study: Proteome coverage

The proteomics field, by definition, being the study of the protein ensemble of an organism faces some challenges. The main challenge to consider is the complexity of the protein sample analysed. Proteins are complex chemical entities embodying a vast range of parameters: abundance, size, charge, physical and chemical characteristics. Often, the same protein mixture can contain peptides at two different extreme values within the same parameter. Hence mass spectrometry faces the demand to ensure the same sensitivity across the vast range of each individual parameter, *i.e.* needs to guarantee a sensitive dynamic range. (The dynamic range in this study has a minimum of 12-orders of magnitude).

Along these lines, the higher abundance proteins can mask or interfere with the accurate expression of the lower abundant proteins. This becomes especially true in the application of the iTRAQ relative quantification. In order to reduce this co-isolation phenomenon, the MS method employed in this thesis used orthogonal two-dimensional chromatography of the protein surrogate tryptic peptides combined with their sensitive and ultra-high resolution mass spectrometry.

Additionally, iTRAQ labelling has been show to constitute a limitation. It has been discovered that the intensity of the iTRAQ ratios can be flattened due to their chemistry, *e.g.* the fold regulation is milder that in the case of microarrays. Another limitation of iTRAQ corresponds to the level of phenylalanine in a sample. The phenylalanine level interferes with the chemistry of the probes. In spite of these limitations iTRAQ still remains the state of the art in quantitative mass spectrometry and although the results might be flatter, the direction of increase is correct (Ow, Salim *et al.*, 2009).

However, iTRAQ LC-MS proteomics has unsurpassed confidence in the global assessment of relative proteome expression even for the lower abundant proteins as evaluated with independent biochemical assay ((Roumeliotis, Halabalaki *et al.*, 2013); (Papachristou, Roumeliotis *et al.*, 2013);(Bouchal, Dvořáková *et al.*, 2015) or immunohistochemistry (Garbis, Tyritzis *et al.*, 2008).

Another approach to further increase analytical accuracy and sensitivity is to undertake a sub-proteome fractionation prior to mass spectrometry (Boisvert, Lam et al., 2010). Hence, subcellular fractionation of the complex protein mixture *a priori*, would decrease sample complexity and increase sensitivity (Seshi, 2006).

Refinement of chromatography method would greatly sharpen the sensitivity of the analysis.

Further limitations, specific to this study encompassed the cell population used, selected by an unspecific marker and cultured *in vitro* for 4 weeks. With the advancement of mass spectrometry, less protein lysate might be needed, hence MS could potentially be run on freshly isolated SSCs versus *in vitro* cultured cells. In the case of cell surface biomarkers, isolation of the membrane proteins prior to proteomics might be of an advantage (Foster, Zeemann *et al.*, 2005; Graneli, Thorfve *et al.*, 2014; Holley, Tai *et al.*, 2015).

A limitation in distinguishing differences between Stro-1 sorted cells and unsorted relies with the fact that the Stro-1 fraction is also encompassed within unsorted cells. Comparing the 2 fractions does not highlight differences due to sorting, as Stro-1 positive cells were also present in the unsorted population. An analysis based on Stro-1 positive versus Stro-1 negative cells would highlight protein differences, although in this type of set up, challenges with culturing the negative fraction could be faced considering Stro-1 negative cells are not plastic adherent.

The limited number of sample (1 patient) and the fact that due to patient data confidentiality there was a lack of knowledge regarding the type of medication that the patient could have been exposed to; might have also affected the MS analysis.

A further limitation that could potentially impact the analysis is cell cycle. The cell cycle could have affected the proteome results as previous studies have revealed dynamic differences in the proteome of mammalian cells attributable to the cell cycle (Sigal, Milo *et al.*, 2006).

Ultimately the value of proteomic data is defined by its biological significance. Proteomics looks only at a ‘temporal defined’ cell type, which alone cannot account for the complexity of the biological phenotype in the long run (Bessarabova, Ishkin *et al.*, 2012).

4.5.3 Data analysis limitation

Throughout the proteomic era, there have been significant advancements in data generating technologies, however our ability to generate data outweigh our ability to analyse the data (Haga and Wu, 2014).

A variety of tools are in place to help filter and manage the vast list of proteins generated by proteomics analysis, with the aim to answer the initial biological quest. The results that these tools can offer often vary greatly, hence the true significance of the data can only be achieved with experimental repeats and further downstream validation experiments. This MS study, presented in this chapter, only analysed the proteome of one patient, therefore as discussed, the results could be patient specific.

4.6 Conclusion

The results in the context of the stated hypothesis in section 4.2 were examined:

I. The hyphenated use of Liquid Chromatography-Mass spectrometry is a sufficiently sensitive approach to compare differences in protein expression between cells treated in basal and osteogenic conditions and sorted by Stro-1. This hypothesis is **true.**

The differences observed between S and U, B and O, offer the potential for new hypothesis to understand the underlining mechanism of protein modulation in SSCs.

I. Stro-1 selected cells exhibit a different set of modulated proteins in comparison to unselected cells. This hypothesis is **false** or the method used was not sufficiently adapted to prove the hypothesis to be true.

Stro-1 is only expressed in native SSCs up to 10 days after *in vitro* culture. The cells used in these experiments were grown for 28 days to allow for enough material for the proteomic study. Within 28 days in culture both Stro-1 sorted and unsorted populations acquired the same proteomic profile. Nevertheless if advancements in mass spectrometry would allow for fewer cells to be analysed, the ideal experiment would be characterised by native SSCs (*ex-vivo*) avoiding *in vitro* expansion. Furthermore, the Stro-1 fraction is also encompassed within the unsorted cells, hence the differences could have been mitigated.

To recapitulate, the current study has established a protocol for proteomic analysis using cultured SSCs. This protocol systematically discovered the proteome of SSC and outlined a new research pathway for the next proteomic study (following chapter). Observations, biologically suggesting a strong protein modulation due to osteo-induction were the main finding of the study.

5. Chapter 5: Mass spectrometry 2

I want to thank Dr Spiro Garbis, Antigoni Manousopoulou and Harvey Johnston for their help and input with the running of the mass spectrometry.

I am grateful to Dr. Jeongmin Woo and Dr. Christopher H. Woelk for their help and input in the MetaCore analysis.

A small *in silico* study analysing the results from chapter 3 and this chapter is contained at the end of this chapter.

5.1 Introduction

Stem cells deriving from the bone marrow (BM), exhibit extraordinary tissue differentiation potential which constitutes an attractive characteristic for the field of regenerative medicine and tissue engineering. SSCs have been utilised for the last 20 years as promising candidates to meet the need of novel therapies related to the increase in bone diseases. One of the major challenges in the utilisation of stem cells for regenerative medicine purposes is the identification and discrimination of the stem cells themselves from other types of cells in the bone marrow, as stated previously. It is critical to separate a pure population of SSCs to ensure the regeneration capacity is maximised (Gothard, Greenhough *et al.*, 2014). Stem cell populations do not exist alone but coexist surrounded by other cell populations, some of which are already differentiated with limited tissue regeneration potential.

Challenges in stem cell identification rely on finding unique identification markers. In this thesis, a novel approach based on mass spectrometry, is employed to unravel the proteomic profile of skeletal stem cells and to create a new hypothesis regarding the identity of markers for SSCs (Eckart, Holthausen *et al.*, 1998).

As previously explained, ideally, identification of unique biomarkers for the isolation of skeletal stem cells could lead to robust and specific isolation protocols ensuring a pure population of cells with high differentiation potential. High differentiation potential is correlated to better prospectives for regenerative medicine and tissue engineering. A copious amount of markers have been found to be expressed on the cell surface but none has to date been specific and exclusive for SSCs. In spite of that, SSCs are isolated according to the research group preference on the choice of marker or a combination of thereof.

With the knowledge acquired on cell expansion, cell harvest and preparation for MS in the previous chapter (MS1), the differential protein expression profile of SSCs cultured in basal and osteogenic media and selected for the Stro-1 marker of 4 patients is explored in this second mass spectrometry study (MS2), in an effort to expose novel cell surface markers for SSCs identification.

As in mass spectrometry study 1 (MS1), the method chosen to achieve the full proteomic study of SSCs was nanospray ionization–ultra high resolution mass spectrometry (LC –nESI-MS). WB was subsequently employed to validate the mass spectrometry results.

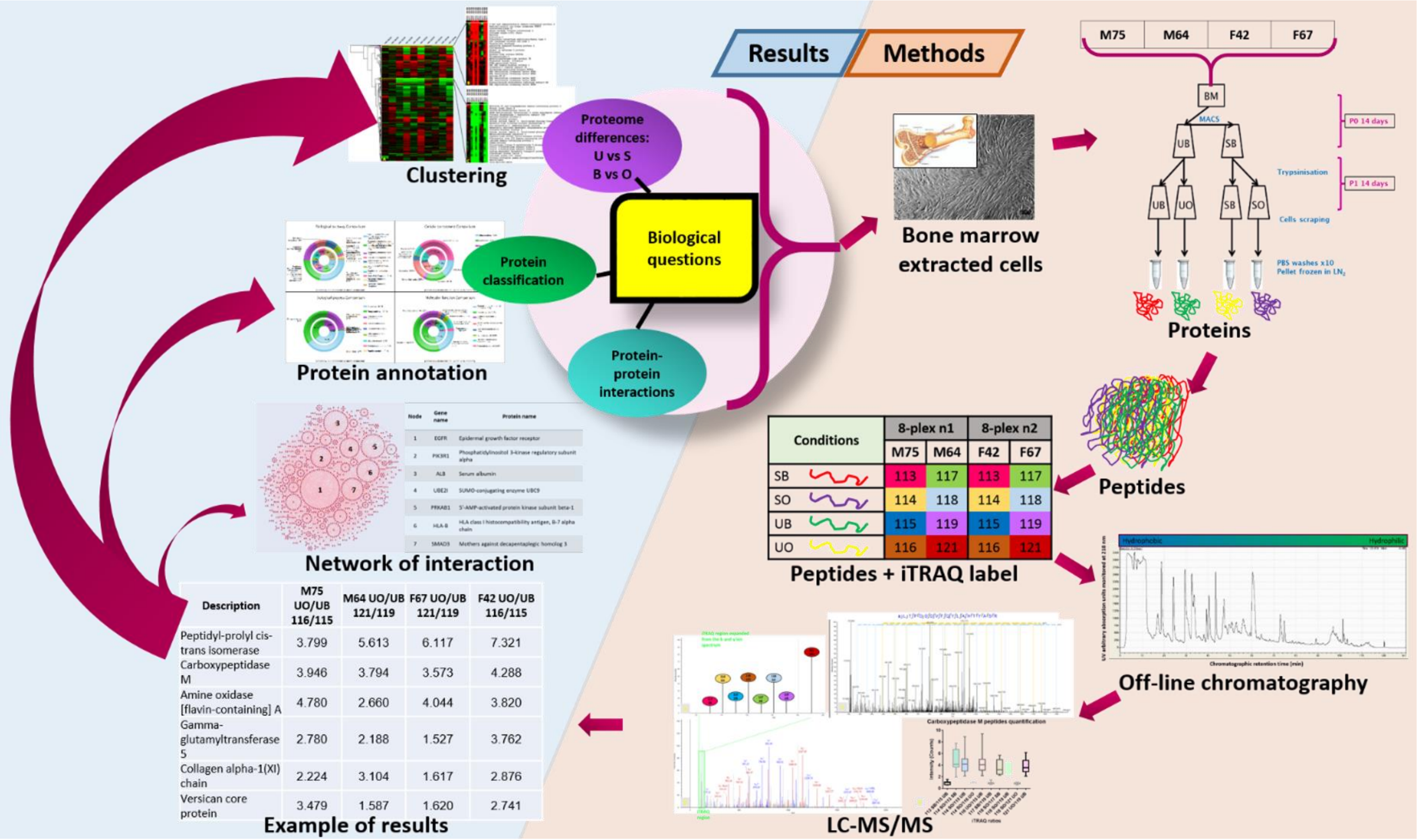
Research aims, methodology and results are pictographically shown in Figure 57, and described and discussed in the following paragraphs.

5.2 Aims and hypothesis

To bypass SSC identification limitations, a second proteomic study was undertaken employing high pressure liquid chromatography followed by iTRAQ 2D-LC-nESI-MS. The aim of the study was to elucidate the proteomic profile of skeletal stem cells.

The hypotheses of the study were:

I. The Stro-1 sorted population of skeletal stem cells presents an enriched and unique proteomic profile.
II. The Stro-1 osteogenic population presents a different proteomic profile from the Stro-1 cells in basal media.


Results
Methods
M75
M64
F42
F67
BM
MACS
UB
SB
UB
UO
SB
SO
P0 14 days
Trypsinisation
P1 14 days
Cells scraping
PBS washes x10
Pellet frozen in LN_2
Proteins
Bone marrow extracted cells
Peptides
Peptides + iTRAQ label
Off-line chromatography
Hydrophobic
Hydrophilic
LC-MS/MS
Carboxypeptidase M peptides quantification
Biological questions
Proteome differences: U vs S B vs O
Protein classification
Protein-protein interactions
Clustering
Protein annotation
Network of interaction
Example of results

Conditions	8-plex n1		8-plex n2	
	M75	M64	F42	F67
SB	113	117	113	117
SO	114	118	114	118
UB	115	119	115	119
UO	116	121	116	121

Node	Gene name	Protein name
1	EGFR	Epidermal growth factor receptor
2	PIK3R1	Phosphatidylinositol 3-kinase regulatory subunit alpha
3	ALB	Serum albumin
4	UBE2I	SUMO-conjugating enzyme UBC9
5	PRKAB1	5'-AMP-activated protein kinase subunit beta-1
6	HLA-B	HLA class I histocompatibility antigen, B-7 alpha chain
7	SMAD3	Mothers against decapentaplegic homolog 3

Description	M75 UO/UB 116/115	M64 UO/UB 121/119	F67 UO/UB 121/119	F42 UO/UB 116/115
Peptidyl-prolyl cis-trans isomerase	3.799	5.613	6.117	7.321
Carboxypeptidase M	3.946	3.794	3.573	4.288
Amine oxidase [flavin-containing] A	4.780	2.660	4.044	3.820
Gamma-glutamyltransferase 5	2.780	2.188	1.527	3.762
Collagen alpha-1(XI) chain	2.224	3.104	1.617	2.876
Versican core protein	3.479	1.587	1.620	2.741

Figure 57: Process flow diagram of mass spectrometry analysis.

The diagram represents a schematic outline of the study aims, experimental design and results. The proteins of bone marrow extracted cells of 4 patients are labelled with iTRAQ probes and processed for LC-MS/MS. Mass spectra are translated into ratios of protein abundance. Protein data is then analysed to create a cluster gene map, protein annotation and a network of interactions map.

5.3 Materials and methods

5.3.1 Reagents and chemicals

The reagents and the chemicals used in this chapter are detailed in the previous chapter (see section 4.3.4).

5.3.2 Differences with mass spectrometry study 1

Unless stated otherwise the materials employed and the method followed are described in mass spectrometry study 1 (section 4.3). The main differences between the two studies encompass the sample population and the steps in the method used, these differences are summarised in Table 16.

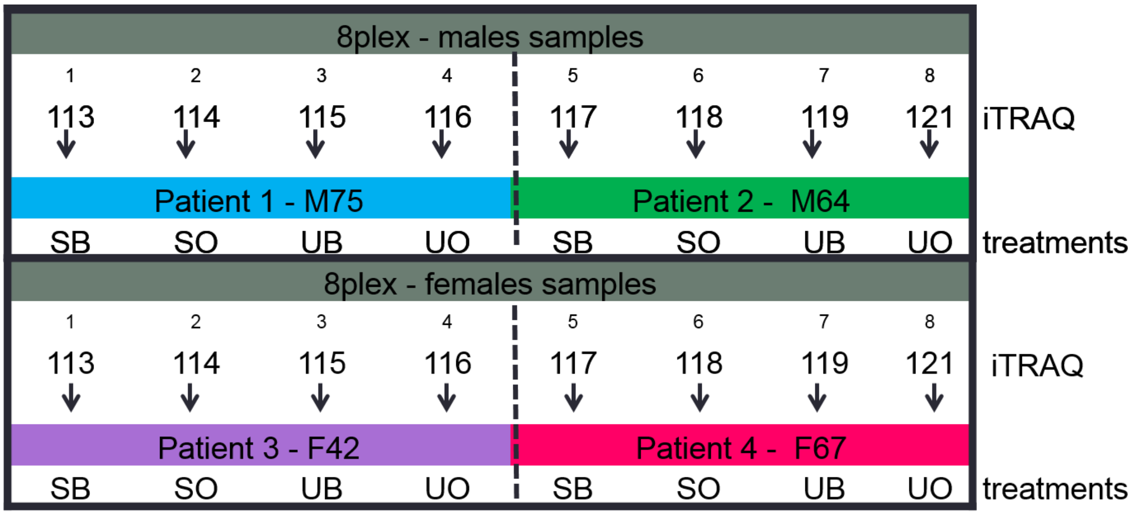
Table 16: Method variation between MS1 and MS2.

The samples were collected at the same manner as MS1. Here 4 patients (2 females and 2 males, no technical replicates) were used instead of 1 patient (with technical replicates) as per MS1. In terms of method, a different chromatography column was used (C_8 instead of C_{18}) which offers a less strict gradient for protein separation and ensures further protein recovery within the middle range. Samples were titrated and once the lysis process was terminated samples were centrifuged and only the supernatant fraction or soluble fraction (SOL) of the cell lysate was considered, whilst in MS1 titration was not adopted and the whole cell lysate (WCL) was used for the study**.** M: male; F: female; C: carbon.

Variation in method (steps)	Mass spectrometry study 1	Mass spectrometry study 2
Patient population	1 patient: F85	4 patients: M75, M64, F42, F67
Cell lysing method	TEAB buffer, WCL fraction used	TEAB buffer, titration, SOL fraction used
Chromatography columns	C_{18}	C_8

5.3.3 Sample preparation

Mass spectrometry samples for this study were deriving from the BM of 4 patients: male aged 75 (M75), male aged 64 (M64), female aged 42 (F42) and female aged 67 (F67); each sorted and unsorted for Stro-1 and cultured in osteogenic or basal conditions for a total of 8 samples (Figure 58). Cell extraction, isolation, expansion and osteogenic differentiation were described in 4.3.2, 4.3.3. Titration with a 21 gauge needle was added to the cell lysis protocol in 4.3.1. Bradford assay, reducing the proteins, blocking cysteine residues, digesting the proteins with trypsin, iTRAQ reagents preparation and labelling and high pH Reverse Phase C_8 chromatography (RP-C_8) were performed as described in section 4.3.5. iTRAQ labels were associated with one unique sample, hence 2 iTRAQ 8plex kits were used to analyse the 8 samples (Figure 58).



8plex - males samples

	1	2	3	4	5	6	7	8
iTRAQ	113	114	115	116	117	118	119	121
	Patient 1 - M75				Patient 2 - M64			
treatments	SB	SO	UB	UO	SB	SO	UB	UO

8plex - females samples

	1	2	3	4	5	6	7	8
iTRAQ	113	114	115	116	117	118	119	121
	Patient 3 - F42				Patient 4 - F67			
treatments	SB	SO	UB	UO	SB	SO	UB	UO

Figure 58: iTRAQ labelling.

Four samples per patient, under different growth and selection conditions (SB, SO, UB and UO), were labelled with a different iTRAQ label. SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.

5.3.4 MS Data processing

Details of peptide identification, protein alignment, raw data processing and algorithm were performed in the same manner of MS1 and are described in sections 4.3.6 and 4.3.7.

The rational of the comparison of the iTRAQ label is given in section 4.3.7. In this study a further comparison is added SO/UB as considered a condition where samples are most different. These samples in fact are the extremity of the bone differentiation pyramid in Figure 59.

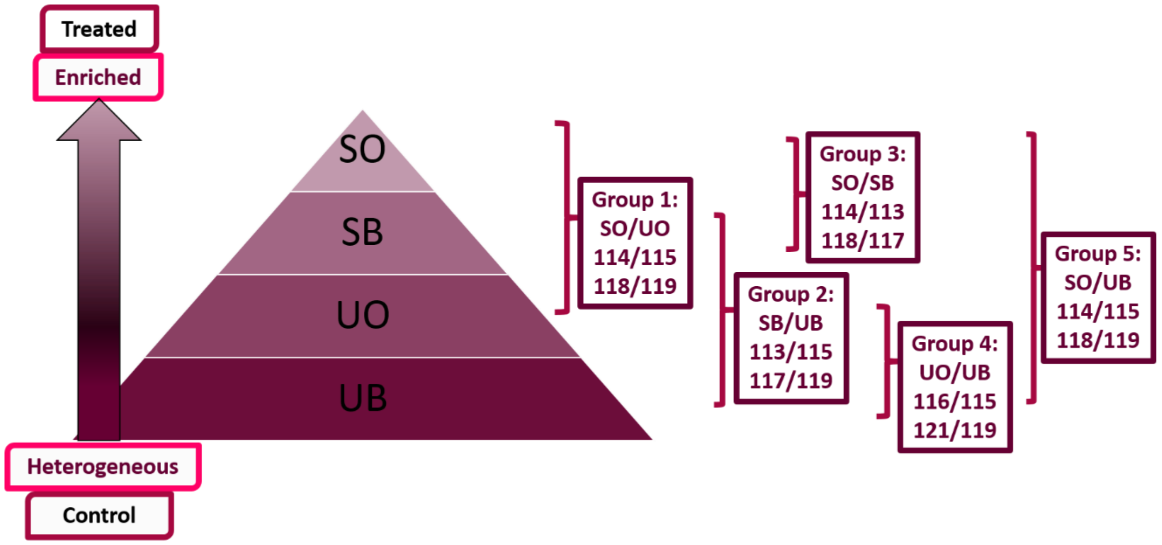

Treated
Enriched
SO
SB
UO
UB
Heterogeneous
Control
Group 1: SO/UO 114/115 118/119
Group 2: SB/UB 113/115 117/119
Group 3: SO/SB 114/113 118/117
Group 4: UO/UB 116/115 121/119
Group 5: SO/UB 114/115 118/119

Figure 59: iTRAQ ratios rationale.

The pyramid represent the bone differentiation hierarchy based on the conditions employed to treat the samples in this study. The bottom of the hierarchy is represented by the unsorted cells in basal media condition (UB). This UB cell population is heterogeneous and it is employed as a baseline control for the iTRAQ ratios. The top of the pyramid is represented by the Stro-1 sorted cells in osteogenic media (SO). This SO population is enriched in respect to the UB, hence is considered as treated. The unsorted osteogenic (UO) and the sorted basal population (SB) are considered as intermediates of this differentiation process. The scope of the ratios is to compare the sorted sample with the unsorted sample under basal and osteogenic conditions and the sorted and unsorted in osteogenic media conditions against the respective sorted and unsorted in basal media conditions.

5.3.5 Downstream data analysis

5.3.5.1 Hierarchical clustering

Hierarchical clustering was performed using the Cluster 3.0 and TreeView 6.0. Node distances were calculated using the Euclidean based metric and then clustered using the complete linkage method. Data were transformed in $\log_2$ prior to analysis. The data set was filtered per gender and ordered from high to low SD. SD was calculated per each row of proteins iTRAQ values, the range used was the 5 iTRAQ ratios per each patient. Proteins with a SD≥0.2 across all patients were employed as the input for this analysis.

5.3.5.2 Upregulated and downregulated proteins

The proteins were filtered on the basis of their expression (upregulation and downregulation). ±1.5 and ±2 fold cut off filters were chosen. Filters were applied to all samples, for each treatment using Microsoft Excel. Proteins passing the filter were employed as input to verify the overlap with the help of a Venn diagram (http://bioinfogp.cnb.csic.es/tools/venny/).

5.3.5.3 Protein annotation and interactions: FunRich and STRING

Data set was filtered per gender and ordered from high to low SD. SD was calculated for each protein in a row-by-row manner. SD range was constituted by the 5 ratios values per each patient in each gender group.

Proteins with a SD≥0.2 were uploaded on the free software FunRich (Functional Enrichment Analysis Tool) (Pathan, Keerthikumar *et al.*, 2015) to create a map of protein interaction, and classified according to biological pathway, biological processes, cellular components and molecular function (Figure 68, Figure 69, Figure 70, and Figure 71 respectively). A map of protein interaction was created and the results were displayed using the ‘packed view’ (Figure 72). The packed view allowed for the visualisation of specifically enriched nodes where a higher degree of distribution corresponded to a higher radius of nodes. Data were emphasized in a circular fashion and placed closely with the other groups having similar degree of distribution.

STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) analysis was attempted to show the complexity of protein-protein interaction (same as in MS1) (Szklarczyk, Franceschini *et al.*, 2015). The genes associated with each of the proteins with a SD≥0.2 between ratios and patients of the same gender were used as the input for the creation of a network of protein interactions map although the list complexity produced was too large for the capacity of the system.

5.3.5.4 MetaCore analysis

Protein–protein interactions were identified using the direct interaction algorithm in MetaCore (Thomson Reuters). The data input for this analysis comprised the top 50 proteins with the highest standard deviation. Ratio values were firstly converted into a logarithm equation (base 2), ensuing the SD was calculated per each ratio and the top 50 proteins identified.

5.3.6 Western blotting validation

The WB method followed was described in section 2.7 and the antibodies used for this chapter were described in Table 17. Alterations to the protocol, described in section 2.7, occurred when the fibrillin-2 (FBN2) antibody was employed. A 7.5% 10 well combo precast gel (Mini-PROTEAN TGX; BioRad) was used due to the big size of the protein (350 kDa). Transfer was run at 40V for 16 h at 4 ˚C. The rest of the process remained the same. Cell lysates employed for this validation were the samples used for the MS1 and this study (MS2). Hela and MG63 cells lysates were used as positive controls for FBN2 and ALP. Human decorin (DCN) synthetic peptide (Abcam ab71694) and MG63 cell lysate were used as positive control for the detection of DCN. 293T cell lysates was used as a positive control for carboxypeptidase M (CPM).

Table 17: Antibodies used for the Western Blotting validation of mass spectrometry studies 1 and 2.

	Antibody type	Antibody/manufacturer	Dilution	Diluent	Incubation time
Primary antibodies	Primary antibody	α-Decorin antibody-mouse monoclonal IgG/ Abcam (ab54728)	1:1,200	5% non-fat milk powder in PBS/tween	Overnight at 4 ˚C
	Primary antibody	α-Alkaline Phosphatase antibody-rabbit monoclonal IgG/ Abcam (ab186422)	1:1,000	5% non-fat milk powder in PBS/tween	Overnight at 4 ˚C
	Primary antibody	α-Alkaline Phosphatase antibody-rabbit polyclonal IgG/ Abcam (ab95462)	1:500	5% non-fat milk powder in PBS/tween	Overnight at 4 ˚C
	Primary antibody	Fibrillin-2 antibody-goat polyclonal IgG/ Santacruz (N-20) sc-28136	1:200	5% non-fat milk powder in PBS/tween	Overnight at 4 ˚C

	Antibody type	**Antibody/ manufacturer**	**Dilution**	**Diluent**	**Incubation time**
Primary antibodies	Primary antibody	Fibrillin-2 antibody-mouse monoclonal IgG/ Santacruz (H-10) sc-393968	1:200	5% non-fat milk powder in PBS/tween	Overnight at 4 °C
	Primary antibody	α-Carboxypeptidase M antibody-rabbit monoclonal/ Abcam ab150405	1:1,500	5% non-fat milk powder in PBS/tween	Overnight at 4 °C
Secondary antibodies	Secondary antibody	Donkey α-mouse IRDye 680RD IgG/ LI-COR, Biosciences 925-68072	1:25,000	5% non-fat milk powder in PBS/tween	1h at RT
	Secondary antibody	Donkey α-goat IRDye 800CW IgG/ LI-COR, Biosciences 925-32214	1:25,000	5% non-fat milk powder in PBS/tween	1h at RT
	Secondary antibody	Goat α-rabbit IRDye 800CW IgG/ LI-COR, Biosciences 925-32211	1:25,000	5% non-fat milk powder in PBS/tween	1h at RT
Housekeeping	Primary antibody	α-β-actin−Peroxidase antibody-mouse monoclonal IgG/Sigma A3854	1:10,000	5% non-fat milk powder in PBS/tween	1h at RT
	Secondary antibody	Rabbit α-HRP IRDye 800CW IgG/ LI-COR, Biosciences 926-32236	1:5,000	5% non-fat milk powder in PBS/tween	1h at RT

5.4 Results

5.4.1 Offline alkaline reverse phase chromatography

As for MS1 the proteins surrogate tryptic peptides were fractionated by RP-C_8 chromatography under alkaline mobile phase conditions (Figure 60). Peptides were separated according to their hydrophilic or hydrophobic retention time behaviour at 218 nm. From left to right, the scale indicates the hydrophobic and hydrophilic index of the tryptic peptides found in the original pooled sample. In this study a total of 63 fractions versus 28 of the MS2 study were collected.

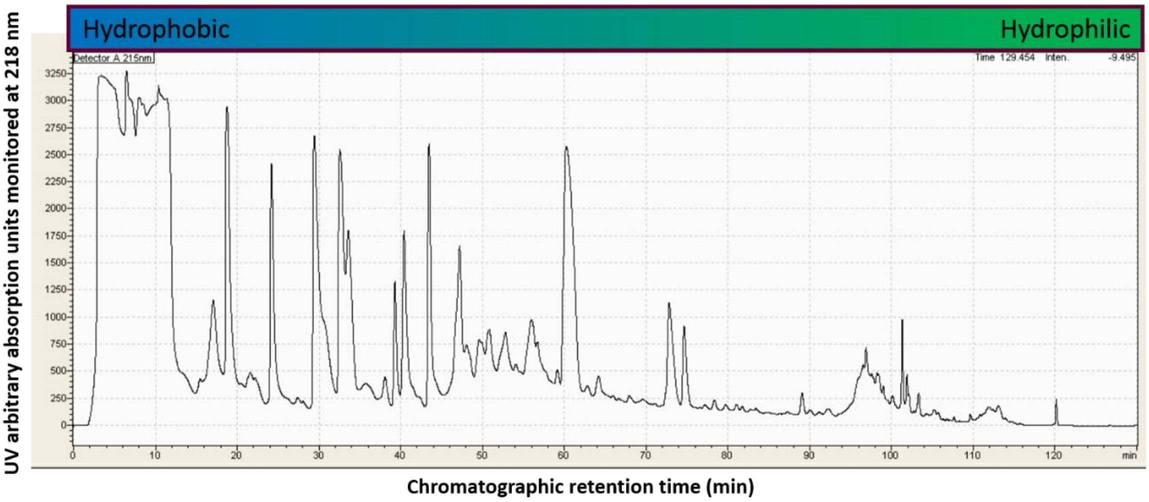


Figure 60: High pH reverse phase chromatography (C_8 column), MS2.

The protein surrogate tryptic peptides were separated in accordance to their differences in hydrophobic properties. A total of 63 peptide fractions were collected for online LC-MS analysis. The y-axis corresponds to UV arbitrary absorption units monitored at 218 nm. The x-axis corresponds the chromatographic retention time in min.

5.4.2 Online LC-MS Analysis

A total of 7594 proteins were discovered at 95% confidence and a false discovery rate (FDR) <0.05% and identified and quantified by nanoRP-ESI-LC-MS/MS-HCD-CID. 6050 of these proteins were common between males and females samples, whilst 814 proteins were only presents in the male samples and 730 only in the female sample (Figure 61).

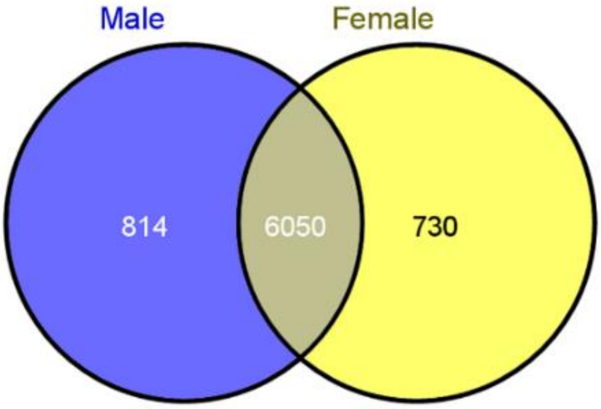


Figure 61: Common proteins between genders.

The Venn diagram represent the overlap of the common proteins between genders. These proteins were identified by nanoRP-ESI-LC-MS/MS-HCD-CID and quantification of thereof was based on the iTRAQ labelled peptides.

Between the multiple spectra found to match the same peptide, the top-scoring spectrum was chosen by default. The protein carboxypeptidase M (CPM) is reported as a representative example of the results (Figure 62). CPM is formed by 443 amino acids which form 39 theoretical peptides if cleaved by trypsin (Figure 62, A). In this MS study 12 peptides (of the theoretical 39) were identified and the ratios of each peptides is represented in the scatter dot plot (Figure 62, B). The scatter dot plot demonstrated that the osteogenic samples presented a higher intensity compared to the samples in basal conditions The peptide SLTPDDDVFQYLAHTYASR of CPM was chosen as a representative example to demonstrate the *b* and *y* ions spectra matches, the iTRAQ region and the overall MS/MS chromatogram (Figure 62, C, D, E).

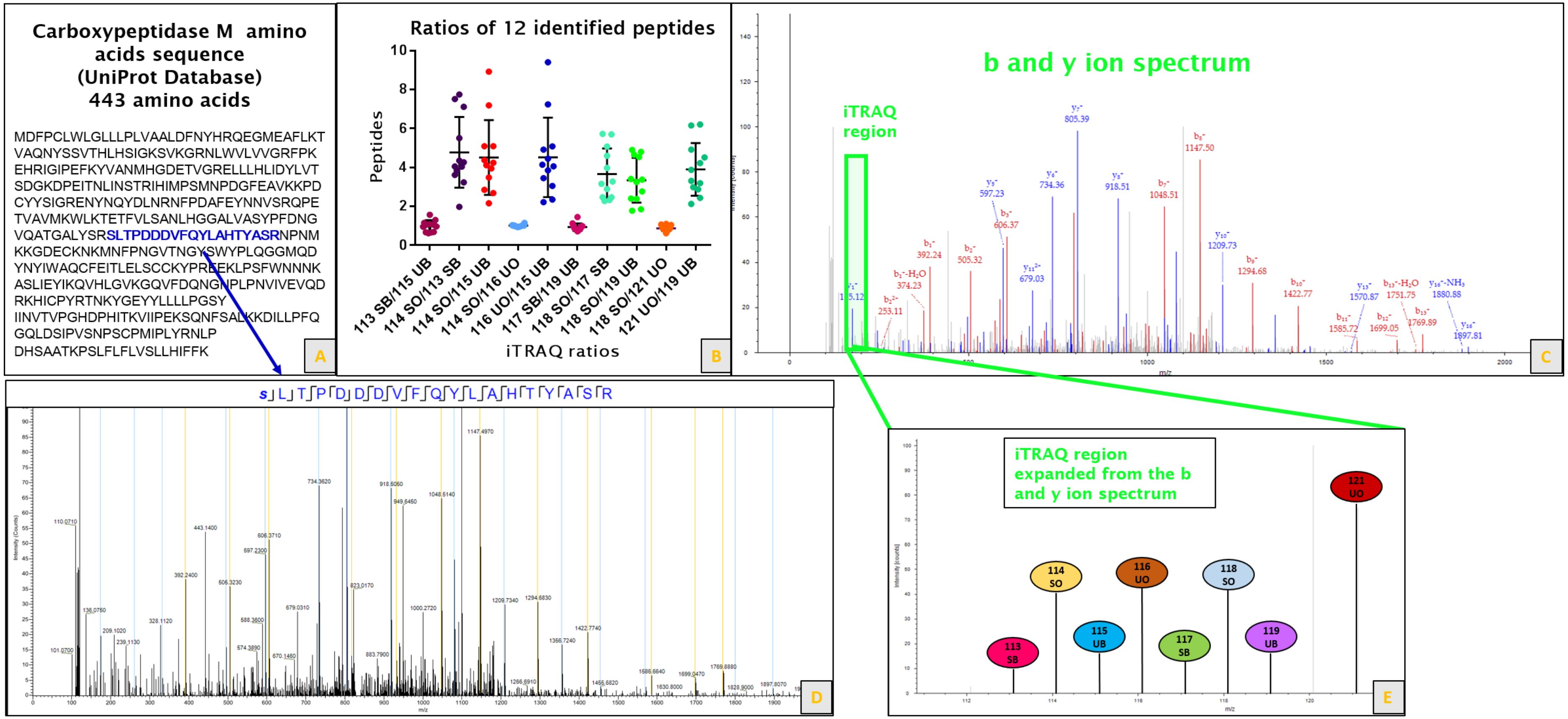


Figure 62: LC/MS representative results of carboxypeptidase M peptide SLTPDDDVFQYLAHTYASR.

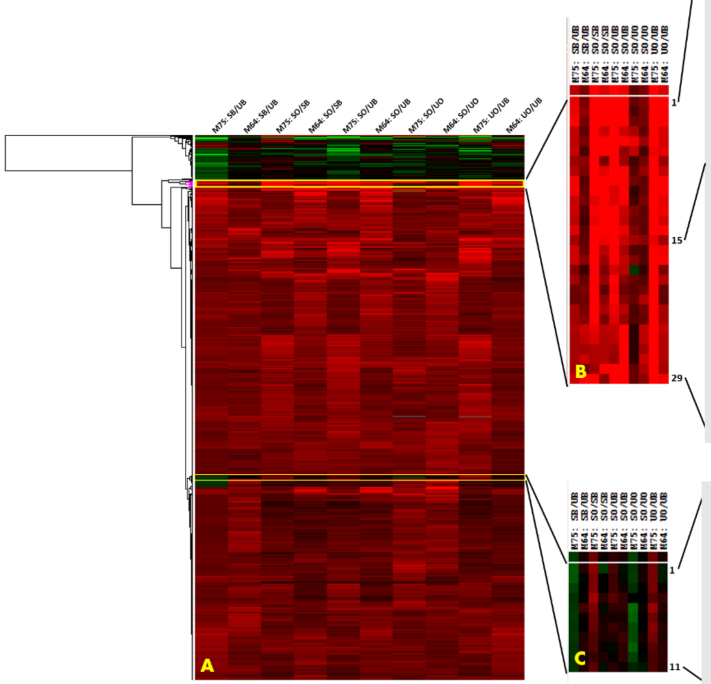
A: Carboxypeptidase M (CPM) amino acids sequence (UniProt database). B: Scatter dot plot of the iTRAQ ratios of 12 peptides within CPM. C: b and y ions spectrum match of the peptide SLTPDDDVFQYLAHTYASR. D: MS/MS spectrum of the peptide SLTPDDDVFQYLAHTYASR. E: Identification of the iTRAQ region of the peptide SLTPDDDVFQYLAHTYASR. SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.

5.4.3 Hierarchical clustering

A comprehensive graphical translation of the proteins expression was created utilising a hierarchical clustering software, Cluster 3.0; and a heat map visualisation software, TreeView. This analysis highlighted interesting patterns in the protein expression. The analysis run on each gender separately, highlighted gender specific differences. Proteins with expression greater than 2SD were extracted from the data set of males and females. For the male dataset a protein list was obtained and analysed by hierarchical clustering analysis (Figure 64). For the female data set, no results were obtained by filtering the protein expression by 2SD.

Overall, these analyses highlighted the different profiles of protein expression between males and females. In the male samples the majority of the proteins were upregulated and no differences between experimental conditions were captured; although, patient specific differences were highlighted for a small set of protein expression (Figure 63, C).

The female hierarchical clustering displayed greater heterogeneity of protein expression. The heterogeneity displayed by the female samples may have been caused by nuclear receptor binding proteins, although this speculation was not verified (Mangelsdorf, Thummel *et al.*, 1995). The sections B and C in Figure 65 highlight differences in protein expression due to osteogenic modulation (green and red patterns).

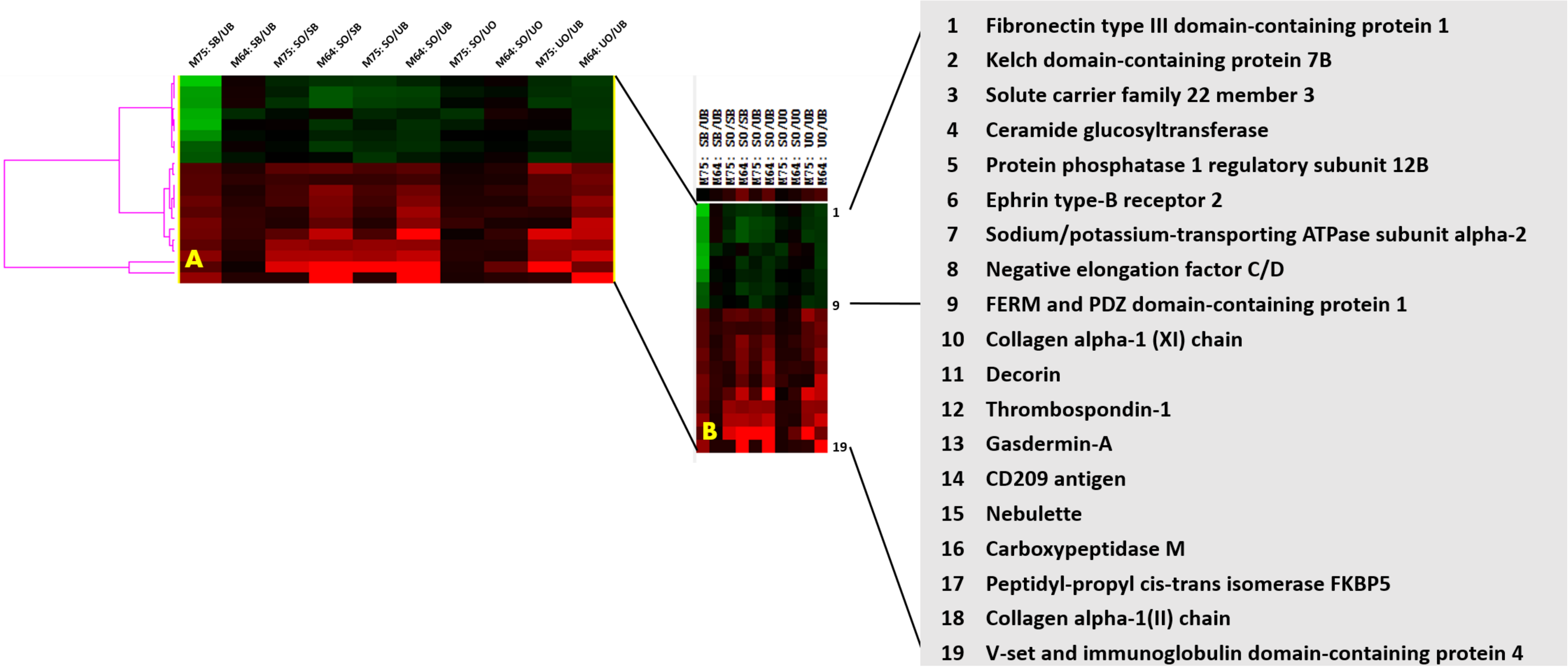


1 Fibrillin-2
2 Amine oxidase [flavin-containing] A
3 Carboxypeptidase M
4 Sushi repeat-containing protein SRPX2
5 Serglycin
6 Aldehyde oxidase
7 N-acylglucosanine 2-epimerase
8 Podocan
9 FERM and PDZ domain-containing protein 1
10 Adipocyte enhancer-binding protein 1
11 Olfactomedin-like protein 3
12 Methyltransferase-like protein 7A
13 Rho GTPase-activating protein 6
14 Tumor necrosis factor alpha-induced protein 8-like protein 1
15 Collagen alpha-1 (VII) chain
16 Gamma-glutamyltransferase 5
17 Insulin receptor substrate 2
18 Sorbin and SH3 domain-containing protein 1
19 Sushi repeat-containing protein SRPX
20 UPF0369 protein C6orrf57
21 Laminin subunit alpha-4
22 Catalase
23 Dimethylaniline monooxygenase [N-oxide-forming] 3
24 Pentraxin-related protein PTX3
25 Inter-alpha-trypsin inhibitor heavy chain H2
26 Versican core protein
27 Hyaluronan and proteoglycan link protein 3
28 A disintegrin and metalloproteinase with thrombospondin motifs
29 Alkaline phosphatase, tissue-nonspecific isozyme

1 Cysteine-rich secretory protein LCCL domain-containing 2
2 Retinal dehydrogenase 2
3 Aldehyde dehydrogenase X, mitochondrial
4 Tetraspanin-18
5 Redox-regulatory protein FAM213A
6 Enoyl-CoA hydratase domain-containing protein 3, mitochondrial
7 Nesprin-2
8 Isocitrate dehydrogenase [NADP], mitochondrial
9 BDNF/NT-3 growth factors receptor
10 Protein-methionine sulfoxide oxidase MICAL1
11 Phosphatidate phosphatase LPIN3

Figure 63: Hierarchical clustering of proteins found in M75 and M64. 2877 proteins.

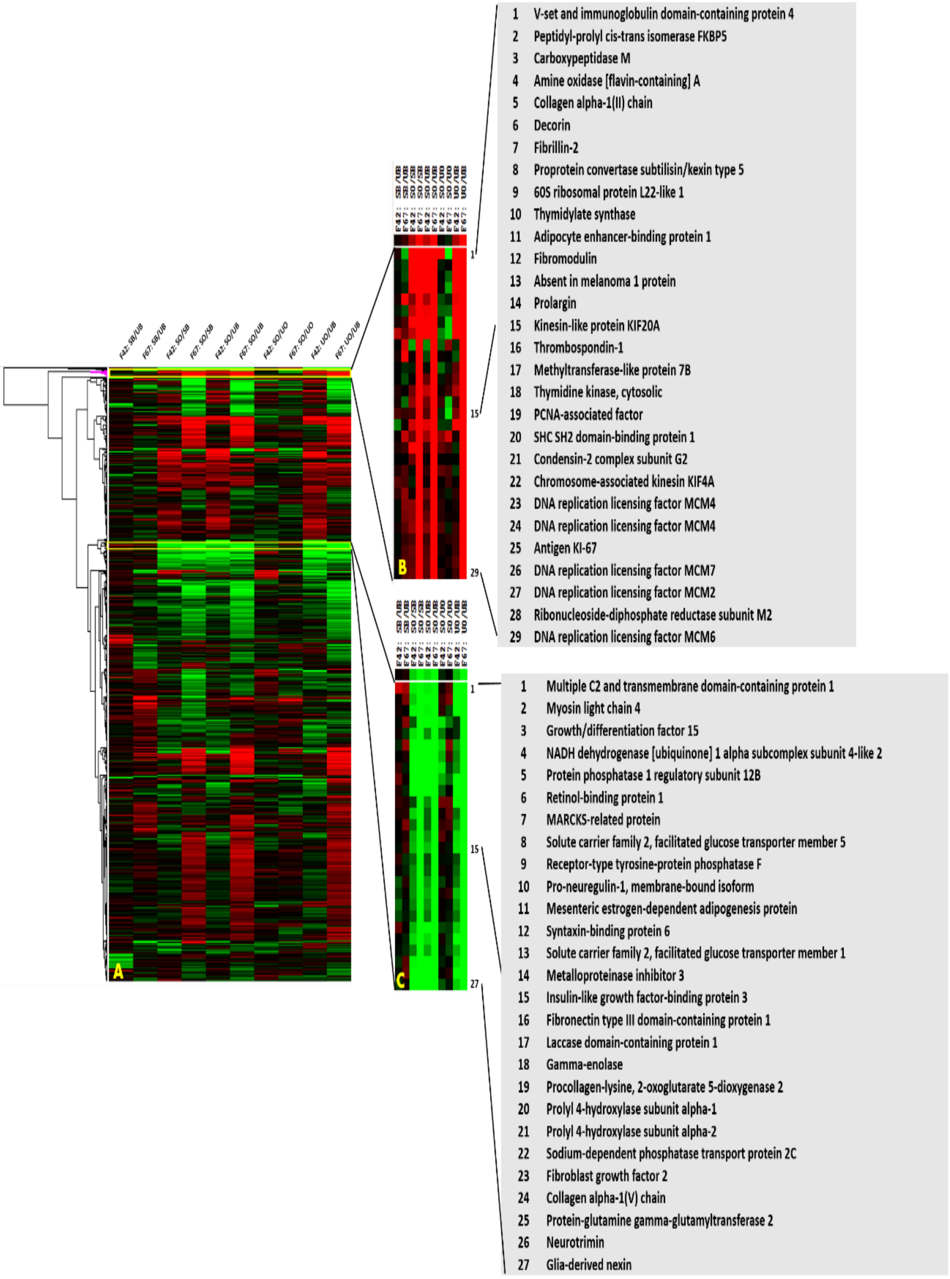
Mass spectrometry results were treated as input for Cluster 3.0 employing the Euclidean correlation as the distance metric for the arrays with complete linkage analysis, to cluster the proteins. TreeView 1.1.6 was used for the visualization of the cluster. The heatmap (left) depicts a heat map with gene linkage of the proteins found in M75 and M64. B and C (right) are a magnified section of the yellow square in A (left). B represents the proteins were upregulation was observed and C represents a section of the analysis where patient specific protein regulation is highlighted. Red indicates upregulation and green downregulation. SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.



1 Fibronectin type III domain-containing protein 1
2 Kelch domain-containing protein 7B
3 Solute carrier family 22 member 3
4 Ceramide glucosyltransferase
5 Protein phosphatase 1 regulatory subunit 12B
6 Ephrin type-B receptor 2
7 Sodium/potassium-transporting ATPase subunit alpha-2
8 Negative elongation factor C/D
9 FERM and PDZ domain-containing protein 1
10 Collagen alpha-1 (XI) chain
11 Decorin
12 Thrombospondin-1
13 Gasdermin-A
14 CD209 antigen
15 Nebulette
16 Carboxypeptidase M
17 Peptidyl-propyl cis-trans isomerase FKBP5
18 Collagen alpha-1(II) chain
19 V-set and immunoglobulin domain-containing protein 4

Figure 64: Most differentially expressed proteins in M75 and M64.

Within the male data set, the proteins which expression was greater than 2SD were extracted and analysed by hierarchical clustering analysis, A. B represents a magnified version of A including the protein description. Red indicates upregulation and green downregulation. SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.



1 V-set and immunoglobulin domain-containing protein 4
2 Peptidyl-prolyl cis-trans isomerase FKBP5
3 Carboxypeptidase M
4 Amine oxidase [flavin-containing] A
5 Collagen alpha-1(II) chain
6 Decorin
7 Fibrillin-2
8 Proprotein convertase subtilisin/kexin type 5
9 60S ribosomal protein L22-like 1
10 Thymidylate synthase
11 Adipocyte enhancer-binding protein 1
12 Fibromodulin
13 Absent in melanoma 1 protein
14 Prolargin
15 Kinesin-like protein KIF20A
16 Thrombospondin-1
17 Methyltransferase-like protein 7B
18 Thymidine kinase, cytosolic
19 PCNA-associated factor
20 SHC SH2 domain-binding protein 1
21 Condensin-2 complex subunit G2
22 Chromosome-associated kinesin KIF4A
23 DNA replication licensing factor MCM4
24 DNA replication licensing factor MCM4
25 Antigen KI-67
26 DNA replication licensing factor MCM7
27 DNA replication licensing factor MCM2
28 Ribonucleoside-diphosphate reductase subunit M2
29 DNA replication licensing factor MCM6

1 Multiple C2 and transmembrane domain-containing protein 1
2 Myosin light chain 4
3 Growth/differentiation factor 15
4 NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4-like 2
5 Protein phosphatase 1 regulatory subunit 12B
6 Retinol-binding protein 1
7 MARCKS-related protein
8 Solute carrier family 2, facilitated glucose transporter member 5
9 Receptor-type tyrosine-protein phosphatase F
10 Pro-neuregulin-1, membrane-bound isoform
11 Mesenteric estrogen-dependent adipogenesis protein
12 Syntaxin-binding protein 6
13 Solute carrier family 2, facilitated glucose transporter member 1
14 Metalloproteinase inhibitor 3
15 Insulin-like growth factor-binding protein 3
16 Fibronectin type III domain-containing protein 1
17 Laccase domain-containing protein 1
18 Gamma-enolase
19 Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2
20 Prolyl 4-hydroxylase subunit alpha-1
21 Prolyl 4-hydroxylase subunit alpha-2
22 Sodium-dependent phosphatase transport protein 2C
23 Fibroblast growth factor 2
24 Collagen alpha-1(V) chain
25 Protein-glutamine gamma-glutamyltransferase 2
26 Neurotrimin
27 Glia-derived nexin

Figure 65: Hierarchical clustering of proteins found in F42 and F67.

Mass spectrometry results were treated as input for Cluster 3.0 employing the Euclidean correlation as the distance metric for the arrays with complete linkage analysis, to cluster the proteins. TreeView 1.1.6 was used for the visualization of the cluster. The image (left) depicts a heat map with gene linkage of the proteins found in F42 and F67. The visual division of the proteins using the hierarchical clustering method reflects differences in protein expression in samples where O is compared to S. B and C are a magnified sections of A. Red indicates upregulation and green downregulation. SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.

5.4.4 Upregulated proteins

Proteins upregulated by 1.5 and 2 fold in both samples and across the genders were visualised in 5 Venn diagrams; one per ratio analysed (Figure 66). Results indicated that the ratios where osteogenic samples were compared to basal samples (*i.e.* O/B), the same 4 proteins repeatedly passed the filters of 1.5 and 2 folds for both males and females. These 4 proteins comprised of peptidyl-prolyl cis-trans isomerase FKBP5 (FKBP5), carboxypeptidase M (CPM), amine oxidase [flavin-containing] A (MAOA), and fibrillin-2 (FBN2). These 4 proteins in conjunction with DCN (DCN) and methyltransferase-like protein 7A (METTL7A) passed the filter, for a total of 6 proteins common to both samples and genders in the UO/UB ratio. When the samples were grown under same media conditions, no proteins upregulated by 1.5 or 2 folds were common across samples and genders.

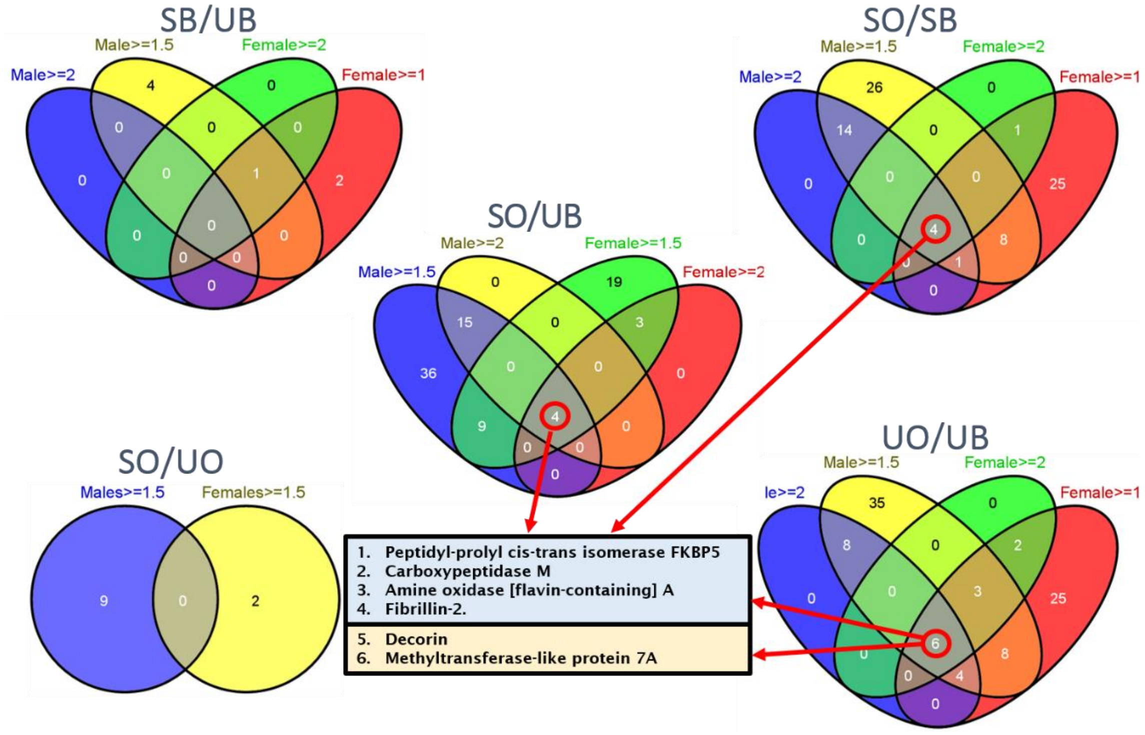

SB/UB
Male>=2
Male>=1.5
Female>=2
Female>=1
4
0
0
0
0
1
2
0
0
0
0
0
0
0
0
SO/SB
Male>=2
Male>=1.5
Female>=2
Female>=1
26
0
14
0
1
0
0
0
25
0
4
8
0
1
0
SO/UB
Male>=1.5
Male>=2
Female>=1.5
Female>=2
0
19
15
0
3
36
0
0
0
9
4
0
0
0
0
SO/UO
Males>=1.5
Females>=1.5
9
0
2
UO/UB
le>=2
Male>=1.5
Female>=2
Female>=1
35
0
8
0
2
0
0
3
25
6
0
8
0
4
0
1. Peptidyl-prolyl cis-trans isomerase FKBP5
2. Carboxypeptidase M
3. Amine oxidase [flavin-containing] A
4. Fibrillin-2.
5. Decorin
6. Methyltransferase-like protein 7A

Figure 66: Upregulated proteins by 2 and 1.5 fold.

The modulated proteins obtained by the LC-MS/MS analysis were filtered applying a fold cut off of 1.5 and 2. The proteins filtered by these criteria were grouped in a Venn diagram which highlighted the common proteins. The same 4 proteins (top box) were found to be more than 2 folds upregulated in the samples modulated by osteogenic differentiation. In UO/UB, 2 extra proteins (bottom box) were found to be upregulated by 2 fold in conjunction with the 4 proteins (top box). SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.

5.4.5 Downregulated proteins

The expression of downregulated proteins was tested employing the same criteria used for upregulation (5.4.4), *i.e.* the same filter for fold cut-off of 1.5 and 2 was applied. Downregulation analysis by 2 fold returned no proteins in any of the samples in either basal or osteogenic conditions; whilst the 1.5 fold criteria returned overall 3 proteins, 1 in group SO/UB (receptor-type tyrosine-protein phosphatase) 1 in group SO/SB (sephrase) and 2 in group UO/UB (plakoglobin and solute carrier family 22 member 3) (Figure 67). Analogously to the upregulation analysis, modulated proteins were a result of osteoinduction.

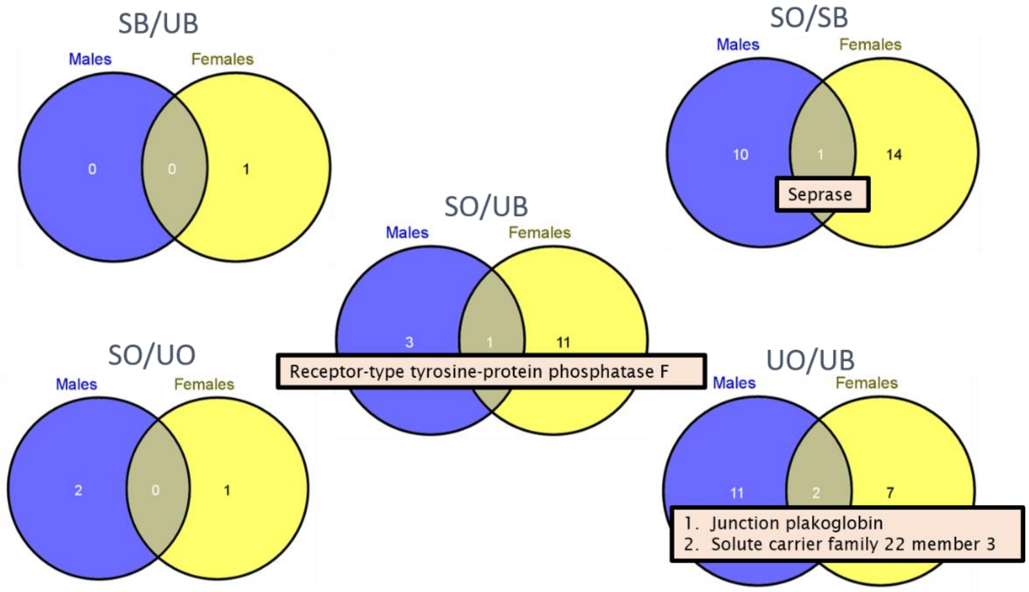

SB/UB
Males
Females
0
0
1
SO/SB
Males
Females
10
1
14
Seprase
SO/UB
Males
Females
3
1
11
Receptor-type tyrosine-protein phosphatase F
SO/UO
Males
Females
2
0
1
UO/UB
Males
Females
11
2
7
1. Junction plakoglobin
2. Solute carrier family 22 member 3

Figure 67: Downregulated proteins by 1.5 fold.

The modulated proteins obtained by the LC-MS/MS analysis were filtered applying a fold cut off of 1.5 and 2 for downregulation. The proteins that passed the filter were then analysed in a Venn diagram which showed the common items. None of the proteins were downregulated by 2 folds. Proteins that passed the 1.5 cut off filter were predominantly present in the samples undergoing osteogenic differentiation. SB: sorted basal; SO: sorted osteogenic; UB: unsorted basal; UO: unsorted osteogenic.

5.4.6 FunRich

Males and females protein data were subjected to classification and pathway analysis to demonstrate their biological context using FunRich free software. The classification of the proteins comprised biological pathways, cellular components, biological process and molecular function.

The analysis of the males and females proteins was displayed in a doughnut style visualisation (Figure 68, Figure 69, Figure 70 and Figure 71). This analysis indicated that the most upregulated biological process was the citric acid (TCA) cycle for both males and females; the majority of the proteins were derived from the cytoplasm of the cells and were mainly involved in the metabolism of the cell. Catalytic activity, in fact, corresponded to the main molecular function in both males and females. According to this classification both males and females displayed similar characteristics. FunRich was also employed to analyse the protein pathways. Pathway analysis allowed for the creation of a protein interaction map (Figure 72). This map showed the intrinsic set of networks between the proteins found in both genders cohorts. The main node for the male sample was centred by the epidermal growth factor (EGF) whilst TNF receptor-associated factor 6 (TRAF6) corresponded to the main node in the female sample (Figure 72).

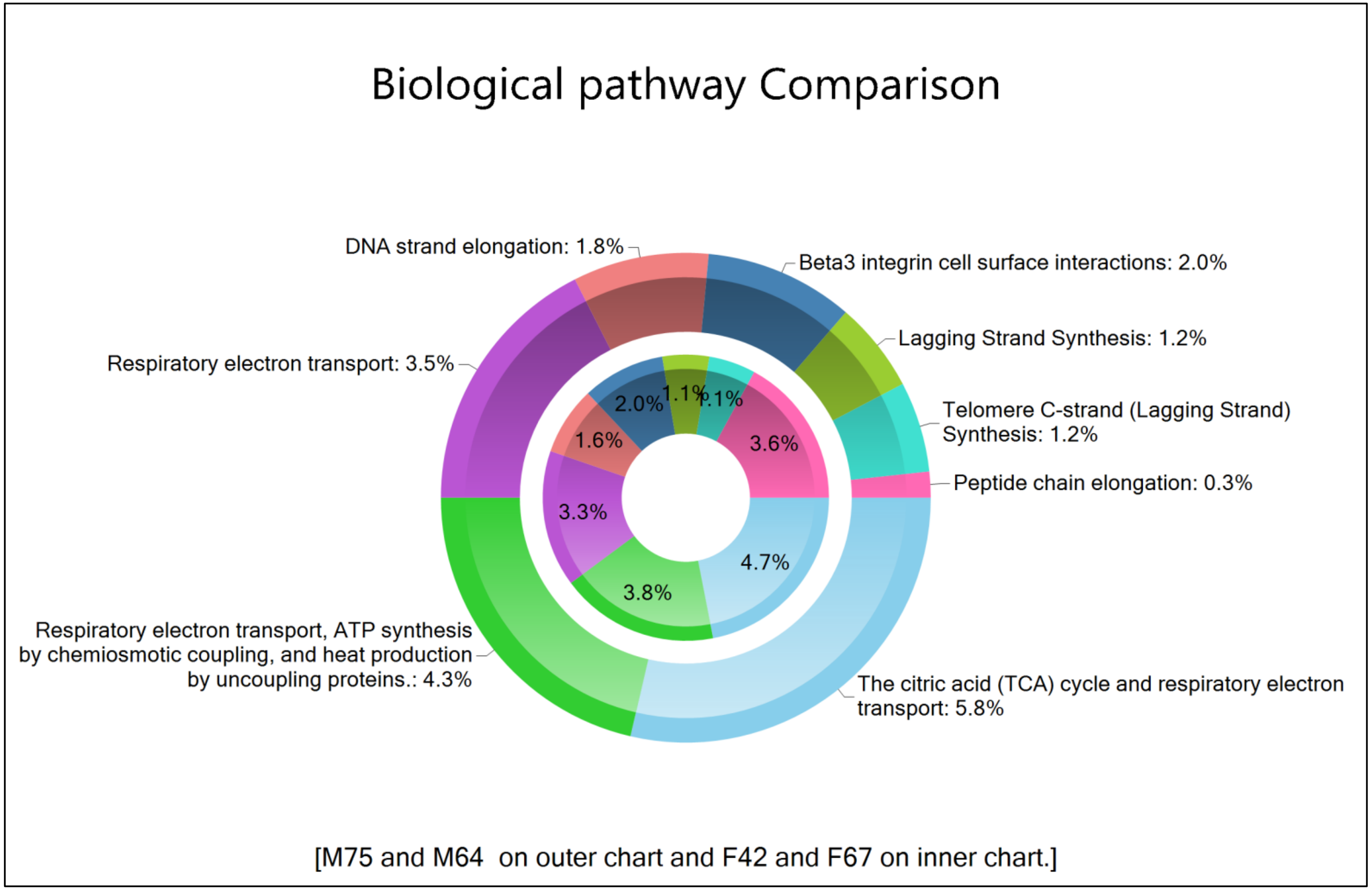


Figure 68: Biological pathway of M75, M64, F42 and F67.

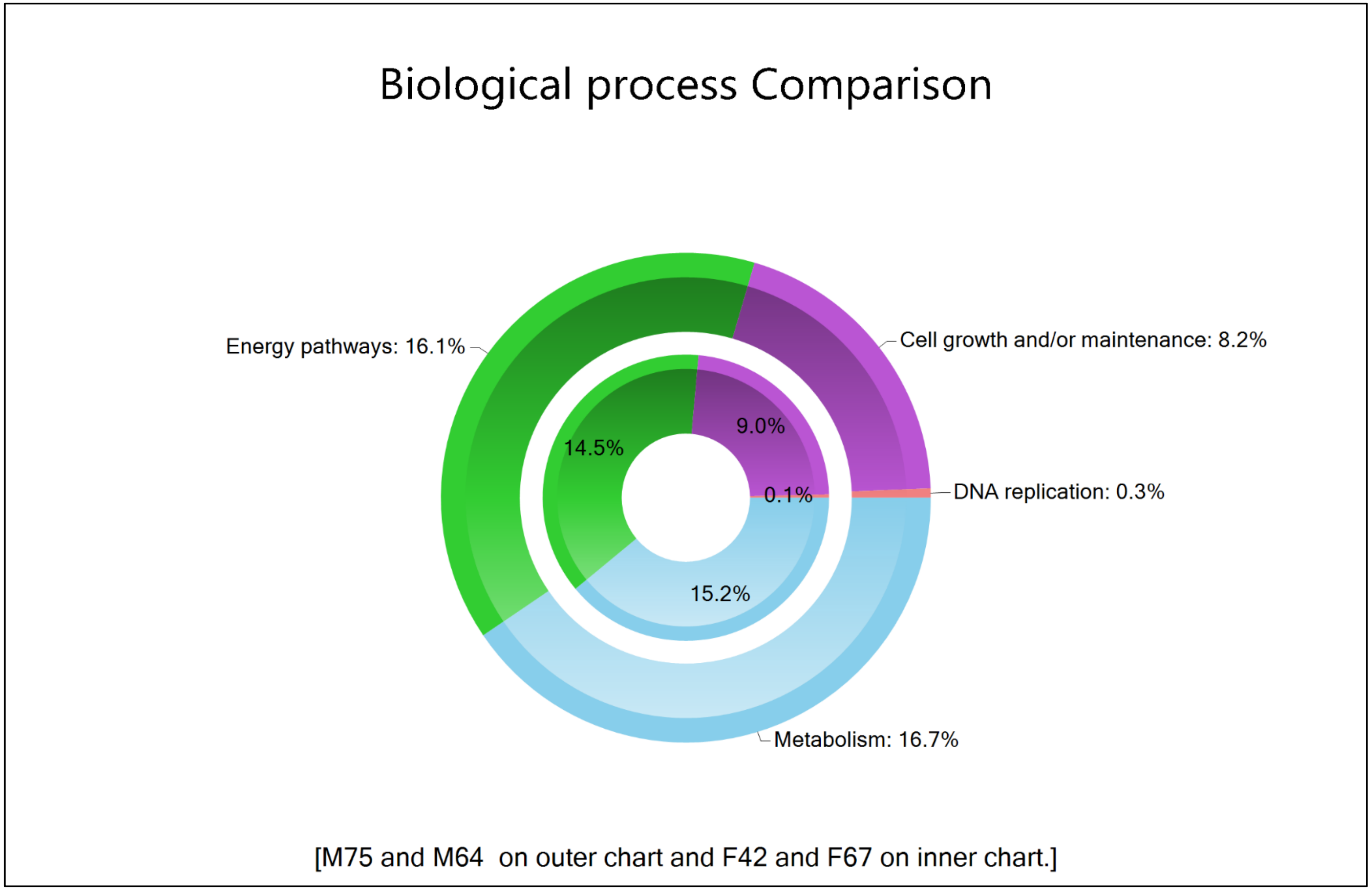


Figure 69: Biological process of M75, M64, F42 and F67.

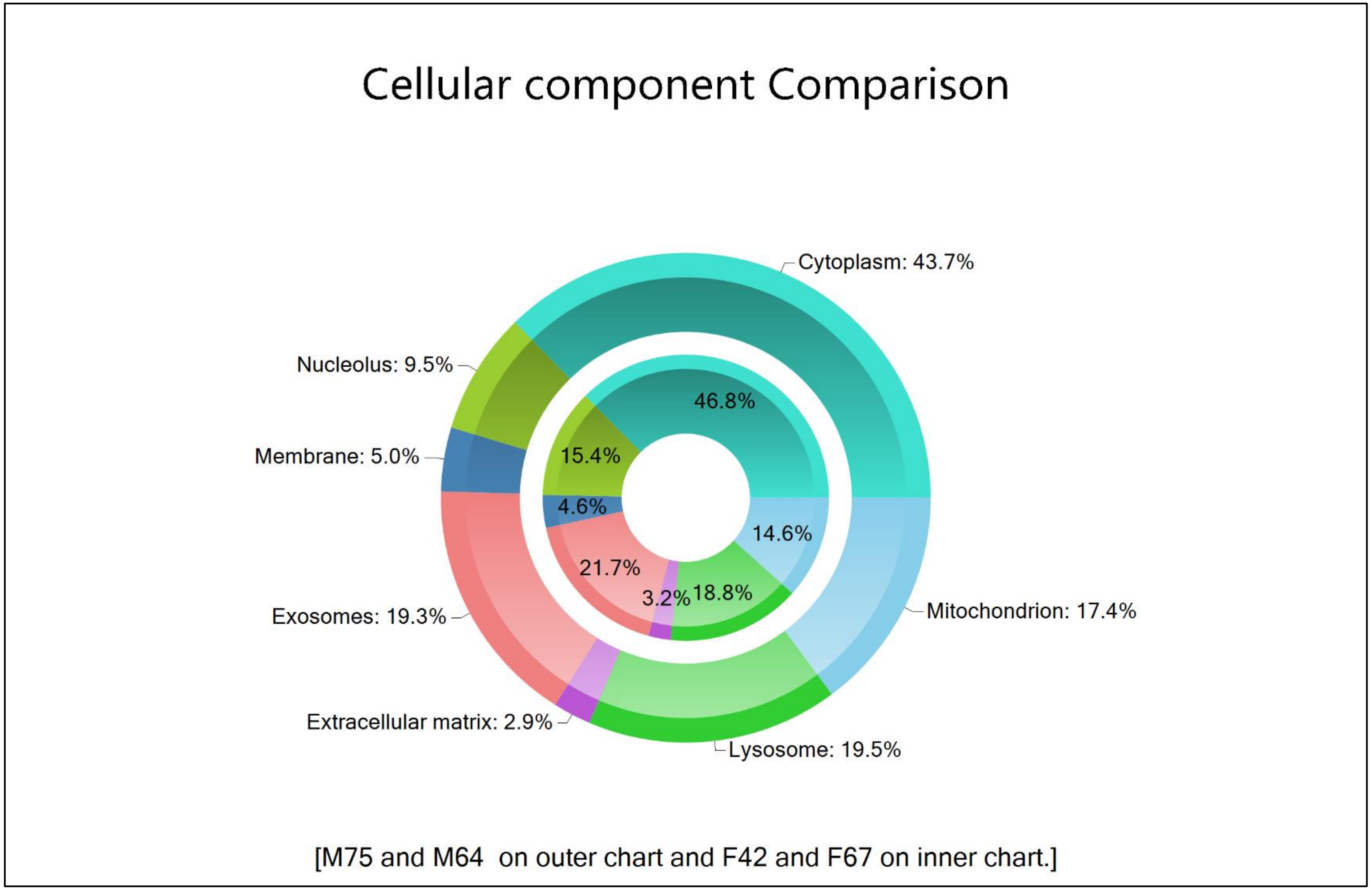


Figure 70: Cellular component of M75, M64, F42 and F67.

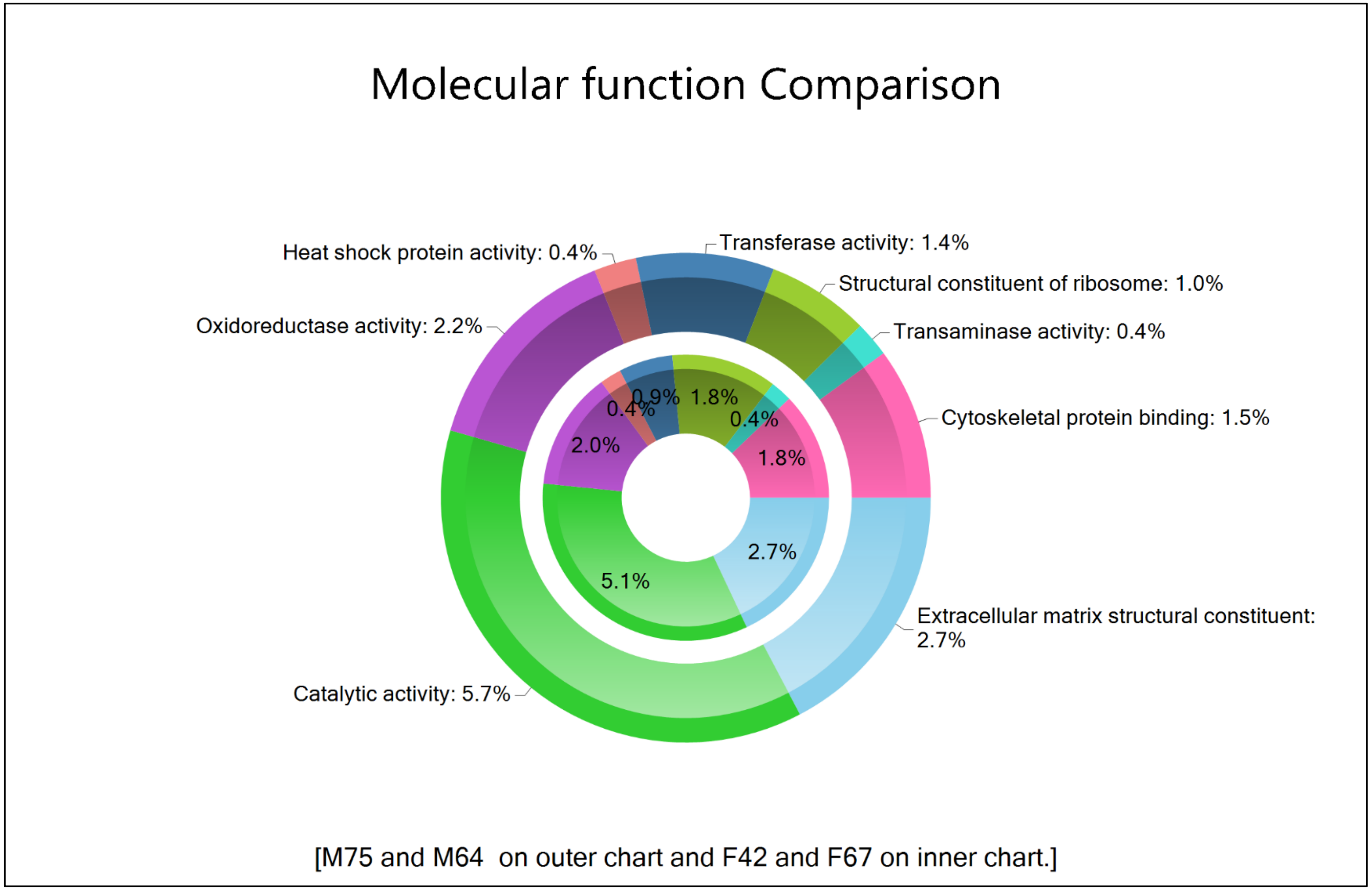
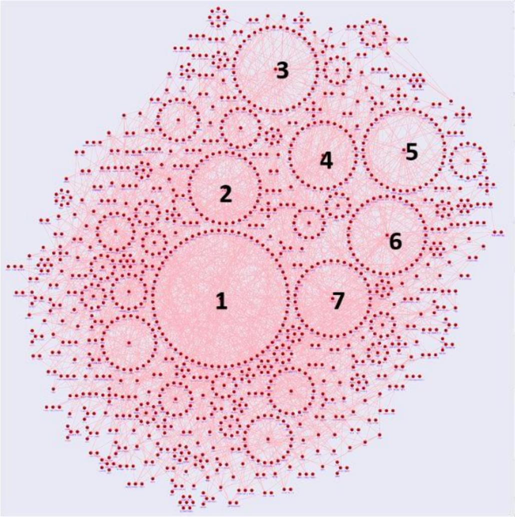


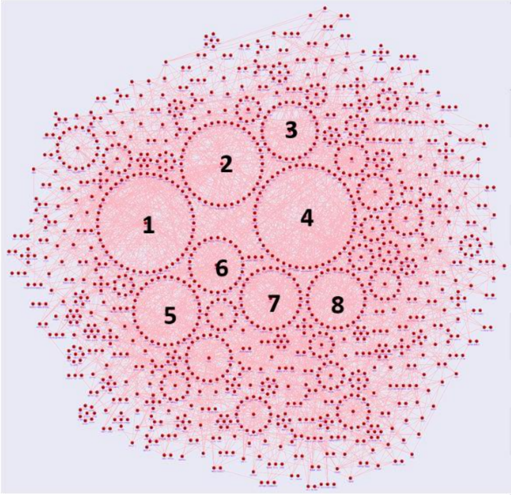
Figure 71: Molecular function of M75, M64, F42 and F67.

M75 and M64 interaction map



Node	Gene name	Protein name
1	EGFR	Epidermal growth factor receptor
2	PIK3R1	Phosphatidylinositol 3-kinase regulatory subunit alpha
3	ALB	Serum albumin
4	UBE2I	SUMO-conjugating enzyme UBC9
5	PRKAB1	5'-AMP-activated protein kinase subunit beta-1
6	HLA-B	HLA class I histocompatibility antigen, B-7 alpha chain
7	SMAD3	Mothers against decapentaplegic homolog 3

F42 and F67 interaction map



Node	Gene name	Protein name
1	TRAF6	TNF receptor-associated factor 6
2	HLA-B	HLA class I histocompatibility antigen, B-48 alpha chain
3	VIM	Vimentin
4	PIK3R1	Phosphatidylinositol 3-kinase regulatory subunit alpha
5	NR3C1	Glucocorticoid receptor
6	CDK1	Cyclin-dependent kinase 17
7	ABL1	Tyrosine-protein kinase ABL1
8	FYN	Tyrosine-protein kinase Fyn

Figure 72: Interaction of the protein processes found in M75, M64, F42 and F67.

These interaction maps show the dense web of processes found between the proteins identified in this study. The main nodes are described in the table adjacent to the map.

5.4.7 MetaCore analysis

As in previous interaction pathway analysis using FunRich (1.1.1), the expression of the common proteins between the genders was queried against the MetaCore database (in previous analysis FunRich was used). MetaCore is a further tool to decipher the connection between proteins. MetaCore analysis is based on protein expression while FunRich does not have the capacity to consider this. Networks of protein interactions were constructed using a direct interaction algorithm. The results indicated the Glucocorticoid Receptor (GCR) - alpha pathway was downregulated following direct inhibition under osteogenic differentiation (Figure 73).

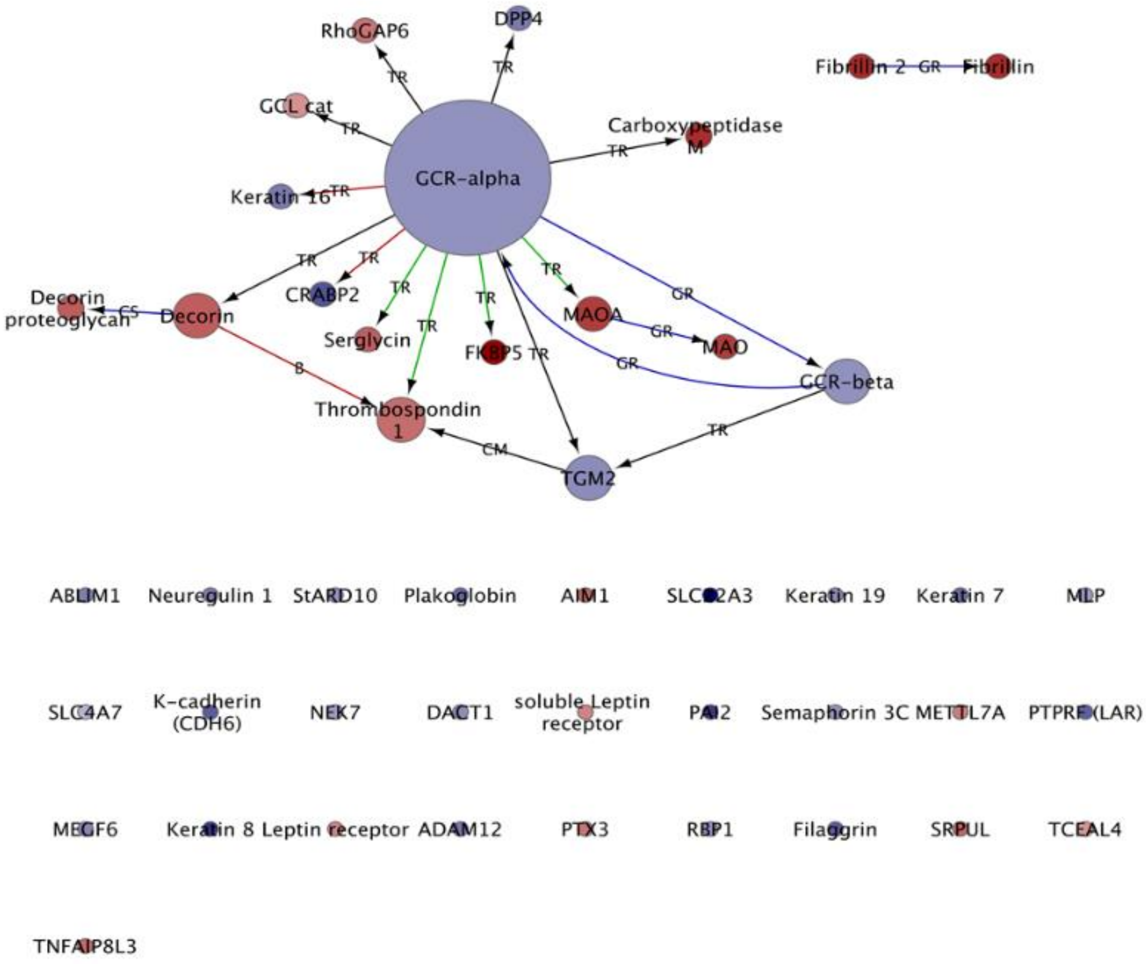

RhoGAP6
DPP4
Fibrillin 2
GR
Fibrillin
GCL cat
Carboxypeptidase M
GCR-alpha
Keratin 16
TR
CS
B
CM
GR
Decorin proteoglycan
Decorin
CRABP2
Serglycin
FKBP5
MAOA
MAO
GCR-beta
Thrombospondin 1
TGM2
ABLIM1
Neuregulin 1
StARD10
Plakoglobin
AIM1
SLC2A3
Keratin 19
Keratin 7
MLP
SLC4A7
K-cadherin (CDH6)
NEK7
DACT1
soluble Leptin receptor
PAI2
Semaphorin 3C
METTL7A
PTPRF (LAR)
MEGF6
Keratin 8
Leptin receptor
ADAM12
PTX3
RBP1
Filaggrin
SRPUL
TCEAL4
TNFAIP8L3

Figure 73: MetaCore Pathway analysis.

Protein interactions were analysed with the MetaCore pathway analysis. The analysis reported a set of upregulated (red) and downregulated proteins (blue) which are connected by the main regulator GCR-alpha.

5.4.8 Western blotting-based validation of selected differentially expressed proteins

Differential expression of the few selected proteins identified in quantitative proteomics analysis was validated using the WB method. The selection of the proteins for validation was performed on the basis of fold changes (2 fold upregulated proteins: CPM, FBN2 and DCN) and on the basis of known association with the skeletal phenotype (ALP). WB analysis of four differentially expressed proteins was carried out using the same lysates that were used for mass spectrometry. Western blot results indicated the upregulation of CPM in the osteogenic sample and a downregulation in the basal samples (Figure 74). The presence of FBN2, DCN and ALP was not validated due to an issue with the antibodies used (both monoclonal and polyclonal). The WB did not show any results, neither on the positive control which indicated an issue with the antibodies used rather than with the samples considering the 2 positive controls (carboxypeptidase and β-actin) worked correctly (Figure 75 and Figure 76). To circumvent this issue different antibodies (polyvalent) were acquired for the same target. However, the results did not show any presence of the examined proteins.

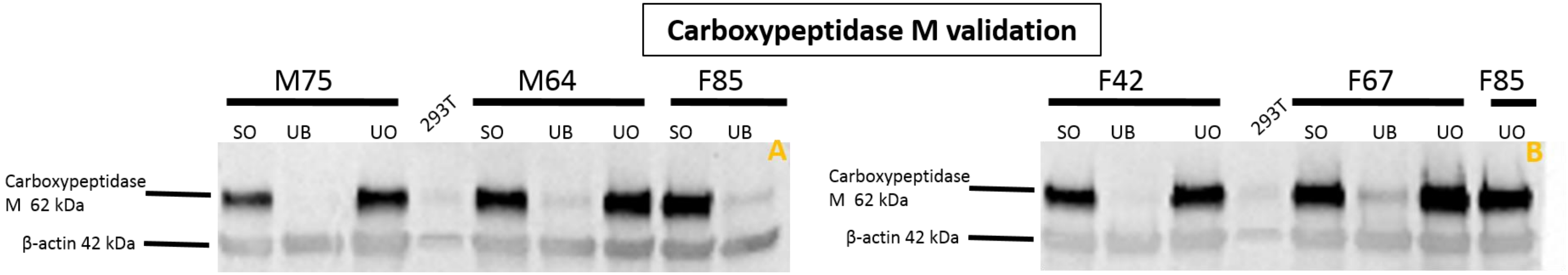

Carboxypeptidase M validation
M75
293T
M64
F85
F42
293T
F67
F85
SO
UB
UO
A
B
Carboxypeptidase M 62 kDa
β-actin 42 kDa

Figure 74: Carboxypeptidase M validation.

Western blotting validation of carboxypeptidase M (CPM) in samples employed for the mass spectrometry analysis. CPM is upregulated in SO and UO of all the samples. It is downregulated in the UB samples. SB: sorted basal, SO: sorted osteogenic, UB: unsorted basal.

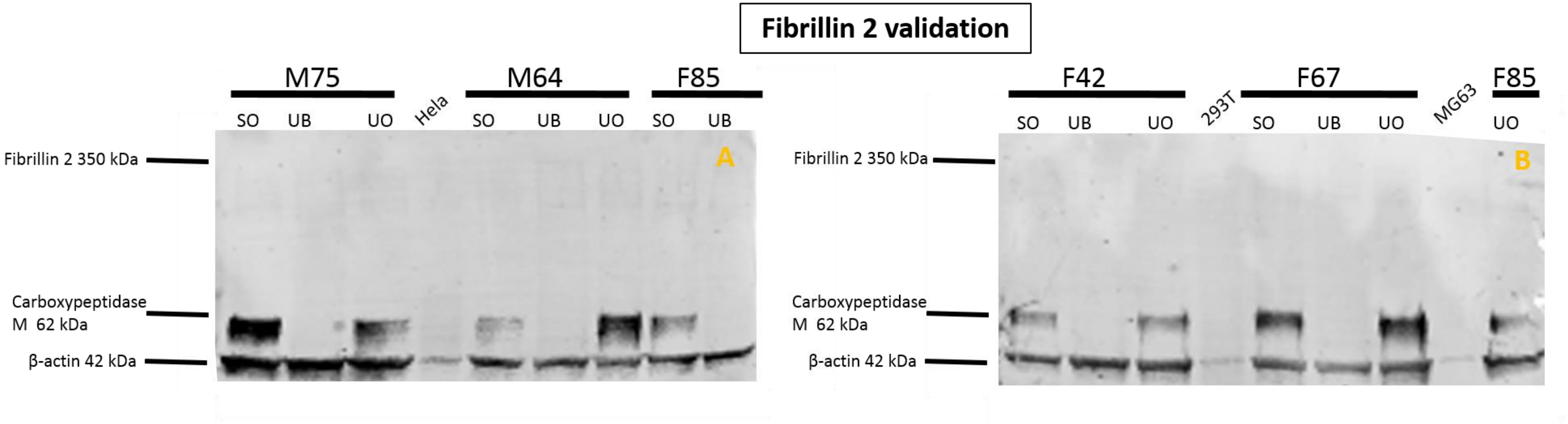

Fibrillin 2 validation
M75
Hela
M64
F85
F42
293T
F67
MG63
F85
SO
UB
UO
A
B
Fibrillin 2 350 kDa
Carboxypeptidase M 62 kDa
β-actin 42 kDa

Figure 75: Fibrillin 2 validation.

Western blotting validation of fibrillin 2 (FBN2) in samples employed for the mass spectrometry analysis. The FBN2 antibody used did not detect any epitope in this gel. CPM and b-actin were used as housekeeping positive controls. SB: sorted basal, SO: sorted osteogenic, UB: unsorted basal.

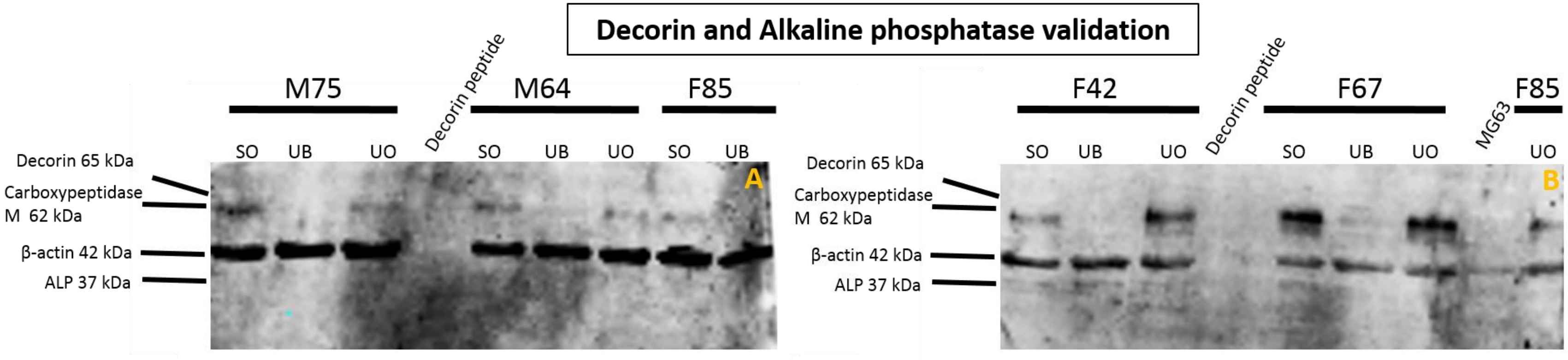

Decorin and Alkaline phosphatase validation
M75
Decorin peptide
M64
F85
SO
UB
UO
SO
UB
UO
SO
UB
A
Decorin 65 kDa
Carboxypeptidase M 62 kDa
β-actin 42 kDa
ALP 37 kDa
F42
Decorin peptide
F67
MG63
F85
SO
UB
UO
SO
UB
UO
UO
B
Decorin 65 kDa
Carboxypeptidase M 62 kDa
β-actin 42 kDa
ALP 37 kDa

Figure 76: Decorin and Alkaline phosphatase validation.

Western blotting validation of decorin (DCN) and alkaline phosphatase (ALP) in samples employed for the mass spectrometry analysis. The antibodies used for DCN and ALP did not react with the correspondent epitopes on these blots. CPM and β-actin were used as housekeeping positive controls. SB: sorted basal, SO: sorted osteogenic, UB: unsorted basal.

5.4.9 *In silico* analysis of 50 kDa proteins for correlation and association to Stro-1 based on immunostaining

Results obtained from WB analysis in chapter 3 indicated Stro-1 to be a 50 kDa antigen. Given the acquired set of proteins from the MS experiments (this chapter), to investigate possible proteins candidates for the Stro-1 antigen, an *in silico* approach was conducted to examine all the discovered proteins of ~50 kDa, with a plasma membrane association and exhibiting similar expression to the Stro-1 antigen. Expression was visually assessed between immunostaining images available on-line of the discovered proteins and images collected for this work.

The following hypothesis was examined:

I. One of the proteins found in the MS study is membrane bound and has a size range between 48 and 52 kDa; it has the potential to be the Stro-1 antigen, if similar immunostaining can be determined.

5.4.10 Method

Proteins between 48 and 52 kDa common between males and females were extracted from the MS data set and entered in the PANTHER Classification System to facilitate the discovery of membrane proteins. PANTHER classified the proteins based on cellular components, results of the protein classification are shown in Figure 77. From this classification, PANTHER was asked to further classify the membrane proteins. Proteins that are integral to the membrane and specifically belonging to the plasma membrane were extracted, yielding the results represented in Figure 78. An on-line search was performed to collect immunostaining data (in the form of microscopy images) of the extracted proteins for comparison with the immunostaining data of Stro-1 undertaken in this thesis.

5.4.11 Results

The proteins between 48 and 52 kDa that resulted to be membrane bound proteins were monocarboxylate transporter 4 (SLC16A3), presenilin 2 (PSEN2), thyroid receptor-interacting protein 6 (TRIP6) and tyrosine-protein kinase (CSK). Figure 79 shows the immunostaining images obtained through the literature search together with the images taken in this thesis. Images from Stro-1 antibody staining (Figure 79; E, F, G, H) show an overall uniform distribution of the protein within the cell membrane, with some structuring features presented in the form of dots (Figure 79; E) or threads (Figure 79; G). Depending on the morphology of the cell lines used, such features are consistent with some of the proteins found within the mass spectrometry results (Figure 79; A, B, C, D). In particular Figure 79, B presents similar dots to the Stro-1 staining in Figure 79 and comparable expression similarities can be observed between Figure 79, D and G.

In spite of the challenge of the analysis, given the different cell types involved in immunostaining and different microscopy configurations (the magnification used appears to be different) between images, the data collected is consistent with PSEN2 and CSK expression exhibiting similar traits to the Stro-1 expression. It is also important to consider the difficulties in comparing membrane staining. Definite answers could be achieved in the case if PSEN2, CSK and Stro-1 antibodies are employed to stain SSC in the same emperiment. Furthermore, ambiguities could be resolved with flow cytometry analysis as a more sensitive method than immunostaining.

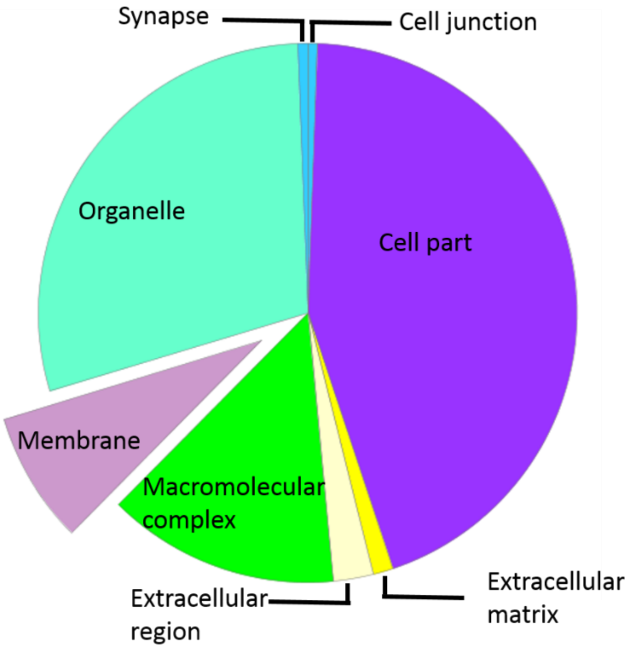


Figure 77: PANTHER protein classification based on cell compartments.

This pie chart classifies the proteins (48-52 kDa in size) according to the cellular component. Membrane proteins are of interest for this analysis hence the membrane part of the chart is highlighted.

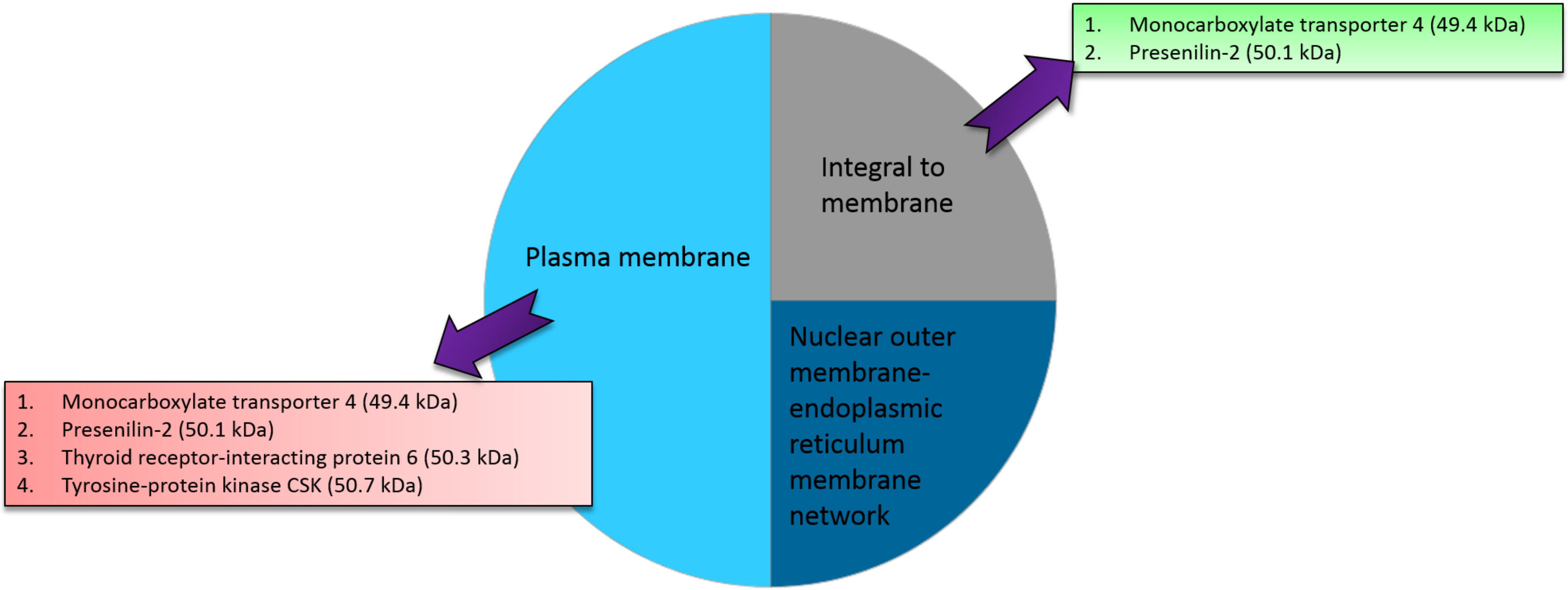


Figure 78: PANTHER protein subgroup of membrane proteins.

The membrane proteins identified in the previous pie chart were sub grouped into three categories by PANTHER: plasma membrane, integral to membrane and nuclear outer membrane-endoplasmic reticulum membrane network. The interest of this classifications relied within cell membrane proteins hence, for plasma membrane category and for integral to membrane category the actual proteins were extracted and reported in the pink and green boxes respectively.

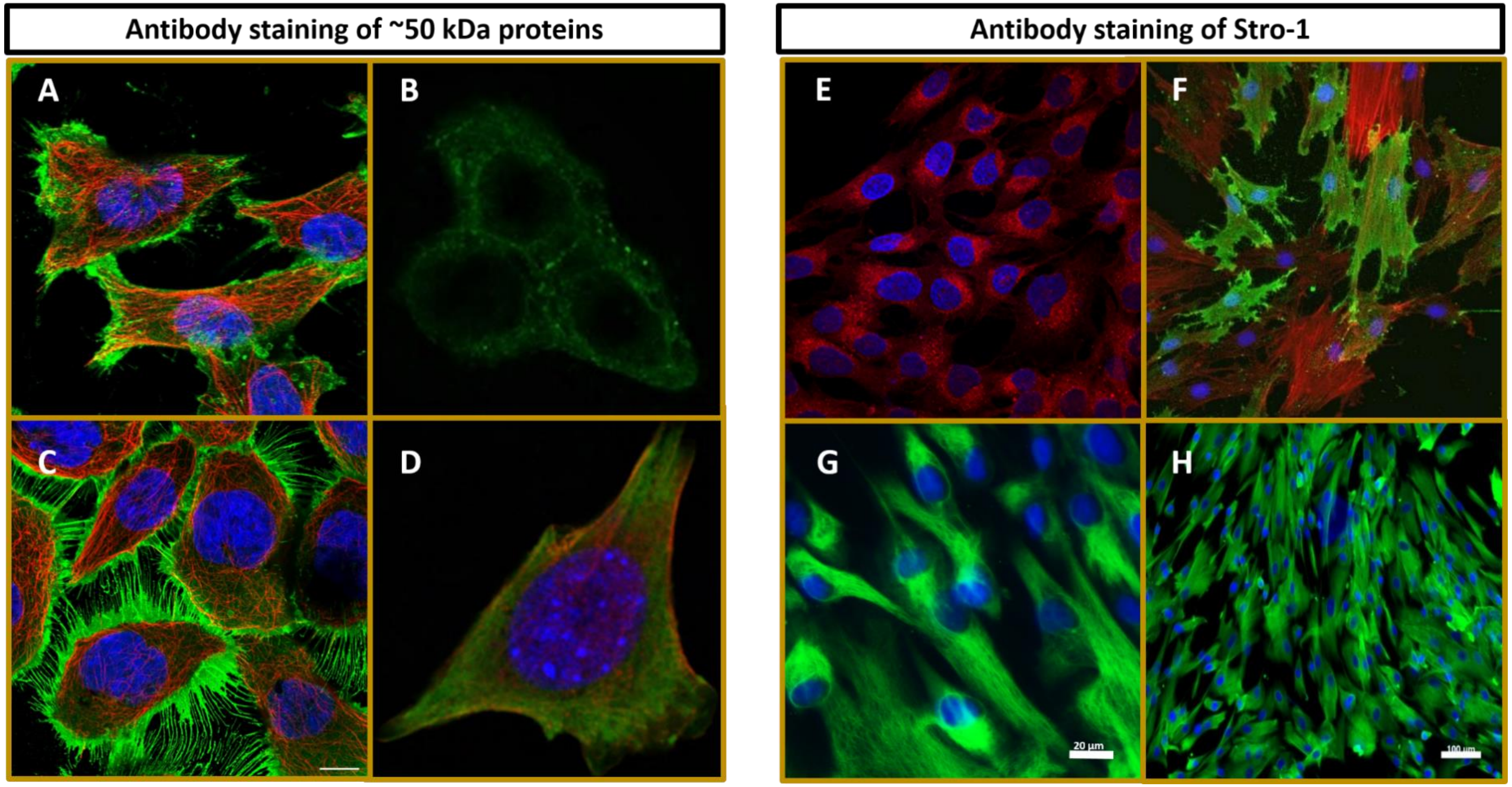

Antibody staining of ~50 kDa proteins
Antibody staining of Stro-1
A
B
C
D
E
F
G
H
20 µm
100 µm

Figure 79: Protein expression of 50 kDa membrane proteins (left) and Stro-1 antibody staining (right).

A): Monocarboxylate transporter 4 (49.4 kDa); cell line: U251; MG Antibody staining: green; source: Atlas antibodies. B): Presenilin-2 (50.1 kDa); cell line: JAR; antibody staining: green; source: Abcam. C): Thyroid receptor-interacting protein 6 (50.3 kDa); cell line: A-431; antibody staining: red; source: The human protein atlas. D): Tyrosine-protein kinase CSK (50.7 kDa); cell line: 3T3-L1; antibody staining: green; source: Antibodies online. E): Stro-1 staining; cell line: MG63; antibody staining: red; source: R&D systems. F): Stro-1 staining of mesenchymal stem cells; antibody staining: green; source: life technologies. G): Stro-1 staining; cell line: HK cells; antibody staining: green; source: 40x image taken for this thesis. H): Stro-1 staining; cell line: HK cells; antibody staining: green; source: 5x image taken for this thesis. Scale bars: G: 20 µm, H: 100 µm

5.5 Discussion

5.5.1 Overview

Bone related diseases continue to exert an ever increasing burden on society's progressively ageing population. However, despite decades of research, no solution has yet emerged to overcome the insidious problem of bone degeneration and related diseases. BM cells have the potential to differentiate into bone, cartilage and fat; a requisite for the field of regenerative medicine and therapeutics application. The challenge that lies in identifying, characterising and selecting stem cells from the BM has, to date, hindered the development of successful skeletal regenerative strategies.

Despite the challenges posed in identifying skeletal stem cells, common isolation methods have emerged utilising a set of markers, although with limited consensus across the research community. It is of primarily importance to identify alternative markers to discriminate SSC from other cell types within the marrow cell population (Bianco, 2014). Identification tools can primarily be associated with proteins, as proteins represent the preponderance of biologically active molecules responsible for the majority of cellular functions. In this regard, the entirety of proteins in a cell, known as the 'proteome', is an enticing area for biomarker discovery. In this thesis the proteome of skeletal stem cells was analysed with advanced mass spectrometry. The power of C_8 reverse phase off-line chromatography together with on-line nano-chromatography coupled with nanospray ionization–ultra high resolution mass spectrometry (LC –nESI-MS) was applied to skeletal stem cells extracted from the BM of 2 male and 2 female patients under two types of growth conditions, basal and osteogenic; and two types of surface marker discrimination, Stro-1 sorted cells and unsorted. The study conducted in this thesis identified over 7,000 proteins differentially modulated across the 4 samples, analysed at 95% FDR<0.05% confidence and identified and quantified by LC –nESI-MS-HCD-CID. The downstream analysis purpose was fourfold:

- Acquire a comprehensive overview of the regulated proteins – Hierarchical clustering.
- Select and analyse the greatest upregulated and downregulated proteins – Upregulation and downregulation analysis by 1.5 and 2 fold.

- Gather an understanding of the protein functions by classifying the proteins – Protein classification on the basis of biological pathways, cellular components, biological process and molecular function.
- Analyse the protein interaction to understand links between processes – Pathway analysis with FunRich and MetaCore software.

5.5.2 Hierarchical clustering provides for a comprehensive overview of the protein expression highlighting patient and gender differences

With over than 7000 proteins, each presenting with 20 different expression values (5 ratios per 4 patients), the results become challenging to interpret. In this instance software that can offer high-throughput analysis are indispensable tools to understand the underlying biology. A useful top level analysis that visualises the differences in protein expression is given by clustering the proteins and creating a heatmap. In this thesis, these steps were achieved by employing the free software, Cluster 3.0 for the clustering and TreeView for building the heatmap. The results showed differences in expression between the male and the female data sets. The male dataset presented less variety of protein modulation and a homogeneous expression between upregulated and downregulated proteins. 11 proteins (Figure 63, C) were found to have a patient specific expression profile. These 11 proteins, presented a downregulated profile in M75, SB/UO and SO/UO ratios; which indicated that these proteins were downregulated upon sorting. Furthermore, an upregulated protein profile was found in M75 SO/SB and UO/UB indicating that the osteogenesis was responsible for the protein modulation. These results represent an example of the analytical potential of hierarchical clustering. Understanding the biological validity of the results would require further validation studies.

Interestingly, the female protein expression profile in clustering analysis appeared highly heterogeneous and the most upregulated and downregulated proteins (Figure 65, B and C) appear to have been influenced by the osteoinduction. Furthermore, gender specific hormone receptors may have contributed to the heterogeneity of results.

5.5.3 Upregulation and downregulation analysis by 1.5 and 2 fold highlighted that protein modulation was driven by osteogenesis

Protein expression results were filtered by fold changes of 2 for upregulation and 1.5 for downregulation. These cut-offs were chosen as they were considered the best suited for the different expression profiles that were present in the samples. Specifically, for upregulation it was intriguing to notice that the same 4 proteins passed the filter for all the samples across the 3 conditions only when osteogenic samples were compared to basal samples (Figure 66). These results indicated that media conditions were the major contributory cause for modulation. The 4 identified proteins included: peptidyl-prolyl cis-trans isomerase, carboxypeptidase M, amine oxidase [flavin-containing] A and fibrillin-2. Association between these proteins and BM mesenchymal stem cells was searched in the literature and discussed in the following paragraphs.

5.5.3.1 Peptidyl-prolyl cis-trans isomerase

Peptidyl-prolyl cis-trans isomerase (FKBP5) is a 51 kDa protein found in the nucleus and in the cytoplasm of the cell. It is a key regulator in the glucocorticoid receptor pathway (Sinclair, Fillman *et al.*, 2013), and in the estrogen signalling pathway (KEGG database). FKBP5 promotes NF-κβ activation, its hyper expression is associated with cancer and resistance to therapy (Romano, Sorrentino *et al.*, 2011). FKBP5 has been found to contribute to Tau isomerisation and stabilisation (Jinwal, Koren *et al.*, 2010). Thus, suggesting a role in degenerative diseases.

In relation to bone, FKBP5 has been found to promote osteoclast differentiation independently from the NF-κβ activation (Kimura, Nagai *et al.*, 2013). A research study suggested FKBP5 is involved with rheumatoid arthritis as it was found to be enhanced at a gene level in a population of BM cells (Matsushita, Hashimoto *et al.*, 2010).

5.5.3.2 Carboxypeptidase M

Carboxypeptidase M (CPM) is a membrane bound enzyme that cleaves arginine and lysine from the COOH-terminal of peptides like other carboxypeptidases. It was named M for its localisation on the cellular membrane (Skidgel, David *et al.*, 1989).

The only evidence of CPM involvement with BM cells describes the cleavage of lysine at the COOH-terminal of the stromal cell-derived factor (SDF)-1α. Cleaving SDF-1α signifies degrading it. This degradation corresponds to facilitating the haematopoietic stem/progenitor cell mobilisation to peripheral blood. Thus, CPM is indirectly associated with the decrease in the chemoattractant activity of SDF-1α (Marquez-Curtis, Jalili *et al.*, 2008).

CPM has not previously been associated with osteogenesis or Stro-1 positive cells. In this work this enzyme was found upregulated in all sorting conditions when grown in osteogenic media. The data suggests that it may play a role in osteogenesis.

5.5.3.3 Amine oxidase [flavin-containing] A

Amine oxidase [flavin-containing] A (MAOA) is an 60 kDa oxidoreductase that is involved in neurotransmitter degradation and synaptic transmission (UniProt database) (Youdim, Edmondson *et al.*, 2006). It is an integral component of the mitochondrion membrane. Mutation of the MAOA gene leads to mental retardation in male subjects (Brunner syndrome) (Brunner, Neien *et al.*, 1993).

The single entry in the literature that associates MAOA with skeletal tissue is described by the association of MAOA to adipocytes formation following the findings of increased MAOA activity during pre-adipocyte differentiation (Bour, Daviaud *et al.*, 2007).

5.5.3.4 Fibrillin-2

The importance of fibrillin-2 (FBN2) in connective tissues has been demonstrated by analysis of FBN2 null mice. These mice displayed a decrease of collagen cross-links in their tendons (Boregowda, Paul *et al.*, 2008).

Other studies revealed the little known importance of FBN2 in micro fibril development. This study revealed that FBN2 is present in postnatal microfibrils and its expression is turned off during foetal development as the epitopes are masked by fibrillin-1. This investigation proposed FBN2 as a ‘biomarker’ in pathogenesis affecting microfibrils fragmentation, such as the Marfan syndrome (Charbonneau, Jordan *et al.*, 2010).

5.5.4 Classification of the proteins highlighted a high catalytic activity

Following the analysis to understand the driving force of protein modulation and classification (hierarchical clustering, upregulation); the analysis of understanding the main biological pathways, cellular components, biological processes and molecular functions between genders was performed. Overall the results of protein classification have shown that the majority of the proteins were involved in the metabolism of the cells, and had a catalytic activity. No differences between the genders were noted in this analysis.

5.5.5 Pathway analysis using FunRich and MetaCore attributed the protein modulation to osteoinduction by downregulation of the GCR-alpha cycle

Interactions, signalling and metabolic pathways occur between intrinsic, highly regulated protein pathways.

To unpick the intrinsic relationship between the proteins discovered by mass spectrometry and to gather biological significance, the entirety of connections between the different proteins pathways was explored using potent *in silico* tools, FunRich and MetaCore (1.1.1, 5.4.7). These tools operate by examining the protein input against a database that in the case of MetaCore especially, is highly curated by a team of experts that scout the literature and generate protein-protein association. FunRich works in a similar fashion although as open software made freely available to the public, (note: MetaCore is proprietary software requiring purchase and subscription) it comes with limitations (*e.g.* it does not consider protein expression values), thus it has not been extensively cited in the literature.

FunRich demonstrated the complex proteins network and highlighted some important nodes, whilst MetaCore highlighted just one node. This node, constituted by the GCR-alpha, was also present in the FunRich analysis of the female samples (gene NR3C1, Figure 71). The MetaCore analysis was run on common proteins between genders whilst the FunRich analysis was run on each gender separately. Ideally the MetaCore analysis should have also been

performed on each gender alone, although limitations in tools availability impeded this process.
The key biological trends observed in this study exhibited the characteristic hallmarks of osteogenic differentiation which is reflected through the consensus of the *in silico* interpretation of the data performed by bioinformatics tools (FunRich and MetaCore) and by observing modulation trends. The FunRich and MetaCore results corroborated with the modulated trends (hierarchical clustering and up-down regulation analysis) demonstrating that osteogenic differentiation was responsible for proteins modulation whilst the sorting of the cells by the Stro-1 marker was not shown to influence protein modulation.

5.5.6 Validation of the mass spectrometry results by Western Blotting

A high throughput of data was obtained by the mass spectrometry analysis of skeletal stem cells. These data were generated with the intention of finding candidates to identify SSC markers. Hence, it was crucial to generate a pipeline for downstream analysis, which could eventually convert into the filtering of few candidates. In the first instance a bioinformatics approach was used to look for proteins with enhanced modulation. This approach assumed that differences between the samples were correlated to highly modulated proteins, purposely neglecting proteins with poor modulation as these were considered not to be primarily responsible for cell processes. WB was adopted as a first step validation of the MS, with the aim to assess the presence of chosen proteins in the sample.

Between the most upregulated proteins, CPM, DCN, and FBN2 were chosen to be validated. ALP was also chosen on the basis of being a well characterised marker for osteogenesis. However, from the chosen proteins, only the validation of CPM was successful, the validation of the remaining proteins (DCN, FBN2and ALP) was not successful. The potential causes of this, were attributed to the poor affinity of the commercial antibodies to the proteins. Due to the controls incorporated in the study, it was clear that the samples were intact and that the technique was run correctly. However, the antibody failed to detect the targets even when a pure peptide was used as a positive control on the gel for the specific case of DCN. Hence, it can be deduced that the validation of DCN, FB2 and ALP was not finalised due to intrinsic reagent's limitations; although, the validation results obtained for CPM indicated agreement with the MS results. Further testing with different antibodies or other molecules (potentially aptamers) need to be employed in the future.

5.5.7 Further observations

In previous studies, 2D gel mass spectrometry was used to elucidate differential expression of proteins between mesenchymal stem cell and non-stem cell mesenchymal cell cultures (olfactory tissue cells). Their expression profiles were surprisingly similar. Between the ~550 proteins identified, actin-cytoplasmic 1 and 2, alpha internexin, alpha enolase, endoplasmin and neurofilament light polypeptide were upregulated in mesenchymal stem cell cultures (Wetzig, Alaiya *et al.*, 2013). These proteins were used to interrogate the results of this thesis. It was discovered that actin, endoplasmin and neurofilament light polypeptide were also identified amongst the protein discovered in this thesis.

In a quantitative approach, iTRAQ labels were employed to highlight the proteome of porcine SSC. 5 proteins were found to be differentially regulated and amongst those vimentin was found to be responsible for cell motility after knock-down validation experiments (Huang, Niu *et al.*, 2015). In this thesis, vimentin was the protein identified in the 50 kDa band analysis (section 3.4.2.4) and it was also observed within the results of MS2.

A similar method employed in this project was used to assess the proteomic profile of a chondrogenic cell line C3H10T1/2. Proteins with enhanced expression profile were found to be: 3-phosphoadenosine 5-phosphosulfate synthase 2 (PAPSS 2), lysil oxidase (LOX), basic transcription factor 3-like 4 (BTF3L4) and fibulin-5 (FBLN5) (Ji, Ji *et al.*, 2010). Amongst these markers, FBLN5 and LOX were also found in this thesis, however the expression profile presented a downregulation. The discrepancy in expression profiles between the published literature and this thesis could be explained by the cells (in this thesis) undergoing osteogenesis (and not chondrogenesis) resulting in a downregulation of proteins, which are actually upregulated in chondrogenesis.

In another study, SILAC tags were employed to tag the membrane proteins of SSC for marker identification. Neprilysin (CD10) and choline transporter-like protein 1 (CD92) resulted to be significantly overexpressed (Graneli, Thorfve *et* *al.*, 2014). Both these proteins were also enclosed in this thesis however their expression was not upregulated.

Biglycan, 4F2 cell-surface antigen heavy chain, PDZ and LIM domain protein 5, tubulointerstitial nephritis antigen-like, and keratin 18 were found upregulated in another study using an iTRAQ 8-plex (Holley, Tai *et al.*, 2015). All these proteins encompassed the results of this thesis, although their expression was not highly modulated in any of the samples.

One of the early mass spectrometry studies on SSC, demonstrated that versican, or chondroitin sulphate proteoglycan, could potentially be employed as an early osteogenic marker. Furthermore, within that study, tenascin and CD44 were detected interacting with versican. The study implied that versican and tenascin worked together to impede cell migration upon osteogenic differentiation (Foster, Zeemann *et al.*, 2005).

As well as a mass spectrometry approach, being employed to study the proteome of SSC, antibodies arrays have also been utilised. In a study screening an array of 228 extracellular antibodies, CD108 and CD24 were identified as upregulated markers in mesenchymal stem cells when compared to non-stem cell mesenchymal cell cultures; moreover CD40 was upregulated in non-stem cell mesenchymal cell cultures when compared to mesenchymal cell culture (Wetzig, Alaiya *et al.*, 2013). CD108 was also identified in this work although no trace of CD24 and CD40 was evident

To recapitulate the proteomic studies published in the literature encompass similarities with the results presented in this thesis. In spite of the fact that a unique specific marker for SSC has not been identified yet, the interest in the field of biomarker discovery will continue to grow.

5.5.8 Considerations

5.5.8.1 Gender differences

A limitation in this approach is characterised by the subdivision of samples by gender per each iTRAQ 8-plex experiment. Thus, samples were analysed in 2 different batches (iTRAQ experiments): one for the male cohort and one for female cohort. The choice of this subdivision was due to the limitation of the iTRAQ set consisting of 8 probes (8 probes=8 samples). Moving forward, an improved technique would employ mixing of the male and the female samples to have more comparability between batches. However, the MS approach is a powerful tool to create new hypothesis led by the data obtained.

5.5.8.2 Differences with MS1 study

In both studies, the overall observation within the results was represented by osteogenic differentiation driving protein expression, in spite of this the number of identified proteins varied greatly between the two studies. The reason for this discrepancy can be attributed to technical differences. For example, different lysis methods and the different chromatography approach taken. The different chromatography column used (C_8 instead of C_{18}) performed a superior separation due to a less drastic hydrophobic gradient. The C_{18} column had the highest degree of hydrophobicity compared to C_8 due to a longer carbon backbone. 67 fractions of peptides were collected from the chromatography experiment in MS2 in contrast to the 28 of MS1. The small pool of patients of MS1 (1 patient in MS1 compared to 4 in MS2) could have also played a role in the smaller number of proteins identified in the MS1 study. It is important to note that in spite of the protein number identified the meaning of the results remained unvaried.

5.5.9 Discussion of *in silico* analysis of 50 kDa proteins for correlation and association to Stro-1 based on immunostaining

In this study I attempted to identify the Stro-1 antigen within the 48 and 52 kDa membrane protein region identified by mass spectrometry. PANTHER was employed to facilitate the classification and identify proteins presented within the cell membrane. Comparison of the immunostained proteins with their corresponding antibodies were visually compared to Stro-1 stained cells. The results indicate similarities with PSEN2 and CSK. PSEN2 is involved in processing the amyloid-β protein in neurons and a mutation in the gene suggests an increase of Alzeihmer disease by increasing the extracellular concentration of the amyloid-β protein which is deposited in senile plaques (Scheuner, Eckman *et al.*, 1996). The enzyme CSK is a negative regulator of the Src family tyrosine kinases and act by phospharylating tyrosin residues (Okada, 2012). The role of PSEN2 in SSC remains unknown. The function of both PSEN2 and CSK does not appear to be directly correlated with SSC.

In summary, this small *in silico* study assessed further the MS results with the rationale derived from previous results discussed in chapter 3. The identification of several potential candidates for Stro-1 protein has created a new hyphothesis (PSEN2 and CSK show the potential to be the Stro-1 antigen) which must be assessed in future studies.

5.6 Conclusion

5.6.1 Hypotheses iteration

The results in the context of the stated hypothesis in section 5.2 were examined:

I. The Stro-1 sorted population of skeletal stem cells presents an enriched and unique proteomic profile. This hypothesis is **false** or the method used was not sufficiently adapted to prove the hypothesis to be true.

As previously discussed in section 4.6, and observed in this study, an evident upregulation of proteins caused by Stro-1 sorting was not captured within the results. Reasons could include the loss of Stro-1 expression within 10 days from extraction, the long time in culture (28 days) could have amortised the characteristic traits of a Stro-1^{+} cell; furthermore the Stro-1^{+} cells are part of the unsorted fraction of the cells, thus, when comparing the two fractions in a ratio equation, a “like for like” situation is presented and hence the results remained unvaried.

II. The osteogenic population presents a different proteomic profile from the cell population in basal media. This hypothesis is **true**.

The results of the different analyses of the mass spectrometry data, can be summarised by saying that for SSCs *in vitro* expanded for 28 days, proteins modulation was highly governed by the chemical stimuli artificially introduced in the culture condition through the addition of dexamethasone and ascorbic acid to the culture media. Sto-1 sorted cells and unsorted cells responded to these stimuli in a similar manner with no evident alteration due to the sorting as expected.

In summary this pilot study reports one of the first multiplex iTRAQ based, MudPit method for the proteomic analysis of SSCs. Understanding the surface proteomic profile of SSC will aid in creating new tools for SSC specific isolation.

III. One or more of the proteins found in the MS study is membrane bound and has a size range between 48 and 52 kDa; it has the potential to be the Stro-1 antigen, if similar immunostaining can be determined. This hypothesis is **true.**

PSN2 and CSK are indeed 50 kDa membrane bound proteins and given their expression (assessed by immunostaining) exhibits similar traits to the Stro-1 expression; thus, PSN2 and CSK encompass the potential of being Stro-1.

5.6.2 Limitation of the MS study

As discussed in the previous chapter the mass spectrometry data and analysis present the following limitations:

- Dynamic range magnitude obscuring very high or low protein expression.
- iTRAQ labelling flattening the protein expression.
- Fixed point analysis and cell cycle influencing the momentary proteins expression.
- Lack of standards of same retention time.
- The use of Stro-1, an unspecific, not well characterised marker to isolate cells.
- *In vitro* expansion of the cells artificially influencing the protein expression.
- Comparison of Stro-1 vs unsorted cells which also encompass Stro-1 cells, “like for like” situation.
- Small patient cohort.
- Patient data confidentiality regarding the exposure to medication.

6. Chapter 6: General discussion

6.1 Setting the scene

The incidence of bone diseases has reached the level of a global epidemic with significantly increased rates of morbidity and mortality and substantial healthcare costs (40 billions) (WHO). The standard treatment for bone conditions such as osteoporosis consists of palliative care with the use of drugs including bisphosphonates (as suggested by NICE guidance 160; http://www.nice.org.uk). In recent years, cell based therapies have improved the potential to alleviate bone diseases, in particular therapies involving the use of skeletal stem cells (SSCs) (Bennett, Bergeron *et al.*, 2007). Currently MSCs are the subject of 506 clinical studies, of which the majority are concentrated in Asia, followed by Europe; details are listed in Table 18 (source: www.clinicaltrials.gov). Interestingly, besides skeletal diseases, SSCs are also involved in the treatment of other conditions such as diabetes, leukaemia, different types of cancer, heart conditions, Alzheimer disease, cerebral infraction, and even infertility (source: www.clinicaltrials.gov).

Table 18: Clinical studies using MSC in the world.

Table is adapted from the information reported on www.clinicatrials.gov when entering 'mesenchymal stem cells' as the search criteria (06/06/2015).

Region name	Number of studies
World	***506***
Africa	13
Central America	9
East Asia	167
Europe	119
Middle East	45
North America	99
North Asia	9
Pacifica	7
South America	12
South Asia	14
Southeast Asia	12

The extent of fields where SSCs are currently employed can be considered as evidence of SSC's clinical potential. However, the clinical potential of SSCs is strictly and primarily correlated with the ability of isolating the SSC in the first instance. Isolating a cell presumes the *a priori* knowledge of the “characteristics” of that cell. Controversy arises even with the characterisation of SSCs; hence the cell specificity with which SSCs are isolated remains debatable. Identification of cells is generally achieved by binding of a specific antibody to a specific antigen expressed on the surface of the cell. Effective isolation of cells would require the knowledge of an antigen specific to one cell type only and a selective antibody with high affinity for this antigen. In the case of the SSC a specific antigen has proved elusive; consequentially no specific antibody is available. On this basis, it is possible to argue that one cannot specifically isolate SSCs with current techniques and knowledge. Obtaining *bona fide* SSCs with high differentiation potential is essential as successful isolation will eventually be reflected in the therapeutic output *i.e.* improving patient healthcare. The methodology currently employed to isolate SSCs is based on a set of markers exhibited by a sub-population of cells residing in the BM stroma. These markers, although non-specific, select a sub-group of SSCs with remarkable differentiation potential and the ability to form CFUs. While this improved selection reduces some of the limitations, the issue of isolating specific stem cells is not fully resolved, but attenuated in the absence of a specific marker.

In this thesis, the general aim was to investigate specific markers for the isolation of SSC using a variety of techniques.

In chapter 3, the osteogenic differentiation potential of SSC was established and antibody selection techniques were elucidated focusing on the Stro-1 marker. In chapter 4 and 5, the identity of SSCs markers was examined by performing a comprehensive study of the proteome of SSC using quantitative mass spectrometry.

Establishing the cell potential to form bone is a widely used and robust technique, which employs staining with alkaline phosphatase (ALP); ALP is considered an early bone marker and is involved in the bone mineralisation process. In this study, cells stained with ALP expressed the marker in different nuances according to the basal or osteogenic culture conditions employed. Samples cultured under osteogenic conditions demonstrated a high level of ALP expression, reflected in the ability of the cells to mineralise bone. ALP expression was quantified by an image-based approach (CellProfiler) and the transcriptome of ALP was analysed by RT-PCR. A late marker for osteogenesis, osteocalcin (OCN) and a marker for chondrogenesis, sex determining region Y-box 9 (SOX9), were chosen as a further positive and negative control for osteogenesis, respectively. Experiments of these markers identified that the Stro-1 cell population employed was capable of osteogenic differentiation.

Analysis of the existing known markers for SSCs was then examined. It was quickly noted that one of the widely used markers, selected by the Stro-1 antibody, is unknown. Furthermore, the Stro-1 antibody encompassed the advantage of being widely available, being in-house produced. Investigations aimed to characterise the Stro-1 antibody and antigen were undertaken employing WB. The results indicated that the Stro-1 antigen is an insoluble protein, demonstrated by the reactivity of the Stro-1 antibody to a 50 kDa antigen in the insoluble fraction of HK cells (Stro-1 positive cells). These results are in agreement with a previous study indicating the same antigen size (Castrechini, Murthi *et al.*, 2010), although a more recent study has claimed the antigen to be of 75 kDa (Ning, Lin *et al.*, 2011). Isolation of the antigen using immunoprecipitation was attempted without success due mainly to the IgM nature of the antibody characterised by high insolubility and size.

The realisation that antibody-based techniques have failed despite various attempts to find a unique isolation marker suggests that alternative technologies to antibodies should be explored. Thus, the prospect of synthetic molecules, aptamers, for the selection of specific surface ligands to substitute antibodies was examined. Aptamer selection fundamentally consist of the selection of high affinity ligands that bind to a given target. This process termed Systematic Evolution of Ligands by Exponential Enrichment, (SELEX) is repeatedly run with the aim of creating a pool of aptamer molecules with high affinity to one or more targets on the cells analysed. The technique was briefly explored in this work. However, significant technical challenges precluded the successful execution of the study. Nevertheless, this attempt has formed the basis of employing aptamer technique in biomarker discovery. Aptamer technology has been previously employed with a similar aim to that proposed in this thesis and an aptamer molecule termed G-8 was found to bind porcine mesenchymal stem cells (Guo, SchAfer *et al.*, 2006). Further characterisation of this aptamer was not found in the literature.

The final stage of this thesis explored the proteome of SSC using quantitative mass spectrometry with the aim of selecting SSC candidate markers. Mass spectrometry was considered a modern, high throughput method which provided better sensitivity than antibody labelling.

Hitherto, previous studies with similar hypotheses have explored a variety of methods to characterise SSC and their surface marker profile. Proteomic methods employed to study SSCs covered 2D gels/MS, iTRAQ-SILAC LC-MS/MS, to phage display, antibodies arrays and aptamers. These methods have been discussed in detail in the discussion in chapter 5 where a comparison between published results and the results of this thesis was described. The majority of the results in the literature reported similar markers to this thesis, although the expression of the proteins in the published results was not in agreement with this thesis. This comparison endorses the method employed as similar proteins were identified, although the expression differences created conflicting observations. The reason behind the observed differences in expression was considered and could be attributed to the different culture methodology and different type of cell source analysed. In order to mitigate these discrepancies further studies combining an increase in sensitivity and homogeneity of cell preparation should be employed. In the mass spectrometry study reported here, I applied iTRAQ LC/MS/MS on cells extracted from 4 different patients in order to gain more insight into the molecular mechanisms underlying osteogenic differentiation. This analysis resulted in the identification of 7594 differentially expressed proteins. When facing this large amount of data, significance of the data can easily be lost unless comprehensive analytical tools, that consider the entirety of the data, are applied. This creates the argument for the use of bioinformatics *in silico* tools analysis to reinforce the argument of the significance of the data set.

The data were interrogated to understand the driving force of protein modulation. Modulation could have been driven by the cell sorting, the osteoinduction process, neither or both. Hierarchical clustering highlighted that protein modulation was driven by the osteoinduction process and not by cell sorting. After this was established proteins were grouped in different categories in accordance to their biological pathways, cellular components, biological processes and molecular pathways. Subsequently the most upregulated proteins were identified and protein-protein interaction and

Metacore analysis were undertaken to facilitate the identification of potential novel osteogenic markers. Between the markers found, alongside the commonly used bone markers, peptidyl-prolyl cis-trans isomerase (FKBP5) and carboxypeptidase M (CPM) were identified as being upregulated in samples which were osteogenically induced.

CPM is a membrane-bound glycoprotein, that degrades extracellular proteins and controls peptide hormones and growth factors on the cell membrane (Craveiro, Ramalho *et al.*, 2006). CPM is also believed to play a role in inflammation as a result of being expressed on macrophages (Denis, Van Acker *et al.*, 2013). CPM presence on cancer cells was postulated to help with the diagnosis of liposarcoma from lipoma (Erickson-Johnson, Seys *et al.*, 2009). Expression of CPM in human BM stromal cells has previously been reported in a study where it was implied that CPM played a role in cell mobilisation (Liu, Martina *et al.*, 2007; Marquez-Curtis, Jalili *et al.*, 2008).

CPM and FKBP5 were found to be upregulated in a study in SSCs undergoing osteogenesis and chondrogenesis. Over-expression of FKBP5 was demonstrated to support and enhance differentiation in all three lineages (Liu, Martina *et al.*, 2007). Hence the results from the published studies are congruent with the results found in this thesis.

In conclusion, the fundamental finding in this work was that by encouraging osteogenesis, cells have shown a highly regulated proteomic profile. Sorting of the cell by the Stro-1 marker did not seem to influence the protein modulation. Regarding the markers identified, more detailed studies are needed to determine the role of these proteins as SSC biomarkers.

6.2 Results in context

To understand the power and importance of mass spectrometry it is important to appreciate its limitations. The UniProt database contains 68,554 proteins and mass spectrometry represents one of the pillars that will contribute to increase this figure (Figure 80) (Zhou, Liotta *et al.*, 2015). A study aimed to create an inventory of proteins was performed by analysing 11 cell lines by mass spectrometry (Geiger, Wehner *et al.*, 2012). It was demonstrated that in one single MS experiment 10,361 ± 120 proteins can be found. Considering that the study identified 11,731 proteins across the 11 cell lines, which corresponds to 10,216 genes in ENSEMBL, this study has identified the 50.5% of the complete human genome (assuming the total number of human genes in ENSEMBL is 20,230).

Similarly, in this thesis 7594 proteins were identified which correspond to the 37.5% of the human genome.

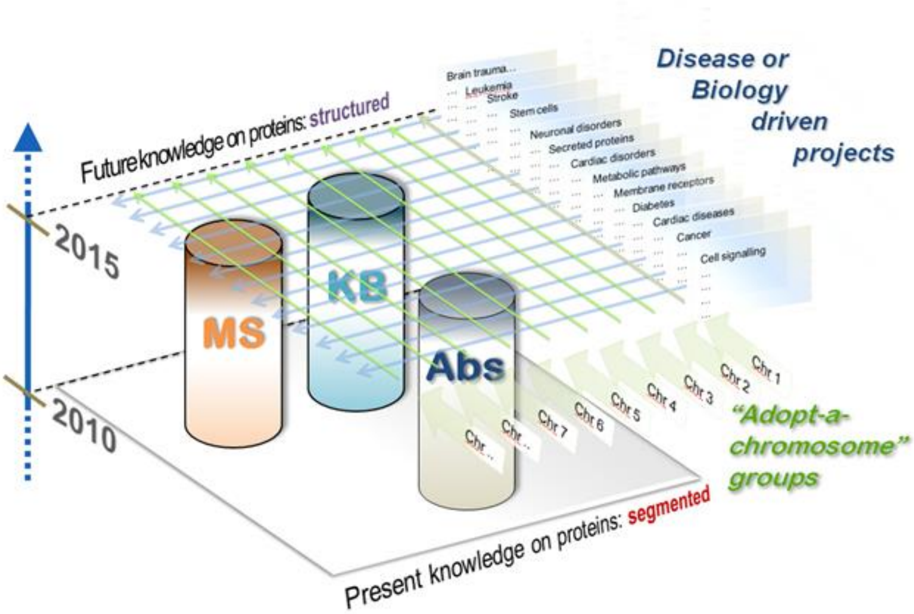

Disease or Biology driven projects
Brain trauma...
Leukemia
Stroke
Stem cells
Neuronal disorders
Secreted proteins
Cardiac disorders
Metabolic pathways
Membrane receptors
Diabetes
Cardiac diseases
Cancer
Cell signalling
Future knowledge on proteins: structured
2015
2010
MS
KB
Abs
Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
Chr 6
Chr 7
Chr...
Chr...
"Adopt-a-chromosome" groups
Present knowledge on proteins: segmented

Figure 80: The three pillars of the Humane Proteome Project (HPP).

The image represents the fundamental pillars of the HPP: Mass Spectrometry (MS), antibodies (Abs) and Knowledgebase (UniProt) (KB). The HPP project aims to map the human proteome by collecting the data from currently available and emerging proteomic techniques. Source: http://www.thehpp.org/overview.php (25/08/2015).

6.3 Limitations

The fundamental aim of this work was to identify potential specific markers for SSC *ex vivo* identification in order to select a homogeneous cell population to be applied for skeletal tissue engineering strategies.

The studies performed in this thesis explored antibodies and aptamer based approaches for the identification of SSC markers and ultimately examined the proteome of SSCs using mass spectrometry.

The main limitations of the study with regards to cell labelling with either antibodies or aptamers resides in the current gap in knowledge regarding the surface marker profile of the SSC. Thus, even approaches that are supposedly targeted, cannot be considered specific, as no targets are unique to SSCs. The drawbacks of the mass spectrometry used in these experiments are primarily based around the inefficiency of the technique to perform an analysis on a relative small number of *ex vivo* cells, therefore the cell populations had to be cultured to ensure adequate protein content for the MS analysis. This cell expansion might have altered the surface markers profile of SSCs, obscuring some markers that could have been potentially employed as selection markers. It was proven to be able to separate SSCs extracted from the bone marrow on the basis of their Stro-1 signal. The 5% of the cells which strongly expressed Stro-1 were separated from the rest by flow cytometry and stored at -80 °C (10.3). This experiment provided the material for the MS analysis on small number of *ex vivo* cell, although, this MS method was never developed due to time restriction and the samples were never assessed.

6.4 Future perspective

The outcome of this study provides important advances in understanding the proteome of SSCs. The work performed provides a basis for future investigations:

1. These studies create the ground to formulate new hypothesis to undertake further proteomic studies.
2. Forms a validated protocol for MS of SSCs.
3. Provides a pipeline for the interpretation of mass spectrometry data.
4. Contribute to the creation of a database of proteins for comparison and correlations with other studies.

Future work should be oriented towards the use of freshly isolated cells; and a range of isolation markers should be explored. Experimental design, which ensures a variation of samples and genders within an 8plex should be sought, to ensure genders comparison between one 8plex study.

Aptamer technology in combination with mass spectrometry should be further investigated as this offers exciting new prospects for identifying unique surface markers.

Future techniques that could be explored include the emerging new technology based on SOMAmers (Slow Off-rate Modified Aptamers), which promise to bring proteomics to the research community, using a simple 96 well-based array, similar to an ELISA but with high-affinity and specificity (Gold, Walker *et al.*, 2012).

At the completion of a fully mapped surface proteome of SSCs, the search for isolation tools will ease, enabling the isolation of a homogeneous cell population and subsequent use within clinical settings.

7. Chapter 7: Conclusion

7.1 Conclusion

The results in the context of the stated hypotheses in section 1.5 were examined:

I. The Stro-1 subpopulation of cells represents the *bona fide* SSCs and as such exhibit high differentiation potential. *Objective:* To assess the differentiation potential of the Stro-1 fraction of cells by molecular techniques. This hypothesis is **true** as the molecular analysis of ALP and OCN demonstrate a higher expression in the Stro-1 sorted cells than in the unsorted cells (Chapter 3).

II. The Stro-1 antigen size can be identified by WB analysis. *Objective:* to develop a WB-based method that allows detection of the Stro-1 antigen. This hypothesis is **true**. A WB method was developed and the Stro-1 antibody demonstrated reactivity to an antigen at 50 kDa. It was further demonstrated the reactivity at 50 kDa in the insoluble fraction of the cell lysate, indicating a presumed hydrophobicity of cell membrane proteins such as Stro-1 (Chapter 3).

III. Within the Stro-1 fraction, resides a subset of cells with a unique proteomic profile that can be revealed by mass spectrometry. *Objective:* to focus on the differences in mass spectrometry profiles between the Stro-1 sorted fraction and the unsorted fraction. This hypothesis was **not verified** as the structure of the results did not allow the testing of this hypothesis. The iTRAQ ratios offer a satisfying understanding of the relative protein expression. This relativity implies that a specific protein is present in all the group analysed. Considering the hypothesis was testing unique proteins in the Stro-1 group only, the hypothesis was not verified due to the intrinsic relativity of the iTRAQ tags (Chapters 4 and 5).

IV. Protein modulation is a result of sorting and not osteoinduction. *Objective:* to assess the source of protein modulation in the mass spectrometry results. This hypothesis is **false**. Protein modulation was found to be driven by osteoinduction. Although the method compared

Sto-1 sorted cells with unsorted cells, which contained Stro-1 cells. This could have caused a dilution effect of the ability of the Stro-1 sorting to drive protein modulation (Chapters 4 and 5).

V. The aptamer SELEX process identifies potential ligands for the isolation of SSC. *Objective:* to run a SELEX experiment on MG63 as mimicking the osteogenic phenotype. This hypothesis was **not verified** due to contamination and time restrictions. Although a protocol to manage growth density of MG63 in concomitance with aptamer selection rounds was developed (Chapter 3).

8. List of References

Axelrad, T. W. and T. A. Einhorn (2009). "Bone morphogenetic proteins in orthopaedic surgery." *Cytokine Growth Factor Rev* **20**(5-6): 481-488.

Baddoo, M., K. Hill, R. Wilkinson, D. Gaupp, C. Hughes, G. C. Kopen and D. G. Phinney (2003). "Characterization of mesenchymal stem cells isolated from murine bone marrow by negative selection." *J Cell Biochem* **89**(6): 1235-1249.

Ball, S. G., C. A. Shuttleworth and C. M. Kielty (2007). "Mesenchymal stem cells and neovascularization: role of platelet-derived growth factor receptors." *J Cell Mol Med* **11**(5): 1012-1030.

Bárcia, R. N., J. M. Santos, M. Filipe, M. Teixeira, J. P. Martins, J. Almeida, A. Água-Doce, S. C. P. Almeida, A. Varela, S. Pohl, K. E. J. Dittmar, S. Calado, S. I. Simões, M. M. Gaspar, M. E. M. Cruz, W. Lindenmaier, L. Graça, H. Cruz and P. E. Cruz (2015). "What Makes Umbilical Cord Tissue-Derived Mesenchymal Stromal Cells Superior Immunomodulators When Compared to Bone Marrow Derived Mesenchymal Stromal Cells?" *Stem Cells International* **2015**: 583984.

Baron, R. (1993). "Anatomy and Ultrastructure of bone." *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism: An Official Publication of the American Society for Bone and Mineral Research*. M. J. Favus, S. Christakos, M. Kleerekoper, E. Shane, R. F. Gagel, C. B. Langman, A. F. Stewart and M. P. Whyte, Lippincott Williams and Wilkins.

Barry, F. P., R. E. Boynton, S. Haynesworth, J. M. Murphy and J. Zaia (1999). "The monoclonal antibody SH-2, raised against human mesenchymal stem cells, recognizes an epitope on endoglin (CD105)." *Biochem Biophys Res Commun* **265**(1): 134-139.

Barry, F. P. and J. M. Murphy (2004). "Mesenchymal stem cells: clinical applications and biological characterization." *Int J Biochem Cell Biol* **36**(4): 568-584.

Battula, V. L., S. Treml, P. M. Bareiss, F. Gieseke, H. Roelofs, P. de Zwart, I. Muller, B. Schewe, T. Skutella, W. E. Fibbe, L. Kanz and H. J. Buhring (2009). "Isolation of functionally distinct mesenchymal stem cell subsets using antibodies against CD56, CD271, and mesenchymal stem cell antigen-1." *Haematologica* **94**(2): 173-184.

Bennett, K. P., C. Bergeron, E. Acar, R. F. Klees, S. L. Vandenberg, B. Yener and G. E. Plopper (2007). "Proteomics reveals multiple routes to the osteogenic phenotype in mesenchymal stem cells." *BMC Genomics* **8**: 380.

Bessarabova, M., A. Ishkin, L. JeBailey, T. Nikolskaya and Y. Nikolsky (2012). "Knowledge-based analysis of proteomics data." *BMC Bioinformatics* **13** (16): S13.

Bianco, P. (2011). "Bone and the hematopoietic nich:a tale of two stem cells." *Blood* **117**: 5281-5288.

Bianco, P. (2014). ""Mesenchymal" Stem Cells." *Annual Review of Cell and Developmental Biology, Vol 30* **30**: 677-704.

Bianco, P., X. Cao, P. S. Frenette, J. J. Mao, P. G. Robey, P. J. Simmons and C.-Y. Wang (2013). "The meaning, the sense and the significance: translating the science of mesenchymal stem cells into medicine." *Nature Medicine* **19**(1): 35-42.

Bianco, P., L. W. Fisher, M. F. Young, J. D. Termine and P. G. Robey (1990). "Expression and localization of the two small proteoglycans biglycan and decorin in developing human skeletal and non-skeletal tissues." *Journal of Histochemistry & Cytochemistry* **38**(11): 1549-1563.

Bianco, P., S. Kuznetsov, M. Riminucci and P. Gehronrobey (2006). "Postnatal Skeletal Stem Cells." *Methods in Enzymology* **419**: 117-148.

Bianco, P., M. Riminucci, S. Gronthos and P. G. Robey (2001). "Bone Marrow Stromal Cells- Nature, Biology and Potential Application." *Stem Cells Coincise Review* **19**: 180-192.

Bianco, P., P. G. Robey and P. J. Simmons (2008). "Mesenchymal stem cells: revisiting history, concepts, and assays." *Cell Stem Cell* **2**(4): 313-319.

Bonewald, L. F. (2011). "The amazing osteocyte." *J Bone Miner Res* **26**(2): 229-238.

Boregowda, R., E. Paul, J. White and T. M. Ritty (2008). "Bone and soft connective tissue alterations result from loss of fibrillin-2 expression." *Matrix Biology* **27**(8): 661-666.

Bouchal, P., M. Dvořáková, T. I. Roumeliotis, Z. Bortlíček, I. Ihnatová, I. Struhárová, J. T. C. Ho, J. Maryáš, H. Imrichová, E. Budinská, R. Vyzula, S. D. Garbis, B. Vojtěšek and R. Nenutil (2015). "Combined proteomics and transcriptomics identifies carboxypeptidase B1 and NF-κB associated proteins as putative biomarkers of metastasis in low grade breast cancer." *Molecular & Cellular Proteomics* **14**: 1814-1830.

Bour, S., D. Daviaud, S. Gres, C. Lefort, D. Prévot, A. Zorzano, M. Wabitsch, J.-S. Saulnier-Blache, P. Valet and C. Carpéné (2007). "Adipogenesis-related increase of semicarbazide-sensitive amine oxidase and monoamine oxidase in human adipocytes." *Biochimie* **89**(8): 916-925.

Brunner, H. G., M. Neien, X. O. Breakefiedl, H. H. Ropers and B. A. van Oost (1993). "Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A." *Science* **262**: 578-580.

Burk, J., C. Gittel, S. Heller, B. Pfeiffer, F. Paebst, A. B. Ahrberg and W. Brehm (2014). "Gene expression of tendon markers in mesenchymal stromal cells derived from different sources." *BioMed Central Research Notes* **7**(826): 1-6.

Caplan, A. I. (1991). "Mesenchymal stem cells." *Journal of Orthopaedic Research* **9**(5): 641-650.

Caplan, A. I. (2007). "Adult Mesenchymal Stem cells for tissue Engineering Versus Regenerative Medicine." *Journal of Cellular Physiology* **213**: 341-347.

Caplan, A. I. (2008). "All MSCs are pericytes?" *Cell Stem Cell* **3**(3): 229-230.

Castrechini, N. M., P. Murthi, N. M. Gude, J. J. Erwich, S. Gronthos, A. Zannettino, S. P. Brennecke and B. Kalionis (2010). "Mesenchymal stem cells in human placental chorionic villi reside in a vascular Niche." *Placenta* **31**(3): 203-212.

Celil, A. B. and P. G. Campbell (2005). "BMP-2 and insulin-like growth factor-I mediate osterix (Osx) expression in human mesenchymal stem cells via the MAPK and protein kinase D signaling pathways." *Journal of Biological Chemistry* **280**(36): 31353-31359.

Chamberlain, G., J. Fox, B. Ashton and J. Middleton (2007). "Concise review: mesenchymal stem cells: their phenotype, differentiation capacity, immunological features, and potential for homing." *Stem Cells* **25**(11): 2739-2749.

Charbonneau, N. L., C. D. Jordan, D. R. Keene, S. Lee-Arteaga, H. C. Dietz, D. B. Rifkin, F. Ramirez and L. Y. Sakai (2010). "Microfibril Structure Masks Fibrillin-2 in Postnatal Tissues." *J Biol Chem* **285**: 20242-20251.

Chen, A. and S. Yang (2015). "Replacing antibodies with aptamers in lateral flow immunoassay." *Biosensors and Bioelectronics* **71**: 230-242.

Chen, D., M. Zhao and G. R. Mundy (2004). "Bone morphogenetic proteins." *Growth Factors* **22**(4): 233-241.

Clarke, B. (2008). "Normal bone anatomy and physiology." *Clin J Am Soc Nephrol* **3** (3): 131-139.

Cohen, D. M. and C. S. Chen (2008). "Mechanical control of stem cell differentiation." *StemBook*. S. B. a. J. Polak. Philadelphia, The Stem Cell Research Community.

Connolly, J., R. Guse, L. Lippiello and R. Dehne (1989). "Development of an osteogenic bone-marrow preparation." *Journal of Bone and Joint Surgery-American Volume* **71A**(5): 684-691.

Coutu, D. L., J. H. Wu, A. Monette, G. E. Rivard, M. D. Blostein and J. Galipeau (2008). "Periostin, a member of a novel family of vitamin K-dependent proteins, is expressed by mesenchymal stromal cells." *J Biol Chem* **283**(26): 17991-18001.

Cox, G., S. A. Boxall, P. V. Giannoudis, C. T. Buckley, T. Roshdy, S. M. Churchman, D. McGonagle and E. Jones (2012). "High abundance of CD271(+) multipotential stromal cells (MSCs) in intramedullary cavities of long bones." *Bone* **50**(2): 510-517.

Coxon, J. P., G. M. Oades, K. W. Colston and R. S. Kirby (2004). "Advances in the use of bisphosphonates in the prostate cancer setting." *Prostate Cancer Prostatic Dis* **7**(2): 99-104.

Cracknell, R. (2010). The ageing population. House of Commons.

Craveiro, R. B., J. D. Ramalho, J. R. Chagas, P. H. M. Wang, D. E. Casarini, J. L. Pesquero, R. C. Araújo and J. B. Pesquero (2006). "High expression of human carboxypeptidase M in Pichia pastoris. Purification and partial characterization." *Brazilian Journal of Medical and Biological Research* **39**(2): 211-217.

Crisan, M., S. Yap, L. Casteilla, C. W. Chen, M. Corselli, T. S. Park, G. Andriolo, B. Sun, B. Zheng, L. Zhang, C. Norotte, P. N. Teng, J. Traas, R. Schugar, B. M. Deasy, S. Badylak, H. J. Buhring, J. P. Giacobino, L. Lazzari, J. Huard and B. Peault (2008). "A perivascular origin for mesenchymal stem cells in multiple human organs." *Cell Stem Cell* **3**(3): 301-313.

Crockett, J. C., M. J. Rogers, F. P. Coxon, L. J. Hocking and M. H. Helfrich (2011). "Bone remodelling at a glance." *Journal of Cell Science* **124**: 991-998.

Curran, J. M., F. Pu, R. Chen and J. A. Hunt (2011). "The use of dynamic surface chemistries to control msc isolation and function." *Biomaterials* **32**(21): 4753-4760.

Da Silva Meirelles, L., P. S. Chagastelles and N. B. Nardi (2006). "Mesenchymal stem cells reside in virtually all post-natal organs and tissues." *Journal of Cell Science*: 2204-2213.

Dalby, M. J., N. Gadegaard and R. O. Oreffo (2014). "Harnessing nanotopography and integrin-matrix interactions to influence stem cell fate." *Nat Mater* **13**(6): 558-569.

Davis, K. A., B. Abrams, Y. Lin and S. D. Jayasena (1998). "Staining of cell surface human CD4 with 2′-F-pyrimidine-containing RNA aptamers for flow cytometry." *Nucleic Acids Research* **26**(17): 3915-3924.

Day, R. M. (2006). "Epithelial Stem Cells and Tissue Engineered Intestine." *Current Stem Cell Research & Therapy* **1**: 113-120.

Day, T. F., X. Guo, L. Garrett-Beal and Y. Yang (2005). "Wnt/beta-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis." *Dev Cell* **8**(5): 739-750.

Delmas, P. D. (1999). "Biochemical Markers of Bone Turnover in Paget’s Disease of bone." *Journal of bone and mineral research* **14**(2): 66-69.

Deng, C. X., A. WynshawBoris, F. Zhou, A. Kuo and P. Leder (1996). "Fibroblast growth factor receptor 3 is a negative regulator of bone growth." *Cell* **84**(6): 911-921.

Denis, C. J., N. Van Acker, S. De Schepper, M. De Bie, L. Andries, E. Fransen, D. Hendriks, M. M. Kockx and A.-M. Lambeir (2013). "Mapping of Carboxypeptidase M in Normal Human Kidney and Renal Cell Carcinoma: Expression in Tumor-Associated Neovasculature and Macrophages." *Journal of Histochemistry & Cytochemistry* **61**(3): 218-235.

Di Girolamo, C. M., D. Stokes, D. Colter, D. G. Phinney, R. Class and D. J. Prockop (1999). "Propagation and senescence of human marrow stromal cells in culture: a simple colony-forming assay identifies samples with greatest potential to propagate and differentiate." *British Journal of Haematology* **107**: 275-281.

Docheva, D., F. Haasters and M. Schieker (2008). "Mesenchymal Stem Cells and Their Cell Surface Receptors." *Current Rheumatology Reviews* **4**(3): 1-6.

Dominici, M., K. Le Blanc, I. Mueller, I. Slaper-Cortenbach, F. Marini, D. Krause, R. Deans, A. Keating, D. Prockop and E. Horwitz (2006). "Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement." *Cytotherapy* **8**(4): 315-317.

Dua, H. and A. Azuara-Blanco (2000). "Limbal stem cells of the corneal epithelium." *Surv Ophthalmol.* **44**(5): 415-425.

Duncan, M. W., R. Aebersold and R. M. Caprioli (2010). "The pros and cons of peptide-centric proteomics." *Nat Biotechnol* **28**(7): 659-664.

Eaton, B. E., L. Gold and D. A. Zichi (1995). "Let's get specific: the relationship between specificity and affinity." *Chemistry & Biology* **2**(10): 633–638.

Eckart, K., M. C. Holthausen, W. Koch and J. Spiess (1998). "Mass spectrometric and quantum mechanical analysis of gas-phase formation, structure, and decomposition of various b2 ions and their specifically deuterated analogs." *Journal of the American Society for Mass Spectrometry* **9**(10): 1002-1011.

Elabd, C., C. Chiellini, A. Massoudi, O. Cochet, L. E. Zaragosi, C. Trojani, J. F. Michiels, P. Weiss, G. Carle, N. Rochet, C. A. Dechesne, G. Ailhaud, C. Dani and

E. Z. Amri (2007). "Human adipose tissue-derived multipotent stem cells differentiate in vitro and in vivo into osteocyte-like cells." *Biochem Biophys Res Commun* **361**(2): 342-348.

Ellington, A. D. and J. W. Szostak (1990). "In vitro selection of RNA molecules that bind specific ligands." *Nature* **346**: 818-822.

Erickson-Johnson, M. R., A. R. Seys, C. W. Roth, A. A. King, R. L. Hulshizer, X. Wang, Y. W. Asmann, R. V. Lloyd, E. K. Jacob and A. M. Oliveira (2009). "Carboxypeptidase M: a biomarker for the discrimination of well-differentiated liposarcoma from lipoma." *Modern Pathology* **22**(12): 1541-1547.

Famulok, M. (1993). "Molecular Recognition of Amino Acids by RNA-aptamers: An l-citrulline binfind RNA motif and its evolution into L-Arginine binder." *Journal of the American Chemical Society* **116**: 1698-1691 1706.

Farrington-Rock, C., N. J. Crofts, M. J. Doherty, B. A. Ashton, C. Griffin-Jones and A. E. Canfield (2004). "Chondrogenic and adipogenic potential of microvascular pericytes." *Circulation* **110**(15): 2226-2232.

Ferrara, N., K. Carver-Moore, H. Chen, M. Dowd, L. Lu, K. S. O'Shea, L. Powell-Braxton, K. J. Hillan and M. W. Moore (1996). "Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene." *Nature* **380**: 439-442.

Foster, L. J., P. A. Zeemann, C. Li, M. Mann, O. N. Jensen and M. Kassem (2005). "Differential expression profiling of membrane proteins by quantitative proteomics in a human mesenchymal stem cell line undergoing osteoblast differentiation." *Stem Cells* **23**(9): 1367-1377.

Friedenstein, A. J., R. K. Chailakhjan and K. S. Lalykina (1970). "The development of fibroblast colonies in monolayer cultures of guines-pig bone marrow and spleen cells." *Cell Proliferation* **3**(4): 393-403.

Friedenstein, A. J., K. V. Petrakova, A. I. Kurolesova and G. P. Frolova (1968). "Heterotopic of bone marrow. Analysis of precursor cells for osteogenic and hematopoietic tissues." *Transplantation* **6**(2): 230-247.

Friedenstein, A. J., I. I. Piatetzky-Shapiro and K. V. Petrakova (1966). "Osteogenesis in transplants of bone marrow cells." *J. Embryol. exp. Morph.,* **16**(3): 381-390.

Gabet, Y. and I. Bab (2011). "Microarchitectural changes in the aging skeleton." *Curr Osteoporos Rep* **9**(4): 177-183.

Gang, E. J., D. Bosnakovski, C. A. Figueiredo, J. W. Visser and R. C. R. Perlingeiro (2007). "SSEA-4 identifies mesenchymal stem cells from bone marrow." *Blood* **109**: 1743-1751.

Ganss, B., R. H. Kim and J. Sodek (1999). "Bone Sialoprotein." *Critical Reviews in Oral Biology & Medicine* **10**(1): 79-98.

Garbis, S., G. Lubec and M. Fountoulakis (2005). "Limitations of current proteomics technologies." *Journal of Chromatography A* **1077**(1): 1-18.

Garbis, S. D., S. I. Tyritzis, T. Roumeliotis, P. Zerefos, E. G. Giannopoulou, A. Vlahou, S. Kossida, J. Diaz, S. Vourekas, C. Tamvakopoulos, K. Pavlakis, D. Sanoudou and C. A. Constantinides (2008). "Search for Potential Markers for Prostate Cancer Diagnosis, Prognosis and Treatment in Clinical Tissue Specimens Using Amine-Specific Isobaric Tagging (iTRAQ) with Two-Dimensional Liquid Chromatography and Tandem Mass Spectrometry." *J Proteome Res* **7**(8): 3146-3158.

Gaur, T., C. J. Lengner, H. Hovhannisyan, R. A. Bhat, P. V. Bodine, B. S. Komm, A. Javed, A. J. van Wijnen, J. L. Stein, G. S. Stein and J. B. Lian (2005). "Canonical WNT signaling promotes osteogenesis by directly stimulating Runx2 gene expression." *J Biol Chem* **280**(39): 33132-33140.

Geiger, T., A. Wehner, C. Schaab, J. Cox and M. Mann (2012). "Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins." *Molecular & Cellular Proteomics* **11**(3).

Gerber, H. P., T. H. Vu, A. M. Ryan, J. Kowalski, Z. Werb and N. Ferrara (1999). "VEGF couples hypertrophic cartilage remodeling, ossification and angiogenesis during endochondral bone formation." *Nature Medicine* **5**(6): 623-628.

Gold, L. and C. Tuerk (1997). Nucleic acid ligands. E. P. Office. **EP 0 533 838 B1:** 1-96.

Gold, L., J. J. Walker, S. K. Wilcox and S. Williams (2012). "Advances in human proteomics at high scale with the SOMAscan proteomics platform." *New Biotechnology* **29**(5): 543-549.

Golub, E. E. and K. Boesze-Battaglia (2007). "The role of alkaline phosphatase in mineralization." *Current opinion in orthopaedics* **18**: 444-448.

Gopinath, S. C. (2007). "Methods developed for SELEX." *Anal Bioanal Chem* **387**(1): 171-182.

Gothard, D., J. Greenhough, E. Ralph and R. O. C. Oreffo (2014). "Prospective isolation of human bone marrow stromal cell subsets: A comparative study between Stro-1-, CD146- and CD105-enriched populations." *Journal of Tissue Engineering* **5**: 1-17.

Graneli, C., A. Thorfve, U. Ruetschi, H. Brisby, P. Thomsen, A. Lindahl and C. Karlsson (2014). "Novel markers of osteogenic and adipogenic differentiation of human bone marrow stromal cells identified using a quantitative proteomics approach." *Stem Cell Research* **12**(1): 153-165.

Gronthos, S., S. E. Graves, S. Ohta and P. J. Simmonds (1994). "The Stro-1+ Fraction of Adult Human Bone Marrow Contains the Osteogenic Precursors." *Blood* **84**(12): 4164-4173.

Gronthos, S., A. C. W. Zannettino, S. J. Hay, S. Shi, S. E. Graves, A. Kortesidis and P. J. Simmons (2003). "Molecular and cellular characterisation of highly purified stromal stem cells derived from human bone marrow." *Journal of Cell Science* **116**: 1827-1835.

Guo, K. T., R. SchAfer, A. Paul, A. Gerber, G. Ziemer and H. P. Wendel (2006). "A new technique for the isolation and surface immobilization of mesenchymal stem cells from whole bone marrow using high-specific DNA aptamers." *Stem Cells* **24**(10): 2220-2231.

Gygi, S. P., B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold (1999). "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." *Nat Biotechnol* **17**: 994-999.

Haga, S. W. and H. F. Wu (2014). "Overview of software options for processing, analysis and interpretation of mass spectrometric proteomic data." *Journal of mass spectrometry* **49**(10): 959-969.

Hammacher, A., K. Mellström, C. Heldin and B. Westermark (1989). "Isoform-specific induction of actin reorganization by platelet-derived growth factor suggests that the functionally active receptor is a dimer." *The Embo journal* **8**(9): 2489-2495.

Harichandan, A. and H. J. Buhring (2011). "Prospective isolation of human MSC." *Best Pract Res Clin Haematol* **24**(1): 25-36.

Harwood, P. J., J. B. Newman and A. L. R. Michael (2010). "An update on fracture healing and non-union." *ORTHOPAEDICS AND TRAUMA* **24**(1): 9-23.

Heo, S. C., W. C. Shin, M. J. Lee, B. R. Kim, I. H. Jang, E. J. Choi, J. S. Lee and J. H. Kim (2015). "Periostin accelerates bone healing mediated by human mesenchymal stem cell-embedded hydroxyapatite/tricalcium phosphate scaffold." *Plos one* **10**(3): 1-17.

Hermann, T. and D. J. Patel (2000). "Biochemistry - Adaptive recognition by nucleic acid aptamers." *Science* **287**(5454): 820-825.

Hicke, B. J., C. Marion, Y. F. Chang, T. Gould, C. K. Lynott, D. Parma, P. G. Schmidt and S. Warren (2001). "Tenascin-C aptamers are generated using tumor cells and purified protein." *J Biol Chem* **276**(52): 48644-48654.

Holley, R. J., G. P. Tai, A. J. K. Williamson, S. Taylor, S. A. Cain, S. M. Richardson, C. L. R. Merry, A. D. Whetton, C. M. Kielty and A. E. Canfield (2015). "Comparative Quantification of the Surfaceome of Human Multipotent Mesenchymal Progenitor Cells." *Stem Cell Reports* **4**(3): 473-488.

Huang, L., C. Niu, B. Willard, W. Zhao, L. Liu, W. He, T. Wu, S. Yang, S. Feng, Y. Mu, L. Zheng and K. Li (2015). "Proteomic analysis of porcine mesenchymal stem cells derived from bone marrow and umbilical cord: implication of the proteins involved in the higher migration capability of bone marrow mesenchymal stem cells." *Stem Cell Research & Therapy* **6**(77).

Ji, Y., J. Ji, F. Sun, Y. Zeng, X. He, J. Zhao, Y. Yu, S. Yu and W. Wu (2010). "Quantitative Proteomics Analysis of Chondrogenic Differentiation of

C3H10T1/2 Mesenchymal Stem Cells by iTRAQ Labeling Coupled with On-line Two-dimensional LC/MS/MS." *Molecular & Cellular Proteomics* **9**(3): 550-564.

Jinwal, U. K., J. Koren, 3rd, S. I. Borysov, A. B. Schmid, J. F. Abisambra, L. J. Blair, A. G. Johnson, J. R. Jones, C. L. Shults, J. C. O'Leary, 3rd, Y. Jin, J. Buchner, M. B. Cox and C. A. Dickey (2010). "The Hsp90 cochaperone, FKBP51, increases Tau stability and polymerizes microtubules." *The Journal of neuroscience* **30**(2): 591-599.

Johnson, I. M., S. G. Kumar and R. Malathi (2003). "RNA binding efficacy of theophylline, theobromine and caffeine." *J Biomol Struct Dyn* **20**(5): 687-692.

Kalli, A., G. T. Smith, M. J. Sweredoski and S. Hess (2013). "Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers." *J Proteome Res*: A-P.

Kapustin, A. N. and C. M. Shanahan (2011). "Osteocalcin: a novel vascular metabolic and osteoinductive factor?" *Arterioscler Thromb Vasc Biol* **31**(10): 2169-2171.

Kassem, M. (2006). "Stem cells: potential therapy for age-related diseases." *Ann N Y Acad Sci* **1067**: 436-442.

Kelly, R. T., A. V. Tolmachev, J. S. Page, K. Tang and R. D. Smith (2010). "The Ion Funnel: Theory, Implementations, and Applications." *Mass Spectrometry Reviews* **29**(2): 294-312.

Kim, H. S., X. Zhang and Y. W. Choi (1994). "Activation and Proliferation of Follicular Dendritic Cell-Like cells by activated T lymphocytes." *The Journal of Immunology* **153**: 2951-2961.

Kimoto, M., R. Yamashige, K. Matsunaga, S. Yokoyama and I. Hirao (2013). "Generation of high-affinity DNA aptamers using an expanded genetic alphabet." *Nat Biotechnol* **31**(5): 453-457.

Kimura, M., T. Nagai, R. Matsushita, A. Hashimoto, T. Miyashita and S. Hirohata (2013). "Role of FK506 binding protein 5 (FKBP5) in osteoclast differentiation." *Modern Rheumatology* **23**(6): 1133-1139.

Kirkham, G. R. and S. H. Cartmell (2007). "Genes and Proteins Involved in the regulation of osteogenesis." *Topics in Tissue Engineering*. N. Ashammakhi, R. Reis and E. Chiellini. **3:** 1-22.

Kolf, C. M., E. Cho and R. S. Tuan (2007). "Mesenchymal stromal cells. Biology of adult mesenchymal stem cells: regulation of niche, self-renewal and differentiation." *Arthritis Res Ther* **9**(204): 1-10.

Komori, T. (2005). "Regulation of skeletal development by the Runx family of transcription factors." *J Cell Biochem* **95**(3): 445-453.

Kong, H. Y. and J. Byun (2013). "Nucleic Acid aptamers: new methods for selection, stabilization, and application in biomedical science." *Biomol Ther (Seoul)* **21**(6): 423-434.

Kouhara, H., Y. R. Hadari, T. Spivak-Kroizman, J. Schilling, D. Bar-Sagi, I. Lax and J. Schlessinger "A Lipid-Anchored Grb2-Binding Protein That Links FGF-Receptor Activation to the Ras/MAPK Signaling Pathway." *Cell* **89**(5): 693-702.

Koyama, N., Y. Okubo, K. Nakao, K. Osawa, K. Fujimura and K. Bessho (2011). "Pluripotency of mesenchymal cells derived from synovial fluid in patients with temporomandibular joint disorder." *Life Sci* **89**(19-20): 741-747.

Krampera, M., M. Franchini, G. Pizzolo and G. Aprili (2007). "Mesenchymal stem cells: from biology to clinical use." *Blood Transfus* **5**(3): 120-129.

Kraus, K. H. and C. Kirker-Head (2006). "Mesenchymal stem cells and bone regeneration." *Vet Surg* **35**(3): 232-242.

Kulasingam, V. and E. P. Diamandis (2008). "Strategies for discovering novel cancer biomarkers through utilization of emerging technologies." *Nat Clin Pract Oncol* **5**(10): 588-599.

Kusser, W. (2000). "Chemically modified nucleic acid aptamers for in vitro selections:evolving evolution." *Reviews in Molecular Biotechnology* **74**: 27-38.

Kuznetsov, S. A., P. H. Krebsbach, K. Satomura, J. Kerr, M. Riminucci, D. Benayahu and P. G. Robey (1997). "Single-colony derived strains of human marrow stromal fibroblast form bone after tranbsplantation in vivo." *Journal of bone and mineral research* **12**(9): 1335-1347.

Laemmli, U. K. (1970). "Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4." *Nature* **227**(5259): 680-685.

Laskar, A., C. N. Mandal and A. Chatterjee (2010). "Protease-Inhibitor Interactions – A structural Insight." *International Journal of Biotechnology and Biochemistry* **6**(2): 231–258.

Lauridsen, L. H. and R. N. Veedu (2012). "Nucleic acid aptamers against biotoxins: a new paradigm toward the treatment and diagnostic approach." *Nucleic Acid Therapeutics* **22**(6): 371-379.

Lecourt, S., J. P. Marolleau, O. Fromigue, K. Vauchez, R. Andriamanalijaona, B. Ternaux, M. N. Lacassagne, I. Robert, K. Boumediene, F. Chereau, P. Marie, J. Larghero, M. Fiszman and J. T. Vilquin (2010). "Characterization of distinct mesenchymal-like cell populations from human skeletal muscle in situ and in vitro." *Exp Cell Res* **316**(15): 2513-2526.

Lee, J.-H. and D. M. Kemp (2006). "Human adipose-derived stem cells display myogenic potential and perturbed function in hypoxic conditions." *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* **341**(3): 882-888.

Lee, K. S., H. J. Kim, Q. L. Li, X. Z. Chi, C. Ueta, T. Komori, J. M. Wozney, E. G. Kim, J. Y. Choi, H. M. Ryoo and S. C. Bae (2000). "Runx2 is a common target of transforming growth factor beta 1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12." *Mol Cell Biol* **20**(23): 8783-8792.

Lee, O. K., T. K. Kuo, W. M. Chen, K. D. Lee, S. L. Hsieh and T. H. Chen (2004). "Isolation of multipotent mesenchymal stem cells from umbilical cord blood." *Blood* **103**(5): 1669-1675.

Li, Y., X. Yu, S. Lin, X. Li, S. Zhang and Y.-H. Song (2007). "Insulin-like growth factor 1 enhances the migratory capacity of mesenchymal stem cells." *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* **356**(3): 780-784.

Liebler, D. C. (2002). "Introduction to proteomics: tools for the new biology." Humana Press.

Lin, G., G. Liu, L. Banie, G. Wang, H. Ning, T. F. Lue and C. S. Lin (2011). "Tissue distribution of mesenchymal stem cell marker Stro-1." *Stem Cells Dev* **20**(10): 1747-1752.

Liu, T. M., M. Martina, D. W. Hutmacher, J. H. Hui, E. H. Lee and B. Lim (2007). "Identification of common pathways mediating differentiation of bone marrow- and adipose tissue-derived human mesenchymal stem cells into three mesenchymal lineages." *Stem Cells* **25**(3): 750-760.

Lo, T., C.-F. Tsai, Y.-R. V. Shih, Y.-T. Wang, S.-C. Lu, T.-Y. Sung, W.-L. Hsu, Y.-J. Chen and O. K. Lee (2012). "Phosphoproteomic Analysis of Human Mesenchymal Stromal Cells during Osteogenic Differentiation." *J Proteome Res* **11**(2): 586-598.

Lu, R., N. F. Neff, S. R. Quake and I. L. Weissman (2011). "Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding." *Nat Biotechnol* **29**(10): 928-933.

Luo, J., J. Chen, Z. L. Deng, X. Luo, W. X. Song, K. A. Sharff, N. Tang, R. C. Haydon, H. H. Luu and T. C. He (2007). "Wnt signaling and human diseases: what are the therapeutic implications?" *Lab Invest* **87**(2): 97-103.

Luo, L., D. Q. Li, A. Doshi, W. Farley, R. M. Corrales and S. C. Pflugfelder (2004). "Experimental dry eye stimulates production of inflammatory cytokines and MMP-9 and activates MAPK signaling pathways on the ocular surface." *Invest Ophthalmol Vis Sci* **45**(12): 4293-4301.

Luria, E. A., A. F. Panasyuk and A. Y. Friedenstein (1971). "Fibroblast Colony Formation from Monolayer Cultures of Blood Cells." *Transfusion* **11**(6): 345-349.

Mackie, E. J., L. Tatarczuch and M. Mirams (2011). "The skeleton: a multi-functional complex organ: the growth plate chondrocyte and endochondral ossification." *J Endocrinol* **211**(2): 109-121.

Majka, S. M., K. A. Jackson, K. A. Kienstra, M. W. Majesky, M. A. Goodell and K. K. Hirschi (2003). "Distinct progenitor populations in skeletal muscle are bone marrow derived and exhibit different cell fates during vascular regeneration." *Journal of Clinical Investigation* **111**(1): 71-79.

Mangelsdorf, D. J., C. Thummel, M. Beato, P. Herrlich, G. Schiitq, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon and R. M. Evan (1995). "The Nuclear Receptor Superfamily - The second decade." *Cell* **83**: 835-839.

Mansukhani, A., P. Bellosta, M. Sahni and C. Basilico (2000). "Signaling by fibroblast growth factors (FGF) and fibroblast growth factor receptor 2 (FGFR2)-activating mutations blocks mineralization and induces apoptosis in osteoblasts." *Journal of Cell Biology* **149**(6): 1297-1308.

Marquez-Curtis, L., A. Jalili, K. Deiteren, N. Shirvaikar, A. M. Lambeir and A. Janowska-Wieczorek (2008). "Carboxypeptidase M expressed by human bone marrow cells cleaves the C-terminal lysine of stromal cell-derived factor-1alpha: another player in hematopoietic stem/progenitor cell mobilization?" *Stem Cells* **26**(5): 1211-1220.

Mason, C. and P. Dunnil (2008). "A brief definition of regenerative medicine." *Regenerative Medicine* **3**(1).

Matsushita, R., A. Hashimoto, T. Tomita, H. Yoshitawa, S. Tanaka, H. Endo and S. Hirohata (2010). "Enhanced expression of mRNA for FK506-binding protein 5 in bone marrow CD34 positive cells in patients with rheumatoid arthritis." *Clin Exp Rheumatol* **28**(1): 87-90.

Maurer, M. H. (2011). "Proteomic definitions of mesenchymal stem cells." *Stem Cells Int* **2011**: 1-9.

McKeague, M. and M. C. Derosa (2012). "Challenges and opportunities for small molecule aptamer development." *J Nucleic Acids*: 748913.

Meirelles Lda, S., A. M. Fontes, D. T. Covas and A. I. Caplan (2009). "Mechanisms involved in the therapeutic properties of mesenchymal stem cells." *Cytokine Growth Factor Rev* **20**(5-6): 419-427.

Mendez-Ferrer, S., T. V. Michurina, F. Ferraro, A. R. Mazloom, B. D. Macarthur, S. A. Lira, D. T. Scadden, A. Ma'ayan, G. N. Enikolopov and P. S. Frenette (2010). "Mesenchymal and haematopoietic stem cells form a unique bone marrow niche." *Nature* **466**(7308): 829-834.

Mindaye, S. T., M. Ra, J. Lo Surdo, S. R. Bauer and M. A. Alterman (2013). "Improved proteomic profiling of the cell surface of culture-expanded human bone marrow multipotent stromal cells." *J Proteomics* **78**: 1-14.

Mindaye, S. T., M. Ra, J. L. Lo Surdo, S. R. Bauer and M. A. Alterman (2013). "Global proteomic signature of undifferentiated human bone marrow stromal cells: Evidence for donor-to-donor proteome heterogeneity." *Stem Cell Research* **11**(2): 793-805.

Moore, K. A. and I. R. Lemischka (2006). "Stem cells and their niches." *Science* **311**(5769): 1880-1885.

Morgan, J. M., A. Wong, C. E. Yellowley and D. C. Genetos (2011). "Regulation of tenascin expression in bone." *J Cell Biochem* **112**(11): 3354-3363.

Nakashima, K., X. Zhou, G. Kunkel, Z. P. Zhang, J. M. Deng, R. R. Behringer and B. de Crombrugghe (2002). "The novel zinc finger-containing transcription factor Osterix is required for osteoblast differentiation and bone formation." *Cell* **108**(1): 17-29.

Ng, E. W., D. T. Shima, P. Calias, E. T. Cunningham, Jr., D. R. Guyer and A. P. Adamis (2006). "Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease." *Nat Rev Drug Discov* **5**(2): 123-132.

Niazi, J. H., S. J. Lee and M. B. Gu (2008). "Single-stranded DNA aptamers specific for antibiotics tetracyclines." *Bioorg Med Chem* **16**(15): 7245-7253.

Niemann, S., C. F. Zhao, F. Pascu, U. Stahl, U. Aulepp, L. Niswander, J. L. Weber and U. Muller (2004). "Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family." *American Journal of Human Genetics* **74**(3): 558-563.

Ning, H., G. Lin, T. F. Lue and C. S. Lin (2011). "Mesenchymal stem cell marker Stro-1 is a 75 kd endothelial antigen." *Biochem Biophys Res Commun* **413**(2): 353-357.

O'Hurley, G., E. Sjöstedt, A. Rahman, B. Li, C. Kampf, F. Pontén, W. M. Gallagher and C. Lindskog (2014). "Garbage in, garbage out: A critical evaluation of strategies used for validation of immunohistochemical biomarkers." *Molecular Oncology* **8**(4): 783-798.

Ohno, I., J. Hashimoto, S. K, K. Takaoka, T. Ochi, K. Matsubara and K. Okubo (1996). "A cDNA Cloning of Human AEBP1 from Primary Cultured Osteoblasts and its Expression in a Differentiating Osteoblastic Cell Line." *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* **228**: 411–414.

Okada, M. (2012). "Regulation of the SRC family kinases by Csk." *International Journal of Biological Sciences* **8**(10): 1385-1397.

Oreffo, R. O., S. Bord and J. T. Triffitt (1998). "Skeletal progenitor cells and ageing human populations." *Clin Sci* **94**(5): 549-555.

Osiak, A., O. Utermohlen, S. Niendorf, I. Horak and K. P. Knobeloch (2005). "ISG15, an interferon-stimulated ubiquitin-like protein, is not essential for STAT1 signaling and responses against vesicular stomatitis and lymphocytic choriomeningitis virus." *Mol Cell Biol* **25**(15): 6338-6345.

Ow, S. Y., M. Salim, J. Noirel, C. Evans, I. Rehman and P. C. Wright (2009). "iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly"." *J Proteome Res* **9**: 5347–5355.

Owen, M. and A. J. Friedenstein (1988). "Stromal stem cells: marrow-derived osteogenic precursors." *Ciba Found Symp* **136**: 42-60.

Owens, K., S. D. Glassman, J. M. Howard, M. Djurasovic, J. L. Witten and L. Y. Carreon (2011). "Perioperative complications with rhBMP-2 in transforaminal lumbar interbody fusion." *European Spine Journal* **20**(4): 612-617.

Ozdemir, T., L. Xu, C. Siedlecki and J. L. Brown (2013). "Substrate curvature sensing through Myosin IIa upregulates early osteogenesis." *Integrative Biology* **5**(11): 1407-1416.

Ozer, A., J. M. Pagano and J. T. Lis (2014). "New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization." *Molecular Therapy-Nucleic Acids* **3**(e183).

Papachristou, E. K., T. I. Roumeliotis, A. Chrysagi, C. Trigoni, E. Charvalos, P. A. Townsend, K. Pavlakis and S. D. Garbis (2013). "The Shotgun Proteomic Study of the Human ThinPrep Cervical Smear Using iTRAQ Mass-Tagging and 2D LC-FT-Orbitrap-MS: The Detection of the Human Papillomavirus at the Protein Level." *J Proteome Res* **12**(5): 2078-2089.

Parisuthiman, D., Y. Mochida, W. R. Duarte and M. Yamauchi (2005). "Biglycan modulates osteoblast differentiation and matrix mineralization." *J Bone Miner Res* **20**(10): 1878-1886.

Park, H. W., J. S. Shin and C. W. Kim (2007). "Proteome of mesenchymal stem cells." *Proteomics* **7**(16): 2881-2894.

Pathan, M., S. Keerthikumar, C. S. Ang, L. Gangoda, C. Y. Quek, N. A. Williamson, D. Mouradov, O. M. Sieber, R. J. Simpson, A. Salim, A. Bacic, A. Hill, D. A. Stroud, M. T. Ryan, J. I. Agbinya, J. M. Mariadasson, A. W. Burgess and S. Mathivanan (2015). "Technical brief funrich: An open access standalone functional enrichment and interaction network analysis tool." *Proteomics*.

Phinney, D. G. (2012). "Functional heterogeneity of mesenchymal stem cells: implications for cell therapy." *J Cell Biochem* **113**(9): 2806-2812.

Phinney, D. G. and D. J. Prockop (2007). "Concise review: mesenchymal stem/multipotent stromal cells: the state of transdifferentiation and modes of tissue repair--current views." *Stem Cells* **25**(11): 2896-2902.

Pittenger, M. F., A. M. Mackay, S. C. Beck, R. K. Jaiswal, R. Douglas, J. D. Mosca, M. A. Moorman, D. W. Simonetti, S. Craig and D. R. Marshak (1999). "Multilineage potential of Adult Human Mesenchymal Stem cells." *Science* **284**: 143-147.

Popova, A. P., P. D. Bozyk, J. K. Bentley, M. J. Linn, A. M. Goldsmith, R. E. Schumacher, G. M. Weiner, A. G. Filbrun and M. B. Hershenson (2010). "Isolation of tracheal aspirate mesenchymal stromal cells predicts bronchopulmonary dysplasia." *Pediatrics* **126**(5): e1127-1133.

Porciani, D., G. Signore, L. Marchetti, P. Mereghetti, R. Nifosi and F. Beltram (2014). "Two Interconvertible Folds Modulate the Activity of a DNA Aptamer Against Transferrin Receptor." *Molecular Therapy-Nucleic Acids* **3**(e144).

Prockop, D. J. (1997). "Marrow Stromal Cells as Stem Cells for Nonhematopoietic Tissues." *Science* **276**(5309): 71-74.

Psaltis, P. J., S. Paton, F. See, A. Arthur, S. Martin, S. Itescu, S. G. Worthley, S. Gronthos and A. C. Zannettino (2010). "Enrichment for STRO-1 expression enhances the cardiovascular paracrine activity of human bone marrow-derived mesenchymal cell populations." *J Cell Physiol* **223**(2): 530-540.

Ramachandran, M. and P. Bates (2007). "Basics of bone." *Basic Orthopaedic Sciences: The Stanmore Guide*. M. Ramachandran. London, Hodder Arnold.

Rawadi, G., B. Vayssiere, F. Dunn, R. Baron and S. Roman-Roman (2003). "BMP-2 controls alkaline phosphatase expression and osteoblast mineralization by a Wnt autocrine loop." *Journal of bone and mineral research* **18**(10): 1842-1853.

Rawlings, N. D., D. P. Tolle and A. J. Barrett (2004). "Evolutionary families of peptidase inhibitors." *Biochem. J.* **378**: 705–716.

Reeves, J. R., J. W. Xuan, K. Arfanis, C. Morin, S. V. Garde, M. T. Riuiz, J. Wisniewski, C. Panchal and J. E. Tanner (2005). "Identification, purification and characterization of a novel human blood protein with binding affinity for prostate secretory protein of 94 amino acids." *Biochem. J.* **385**: 105-114.

Rittling, S. R., H. N. Matsumoto, M. D. McKee, A. Nanci, X. R. An, K. E. Novick, A. J. Kowalski, M. Noda and D. T. Denhardt (1998). "Mice lacking osteopontin show normal development and bone structure but display altered osteoclast formation in vitro." *Journal of bone and mineral research* **13**(7): 1101-1111.

Romano, S., A. Sorrentino, A. L. Di Pace, G. Nappo, C. Mercogliano and M. F. Romano (2011). "The Emerging Role of Large Immunophilin FK506 Binding Protein 51 in Cancer." *Current Medicinal Chemistry* **18**: 5424-5429.

Ross, P. L., Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson and D. J. Pappin (2004). "Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents." *Mol Cell Proteomics* **3**(12): 1154-1169.

Roubelakis, M. G., K. I. Pappa, V. Bitsika, D. Zagoura, A. Vlahou, H. A. Papadaki, A. Antsaklis and N. P. Anagnou (2007). "Molecular and proteomic characterization of human mesenchymal stem cells derived from amniotic fluid: comparison to bone marrow mesenchymal stem cells." *Stem Cells Dev* **16**(6): 931-952.

Roumeliotis, T. I., M. Halabalaki, X. Alexi, D. Ankrett, E. Giannopoulou, A. L. Skaltsounis, B. S. Sayan, M. N. Alexis, P. A. Townsend and S. D. Garbis (2013). "Pharmacoproteomic study of the natural product Ebenfuran III in DU-145 prostate cancer cells: The quantitative and temporal interrogation of chemically induced cell death at the protein level." *J Proteome Res.*

Rucci, N. (2008). "Molecular biology of bone remodelling." *Clinical Cases in Mineral and Bone Metabolism* **5**(1): 49-56.

Sacchetti, B., A. Funari, S. Michienzi, S. Di Cesare, S. Piersanti, I. Saggio, E. Tagliafico, S. Ferrari, P. G. Robey, M. Riminucci and P. Bianco (2007). "Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment." *Cell* **131**(2): 324-336.

Samartzis, D., N. Khanna, F. H. Shen and H. S. An (2005). "Update on bone morphogenetic proteins and their application in spine surgery." *J Am Coll Surg* **200**(2): 236-248.

Sassanfar, M. and J. W. Szostak (1993). "An RNA motif that binds ATP." *Nature* **364**: 550-553.

Savage, J. W., M. P. Kelly, S. A. Ellison and P. A. Anderson (2015). "A population-based review of bone morphogenetic protein: associated complication and reoperation rates after lumbar spinal fusion." *Neurosurgical Focus* **39**(4): E13-E13.

Scheuner, D., C. Eckman, M. Jensen, X. Song, M. Citron, N. Suzuki, T. D. Bird, J. Hardy, M. Hutton, W. Kukull, E. Larson, E. LevyLahad, M. Viitanen, E. Peskind, P. Poorkaj, G. Schellenberg, R. Tanzi, W. Wasco, L. Lannfelt, D. Selkoe and S. Younkin (1996). "Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease." *Nature Medicine* **2**(8): 864-870.

Sekiya, I., M. Ojima, S. Suzuki, M. Yamaga, M. Horie, H. Koga, K. Tsuji, K. Miyaguchi, S. Ogishima, H. Tanaka and T. Muneta (2012). "Human mesenchymal stem cells in synovial fluid increase in the knee with degenerated cartilage and osteoarthritis." *Journal of Orthopaedic Research* **30**(6): 943-949.

Sell, S. (2004). "Stem cell origin of cancer and differentiation therapy." *Crit Rev Oncol Hematol* **51**(1): 1-28.

Seshi, B. (2006). "An integrated approach to mapping the proteome of the human bone marrow stromal cell." *Proteomics* **6**(19): 5169-5182.

Shapiro, F. (2008). "Bone development and its relation to fracture repair. The role of mesenchymal osteoblasts and urface osteoblasts." *European Cells and Materials* **15**: 53-76.

Shepro, D. and N. M. Morel (1993). "Pericyte physiology." *The FASEB Journal* **7**(11): 1031-1038.

Shi, S. and S. Gronthos (2003). "Perivascular Niche of Postnatal Mesenchymal Stem Cells in Human Bone Marrow and Dental Pulp." *Journal of bone and mineral research* **18**(4): 696-704.

Sigal, A., R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, I. Alaluf, N. Swerdlin, N. Perzov, T. Danon, Y. Liron, T. Raveh, A. E. Carpenter, G. Lahav and U. Alon (2006). "Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins." *Nature Methods* **3**(7): 525-531.

Simmons, P. J. and B. Torok-Storb (1991). "Identification of stromal Cell precursors in Human bone marrow by a Novel Monoclonal Antobody, STRO-1." *Blood* **78**: 55-62.

Simonsen, J. L., C. Rosada, N. Serakinci, J. Justesen, K. Stenderup, S. I. Rattan, T. G. Jensen and M. Kassem (2002). "Telomerase expression extends the proliferative life-span and maintains the osteogenic potential of human bone marrow stromal cells." *Nat Biotechnol* **20**(6): 592-596.

Sims, N. A. and T. J. Martin (2014). "Coupling the activities of bone formation and resorption: a multitude of signals within the basic multicellular unit." *BoneKEy Rep* **3**.

Sinclair, D., S. G. Fillman, M. J. Webster and C. S. Weickert (2013). "Dysregulation of glucocorticoid receptor co-factors FKBP5, BAG1 and PTGES3 in prefrontal cortex in psychotic illness." *Scientific Reports* **3**.

Skidgel, R. A., R. M. David and F. Tan (1989). "Human Carboxypeptidase M." *J Biol Chem*: 2236-2241.

Sohni, A. and C. M. Verfaillie (2013). "Mesenchymal stem cells migration homing and tracking." *Stem Cells International* **2013**.

Song, K. M., S. Lee and C. Ban (2012). "Aptamers and their biological applications." *Sensors (Basel)* **12**(1): 612-631.

Srisawat, C., I. J. Goldstein and D. R. Engelke (2001). "Sephadex binding RNA ligands-rapid affinity purification of RNA from complex mixtures." *Nucleic Acids Research* **29**(2): 1-5.

Standal, T., M. Borset and A. Sundan (2004). "Role of osteopontin in adhesion, migration, cell survival and bone remodelling." *Experimental oncology* **26**(3): 179-184.

Stewart, K., S. Walsh, J. Screen, C. M. Jefferiss, J. Chainey, G. R. Jordan and J. N. Beresford (1999). "Further Characterization of Cells Expressing STRO-1 in Cultures of Adult Human Bone Marrow Stromal Cells." *Journal of bone and mineral research* **14**(8): 1345-1355.

Su, X., K. Sun, F. Z. Cui and W. J. Landis (2003). "Organization of apatite crystals in human woven bone." *Bone* **32**(2): 150-162.

Sun, H., X. Zhu, P. Y. Lu, R. R. Rosato, W. Tan and Y. Zu (2014). "Oligonucleotide Aptamers: New Tools for Targeted Cancer Therapy." *Molecular Therapy-Nucleic Acids* **3**(e182).

Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen and C. Von Mering (2015). "STRING v10: protein–protein interaction networks, integrated over the tree of life." *Nucleic Acids Research* **43**(Database issue): D447-D452.

Tamama, K., V. H. Fan, L. G. Griffith, H. C. Blair and A. Wells (2006). "Epidermal growth factor as a candidate for ex vivo expansion of bone marrow-derived mesenchymal stem cells." *Stem Cells* **24**(3): 686-695.

Thisse, B. and C. Thisse (2005). "Functions and regulations of fibroblast growth factor signaling during embryonic development." *Developmental Biology* **287**(2): 390-402.

Troyer, D. L. and M. L. Weiss (2008). "Wharton's jelly-derived cells are a primitive stromal cell population." *Stem Cells* **26**(3): 591-599.

Tu, Q. S., P. Valverde and J. Chen (2006). "Osterix enhances proliferation and osteogenic potential of bone marrow stromal cells." *Biochemical and biophysical reserach communications* **341**(4): 1257-1265.

Tuerk, C. and L. Gold (1990). "Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase." *Science* **249**(4968): 505-510.

Tufvesson, E. and G. Westergren-Thorsson (2003). "Biglycan and decorin induce morphological and cytoskeletal changes involving signalling by the small GTPases RhoA and Rac1 resulting in lung fibroblast migration." *Journal of Cell Science* **116**: 4857-4864.

Visvader, J. E. and G. J. Lindeman (2011). "The unmasking of novel unipotent stem cells in the mammary gland." *Embo Journal* **30**(24): 4858-4859.

Wakitani, S., T. Goto, S. J. Pineda, R. G. Young, J. M. Mansour, A. I. Caplan and V. M. Goldberg (1994). "Mesenchymal cell-based repair of large, full-thickness defects of articular cartilage " *Journal of Bone and Joint Surgery-American Volume* **76A**(4): 579-592.

Walker, K. E. (2008). "*Effects of Isolation Methods on proliferation and GD2 expression by porcine umbilical cords stem cells*". Master of Science, Kansas state university.

Wang, X. H., D. O'Connor, M. Brimmell and G. Packham (2009). "The BAG-1 cochaperone is a negative regulator of p73-dependent transcription." *Br J Cancer* **100**(8): 1347-1357.

Watson, J. T. and O. D. Sparkman (2008). "Inroduction to Mass Spectrometry. Instrumenattion, Applications and Strategies for Data Interpretations." Wiley.

Weigand, J. E. and B. Suess (2009). "Aptamers and riboswitches: perspectives in biotechnology." *Applied Microbiology and Biotechnology* **85**(2): 229-236.

Weissman, I. L. (2000). "Stem Cells: Units of Development, Units of Regeneration and Units in Evolution." *Cell* **100**: 157-168.

Werstuck, G. and M. R. Green (1998). "Controlling Gene Expression in Living Cells Through Small Molecule-RNA Interactions." *Science* **282**(5387): 296-298.

Wetzig, A., A. Alaiya, M. Al-Alwan, C. B. Pradez, M. S. Pulicat, A. Al-Mazrou, Z. Shinwari, G. M. Sleiman, H. Ghebeh, H. Al-Humaidan, A. Gaafar, I. Kanaan and C. Adra (2013). "Differential marker expression by cultures rich in mesenchymal stem cells." *Bmc Cell Biology* **14**.

Wilkins, M. R., C. Pasquali, R. D. Appel, K. Ou, O. Golaz, J. C. Sanchez, J. X. Yan, A. A. Gooley, G. Hughes, I. Humphery-Smith, K. L. Williams and D. F. Hochstrasser (1996). "From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis." *Biotechnology (NY)* **14**(1): 61-65.

Wilson, C. and J. W. Szostak (1998). "Isolation of a fluorophore-specific DNA aptamer with weak redox activity." *Chemistry & Biology* **5**(11): 609-617.

Wozney, J. M., V. Rosen, A. J. Celeste, L. M. Mitsock, M. J. Whitters, R. W. Kriz, R. M. Hewick and E. A. Wang (1988). "Novel regulators of bone formation:molecular clones and activities." *Science* **242**: 1528-1534.

Xiang, D., S. Shigdar, G. Qiao, T. Wang, A. Z. Kouzani, S. F. Zhou, L. Kong, Y. Li, C. Pu and W. Duan (2015). "Nucleic acid aptamer-guided cancer therapeutics and diagnostics: the next generation of cancer medicine." *Theranostics* **5**(1): 23-42.

Yamamoto, K., N. Ohga, Y. Hida, N. Maishi, T. Kawamoto, K. Kitayama, K. Akiyama, T. Osawa, M. Kondoh, K. Matsuda, Y. Onodera, M. Fujie, K. Kaga, S. Hirano, N. Shinohara, M. Shindoh and K. Hida (2012). "Biglycan is a specific marker and an autocrine angiogenic factor of tumour endothelial cells." *Br J Cancer* **106**(6): 1214-1223.

Yao, C., T. Zhu, Y. Qi, Y. Zhao, H. Xia and W. Fu (2010). "Development of a quartz crystal microbalance biosensor with aptamers as bio-recognition element." *Sensors (Basel)* **10**(6): 5859-5871.

Yarden, Y., J. A. Escobedo, W.-J. Kuang, T. L. Yang-Feng, T. O. Daniel, P. M. Tremble, E. Y. Chen, M. E. Ando, R. N. Harkins, U. Francke, V. A. Fried, A. Ullrich and L. T. Williams (1986). "Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors." *Nature* **323**: 226-232.

Youdim, M. B. H., D. Edmondson and K. F. Tipton (2006). "The therapeutic potential of monoamine oxidase inhibitors." *Nature Reviews Neuroscience* **7**(4): 295-309.

Young, M. F., J. M. Kerr, K. Ibaraki, A. Heegaard and P. M. Robey (1992). "Structure, expression, and regulation of the major noncollagenous matrix proteins of bone." *Clinical Orthopedics and related research* **281**: 275-293.

Zhou, G., Q. P. Zheng, F. Engin, E. Munivez, Y. Q. Chen, E. Sebald, D. Krakow and B. Lee (2006). "Dominance of SOX9 function over RUNX2 during skeletogenesis." *Proc Natl Acad Sci U S A* **103**(50): 19004-19009.

Zhou, W., L. A. Liotta and E. F. Petricoin (2015). "Cancer metabolism and mass spectrometry-based proteomics." *Cancer Letters* **356**(2, Part A): 176-183.

Zimbres, F. M., A. Tárnok, H. Ulrich and C. Wrenger (2013). "Aptamers: Novel Molecules as Diagnostic Markers in Bacterial and Viral Infections?" *BioMed Research International*.

9. Appendix 1

9.1 ALP staining of 3 patients SSC Stro-1 sorted and unsorted cells.

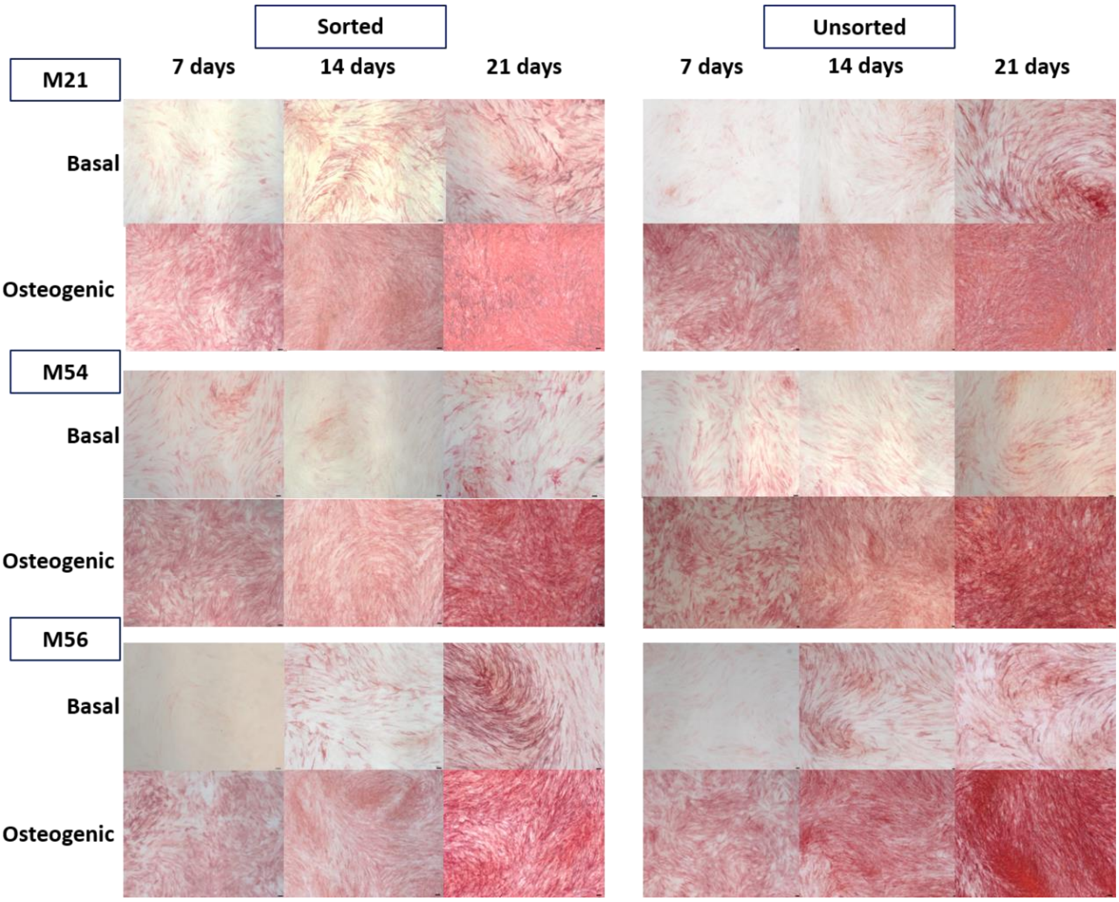

Sorted
Unsorted
7 days
14 days
21 days
7 days
14 days
21 days
M21
Basal
Osteogenic
M54
Basal
Osteogenic
M56
Basal
Osteogenic

Figure 81: ALP staining panel of 3 patients Stro-1 sorted and unsorted SSC.

Cells were extracted from the bone marrow of three patients M21, M54 and M56. Cells were cultured in basal and osteogenic media for 21 days. ALP staining was performed as per 3.3.1.2.

9.2 Flow cytometry

9.2.1 Sample preparation

1,000,000 cells were fixed with 4% PFA in PBS for 20 min, centrifuged at 5,000 rpm at 4 °C for 5 min, washed in PBS twice and re-suspended in PBS. Cells were permeabilised and blocked in blocking buffer (3% triton, 5% goat serum and 1% BSA in PBS) for 15 min, then washed once with PBS and resuspended in primary antibodies:

- Stro-1 hybridoma supernatant undiluted (in-house produced);
- Stro-1 hybridoma supernatant diluted 1:100 (in-house produced);
- Stro-1 diluted 1:100 (R&D Systems);
- IgM Lambda isotype (Sigma) diluted 1:100;
- No antibody control

Cells were incubated for 1 h at 4 °C. Cells were washed 3 times in PBS/tween and incubated with fluorescent secondary antibody specific for IgM (1:200 Alexa Fluor® 488 goat α-mouse IgM), for 1 h at 4 °C. At the end of the incubation, cells were washed 3 times in PBS/tween and the fluorescence was measured with a flow cytometry device Guava easyCyte 8HT Benchtop (Millipore).

9.2.2 Flow cytometry measurements

Flow cytometry was employed with non-adherent cell populations. The GuavaExpressPro (version 4.2.1) software program was used for sample data acquisition and analysis. Gating strategy and protocol setup was the same for all samples. 5,000 events were collected for each reading. When acquiring data, a gate was used to identify the cell population of interest based on side versus forward scatter. The cell population gate and Alexa Fluor® fluorescence calibration settings were optimised with the isotype control of each cell type. Every sample was analysed in triplicates.

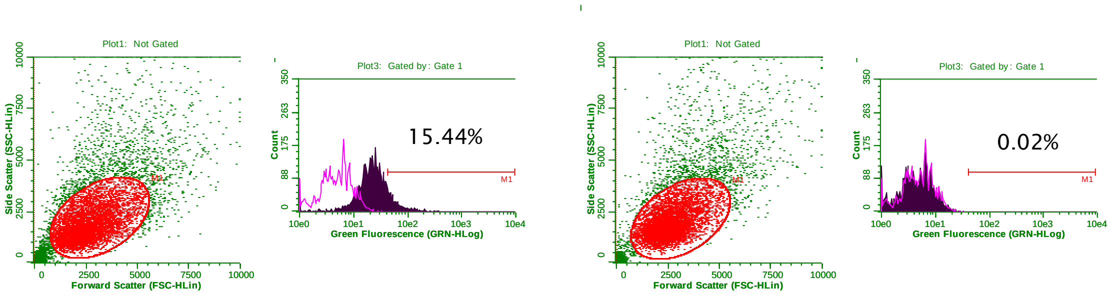
9.2.3 Results

Raji, 721.221, DU145, PC3 and C28I2 cells were examined for their reactivity with the Stro-1 antibody by flow cytometry. An exceedingly wide level of binding (from 15.44% of Raji to 98.10% of 721.221) was detected when the antibody produced in-house was used undiluted. The binding detected by flow cytometry was revealed to be an artefact due to cell debris when the same samples were imaged with a fluorescent microscope (Figure 87 and Figure 91). Hence results indicate cells were consistently negative for the Stro-1 antibody.

Raji

IgM isotype overlaid to Stro-1 Hybridoma supernatant undiluted

IgM isotype overlaid to sample with no primary antibody



IgM isotype overlaid to Stro-1 Hybridoma supernatant 1:100 dilution

IgM isotype overlaid to Stro-1 R&D systems 1:100 dilution

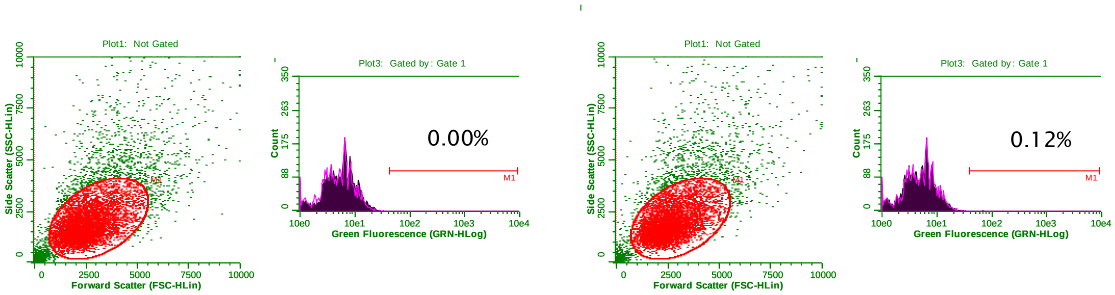


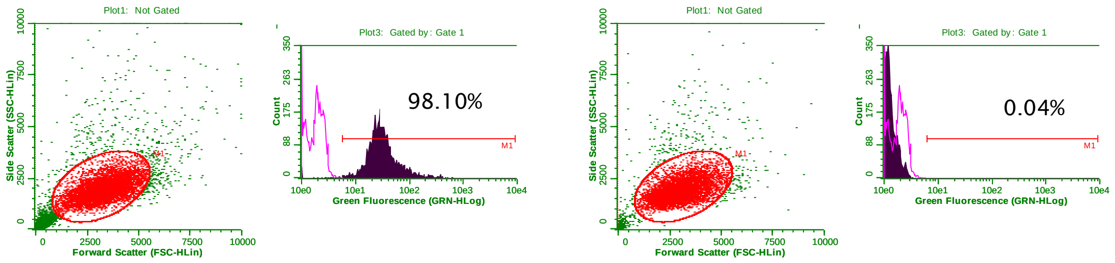
Figure 82: Stro-1 expression in Raji cells analysed by flow cytometry.

Raji cells were cultured under appropriate conditions (3.3.2.1), harvested and stained with Stro-1 antibody from hybridoma supernatant, neat and diluted 1:100. Commercial Stro-1 antibody sourced by R&D was used 1:100 to stain the cells. Isotype and secondary antibody controls were also included in the experiment. The dot plot (left) shows the distribution of the cell population and the gate shows the cell population considered in the analysis. The histogram (right) shows the Stro-1 expression (purple) normalised to the IgM isotype (pink). The percentage above the M1 marker indicates the percentage of positive cells.

721.221

IgM isotype overlaid to Stro-1 Hybridoma supernatant undiluted

IgM isotype overlaid to sample with no primary antibody



IgM isotype overlaid to Stro-1 Hybridoma supernatant 1:100 dilution

IgM isotype overlaid to Stro-1 R&D systems 1:100 dilution

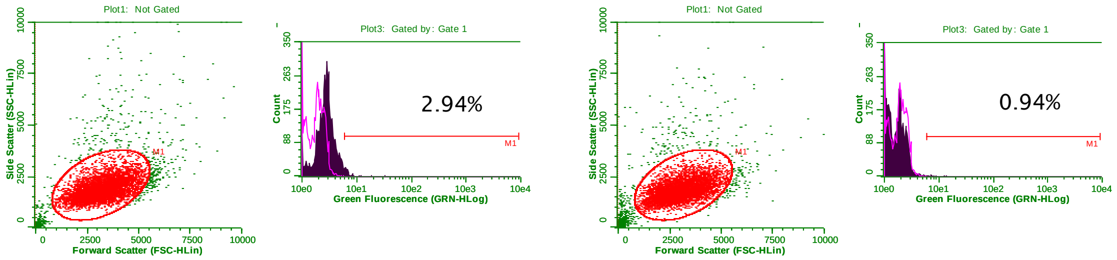


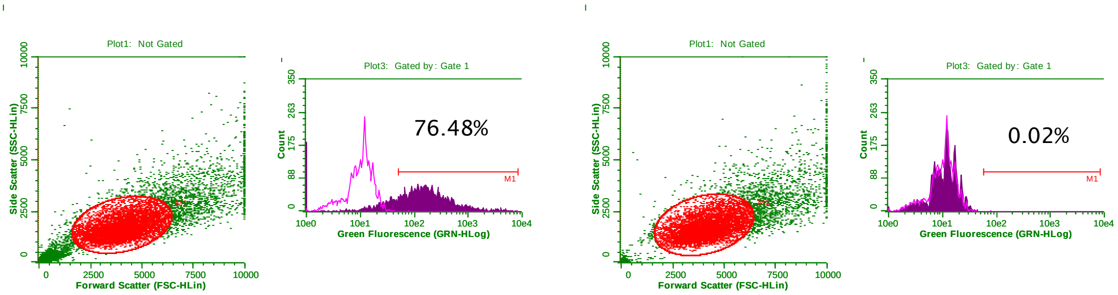
Figure 83: Stro-1 expression in 721.221 cells analysed by flow cytometry

721.221 cells were cultured under appropriate conditions (3.3.2.1), harvested and stained with Stro-1 antibody from hybridoma supernatant, neat and diluted 1:100. Commercial Stro-1 antibody sourced by R&D was used 1:100 to stain the cells. Isotype and secondary antibody controls were also included in the experiment. The dot plot (left) shows the distribution of the cell population and the gate shows the cell population considered in the analysis. The histogram (right) shows the Stro-1 expression (purple) normalised to the IgM isotype (pink). The percentage above the M1 marker indicates the percentage of positive cells.

DU145

IgM isotype overlaid to Stro-1 Hybridoma supernatant undiluted

IgM isotype overlaid to sample with no primary antibody



IgM isotype overlaid to Stro-1 Hybridoma supernatant 1:100 dilution

IgM isotype overlaid to Stro-1 R&D systems 1:100 dilution

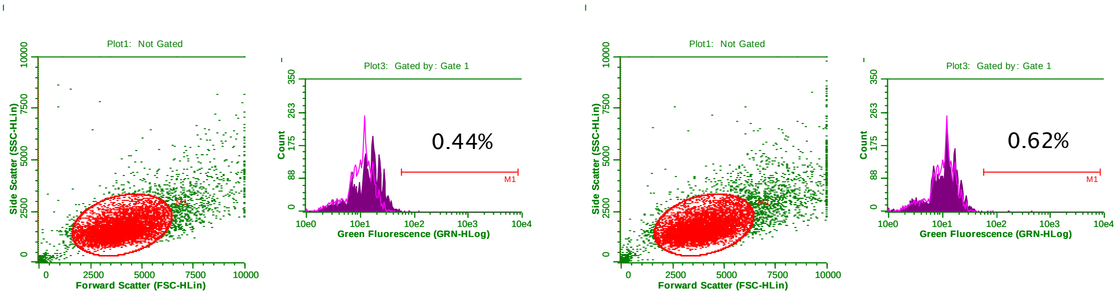


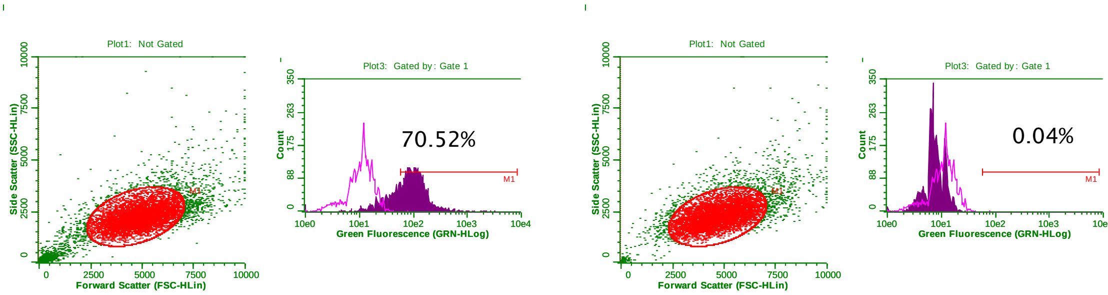
Figure 84: Stro-1 expression in DU145 cells analysed by flow cytometry

DU145 cells were cultured under appropriate conditions (3.3.2.1), harvested and stained with Stro-1 antibody from hybridoma supernatant, neat and diluted 1:100. Commercial Stro-1 antibody sourced by R&D was used 1:100 to stain the cells. Isotype and secondary antibody controls were also included in the experiment. The dot plot (left) shows the distribution of the cell population and the gate shows the cell population considered in the analysis. The histogram (right) shows the Stro-1 expression (purple) normalised to the IgM isotype (pink). The percentage above the M1 marker indicates the percentage of positive cells.

PC3

IgM isotype overlaid to Stro-1 Hybridoma supernatant undiluted

IgM isotype overlaid to sample with no primary antibody



IgM isotype overlaid to Stro-1 Hybridoma supernatant 1:100 dilution

IgM isotype overlaid to Stro-1 R&D systems 1:100 dilution

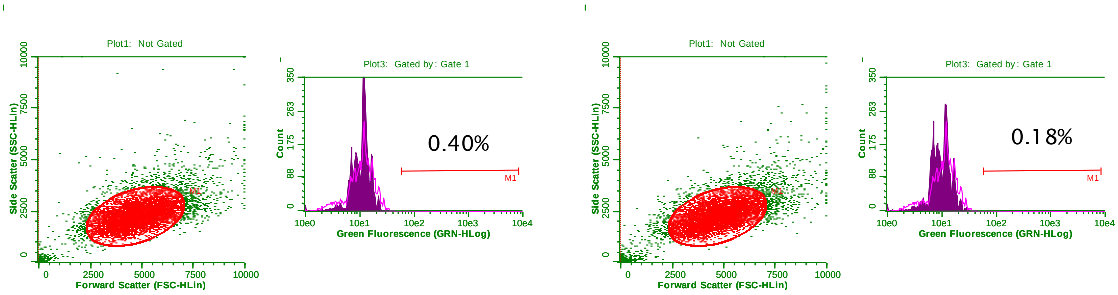


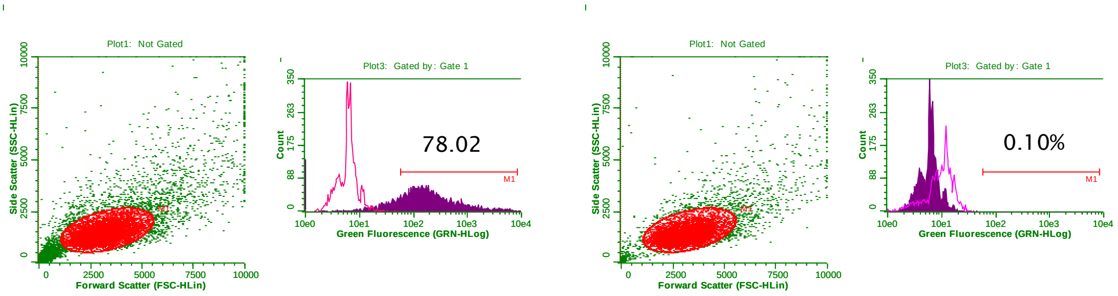
Figure 85: Stro-1 expression in PC3 cells analysed by flow cytometry.

PC3 cells were cultured under appropriate conditions (3.3.2.1), harvested and stained with Stro-1 antibody from hybridoma supernatant, neat and diluted 1:100. Commercial Stro-1 antibody sourced by R&D was used 1:100 to stain the cells. Isotype and secondary antibody controls were also included in the experiment. The dot plot (left) shows the distribution of the cell population and the gate shows the cell population considered in the analysis. The histogram (right) shows the Stro-1 expression (purple) normalised to the IgM isotype (pink). The percentage above the M1 marker indicates the percentage of positive cells.

C28I2

IgM isotype overlaid to Stro-1 Hybridoma supernatant undiluted

IgM isotype overlaid to sample with no primary antibody



IgM isotype overlaid to Stro-1 Hybridoma supernatant 1:100 dilution

IgM isotype overlaid to Stro-1 R&D systems 1:100 dilution

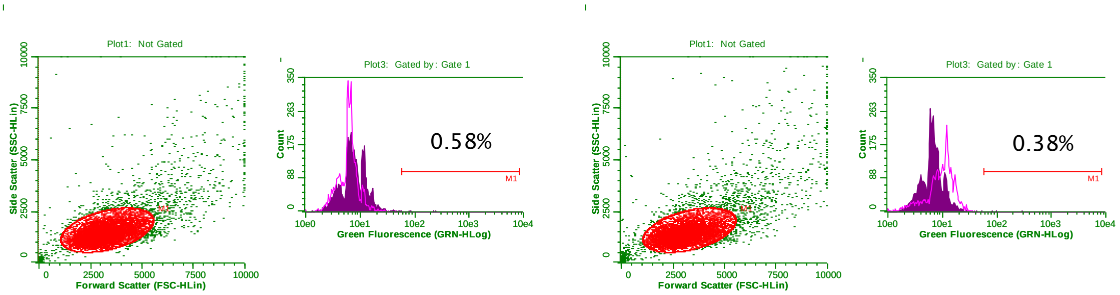


Figure 86: Stro-1 expression in C28I2 cells analysed by flow cytometry.

C28I2 were cultured under appropriate conditions (3.3.2.1), harvested and stained with Stro-1 antibody from hybridoma supernatant, neat and diluted 1:100. Commercial Stro-1 antibody sourced by R&D was used 1:100 to stain the cells. Isotype and secondary antibody controls were also included in the experiment. The dot plot (left) shows the distribution of the cell population and the gate shows the cell population considered in the analysis. The histogram (right) shows the Stro-1 expression (purple) normalised to the IgM isotype (pink). The percentage above the M1 marker indicates the percentage of positive cells.

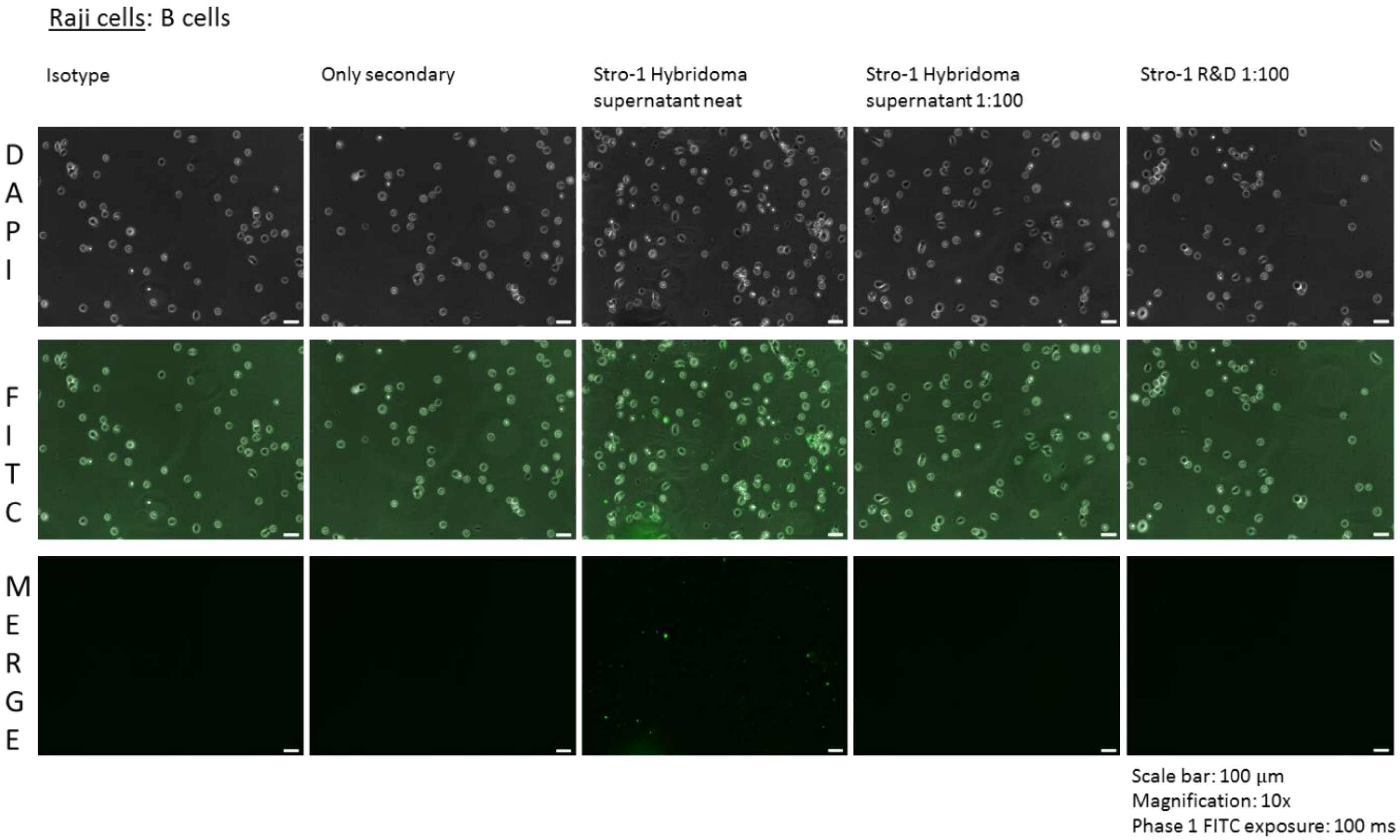

Raji cells: B cells
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
DAPI
FITC
MERGE
Scale bar: 100 µm
Magnification: 10x
Phase 1 FITC exposure: 100 ms

Figure 87: Raji cells immunostaining for Stro-1.

Stro-1 expression (green) in suspension cells previously analysed with flow cytometry. An IgM isotype and a test with just secondary antibody was run as a negative control. Stro-1 hybridoma supernatant (produced in-house) was used undiluted and diluted 1 in 100. Stro-1 antibody sourced from R&D systems was used diluted 1 in 100. Scale bar: 100 µm.

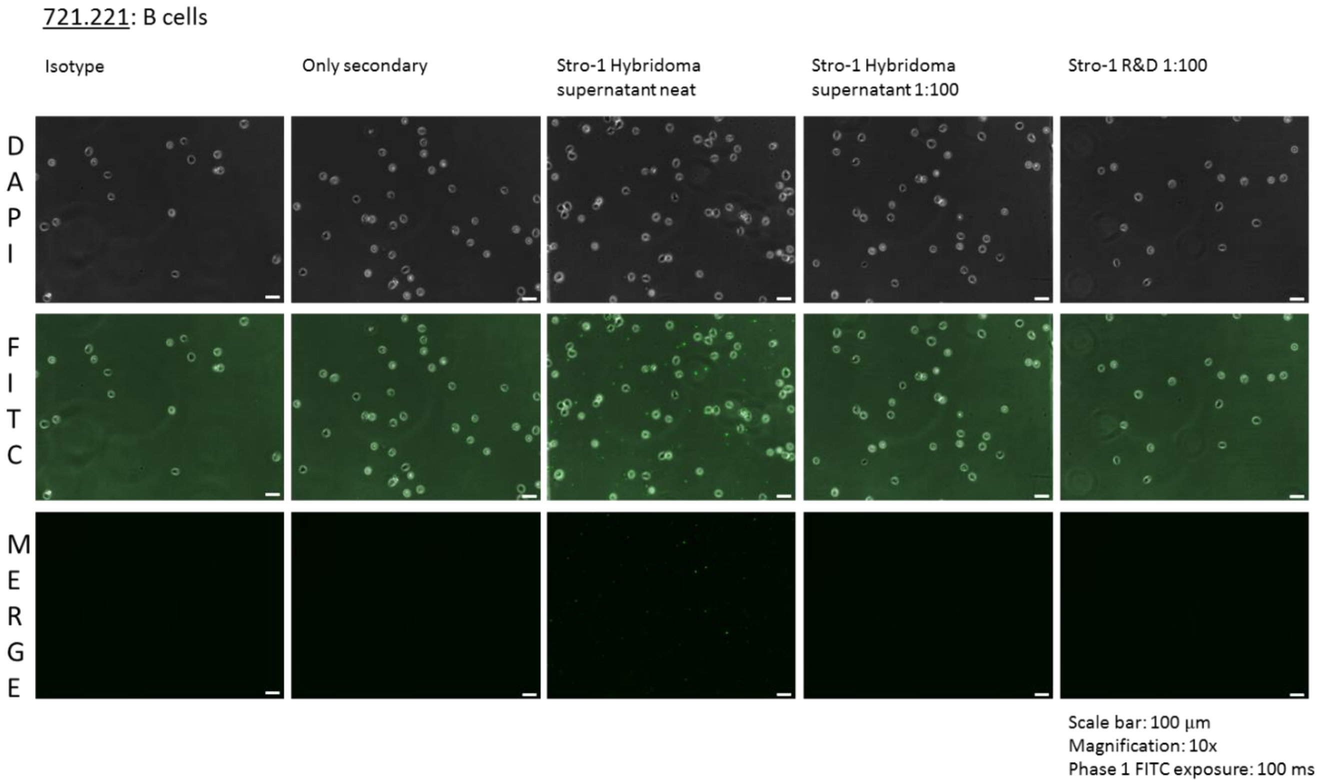

721.221: B cells
Isotype
Only secondary
Stro-1 Hybridoma supernatant neat
Stro-1 Hybridoma supernatant 1:100
Stro-1 R&D 1:100
DAPI
FITC
MERGE
Scale bar: 100 µm
Magnification: 10x
Phase 1 FITC exposure: 100 ms

Figure 88: 721.221 cells immunostaining for Stro-1.

Stro-1 expression (green) in suspension cells previously analysed with flow cytometry. An IgM isotype and a test with just secondary antibody was run as a negative control. Stro-1 hybridoma supernatant (produced in-house) was used undiluted and diluted 1 in 100. Stro-1 antibody sourced from R&D systems was used diluted 1 in 100. Scale bar: 100 µm.

9.3 Cell fractionation

A specific cell fractionation by solvent extraction was attempted with the aim to reduce the complexity of the protein and verify which cell fraction would react to the Stro-1 antibody. HK cells were fractionated using the ProteoExtract® Subcellular Proteome Extraction Kit (Millipore) according to manufacturer instructions. HK cells were used at 80% confluence in a T75. Media was removed and cell were washed twice in 3 mL of PBS. Cells were incubated twice with 2mL of wash buffer for 5 min. Subsequently, 500 μL of buffer I, II, III and IV were added to the cell surface once at a time with a cocktail of protease inhibitors (Sigma), incubated for the time stated in Table 19. After each buffer incubation the supernatant was collected in a clean tube and analysed by WB.

Table 19: Subcellular proteome extraction buffers.

Buffer	Incubation conditions	Cell compartment extraction
I	10 min at 4 °C with gentle agitation	Cytosolic proteins
II	30 min at 4 °C with gentle agitation	Membrane/organelle
III	10 min at 4 °C with gentle agitation	Nuclear proteins
IV	10 min at 4 °C with gentle agitation	Cytoskeletal matrix

Subsequentialy each proteic fraction was tested for reactivity to Stro-1 by WB with the aim to link cellular localisation of Stro-1 with the 50 kDa findings. The cell fractions showing higher reactivity to Stro-1 were fractions 2 and 4, respectively the membrane fraction and the cytoskeletal fraction (Figure 89). A less weak reactivity was also shown in the nucleic protein fraction whilst no reactivity was present in the cytosolic protein extraction.

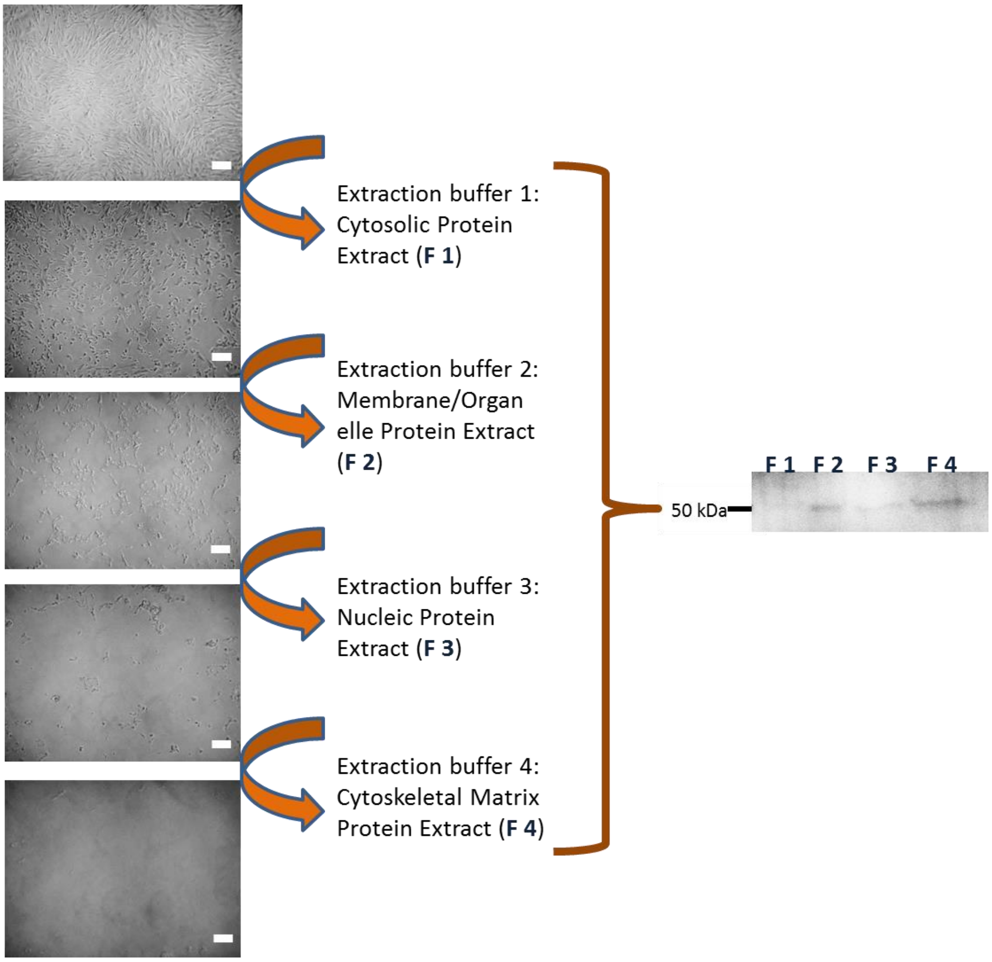

Extraction buffer 1: Cytosolic Protein Extract (F 1)
Extraction buffer 2: Membrane/Organelle Protein Extract (F 2)
Extraction buffer 3: Nucleic Protein Extract (F 3)
Extraction buffer 4: Cytoskeletal Matrix Protein Extract (F 4)
F 1
F 2
F 3
F 4
50 kDa

Figure 89: HK cells fractionation by solvent extraction.

Different buffers from the ProteoExtract® Subcellular Proteome Extraction Kit (Millipore) were used to extract different compartments of the HK cells. 50 µg of protein deriving from the cell fractionation was loaded onto a gel. F1) Cytosolic protein Extract; F2) Membrane/Organelles Protein extract; F3) Nucleic protein extract; F4) Cytoskeletal Matrix Protein. After gel electrophoresis, proteins were transferred onto a PVDF membrane which was probed for the Stro-1 antibody and then re-probed with α-mouse IgM secondary antibody. WB was imaged with the ECL method. Bands appeared are at 50 kDa. Scale bar: 250 µm.

9.4 MACS IP

Immunoprecipitation (IP) was attempted employing the MACS protocol for cell separation. The protocol followed was the same as described in section 2.2, although cell lysate rather than whole cells was employed. The protocol steps is graphically illustrated in Figure 90. The WB results indicated that Stro-1 reacted at 50 kDa with HK membrane fraction, WCL, UNS and to a lesser degree to the SOL fraction of HK cells. MACS IP shows reactivity at 75 and 50 kDa. Considering that the Stro-1 antibody exhibits the same reactivity as demonstrated in Figure 33, concluding that immunoprecipitation has been successful is not possible.

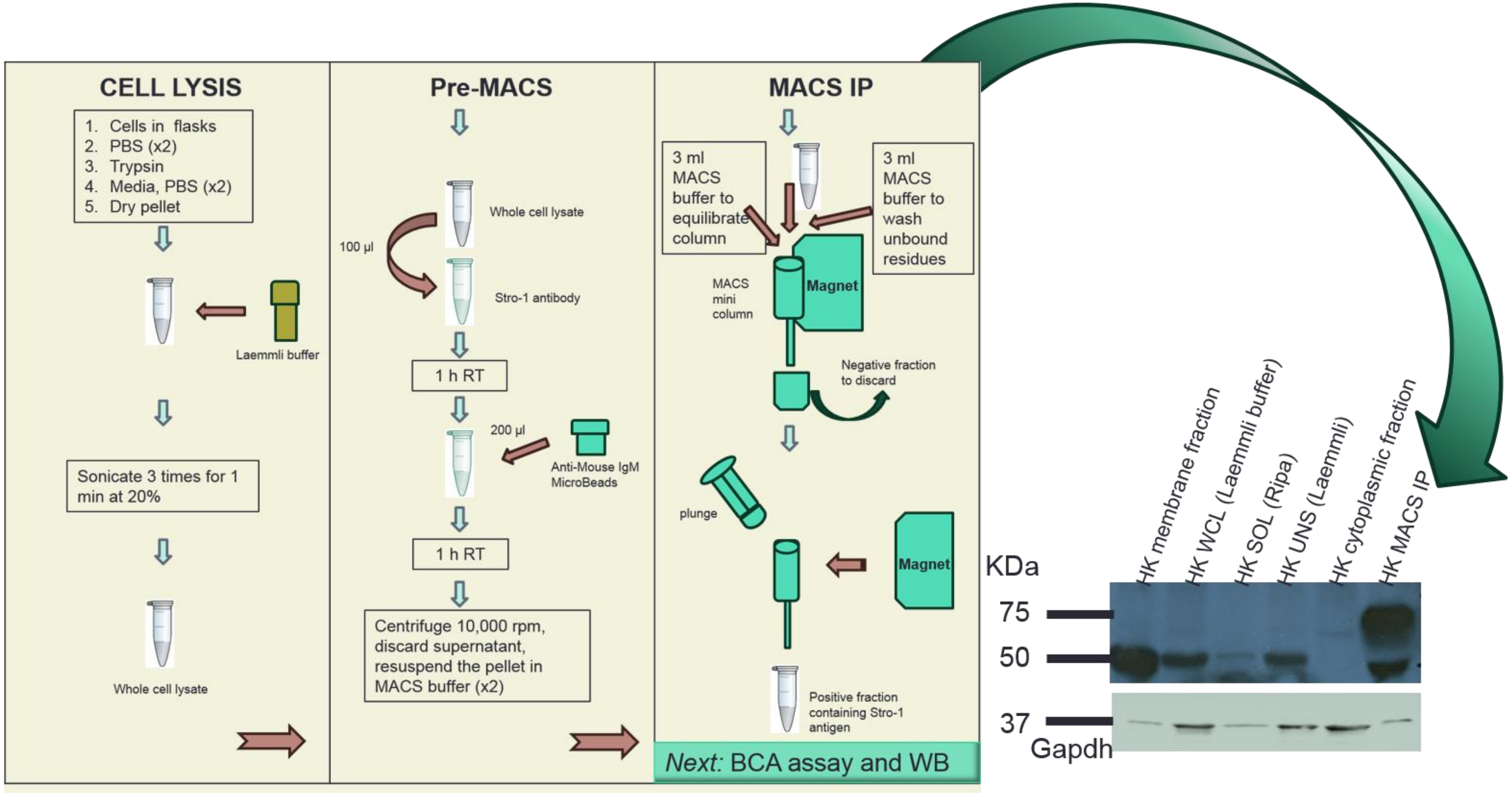

CELL LYSIS
1. Cells in flasks
2. PBS (x2)
3. Trypsin
4. Media, PBS (x2)
5. Dry pellet
Laemmli buffer
Sonicate 3 times for 1 min at 20%
Whole cell lysate
Pre-MACS
Whole cell lysate
100 µl
Stro-1 antibody
1 h RT
200 µl
Anti-Mouse IgM MicroBeads
1 h RT
Centrifuge 10,000 rpm, discard supernatant, resuspend the pellet in MACS buffer (x2)
MACS IP
3 ml MACS buffer to equilibrate column
3 ml MACS buffer to wash unbound residues
MACS mini column
Magnet
Negative fraction to discard
plunge
Magnet
Positive fraction containing Stro-1 antigen
Next: BCA assay and WB
KDa
75
50
37
Gapdh
HK membrane fraction
HK WCL (Laemmli buffer)
HK SOL (Ripa)
HK UNS (Laemmli)
HK cytoplasmic fraction
HK MACS IP

Figure 90: MACS IP and WB.

The diagram graphically indicates the steps followed during the MACS IP. The WB on the right was performed by loading 100 µg of proteins, Stro-1 hybridoma supernatant was employed neat as primary antibody and α-mouse HRP conjugated secondary antibody (Sigma) was employed (1:40,000). Luminata crescendo (Millipore) was employed as HRP substrate. Image were developed on film (Sigma) for 20s with 2 layers of film.

9.5 Blue native page (BN-PAGE)

Stro-1 hybridoma supernatant was centrifuged for 10 min at 13,000 rpm and concentrated in Vivaspin 500 MWCO 3000 (GE Healthcare) for 10 min at 13,000 rpm. 10 µL of the concentrated supernatant was loaded onto a HYDRAGEL PROTEIN(E) precast gel (Sebia) and run for 40 min employing the HYDRASYS instrument (Sebia). This methodology is specific for the separation of serum proteins contained in the sample (the Stro-1 antibody is collected from hybridoma cell culture supernatant which contains FCS, see section 2.5). The gel was oven dried at 100 ˚C and stained for 4 min in amidoblack Sebia stain. This method was kindly run by Mrs Christine Penfold, tenovus building of the University of Southampton under the assistance of the author.

9.6 WB Band Analysis (documentation given by Dr. Omar Jallow, St. Georges University of London)

Proteomics Results Report

Protein digestion: Protein bands were digested with trypsin using a digesting robotic system (Projester, Genomic Solutions). In brief, gel bands were excised from the gel and placed in a microtitre tray, the Projester was set to first wash and shrink the gel plugs using 50mM Ammonium Bicarbonate (Ambic) and acetonitrile (ACN). The proteins were then reduced with DTT and alkylated using IAA. Digestion was performed using trypsin solution (Sigma sequencing grade; 20ug resuspended in 540ul 2mM HCL/60ul ACN). The peptide containing supernatant was pooled into an additional microtitre tray and digested peptides in the gel extracted with 10% formic acid and pooled with the initial supernatant. After this, samples were frozen (-80oC) freeze dried and then resuspended with 25ul of 5% ACN/ 0.1% formic acid for MS analysis. A gel plug containing BSA was also digested alongside the experimental samples as a system control.

LC/MS/MS analysis: Peptides were analysed by LC/MS/MS using a Surveyor LC system and LCQ Deca XP Plus (ThermoScientific). Briefly, peptides were resolved by reverse phase chromatography (Biobasic column, ThermoScientific; 180uM x 15mm) over a 30min ACN gradient at a flow rate of 2ul/min. Peptides were ionised by electrospray ionisation and MS/MS was acquired on ions dependant on their charge state and intensity. Quality control checks for the optimal performance of the instrumentation were in place. Mass accuracy and sensitivity of the MS was confirmed with the direct infusion of glufibrinopeptide (2.5pmoles/ul) and LC/MS/MS performance was assessed with a digest of BSA. Sensitivity, retention time, peptides identified and protein sequence coverage were all within the specified ranges. BSA quality control checks were performed prior to the analysis of the samples and post acquisition.

Data processing: The data files (.raw) were converted into mascot generic files using the MassMatrix File Conversion Tool (Version 2.0; http://www.massmatrix.net) for input into the Mascot searching algorithm (Matrix Science). The data files were searched against NCBInr (v. 20080527) with human taxonomy using the following search criteria: tryptic peptides with up to one missed cleavage and carbamidomethylation of cysteines and oxidation of methionines, which were set as variable modifications. A cut off score of 34 was used all proteins that had a score less than 100 were manually verified.

Results: Proteins identified from the gel bands are listed in the Results Table. The proteins identified in each sample are ranked with the protein with the highest Mascot score listed first (protein hit No.) for each sample. Results table not attached.

9.7 Solubility test with different lysis buffers

A variety of lysis buffers, detailed in Table 20 were adopted to verify the shift in protein expression between the WCL, SOL and INS fractions for the 50 kDa band identified. The buffers solubilisation strength included harsh (Laemmli), semi-harsh (Ripa), mild (Pierce IP), and very mild (Hmken) (Figure 91). It was shown that enrichment of the INS fraction was obtained with the Ripa, Pierce IP and Hmken buffer (Figure 92). In contrast, for each of the three fractions lysed with Laemmli buffer, the same reactivity to the antibody was established. Laemmli buffer solubilises the entirety of the proteins in the cell.

Table 20: Lysis buffers.

Lyses buffers were manually prepared with the exception of the Pierce immunoprecipitation buffer which was purchased.

Buffer ID	**Composition**
Hmken (Wang, O'Connor *et al.*, 2009).	HEPES (4-(2-hydroxy-ethyl)-piperazine-1-ethane-sulphonic acid) pH 7.2, 10 mM, MgCl2 5 mM, KCl 142 mM, EGTA (ethylene glycol-bis (2-aminoethyl)-N,N,N',N'-tetraacetic acid) 2 mM, and nonidet P40 0.2% (v/v).
Pierce immunoprecipitation (IP) (Pierce Biotechnology Inc.).	0.025 M Tris, 0.15 M NaCl, 0.001M EDTA, and nonidet P40 1% (v/v), 5% glycerol, pH7.4.
Ripa (Luo, Li *et al.*, 2004).	50 mM Tris HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate and 0.1 % of sodium dodecyl sulphate (SDS).
Laemmli(Laemmli, 1970).	2% (w/v) SDS, 10% glycerol, 63 mM Tris HCl, pH 6.8.

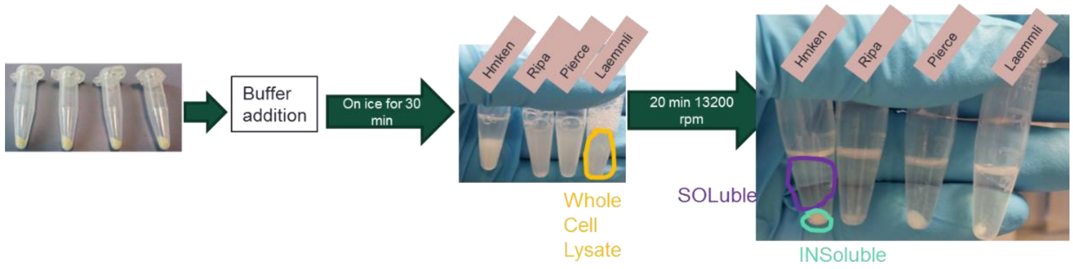

Buffer addition
On ice for 30 min
Hmken
Ripa
Pierce
Laemmli
Whole Cell Lysate
20 min 13200 rpm
Hmken
Ripa
Pierce
Laemmli
SOLuble
INSoluble

Figure 91: HK cells solubility under different lysis buffers.

HK cells were lysed with different buffers, harsh (Laemmli), semi-harsh (Ripa), mild (Pierce IP) and very mild (Hmken) solubilisation. After lysates were centrifuged for 30 min the size of the pellet correlated with the solubilisation strength of each buffer.

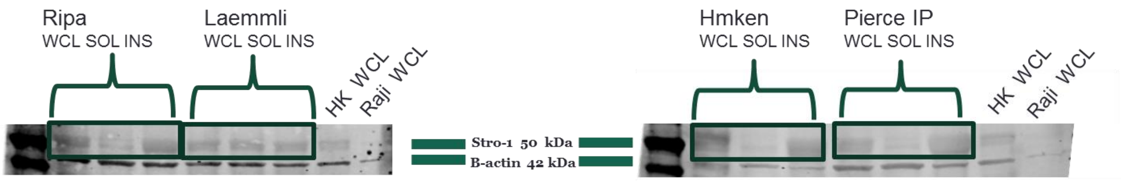

Ripa
WCL SOL INS
Laemmli
WCL SOL INS
HK WCL
Raji WCL
Stro-1 50 kDa
B-actin 42 kDa
Hmken
WCL SOL INS
Pierce IP
WCL SOL INS
HK WCL
Raji WCL

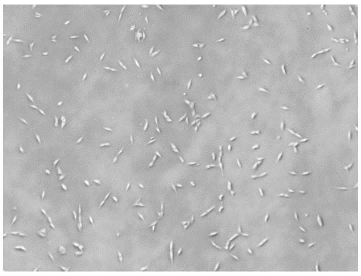
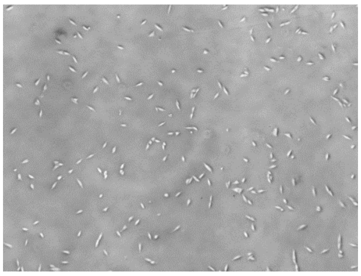
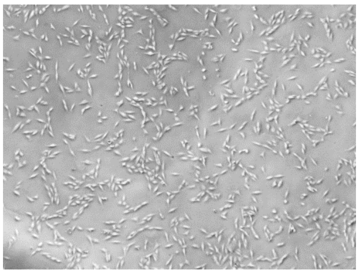
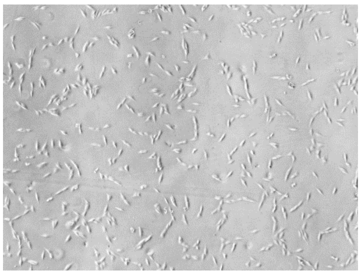
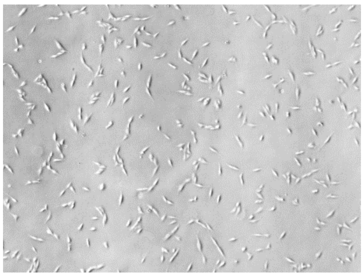
Figure 92: WB with different lysis buffers.

Ripa, Laemmli, Hmken and Pierce IP buffers were used to lyse HK cells soluble (SOL) and insoluble (INS) fractions. 50 µg of protein were loaded onto a gel and probed with Stro-1 antibody overnight and then re-probed with IRDye 680 Goat α-mouse IgM (µ chain specific) secondary antibody (1:30,000). β-actin was used as a loading control. HK and Raji WCL were employed as a positive and a negative control respectively. WBs were imaged with an Odyssey fluorescent imaging unit.

9.8 Preliminary experiments for the aptamer process

Cell seeding density was crucial for the aptamer SELEX process. Cells were required to be at the same density prior to the selection to ensure that the ratio of aptamer molecules to cells was constant throughout the process. Furthermore, the cell employed, MG63 were known to change morphology, hence surface markers if confluent. On this basis, few preliminary experiments were run to gain a better understanding of cell seeding number and cell density after a period of time.

Table 21: Cell density preliminary study.

	Cells image	**Seeded**	**Counted after 20h incubation**	**% cells adhered in 20h**	**Seeding density (cells/cm^2)**
1		5,0000 (1 well of 6 well plate)	30,750	61.5%	5,263.158
2		54,000 (1 well of 6 well plate)	29,500	54.6%	5,684.211
3		216,000 (1 well of 6 well plate)	185,000	85.6%	22,736.84
4		360,000 (T25)	213,000	59.2%	1,4400
5		720,000 (T75)	472,000	65.6%	9,600

10.Appendix 2

The following set of experiments served the purpose to validate the methods underlying mass spectrometry analysis of SSC extracted from human bone marrow. These parameters (cell count, protein quantification and lysis buffer efficiency) were investigated as critical for the mass spectrometry sample preparation therefore a control over these values was sought.

10.1 Protein concentration correlates with cell number

For the mass spectrometry study it was of primary importance to determine the protein concentration of the samples prepared. In this instance an experiment was run with the purpose of gaining an understanding of the protein content at distinct cell densities. For this experiment C2C12 cells were used. The experiment confirmed a linear correlation between cell number and protein content. To obtain 200 µg of proteins from C2C12 cells, 1,070,000 of cells were needed. This number is known to differ between cell types and is correlated to the efficacy of the cell harvest and protein extraction method.

A similar experimental design was used to calculate the correlation between two different lysis buffers.

10.2 Correlation between lysis buffers

Following confirmation of the cell number working range, the two lysis buffers: TEAB and Ripa (9.7) were compared. TEAB buffer was used in mass spectrometry sample preparation protocol while RIPA buffer is widely employed for WB. From the experimental results it was observed that there were no major differences between the two buffers.

10.2.1 Guava cell count versus haemocytometer count

Two different cell counting methods were examined. The aim of this experiment was to guarantee that the counting methods used were consistent and comparable. The importance of this study relied on the fact that protein content is related to cell number. C2C12 cells in PBS solution were used in this experiment. Cells were initially counted with a flow cytometry device, Guava easyCyte 8HT Benchtop (Millipore) and then counted using a disposable haemocytometer chamber (FastRead Counting Slides; ISL, Immune Systems Ltd.). The experiment was run in triplicate. From the results it was deduced that both counting methodologies were consistent within the experimental range under study. Results are shown in Figure 93.

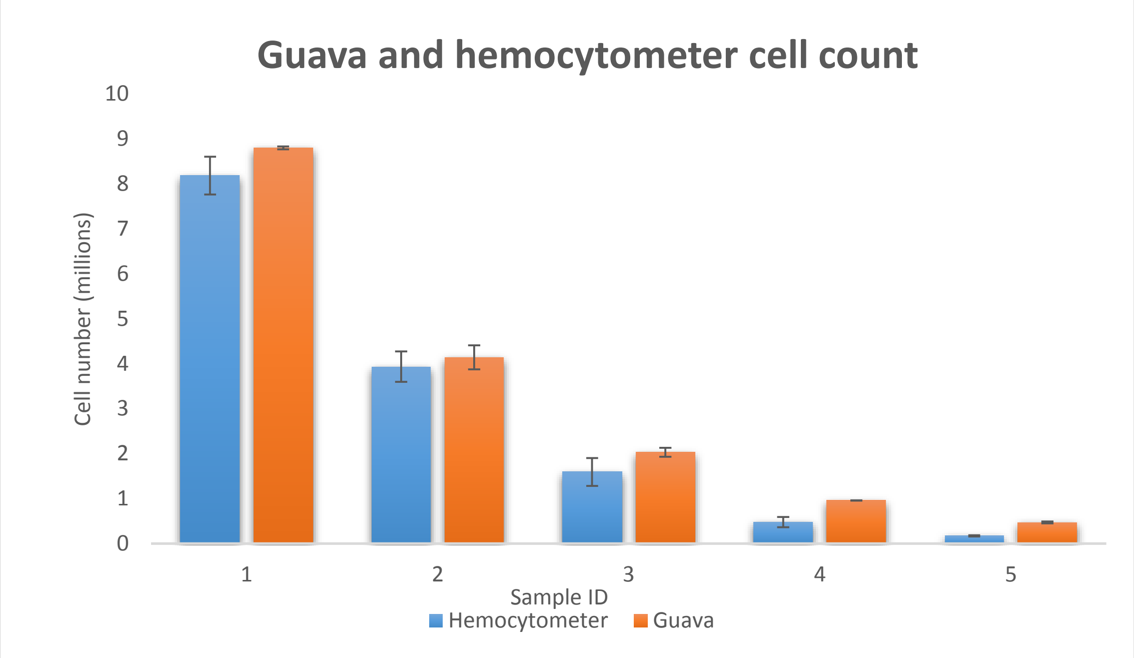


Figure 93: Flow cytometry cell counts against manual cell count with a haemocytometer.

5 different C2C12 cell concentrations were counted automatically by a flow cytometry device and manually with a disposable haemocytometer chamber. Data presented as mean ± standard deviation (SD); n=3.

10.2.2 Protein quantification according to cell number

Prior to sample preparation for mass spectrometry, the protein amount extracted from the cell samples was confirmed. 100 µg of total protein was the required amount for the mass spectrometry studies (100 µg of proteins is the maximum binding capacity of the iTRAQ tags as per manufacturer's instructions). For this reason different cells concentration were prepared and proteins extracted. The experiment was performed with C2C12 cells. Tissue culture flasks containing C2C12 cells were washed 3 times in PBS, and then cells were removed by trypsin treatment and filtered using a 70 µm filter to provide a single cell suspension. Cells were resuspended in PBS and cell number was quantified using a flow cytometry device Guava easyCyte 8HT Benchtop (Millipore). Cells were tested for their protein content using a BCA assay. The experiment was run in triplicate. Results are reported below in Figure 94.

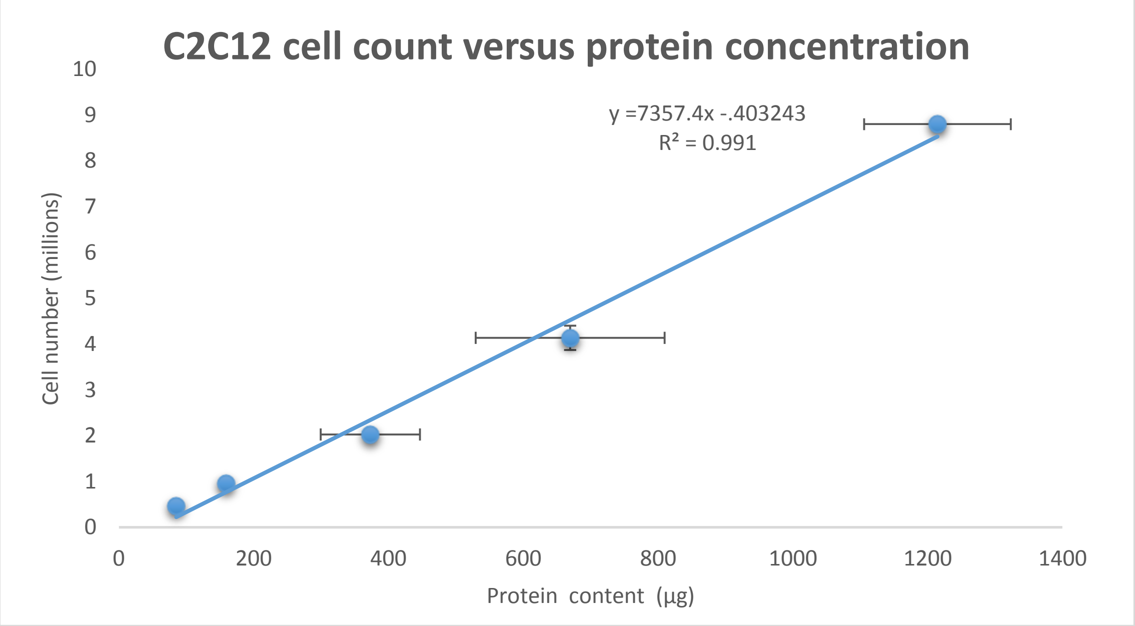


Figure 94: C2C12 protein concentration at different cell densities.

C2C12 cells were harvested and counted with a flow cytometry device (Guava easyCyte 8HT Benchtop, Millipore) then each sample was diluted to obtain different cell concentrations. Each different cell concentration was then pelleted and lysed using RIPA buffer. Protein quantification was obtained with the use of Pierce BCA assay kit. Data presented as mean ± standard deviation (SD); n=3.

10.2.3 Correlation between lysis buffers

As a pre-validation work for the MS project, the efficacy of the TEAB (Table 22) lysis buffer used in the mass spectrometry procedure in comparison to RIPA (Table 22) buffer was examined. SSCs were grown as described in the above protocol and trypsinised after 3 PBS washes. Cells were filtered using a 70 µm filter to provide a single cell suspension. Cells were counted with a flow Cytometry device (Guava easyCyte 8HT Benchtop, Millipore) then pelleted. Cell pellets were lysed with RIPA and TEAB buffer independently, incubated on ice for 30 min and sonicated 3 times for 30 s with a tip sonicator at an amplification of 20%. BCA assay protocol detailed in section 2.8 was performed.

Table 22: Description of adopted lysis buffers for protein determination.

Each lysis buffer reagent is listed in the buffer composition column. According to the experiment protocol, the buffer chosen was added to a cell pellet and incubated on ice for 30 min.

Buffer abbreviated name	**Buffer extended name**	**Buffer composition**
RIPA (Luo, Li *et al.*, 2004)	RIPA	0.75 M NaCl (Fisher Scientific), 5% v/v IgePal CA_{630} (Sigma) 2.5% w/v sodium deoxycholate (Sigma), 2.5% w/v SDS (Fisher Scientific), 0.25 M Tris (pH 8)
TEAB (Roumeliotis, Halabalaki *et al.*, 2013)	Triethylammonium bicarbonate	0.5 M triethylammonium bicarbonate

10.2.4 Correlation between lysis buffers

Two different lysis buffers were examined in this experiment: TEAB and RIPA. Both buffers were used to lyse different concentration of SSCs. The results indicate that both buffers successfully lyse the cells (Figure 95).

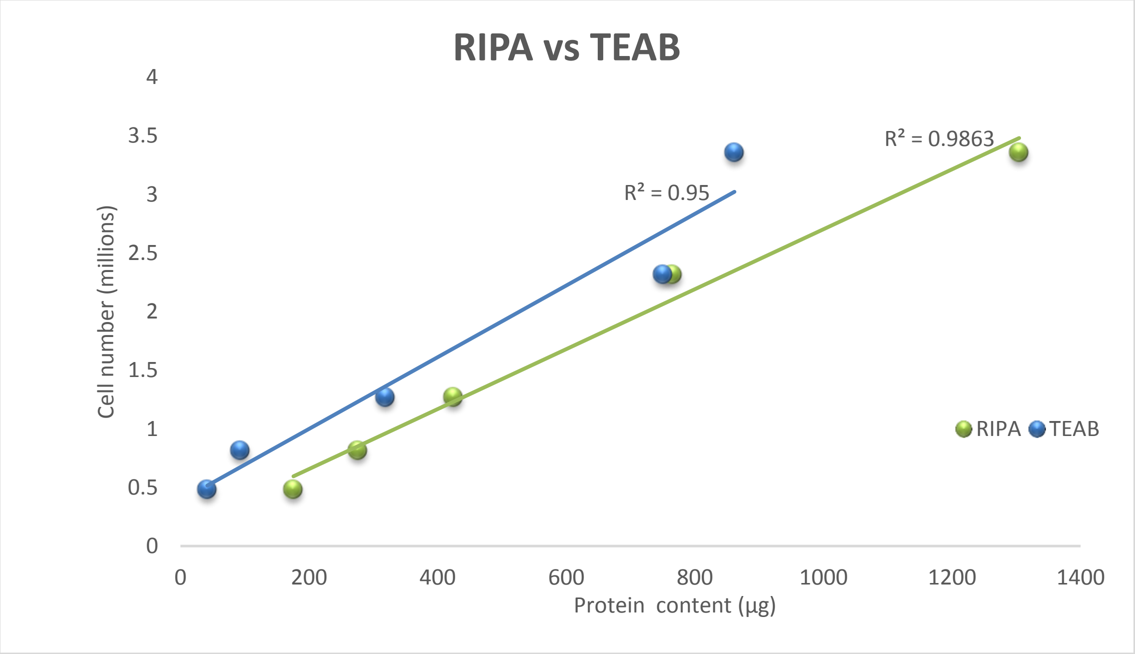


Figure 95: Comparisons of proteins extracted from hBM cells using two different lysis buffers.

5 different cell concentration samples were made in duplicates and one duplicate was lysed with RIPA buffer and the other duplicate was lysed with TEAB buffer. Cells were counted with Guava easyCyte 8HT Benchtop flow cytometry device (Millipore) prior to lysis. Proteins were quantified with Pierce BCA assay kit. Data presented as mean ± standard deviation (SD); n=3.

10.3 Separation of the Stro-1 bright fraction by flow cytometry

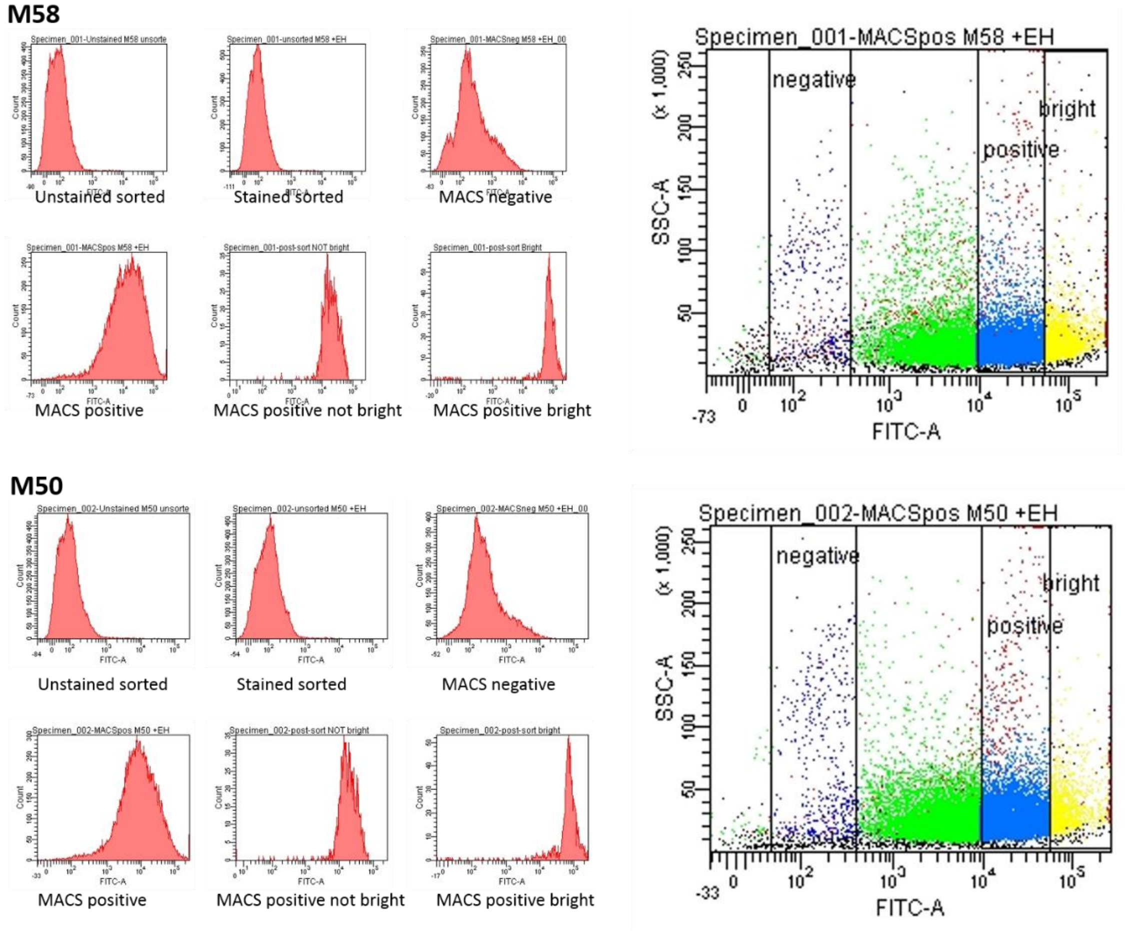


Figure 96: Stro-1 bright cells separation.

Freshly isolated SSCs were separated by flow cytometry according to their expression of the Stro-1 antigen.

The 5% of cells which fluoresced the most for Stro-1 in the MACS positive, the MACS negative and the MACS positive not bright fractions were collected in different tubes. The aim of this experiment was to run MS analysis on freshly isolated *i.e ex vivo* cells to circumvent the unknown effect on protein expression of *in vitro* expansion. When this experiment was planned and run, there was a plan to develop a MS analysis based on small protein amount (small cell number). This method was never developed and the samples were never analysed. This results show how cells could be freshly isolated for further studies.