

**Deconstructing the soft tissue
sarcoma matrixome and adhesome for
drug target and biomarker discovery**

Valeriya Pankova

A thesis submitted for the degree of
Doctor of Philosophy

October 2023

The Institute of Cancer Research
University of London

Declaration

The work presented in this thesis was completed under the supervision of Dr. Paul Huang in the Molecular and Systems Oncology team at the Institute of Cancer Research, London, United Kingdom.

I, Valeriya Pankova, confirm that the work presented within this thesis is my own. Information that has been derived from other sources is clearly indicated within the thesis.

Signature:

A handwritten signature in black ink, appearing to read 'Valeriya Pankova', written in a cursive style.

Abstract

Soft tissue sarcomas (STS) are a rare and diverse group of mesenchymal malignancies that present significant challenges in clinical management due to their extensive clinical and biological heterogeneity. STS have a dismal prognosis, and an incomplete understanding of the underlying biology of these tumours impedes progress in clinical management. Previous research efforts have primarily focused on the molecular characteristics of STS tumour cells, leaving a critical gap in our knowledge regarding the role of the tumour microenvironment (TME), specifically the extracellular matrix (ECM), in STS pathobiology. This thesis aimed to address this knowledge gap by leveraging a recent large-scale proteomic study of STS to comprehensively analyse the composition of the ECM and its associated integrin adhesion signalling across multiple STS subtypes. The findings revealed substantial intra-subtype heterogeneity in matrix signalling among different STS subtypes, such as leiomyosarcoma (LMS), dedifferentiated liposarcoma (DDLPS), and undifferentiated pleomorphic sarcoma (UPS). This matrix signalling heterogeneity was associated with some of the clinical diversity observed within these STS subtypes. Throughout the analysis, the study identified both subtype-agnostic and subtype-specific prognostic biomarkers, shedding light on potential targets for improving clinical outcomes. To address the lack of STS-specific preclinical models of the ECM, the thesis introduced a workflow for generating and characterising patient-derived ECM. Additionally, the analysis uncovered zyxin as a previously unrecognised protein implicated in LMS pathogenesis. In summary, this thesis contributes to our understanding of STS pathobiology by identifying several ECM-related prognostic biomarkers with potential for future development and suggesting putative anti-stroma therapy targets tailored to specific STS subtypes. These insights can potentially guide the development of more effective treatment strategies for STS patients and advance our overall understanding of these complex and heterogeneous malignancies.

Acknowledgements

My PhD journey has been both a challenging and rewarding experience, made possible by the support and encouragement of many people.

Firstly, I would like to thank my primary supervisor, Dr Paul Huang, for providing me with the opportunity to work on this project and for his continuous guidance and encouragement throughout the project. Thank you for always being reliable, enthusiastic and excited about the project's directions and motivating me on challenging days.

Additionally, I want to express my appreciation to my past and present colleagues within the Molecular and Systems Oncology at the ICR. Their support, scientific advice and personal encouragement have been important to me during my project.

Outside of work, I would like to thank all my friends who supported and encouraged me throughout my PhD. I cannot find enough words to express my deep gratitude to my fiancé, Daniele, for his extraordinary patience and unwavering support. Thank you for looking after me, encouraging me when I felt overwhelmed and doubted myself, and keeping me on track during writing.

Last but not least, I would like to thank my family, whose belief in me, love and support were invaluable and motivated me.

Table of Contents

Declaration	2
Abstract	3
Acknowledgements.....	4
Table of Contents	5
List of Figures	12
List of Supplementary Figures	15
List of Tables	17
List of Supplementary Tables	19
Abbreviations and acronyms.....	22
CHAPTER 1 INTRODUCTION	25
1.1.Soft tissue sarcoma	25
1.1.1. Overview	25
1.1.2. Development and aetiology.....	26
1.1.3. Risk factors	26
1.1.3.1. Ionising radiation	26
1.1.3.2. Family history	27
1.1.3.3. Viral infection	27
1.1.4. Classification	28
1.1.4.1. Histology-based classification.....	28
1.1.4.2. Genetic classification	28
1.2.Diagnosis and prognosis of STS	29
1.2.1. Diagnosis	29
1.2.2. Prognostication	29
1.2.2.1. Grading systems	30
1.2.2.2. Prognosis prediction tools: Staging and Nomograms.....	32
1.3.Clinical and biological overview of select STS subtypes	34
1.3.1. Leiomyosarcoma	34

1.3.1.1. Tumour biology.....	35
1.3.2. Dedifferentiated Liposarcoma	38
1.3.2.1. Tumour biology.....	39
1.3.3. Undifferentiated Pleomorphic Sarcoma	40
1.3.3.1. Tumour biology.....	41
1.4.Tumour microenvironment	43
1.4.1. Immune TME in STS	44
1.4.2. The extracellular matrix	47
1.4.2.1. Matrisome definition.....	49
Collagens.....	50
Glycoproteins.....	53
Proteoglycans.....	54
ECM-affiliated factors	55
ECM regulators	56
Secreted factors	57
1.4.2.2. Basement membrane	58
1.4.2.3. Matrisome prognostic scores	60
1.5.Integrin signalling.....	66
1.5.1. Adhesome definition	71
1.6.Matrisome and adhesome components in STS.....	72
1.6.1. ECM in STS.....	72
1.6.1.1. Mass spectrometry-based analysis.....	77
1.6.1.2. Transcriptomic studies	77
1.6.1.3. IHC and IF studies	78
1.6.2. Integrins in STS.....	81
1.7.Preclinical models mimicking tumour ECM.....	85
1.7.1. Matrix material commonly used in cancer research.....	85
1.7.2. Decellularised ECM scaffolds and tissue-derived ECM hydrogels for cancer research.....	87
1.7.2.1. Decellularisation methods overview	88
1.7.3. Tissue-derived ECM hydrogels	90
1.7.3.1. Pepsin-based preparation of ECM hydrogel.....	92
1.8.Hypothesis and aims	92
 CHAPTER 2 MATERIALS AND METHODS.....	 94

2.1. Bioinformatics and statistical methods	94
2.1.1. Patient cohort and clinical data	94
2.1.2. Proteomic data filtering and renormalisation	94
2.1.3. Clustering	95
2.1.4. Differential expression analysis	96
2.1.5. Overrepresentation analysis	96
2.1.6. Survival analyses	96
2.1.7. Matrix scores	97
2.1.8. Matrisome and adhesome networks correlation analysis	97
2.1.9. Protein-protein interaction network analysis	98
2.1.10. Statistical analyses	98
2.2. Immunohistochemistry for tumour infiltrating lymphocytes	98
2.3. Generation of preclinical ECM models	100
2.3.1. Generation of decellularised ECM scaffolds	100
2.3.2. ECM hydrogel generation	101
2.4. Characterisation of preclinical ECM models	101
2.4.1. Protein digestion and sample preparation for mass spectrometry (frozen tumours, scaffolds and hydrogels)	101
2.4.2. Liquid chromatography and mass spectrometry (frozen tumours, scaffolds and hydrogels)	102
2.5. Cell line maintenance	103
2.6. Molecular biology techniques	103
2.6.1. Short hairpin RNA (shRNA) oligo design and cloning	103
2.6.2. Bacterial transformation	104
2.6.3. Plasmid DNA preparation	105
2.6.4. Lentiviral plasmid preparation, transfection and transduction	105
2.7. Phenotypic assays	106
2.7.1. ECM-coated plate preparation	106
2.7.2. Cell adhesion assay	106
2.7.3. Growth curve assay	106
2.7.4. Colony formation assay	107
2.7.5. Dose response proliferation assay	107
2.7.6. Small molecule inhibitor screen	108
2.7.7. Live single-cell imaging	109

2.8. Protein analysis	112
2.8.1. Immunoblotting	112
CHAPTER 3 OVERVIEW OF STS MATRISOME AND ADHESOME	114
3.1. Background and objectives	114
3.2. Results	115
3.2.1. Cohort characteristics	115
3.2.2. Matrisome and adhesome overview in STS	117
3.2.3. Matrisome and adhesome networks in STS.....	126
3.2.4. Identification of a prognostic matrix score for STS	134
3.2.4.1. Combination of proteoglycan score with TILs counts	138
CD3+ TILs.....	138
CD4+ TILs.....	140
CD8+ TILs.....	142
3.3. Discussion	146
3.3.1. Matrisome and adhesome overview in STS	146
3.3.2. Proteoglycan score in STS	150
3.3.3. Tumour infiltrating lymphocytes in STS	151
3.4. Supplementary Material	153
3.4.1. Supplementary Figures	153
3.4.2. Supplementary Tables.....	160
CHAPTER 4 UNDERSTANDING HETEROGENEITY IN MATRIX SIGNALLING IN LMS, DDLPS AND UPS	169
4.1. Background and objectives	169
4.2. Results	170
4.2.1. Intra-subtype heterogeneity in LMS	170
4.2.1.1. Clinicopathological features of LMS cohort.....	170
4.2.1.2. Identification of LMS subtypes	171
4.2.1.3. Biological characterisation of LMS subgroups.....	172
Characterisation of ECM and integrin adhesion networks in LMS subgroups.....	172
Characterisation of functional proteomic networks in LMS subgroups	176
Clinical characterisation of LMS subgroups	178
4.2.2. Intra-subtype heterogeneity in DDLPS.....	179

4.2.2.1.	Clinicopathological features of DDLPS cohort	179
4.2.2.2.	Identification DDLPS subgroups	180
4.2.2.3.	Biological characterisation of DDLPS subgroups	181
	Characterisation of ECM and integrin adhesion networks in DDLPS subgroups	181
	Characterisation of functional proteomic networks in DDLPS subgroups	183
	Overview of tumour infiltrating lymphocytes in DDLPS subgroups	184
4.2.2.4.	Clinical characterisation of DDLPS subgroups	185
4.2.3.	Intra-subtype heterogeneity in UPS.....	188
4.2.3.1.	Cohort characteristics.....	188
4.2.3.2.	Identification of UPS subgroups	189
4.2.3.3.	Biological characterisation of UPS subgroups	190
	Characterisation of ECM and integrin adhesion networks in the UPS subgroups	190
	Overview of tumour infiltrating lymphocytes in the UPS subgroups	193
	Clinical characterisation of the UPS subgroups	197
4.3.	Discussion	197
4.3.1.	Matrix signalling heterogeneity in LMS.....	198
4.3.2.	Matrix signalling heterogeneity in DDLPS	202
4.3.3.	Matrix signalling heterogeneity in UPS	205
4.3.4.	Limitations.....	207
4.4.	Supplementary Material	210
4.4.1.	Supplementary Figures	210
4.4.2.	Supplementary Tables.....	216

CHAPTER 5 GENERATION AND CHARACTERISATION OF ECM

PRECLINICAL MODELS..... 228

5.1. Background and objectives 228

5.2. Results 229

5.2.1. Generation of preclinical ECM models..... 229

5.2.2. Characterisation of the composition of preclinical ECM models

5.2.3. The biological role of LMS ECM..... 240

5.2.3.1. Adhesion of SK-UT-1 cells to LMS ECM..... 241

5.2.3.2. Drug response of SK-UT-1 cells grown on LMS ECM

5.2.3.3. Migration of LMS cells on LMS ECM

 Baseline migratory characteristics of LMS cells

 Influence of LMS ECM matrix on cell speed.....

Influence of LMS ECM on cell directionality	250
5.2.3.4. Comparison of LMS migration on LMS ECM and commercially purified matrices	251
5.3.Discussion	254
5.3.1. Generation and characterisation of preclinical ECM models.....	254
5.3.1.1. Decellularised ECM scaffolds.....	254
5.3.1.2. ECM hydrogels.....	255
5.3.1.3. Limitations of generated preclinical ECM models	257
5.3.2. The biological role of LMS ECM.....	258
5.4.Supplementary Material	262
5.4.1. Supplementary Figures	262
 CHAPTER 6 IDENTIFICATION AND FUNCTIONAL VALIDATION OF MATRIX SIGNALLING HITS FROM PROTEOMIC DATA	 264
6.1.Background and objectives	264
6.2.Results	264
6.2.1. Identification of the ECM and integrin adhesion signalling proteins for functional validation.....	264
6.2.2. ZYX biology	270
6.2.3. Functional assessment of the role of ZYX in LMS cell lines	272
6.2.3.1. Selection of LMS cell lines for genetic knockdown of ZYX.....	273
6.2.3.2. Confirmation of ZYX knockdown in LMS cell lines	274
6.2.3.3. Role of ZYX in cell growth and migration of LMS.....	276
ZYX knockdown reduces LMS cell growth.....	276
ZYX knockdown increases SHEF-LMS ws cell migration.....	282
6.3.Discussion	285
6.3.1. ZYX biology	286
6.3.2. ZYX role in cell growth.....	288
6.3.3. ZYX role in migration	289
6.4.Supplementary Material	291
6.4.1. Supplementary Tables.....	291
 CHAPTER 7 CONCLUSIONS AND FUTURE DIRECTIONS.....	 292
7.1.Proteoglycan and CD8+ TILs score is prognostic in LMS, DDLPS and UPS cohort.....	293

7.2.Implications of matrix signalling heterogeneity in DDLPS and UPS.....	296
7.2.1. DDLPS.....	296
7.2.2. UPS.....	298
7.3.Future work related to preclinical ECM models.....	301
7.3.1. Synthetic extracellular matrices.....	303
7.4.Zyxin in LMS metastases	304
7.4.1. Therapeutic targeting of ZYX.....	305
7.4.2. Proteolysis-targeting chimaera technology	306
7.5.Concluding remarks.....	307
 CHAPTER 8 REFERENCES	 308

List of Figures

FIGURE 1.1 COMPOSITION OF THE TUMOUR MICROENVIRONMENT (TME).....	43
FIGURE 1.2 A SIMPLIFIED DIAGRAM OF THE ECM.	47
FIGURE 1.3 FUNCTIONAL ROLES OF THE ECM.....	48
FIGURE 1.4 COMPOSITION OF HUMAN <i>IN SILICO</i> MATRISOME.....	50
FIGURE 1.5 LOCALISATION AND COMPOSITION OF THE BASEMENT MEMBRANE.	58
FIGURE 1.6 THREE CONFORMATIONAL STATES OF AN INTEGRIN RECEPTOR.....	67
FIGURE 1.7 INTEGRIN ACTIVATION AND ‘INSIDE-OUT’ SIGNALLING.....	68
FIGURE 1.8 FORMATION OF INTEGRIN-MEDIATED NASCENT AND FOCAL ADHESION AND THE ‘OUTSIDE-IN SIGNALLING’.	69
FIGURE 1.9 COMPOSITION OF LITERATURE-CURATED HUMAN ADHESOME.....	70
FIGURE 1.10 SCHEMATIC OF A TYPICAL HYDROGEL.....	85
FIGURE 1.11 TISSUE DECELLULARISATION.	87
FIGURE 1.12 POST-PROCESSING OF DECELLULARISED TISSUE TO GENERATE TISSUE-DERIVED ECM HYDROGEL.	91
FIGURE 2.1 A FLOWCHART SHOWING THE STEP-BY-STEP PROCEDURE TO TRACK GFP+ CELLS USING THE TRACKMATE PLUGIN IN FIJI/IMAGEJ.	111
FIGURE 3.1 COVERAGE OF MATRISOME AND ADHESOME WITHIN THE SOFT TISSUE SARCOMA PROTEOMIC DATASET.....	118
FIGURE 3.2 THE MATRISOME AND ADHESOME PROFILES IN SOFT TISSUE SARCOMA (STS).	121
FIGURE 3.3 UNIFORM MANIFOLD APPROXIMATION AND PROJECTION (UMAP) OF MATRISOME AND ADHESOME PROFILES.	122
FIGURE 3.4 UNIQUE MATRISOME AND ADHESOME COMPONENTS IN THE SIX MOST COMMON SUBTYPES OF SOFT TISSUE SARCOMA (STS).	124
FIGURE 3.5 MATRISOME AND ADHESOME NETWORKS IN SOFT TISSUE SARCOMA (STS).....	127
FIGURE 3.6 BIOLOGICAL CHARACTERISATION OF CLUSTERS OF CO-REGULATED MATRISOME AND ADHESOME PROTEINS IN SOFT TISSUE SARCOMA (STS).	129
FIGURE 3.7 HISTOLOGICAL SUBTYPES ASSOCIATED WITH MATRISOME AND ADHESOME NETWORKS IN SOFT TISSUE SARCOMA (STS).	131
FIGURE 3.8 ANATOMICAL LOCATION ASSOCIATED WITH MATRISOME AND ADHESOME NETWORKS IN SOFT TISSUE SARCOMA (STS).	132
FIGURE 3.9 HISTOLOGICAL GRADE ASSOCIATED WITH MATRISOME AND ADHESOME NETWORKS IN SOFT TISSUE SARCOMA (STS).	133
FIGURE 3.10 PROTEOGLYCAN SCORE IS PROGNOSTIC OF OVERALL SURVIVAL IN A SUB-COHORT OF SOFT TISSUE SARCOMA (STS).	135
FIGURE 3.11 CD3+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) ARE INVERSELY ASSOCIATED WITH PROTEOGLYCAN GROUPS.	139

FIGURE 3.12 CD4+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) ARE INVERSELY ASSOCIATED WITH PROTEOGLYCAN GROUPS.	141
FIGURE 3.13 CD8+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) ARE ASSOCIATED WITH PROTEOGLYCAN GROUPS, AND PROTEOGLYCAN/CD8+ SCORE IDENTIFIES A HIGH-RISK STS GROUP.	143
FIGURE 4.1 MATRISOME AND ADHESOME COMPONENTS SPECIFIC FOR EACH LEIOMYOSARCOMA (LMS) SUBGROUP.	173
FIGURE 4.2 EXTRACELLULAR MATRIX AND INTEGRIN ADHESION NETWORKS IN LEIOMYOSARCOMA (LMS) SUBTYPES.	174
FIGURE 4.3 HALLMARKS IN LEIOMYOSARCOMA (LMS) SUBTYPES.	176
FIGURE 4.4 CLINICAL CHARACTERISATION OF LEIOMYOSARCOMA (LMS) SUBGROUPS	177
FIGURE 4.5 BIOLOGICAL CHARACTERISATION OF DEDIFFERENTIATED LIPOSARCOMA (DDLPS) SUBGROUPS.	182
FIGURE 4.6 HALLMARKS IN DEDIFFERENTIATED LIPOSARCOMA (DDLPS) SUBGROUPS.	183
FIGURE 4.7 TUMOUR INFILTRATING LYMPHOCYTES (TILS) IN DEDIFFERENTIATED LIPOSARCOMA (DDLPS) SUBGROUPS.	185
FIGURE 4.8 CLINICOPATHOLOGICAL CHARACTERISTICS SPECIFIC TO EACH DEDIFFERENTIATED LIPOSARCOMA (DDLPS) SUBGROUPS AND ASSOCIATION OF THE DDLPS SUBGROUPS WITH SURVIVAL OUTCOMES.	186
FIGURE 4.9 IDENTIFICATION OF UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS BASED ON MATRISOME AND ADHESOME PROTEINS EXPRESSION.	191
FIGURE 4.10 MATRISOME AND ADHESOME COMPONENTS SPECIFIC FOR EACH CONSENSUS CLUSTERS IN UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS).	192
FIGURE 4.11 TUMOUR INFILTRATING LYMPHOCYTES (TILS) IN UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS.	194
FIGURE 4.12 HALLMARKS IN UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS.	195
FIGURE 4.13 CLINICOPATHOLOGICAL CHARACTERISTICS SPECIFIC TO EACH UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUP AND ASSOCIATION OF THE SUBGROUPS WITH SURVIVAL OUTCOMES.	196
FIGURE 5.1 A PIPELINE FOR GENERATION AND CHARACTERISATION OF LMS PATIENT-DERIVED ECM SCAFFOLDS AND LMS PATIENT-DERIVED ECM HYDROGELS.	230
FIGURE 5.2 A WORKFLOW FOR RESEARCH APPLICATION OF ECM SCAFFOLDS AND PRE-GEL ECM SOLUTION.	232
FIGURE 5.3 LEIOMYOSARCOMA (LMS) ECM SCAFFOLDS AND HYDROGELS RETAIN MATRISOME COMPONENTS FROM THEIR NATIVE COUNTERPARTS.	234
FIGURE 5.4 DECELLULARISATION AND HYDROGEL PREPARATION ENRICH FOR MATRISOME CONTENT.	235

FIGURE 5.5 DECELLULARISATION AND ECM HYDROGEL PREPARATION RETAIN COLLAGEN CHAINS.	236
FIGURE 5.6 CONSENSUS MATRISOME PROTEINS IN LEIOMYOSARCOMA (LMS).	238
FIGURE 5.7 CONSENSUS MATRISOME PROTEINS ARE LEIOMYOSARCOMA (LMS) SPECIFIC.	239
FIGURE 5.8 ADHESION OF SK-UT-1 CELLS TO PLASTIC AND LEIOMYOSARCOMA DERIVED PRE-GEL EXTRACELLULAR MATRIX SOLUTION (LMS ECM).....	241
FIGURE 5.9 SK-UT-1 CELL ADHESION AT 24 HOUR AND CELL GROWTH ON PLASTIC COMPARED TO LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM).	242
FIGURE 5.10 PEARSON'S CORRELATION ANALYSIS OF BIOLOGICAL REPLICATES.	243
FIGURE 5.11 SMALL MOLECULE INHIBITOR SCREEN OF SK-UT-1 ON PLASTIC AND LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM).	244
FIGURE 5.12 CONFIRMATION OF SK-UT-1 SENSITIVITY TO DASATINIB ON LEIOMYOSARCOMA PRE- GEL EXTRACELLULAR MATRIX (LMS ECM).	245
FIGURE 5.13 CONFIRMATION OF SK-UT-1 SENSITIVITY TO DASATINIB ON LEIOMYOSARCOMA PRE- GEL EXTRACELLULAR MATRIX (LMS ECM).	246
FIGURE 5.14 BASELINE MIGRATION PROPERTIES OF LEIOMYOSARCOMA (LMS) CELL LINES.	247
FIGURE 5.15 AVERAGE CELL SPEED OF LEIOMYOSARCOMA (LMS) CELL LINES ON LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM) AND PLASTIC.	249
FIGURE 5.16 THE DIRECTIONALITY OF LEIOMYOSARCOMA (LMS) CELL LINES ON LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM) AND PLASTIC.	251
FIGURE 6.1 SUMMARY OF A WORKFLOW TO IDENTIFY ZYX AS A PROTEIN FOR FUNCTIONAL VALIDATION IN CELL VIABILITY AND MIGRATION ASSAYS IN LMS CELL LINES.	266
FIGURE 6.2 SIGNIFICANCE ANALYSIS OF MICROARRAY (SAM) ANALYSIS FOR (A) LOCAL RECURRENCE- FREE SURVIVAL (LRFS), (B) METASTASIS-FREE SURVIVAL (MFS) AND OVERALL SURVIVAL (OS) IN LEIOMYOSARCOMA.	267
FIGURE 6.3 ZYX IS ASSOCIATED WITH METASTASIS-FREE SURVIVAL IN LEIOMYOSARCOMA.	269
FIGURE 6.4 ZYX PROTEIN-PROTEIN INTERACTION NETWORK.....	271
FIGURE 6.5 BIOLOGICAL PATHWAYS ENRICHED IN ZYX HIGH COMPARED TO ZYX LOW LEIOMYOSARCOMA (LMS) SAMPLES.	272
FIGURE 6.6 ZYX EXPRESSION IN ESTABLISHED AND PATIENT-DERIVED LEIOMYOSARCOMA CELL LINES.	274
FIGURE 6.7 CONFIRMATION OF ZYX KNOCKDOWN IN LEIOMYOSARCOMA CELL LINES.....	275
FIGURE 6.8 ZYX KNOCKDOWN REDUCED SHEF-LMS WS CELL POPULATION GROWTH.	277
FIGURE 6.9 ZYX KNOCKDOWN REDUCED SHEF-LMS W1 CELL POPULATION GROWTH.	279
FIGURE 6.10 ZYX KNOCKDOWN IN SK-UT-1 CELL POPULATION GROWTH.....	281
FIGURE 6.11 ZYX KNOCKDOWN INCREASES CELL SPEED OF SHEF-LMS WS CELLS.....	282
FIGURE 6.12 CELL SPEED AND DIRECTIONALITY OF SK-UT-1 ZYX KNOCKDOWN CELLS.	284
FIGURE 6.13 CELL SPEED AND DIRECTIONALITY OF SHEF-LMS-W1 ZYX KNOCKDOWN CELLS.	285

List of Supplementary Figures

SUPPLEMENTARY FIGURE 3.1 KAPLAN-MEIER PLOTS OF LOCAL RECURRENCE-FREE SURVIVAL (LRFS), METASTASIS-FREE SURVIVAL (MFS) AND OVERALL SURVIVAL (OS) IN A COHORT OF N=321 SOFT TISSUE SARCOMA (STS). HORIZONTAL AND VERTICAL LINES INDICATE THE MEDIAN MFS AND OS.	153
SUPPLEMENTARY FIGURE 3.2 SIGNIFICANTLY DIFFERENTIALLY EXPRESSED MATRISOME AND ADHESOME PROTEINS IN HISTOLOGICAL SUBTYPES OF SOFT TISSUE SARCOMA (STS).	154
SUPPLEMENTARY FIGURE 3.3 IDENTIFICATION OF MATRISOME AND ADHESOME NETWORKS IN SOFT TISSUE SARCOMA (STS).	155
SUPPLEMENTARY FIGURE 3.4 ASSOCIATION OF NABA-DEFINED MATRISOME RELATED GENE SETS WITH SURVIVAL OUTCOMES IN A SUB-COHORT OF SOFT TISSUE SARCOMA (STS)	156
SUPPLEMENTARY FIGURE 3.5 ASSOCIATIONS OF A PROTEOGLYCAN/ CD3+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) SCORE WITH SURVIVAL OUTCOMES IN A SUB COHORT OF SOFT TISSUE SARCOMA (STS).	157
SUPPLEMENTARY FIGURE 3.6 ASSOCIATIONS OF A PROTEOGLYCAN/ CD4+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) SCORE WITH SURVIVAL OUTCOMES IN A SUB-COHORT OF SOFT TISSUE SARCOMA (STS).	158
SUPPLEMENTARY FIGURE 3.7 ASSOCIATIONS OF A PROTEOGLYCAN/ CD8+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) SCORE WITH SURVIVAL OUTCOMES IN A SUB-COHORT OF SOFT TISSUE SARCOMA (STS).	159
SUPPLEMENTARY FIGURE 4.1 KAPLAN-MEIER PLOTS OF LOCAL RECURRENCE-FREE SURVIVAL (LRFS), METASTASIS-FREE SURVIVAL (MFS) AND OVERALL SURVIVAL (OS) IN A COHORT OF N = 80 LEIOMYOSARCOMA (LMS).	210
SUPPLEMENTARY FIGURE 4.2 IDENTIFICATION OF LEIOMYOSARCOMA (LMS) SUBGROUPS BASED ON THE PROTEIN EXPRESSION OF MATRISOME AND ADHESOME.	211
SUPPLEMENTARY FIGURE 4.3 SUPPLEMENTARY FIGURE 4.3 KAPLAN-MEIER PLOTS OF LOCAL RECURRENCE-FREE SURVIVAL (LRFS), METASTASIS-FREE SURVIVAL (MFS) AND OVERALL SURVIVAL (OS) IN A COHORT OF N = 39 DEDIFFERENTIATED LIPOSARCOMA (DDLPS).	212
SUPPLEMENTARY FIGURE 4.4 IDENTIFICATION OF DEDIFFERENTIATED LIPOSARCOMA (DDLPS) SUBGROUPS BASED ON THE PROTEIN EXPRESSION OF MATRISOME AND ADHESOME.	213
SUPPLEMENTARY FIGURE 4.5 KAPLAN-MEIER PLOTS OF LOCAL RECURRENCE-FREE SURVIVAL (LRFS), METASTASIS-FREE SURVIVAL (MFS) AND OVERALL SURVIVAL (OS) IN A COHORT OF N = 53 UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS).	214
SUPPLEMENTARY FIGURE 4.6 IDENTIFICATION OF UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS BASED ON MATRISOME AND ADHESOME PROTEIN EXPRESSION.	215

SUPPLEMENTARY FIGURE 5.1 COMPARISON OF CELL SPEED IN LEIOMYOSARCOMA (LMS) CELL LINES PLATED ON A COATING OF LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM) AND COMMERCIAL PURIFIED MATRICES.....	262
SUPPLEMENTARY FIGURE 5.2 COMPARISON OF DIRECTIONALITY INDICES IN LEIOMYOSARCOMA (LMS) CELL LINES PLATED ON A COATING OF LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM) AND COMMERCIAL PURIFIED MATRICES.	263

List of Tables

TABLE 1.1 DEFINITION OF GRADING PARAMETERS FOR FÉDÉRATION NATIONALE DES CENTRES DE LUTTE CONTRE LE CANCER (FNCLCC).	30
TABLE 1.2 SUMMARY OF KEY BIOLOGICAL AND CLINICAL FEATURES OF LMS, DDLPS AND UPS.....	42
TABLE 1.3 SUMMARY OF STUDIES DESCRIBING THE GENERATION OF PROGNOSTIC MATRIX SCORE IN CARCINOMAS.....	61
TABLE 1.4 SUMMARY OF THE STUDIES ON ECM COMPONENTS IN STS.	73
TABLE 1.5 SUMMARY OF THE STUDIES ON INTEGRINS IN STS.....	82
TABLE 2.1 SUMMARY OF THE NUMBER OF DDLPS, UPS AND LMS PATIENTS WITH AVAILABLE IHC DATA FOR CD3/4/8+ TILS.....	99
TABLE 2.2 SHORT HAIRPIN RNA (SHRNA) OLIGO SEQUENCES USED FOR CLONING.	104
TABLE 2.3 SMALL MOLECULE INHIBITOR SCREEN COMPONENTS.	108
TABLE 2.4 PRIMARY AND SECONDARY ANTIBODIES USED FOR IMMUNOBLOTTING WITH ASSOCIATED DILUTIONS AND SUPPLIER INFORMATION.	113
TABLE 3.1 CLINICOPATHOLOGICAL CHARACTERISTICS OF N=321 SOFT TISSUE SARCOMA (STS) CASES	116
TABLE 3.2 MATRISOME GENE SETS FROM THE MOLECULAR SIGNATURES DATABASE (MSIGDB) EVALUATED FOR PROGNOSTIC SCORE.	136
TABLE 3.3 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND PROTEOGLYCAN SCORE WITH OVERALL SURVIVAL (OS).	137
TABLE 3.4 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND PROTEOGLYCAN/CD8+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) SCORE WITH OVERALL SURVIVAL (OS).	145
TABLE 4.1 CLINICOPATHOLOGICAL CHARACTERISTICS OF N=80 LEIOMYOSARCOMA (LMS) CASES..	171
TABLE 4.2 CLINICOPATHOLOGICAL CHARACTERISTICS OF N = 39 DEDIFFERENTIATED LIPOSARCOMA (DDLPS) CASES.	180
TABLE 4.3 CLINICOPATHOLOGICAL CHARACTERISTICS OF N = 53 UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) CASES.....	189
TABLE 4.4 SUMMARY OF CHAPTER 4 ANALYSIS.....	198
TABLE 5.1 NUMBERS OF MATRISOME, NON-MATRISOME AND TOTAL NUMBER OF PROTEIN IDENTITIES (IDS) IDENTIFIED ACROSS 7 TUMOURS, 7 SCAFFOLDS AND 7 HYDROGELS.....	233
TABLE 5.2 SUMMARY OF ESTABLISHED AND PATIENT-DERIVED UTERINE LEIOMYOSARCOMA CELL LINES.....	241

TABLE 5.3 COMPARISON OF THE CELL SPEED OF LEIOMYOSARCOMA (LMS) CELL LINES PLATED ON A COATING OF LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM) AND COMMERCIAL PURIFIED MATRICES.....	252
TABLE 5.4 COMPARISON OF THE CELL DIRECTIONALITY OF LEIOMYOSARCOMA (LMS) CELL LINES PLATED ON A COATING OF LEIOMYOSARCOMA PRE-GEL EXTRACELLULAR MATRIX (LMS ECM) AND COMMERCIAL PURIFIED MATRICES.....	253
TABLE 6.1 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSIS ASSESSING THE ASSOCIATION OF SIGNIFICANCE ANALYSIS OF MICROARRAY (SAM) SIGNIFICANT PROTEINS WITH LOCAL RECURRENCE-FREE SURVIVAL, METASTASIS-FREE SURVIVAL AND OVERALL SURVIVAL IN LEIOMYOSARCOMA PATIENTS.	268
TABLE 6.2 UNIVARIABLE COX REGRESSION ANALYSIS FOR ZYX IN SOFT TISSUE SARCOMA N = 271 COHORT.	270
TABLE 6.3 SUMMARY OF ESTABLISHED AND PATIENT-DERIVED LEIOMYOSARCOMA CELL LINES. ...	273

List of Supplementary Tables

SUPPLEMENTARY TABLE 3.1 CLINICOPATHOLOGICAL CHARACTERISTICS OF N=321 SOFT TISSUE SARCOMA (STS) CASES.....	160
SUPPLEMENTARY TABLE 3.2 REACTOME ENRICHED PATHWAYS IN MATRISOME AND ADHESOME NETWORKS OF SOFT TISSUE SARCOMA (STS).	161
SUPPLEMENTARY TABLE 3.3 STATISTICAL COMPARISONS BETWEEN HISTOLOGICAL SUBTYPES IN CONSENSUS CLUSTER 1 (CC1).....	162
SUPPLEMENTARY TABLE 3.4 CLINICOPATHOLOGICAL CHARACTERISTICS OF N = 148 SOFT TISSUE SARCOMA (STS) CASES.....	163
SUPPLEMENTARY TABLE 3.5 STATISTICAL ASSOCIATION OF CLINICOPATHOLOGICAL FEATURES WITH PROTEOGLYCAN GROUPS.	164
SUPPLEMENTARY TABLE 3.6 SUMMARY OF MULTIVARIABLE (MVA) COX REGRESSION ANALYSIS ASSESSING BASEMENT MEMBRANE SCORE WITH LOCAL RECURRENCE-FREE SURVIVAL (LRFS) AND OVERALL SURVIVAL (OS).....	165
SUPPLEMENTARY TABLE 3.7 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSIS ASSESSING OVERALL SURVIVAL (OS), LOCAL RECURRENCE-FREE SURVIVAL (LRFS) AND METASTASIS-FREE SURVIVAL (MFS) IN N = 145 STS PATIENTS WITH CD3+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) COUNTS.....	166
SUPPLEMENTARY TABLE 3.8 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSES ASSESSING OVERALL SURVIVAL (OS), LOCAL RECURRENCE-FREE SURVIVAL (LRFS) AND METASTASIS-FREE SURVIVAL (MFS) IN N = 145 STS PATIENTS STRATIFIED BY PROTEOGLYCAN AND CD3+ TUMOUR INFILTRATING LYMPHOCYTES SCORE.	166
SUPPLEMENTARY TABLE 3.9 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSIS ASSESSING OVERALL SURVIVAL (OS), LOCAL RECURRENCE-FREE SURVIVAL (LRFS) AND METASTASIS-FREE SURVIVAL (MFS) IN N = 142 STS PATIENTS WITH CD4+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) COUNTS.....	167
SUPPLEMENTARY TABLE 3.10 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSES ASSESSING OVERALL SURVIVAL (OS), LOCAL RECURRENCE-FREE SURVIVAL (LRFS), AND METASTASIS-FREE SURVIVAL (MFS) IN N = 142 STS PATIENTS STRATIFIED BY PROTEOGLYCAN AND CD4+ TUMOUR INFILTRATING LYMPHOCYTES SCORE.	167
SUPPLEMENTARY TABLE 3.11 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSIS ASSESSING OVERALL SURVIVAL (OS), LOCAL RECURRENCE-FREE SURVIVAL (LRFS) AND METASTASIS-FREE SURVIVAL (MFS) IN N = 148 STS PATIENTS WITH CD8+ TUMOUR INFILTRATING LYMPHOCYTES (TILS) COUNTS.....	168
SUPPLEMENTARY TABLE 3.12 SUMMARY OF UNIVARIABLE (UVA) COX REGRESSION ANALYSES ASSESSING OVERALL SURVIVAL (OS), LOCAL RECURRENCE-FREE SURVIVAL (LRFS), AND	

METASTASIS-FREE SURVIVAL (MFS) IN N = 148 STS PATIENTS STRATIFIED BY PROTEOGLYCAN AND CD8+ TUMOUR INFILTRATING LYMPHOCYTES SCORE.	168
SUPPLEMENTARY TABLE 4.1 A SUMMARY OF 112 DIFFERENTIALLY EXPRESSED MATRISOME AND ADHESOME PROTEINS (DEPS) UNIQUELY UPREGULATED IN EACH OF THE THREE LMS SUBGROUPS.....	216
SUPPLEMENTARY TABLE 4.2 SUMMARY OF STATISTICAL TESTS TO ASSESS THE ASSOCIATION BETWEEN CLINICOPATHOLOGICAL FEATURES AND LEIOMYOSARCOMA (LMS) SUBGROUPS.	217
SUPPLEMENTARY TABLE 4.3 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND LEIOMYOSARCOMA (LMS) SUBGROUPS WITH METASTASIS-FREE SURVIVAL (MFS).	218
SUPPLEMENTARY TABLE 4.4 SUMMARY OF MULTIVARIABLE (MVA) COX REGRESSION ANALYSIS ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND LEIOMYOSARCOMA (LMS) SUBGROUPS WITH LOCAL RECURRENCE-FREE SURVIVAL (LRFS) AND OVERALL SURVIVAL (OS).	219
SUPPLEMENTARY TABLE 4.5 SUMMARY OF STATISTICAL TESTS TO ASSESS THE ASSOCIATION BETWEEN CLINICOPATHOLOGICAL FEATURES AND MATRISOME AND ADHESOME CLUSTERS IN DEDIFFERENTIATED LIPOSARCOMA (DDLPS).....	220
SUPPLEMENTARY TABLE 4.6 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND DEDIFFERENTIATED LIPOSARCOMA (DDLPS) SUBGROUPS WITH LOCAL RECURRENCE-FREE SURVIVAL (LRFS).	221
SUPPLEMENTARY TABLE 4.7 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND DDLPS SUBGROUPS WITH OVERALL SURVIVAL (OS).	222
SUPPLEMENTARY TABLE 4.8 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND DDLPS SUBGROUPS WITH METASTASIS-FREE SURVIVAL (MFS).	223
SUPPLEMENTARY TABLE 4.9 SUMMARY OF STATISTICAL TESTS TO ASSESS THE ASSOCIATION BETWEEN CLINICOPATHOLOGICAL FEATURES AND UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS.	224
SUPPLEMENTARY TABLE 4.10 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS WITH METASTASIS-FREE SURVIVAL (MFS).	225
SUPPLEMENTARY TABLE 4.11 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS	

AND UNDIFFERENTIATED PLEOMORPHIC (UPS) SUBGROUPS WITH LOCAL RECURRENCE-FREE SURVIVAL (LRFS).	226
SUPPLEMENTARY TABLE 4.12 SUMMARY OF UNIVARIABLE (UVA) AND MULTIVARIABLE (MVA) COX REGRESSION ANALYSES ASSESSING THE ASSOCIATION OF CLINICOPATHOLOGICAL FACTORS AND UNDIFFERENTIATED PLEOMORPHIC SARCOMA (UPS) SUBGROUPS WITH OVERALL SURVIVAL (OS).	227
SUPPLEMENTARY TABLE 6.1 SUMMARY OF MULTIVARIABLE (MVA) COX REGRESSION ANALYSIS ASSESSING THE ASSOCIATION OF LEIOMYOSARCOMA CLINICOPATHOLOGICAL FACTORS AND ZYX EXPRESSION WITH METASTASIS-FREE SURVIVAL (MFS).	291

Abbreviations and acronyms

2D	Two-dimensional
3D	Three-dimensional
ACN	Acetonitrile
ADAM	A disintegrin and metalloproteinase
AIDS	Acquired immunodeficiency syndrome
AJCC	American Joint Committee on Cancer
ANOVA	Analysis of variance
ARMs	antibody recruiting molecules
AS	Angiosarcoma
ASPS	Alveolar soft part sarcoma
BCA	Bicinchoninic acid
BME	Basement membrane extract
CAR T-cell	Chimeric antigen receptor T-cell
CCS	Clear cell sarcoma
CDF	Consensus cumulative distribution function
CI	Confidence Intervals
CSPG4	Chondroitin sulfate proteoglycan 4
DDLPS	Dedifferentiated liposarcoma
DEPs	Differentially expressed proteins
DES	Desmoid tumours
DFS	Disease-free survival
DMSO	Dimethyl sulfoxide
DRR	DNA damage response repair
DSRCT	Desmoplastic small round cell tumours
DSS	Disease-specific survival
DTF	Desmoid-type fibromatosis
ECM	Extracellular matrix
ECOG	Eastern Cooperative Oncology Group
EDTA	Ethylenediaminetetraacetic acid
eLMS	Extra-uterine LMS
EMT	Epithelial-to-mesenchymal transition
ES	Epithelioid sarcoma
FA	Formic acid
FACIT	Fibril-associated collagens with interrupted triple helices
FDR	False discovery rate
FFPE	Formalin-fixed paraffin embedded
FISH	Fluorescent in situ hybridisation
FNCLCC	Federation Nationale des Centres de Lutte Contre le Cancer
GAG	Glycosaminoglycan
GFP	Green fluorescent protein
HCL	Hydrochloric acid

HGSOC	High-grade serous ovarian cancer
HHV-8	Human Herpesvirus-8
HIV	Human immunodeficiency virus
HR	Hazard ratio
HRP	Horseradish peroxidase
HRR	Homologous recombination repair
HSPG2	Heparan Sulfate Proteoglycan 2
IACs	Integrin adhesion complexes
ICI	Immune checkpoint inhibitor
IF	Immunofluorescence
IHC	Immunohistochemistry
ILK	Integrin-linked kinase
INT	Istituto Nazionale dei Tumori
IQR	Interquartile range
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LMS	Leiomyosarcoma
LOX	Lysyl oxidase
LRFS	Local recurrence free survival
MALDI-IMS	Matrix-assisted laser desorption/ionisation imaging mass spectrometry
MatrisomeDB	Matrisome database
MFH	Malignant fibrous histiocytoma
MFS	Metastasis free survival
MMP	Matrix metalloproteinase
MSCs	Mesenchymal stem cells
MSigDB	The Molecular Signatures Database
mTOR	Mechanistic target of rapamycin
MyxFS	Myxofibrosarcoma
NA	Not applicable
NCI	National Cancer Institute
OS	Overall survival
PAK	RAC1 activated kinase
PARP	Poly ADP-ribose polymerase
PBS	Phosphate-buffered saline
PEG	polyethylene glycol
PFS	Progression-free survival
PI3K	phosphatidylinositol 3-kinase
PLOD	Procollagen-lysine,2-oxoglutarate 5-dioxygenase
PPI	Protein-protein interaction
PROSPECTUS	PROgnoStic and PrEdiCTive ImmUnoprofiling of Sarcomas
PS	Performance status
RAC1	Ras-related C3 botulinum toxin substrate 1
RhoA	Ras homolog family member A
RMH	Royal Marsden Hospital

RPPA	Reverse phase protein array
RT	Rhabdoid tumours
RT-PCR	Reverse transcription-polymerase chain reaction
RTK	Receptor tyrosine kinase
SAM	Significance Analysis of Microarrays
SDS	Sodium dodecyl sulfate
shRNA	Short hairpin RNA
SigClust	Significance of Clustering
siRNA	Small interfering RNA
SS	Synovial sarcoma
ssGSEA	Single-sample Gene Set Enrichment Analysis
STRINGdb	Search Tool for Retrieval of Interacting Genes/Proteins database
STS	Soft tissue sarcomas
SWATH	Sequential window acquisition of all theoretical mass spectra
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
TCGA	The Cancer Genome Atlas
TCGA-SARC	The Cancer Genome Atlas Sarcoma
TFA	Trifluoroacetic acid
TILs	Tumour infiltrating lymphocytes
TMA	Tissue microarrays
TME	Tumour microenvironment
TMT	Tandem mass tag
TNM	Tumour, node, metastasis
uLMS	Uterine LMS
UMAP	Uniform manifold approximation and projection
UMAP	Uniform manifold approximation and projection
uPAR	Urokinase-type plasminogen activator receptor
UPS	Undifferentiated pleomorphic sarcoma
UV	Ultraviolet
WDLPS	Well-differentiated LPS
WGD	Whole genome duplication
WHO	World health organisation
ZYX	Zyxin

Chapter 1 Introduction

1.1. Soft tissue sarcoma

1.1.1. Overview

Soft tissue sarcomas (STS) are a group of rare and heterogeneous cancers that account for 1% of all adult cancer diagnoses each year¹. STS arise from cells of mesenchymal stem cells (MSCs) lineage, which, during development, differentiate to form connective tissues in the body such as muscles, fat, blood vessels, nerves and tendons². In line with that, STS can manifest in various body regions, the extremities being the most common location³. According to the most recent 2020 World Health Organisation (WHO) STS classification, over 80 different histological subtypes of STS have been identified, reflecting histological heterogeneity of the disease⁴. In addition, STS are characterised by clinical heterogeneity. Patients diagnosed with STS exhibit significant variations in clinical prognoses, response to therapies, local and distant relapse rate, and survival outcomes, even within the same histological subtype. Clinical presentation of patients with STS is highly variable, and no specifically defined clinical features are characteristic of all STS cases, often leading to misdiagnosis or delayed diagnosis⁵. To address this complexity and improve patient outcomes, managing STS by multidisciplinary sarcoma specialists, including medical oncologists, surgeons, radiologists, and pathologists, is essential⁶. Despite the advancement in sarcoma management, the prognosis for patients with STS is dismal, with a 5-year survival rate of ~55-65%⁷. Surgery remains the mainstay of clinical management for localised sarcomas; however, despite wide surgical margins, up to 50% of patients relapse after surgery⁸. The treatment in the advanced setting is usually palliative, with a median overall survival (OS) of less than 2 years, and only 15% of advanced STS patients remain alive 5 years after initial diagnosis^{7,9,10}. These statistics underscore the urgent need for improved therapeutic strategies and a better understanding of STS pathobiology.

1.1.2. Development and aetiology

The exact cell of origin remains unknown for most STS histological subtypes. However, there is evidence that MSCs or MSC-derived progenitor cells represent putative sarcoma-initiating cells for certain STS subtypes¹¹. It is believed that oncogenic hits occurring in MSCs or their progenitor cells could lead to the development of specific subtypes of STS². Liposarcoma (LPS) is believed to arise from the precursor lipoblasts of the adipocytes (i.e. fat cells) in adipose (i.e. fat) tissues¹². Angiosarcoma (AS) develops in the inner lining of blood vessels and lymph vessels from endothelial cells¹³. Leiomyosarcoma (LMS) originates from the smooth muscle cells and can arise in smooth muscle layers of the uterus, retroperitoneum or walls of blood vessels¹⁴. One example of STS where the cell of origin is unknown is undifferentiated pleomorphic sarcoma (UPS).

1.1.3. Risk factors

It is currently unknown what causes malignant transformation in cells giving rise to STS. Nevertheless, some risk factors are known to increase the chance of developing STS. Established risk factors that increase the likelihood of developing STS are radiation, family cancer syndromes, chronic lymphedema and exposure to certain chemicals and viruses¹⁵.

1.1.3.1. *Ionising radiation*

Ionising radiation has been linked to STS developing in atomic bomb survivors and patients previously treated with radiotherapy for other conditions^{16,17}. Radiation-associated AS commonly occurs secondary to radiotherapy-treated breast cancers and radiation-associated UPS occurs subsequent to lymphoma treated with radiotherapy¹⁸. Although rare, radiation-associated STS have aggressive clinical behaviour and a poorer prognosis than sporadic STS of the same histological subtype, and therefore, is an important clinical STS group^{18–20}.

1.1.3.2. Family history

Cancer predisposition syndromes like Li-Fraumeni syndrome caused by germline loss-of-function mutations in the p53 tumour suppressor gene (*TP53*) increases the chance of rhabdomyosarcoma, LPS and LMS^{21–23}. Neurofibromatosis type 1 with the loss-of-function mutations in the neurofibromin (*NF1*) gene increases the risk of malignant peripheral nerve sheath tumour development^{24,25}. A recent comprehensive genetic study of four international sarcoma cohorts (n = 1162) found that approximately 55% of sarcoma patients carried pathogenic germline variants²⁶. This study confirmed the presence of known risk factor genes and identified novel genes associated with the DNA damage response, such as ataxia-telangiectasia mutated (*ATM*), ataxia telangiectasia and Rad3-related (*ATR*), and Xeroderma pigmentosum group D helicase (*ERCC2*), as well as genes involved in homologous recombination, such as breast cancer gene 2 (*BRCA2*). These findings underscore the potential for integrating genetic screening and early detection strategies for STS, even in cases with no history of familial cancer syndromes²⁷.

1.1.3.3. Viral infection

The most well-studied association between viral infection and STS is the case of Kaposi's sarcoma^{28–30}. The accepted mechanism of Kaposi's sarcoma development is an oncogenic transformation of endothelial cells with Human Herpesvirus-8 (HHV-8) that leads to cutaneous lesions. However, HHV-8 infection alone rarely induces Kaposi's sarcoma in the general population. Groups at the highest risk of Kaposi's sarcoma include people infected with human immunodeficiency virus (HIV) and living with acquired immunodeficiency syndrome (AIDS), individuals who have undergone organ transplants or are taking immunosuppressive medications.

1.1.4. Classification

1.1.4.1. Histology-based classification

Understanding the classification of STS is crucial for accurate diagnosis, appropriate management and prognostication. The 5th Edition of the WHO classification of soft tissue and bone tumours describes benign, intermediate and malignant tumours and divides STS histological entities into the 'differentiated tumours', tumours of 'uncertain differentiation' and 'undifferentiated small round cell STS'⁴. Tumours described as differentiated demonstrate morpho-phenotypic differentiation that resembles a normal mesenchymal cell. For instance, adipocytic tumours are a cohort of tumours that resemble adipocytes (e.g. well-differentiated LPS (WDLPS), dedifferentiated LPS (DDLPS), myxoid LPS and pleomorphic LPS). Fibroblastic and myofibroblastic tumours develop from fibroblasts and myocytes (muscle progenitor cells), including entities such as fibrosarcoma and desmoid-type fibromatosis. Vascular tumours form from blood vessels or lymph vessels (e.g. angiosarcoma). Entities 'of uncertain differentiation' do not readily resemble any normal mesenchymal cells and include well-characterised entities such as synovial and epithelioid sarcomas.

1.1.4.2. Genetic classification

In addition to histology-based classification, the new WHO classification recognises the significance of molecular features and genetic alterations in the classification and diagnosis of STS. Broadly, STS can be divided into two categories based on genetic characteristics³¹. The first group is the STS with simple genetic alterations and near-diploid karyotypes. The second group has complex and unbalanced karyotypes. 'Simple' STS harbour genetic translocations that encode aberrant transcriptional proteins and have few mutations. Synovial sarcomas contain a fusion of SS18 subunit of BAF chromatin remodelling complex (*SS18*) with either Synovial Sarcoma, X Breakpoint 1 (*SSX1*), *SSX2* or *SSX4*^{32,33}. Another example of translocation-driven STS is alveolar soft part sarcoma, characterised by the *ASPSCR1-TFE3* fusion gene³⁴. Testing for characteristic genomic alterations such as

translocations can provide a more accurate diagnosis, as diagnosis based on morphology alone is challenging. Genomically 'complex' STS show a higher frequency of copy-number alterations and recurrent mutations than the 'simple' STS. Alterations typically affect key regulators of genome stability such as *TP53*, retinoblastoma gene (*RB1*), cyclin-dependent kinase inhibitor 2A (*CDKN2A*), mouse double minute 2 homolog (*MDM2*) and cyclin-dependent kinase 4 (*CDK4*) genes^{35,36}. Common subtypes with complex karyotypes are DDLPS presenting with *MDM2* and *CDK4* amplifications, UPS with point mutations in *TP53*, truncating mutations of alpha-thalassemia/mental retardation, X-linked (*ATRX*) and *RB1*, and LMS with frequent point mutations in or deletions of *TP53*, *RB1* and phosphatase and tensin homolog (*PTEN*)^{35,37}. Overall, oncogenic drivers are rare in STS, and most genetic alterations are deletions and point mutations in tumour suppressor genes.

1.2. Diagnosis and prognosis of STS

1.2.1. Diagnosis

Due to the rarity and unspecific symptoms of STS, diagnosis is often delayed. It is recommended to refer all patients with suspicious soft tissue lesions (with a diameter ≥ 5 cm) to a sarcoma specialist³⁸. STS is diagnosed by a combination of clinical examination and imaging tests. Following that, multiple core needle biopsies and immunohistochemistry (IHC) are the standard approaches to confirm the diagnosis. A sarcoma expert pathologist is required to make a diagnosis according to the latest WHO classification. Fluorescent in situ hybridisation (FISH) and reverse transcription–polymerase chain reaction (RT–PCR) tests for characteristic genomic alterations are employed in some instances when a diagnosis is uncertain or a clinical presentation is unusual.

1.2.2. Prognostication

Accurate prognosis is essential for tailoring the treatment for STS patients. STS prognosis depends on several factors, including patient characteristics

(age, sex and performance status) and tumour characteristics (grade, size, anatomical site and histological subtype of STS)^{39,40}. Typically, male STS patients with large, high-grade tumours, with incomplete surgical resection and Eastern Cooperative Oncology Group (ECOG) performance status of ≥ 2 tend to have an unfavourable prognosis^{39,40}. Several studies with large cohorts of STS patients determined that histological grade is the most important prognostic factor for metastasis-free survival (MFS), disease-free survival (DFS) and OS, with high-grade tumours associated with worse outcome^{3,41}. It should be noted that grade does not predict local recurrence.

1.2.2.1. Grading systems

Various STS grading systems have been developed, each with different tumour-descriptive parameters. The two widely established systems that the WHO accepts are the National Cancer Institute (NCI)⁴² and the Federation Nationale des Centres de Lutte Contre le Cancer (FNCLCC)⁴³ systems.

Table 1.1 Definition of grading parameters for Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC).

HPF: high power field.

Parameter	Criteria
Tumour differentiation	
Score 1	Well differentiated, sarcoma resembling normal adult tissue
Score 2	Moderately differentiated, confirmed histological diagnosis
Score 3	Undifferentiated, embryonic sarcoma, synovial sarcoma, epithelioid sarcoma, clear-cell sarcoma, soft tissue alveolar sarcoma, undifferentiated sarcoma and uncertain histology
Mitotic count (per 10 HPF)	
Score 1	0-9
Score 2	10-19
Score 3	>20
Tumour necrosis	
Score 1	No necrosis
Score 2	<50%
Score 3	$\geq 50\%$
Histologic grade (cumulative score)	
Grade 1: 2-3 points	Low risk of metastasis
Grade 2: 4-5 points	Real risk of metastasis
Grade 3: 6-8 points	High risk of metastasis

To construct the grading system, FNCLCC evaluated tumour differentiation, cellularity, number of atypical nuclei, giant tumour cells, mitotic count, tumour necrosis and presence of vascular emboli⁴³. Multivariate analysis using a Cox model established that the three parameters (tumour differentiation, mitotic count and tumour necrosis) remained independent prognostic factors for MFS and OS, and those were used to construct the grading system. Table 1.1 describes parameters that determine the FNCLCC tumour grades. According to the system, a score from 1 to 3 is independently assigned for each of the three parameters. Tumours with a total score of 2 – 3 are classified as grade 1 (low), 4 – 5 as grade 2 (intermediate), and 6 – 8 as grade 3 (high). In general, grading systems aim to stratify patients into low grade (grade 1) with favourable prognoses of ~90% 5-year MFS rate and into high grade (grade 3) with a poor prognosis of ~40% MFS rate⁴¹. An effective grading system aims to minimise the number of patients in the intermediate grade 2 group, which does not provide precise prediction on prognosis and is less useful in clinical decision-making.

A comparative study of the NCI and the FNCLCC systems with n = 410 adult STS cases showed a significantly lower number of tumours classified as grade 2 and a higher number of grade 3 by the FNCLCC versus the NCI⁴⁴. Additionally, compared to the NCI, the FNCLCC system showed a better prediction of MFS between tumour grades. Consequently, the FNCLCC was recommended over the NCI for clinical use in STS management.

Although grading is very meaningful for STS prognosis, its use has several limitations. The prognostic value of grade varies between STS subtypes. It is most useful for UPS, LMS and LPS but less informative in other subtypes such as rhabdomyosarcoma, epithelioid sarcoma and alveolar soft part sarcoma. Grading has no prognostic value in malignant peripheral nerve sheath tumours^{41,45}. Additionally, grading scores show high discrepancy rates even among experienced sarcoma pathologists, particularly in the differentiation score^{46,47}. Notably, scores from core needle biopsies should be used cautiously as small biopsies do not represent the whole tumour and tend to undergrade STS lesions, particularly in LMS⁴⁸⁻⁵¹. Overall, grading is not the

only prognostic factor; it has certain limitations, and relying solely on the grade for STS prognostication is not recommended.

1.2.2.2. Prognosis prediction tools: Staging and Nomograms

For many solid cancer types, The American Joint Committee on Cancer/Union for International Cancer Control (AJCC/UICC) tumour, node, metastasis (TNM) staging system is an internationally accepted standard for cancer staging. The system is based on assessing tumour size (T), involvement of regional lymph nodes (N) and presence of distant metastasis (M). Despite the efforts to tailor the TNM staging for STS, the system is suboptimal for prognosis prediction and has low utility in STS clinical practice^{38,52,53}. The system has shortcomings for the staging of STS tumours. Firstly, tumour size is considered a categorical variable by the TNM system, although it has been shown that size might have a nonlinear relationship with prognosis⁵⁴. The STS prognosis greatly depends on a tumour's anatomical location and histological subtypes. Even though a recent 8th AJCC edition includes site-specific systems, it cannot distinguish the prognosis between tumours of different histologies arising at the same anatomical site⁵⁵. To overcome the limitations of the TNM staging system, nomograms were developed for STS.

In medical oncology, a nomogram is a visualisation tool for a statistical model that calculates the probability of a survival outcome event (e.g. OS or recurrence). The nomogram risk calculations consider several clinical or biological prognostic covariates. As prespecified by the model, each covariate has a different weight regarding its contribution to a predicted outcome (based on its importance in predicting the outcome). A score is assigned to each covariate, and the sum of scores is converted to the probability of an outcome. The probability is calculated using a graphic representation or web-based calculators. Site- and histology-specific nomograms for STS were developed by several research groups⁴⁷. To provide accurate predictions and to be incorporated into clinical practice, the nomograms must have good predictive accuracy and be externally validated. A standard method to measure nomogram's accuracy (discrimination) is Harrell's concordance index (C-

index), which calculates the concordance of the predicted survival outcome with the actual observed survival⁵⁶. C-index has a value between 0.5-1.0, and it is accepted that values above 0.7 correspond to a good predictive value of a nomogram. Herein, I will discuss two site-specific nomograms that have been extensively validated and implemented in STS clinical practice.

Nomograms for STS of extremities (eSTS)

Half of STS occur in the extremities³. Local recurrence of eSTS can usually be managed with limb-sparing resections or amputation, whilst metastatic disease in eSTS is fatal⁵⁷. As such, it is important to predict the risk of distant metastases in those patients to provide the best course of treatment.

The two nomograms to predict OS and risk of metastasis (at 5 and 10 years) were developed by Istituto Nazionale dei Tumori (INT) on n = 1452 patients who received surgery for eSTS⁵⁸. The nomogram includes age, tumour size, grading and 9 histological categories as covariables. This study was a collaborative effort between 4 research institutions. The nomogram was validated on three independent external cohorts (n = 420 from the French validation cohort, n = 1436 from the Canadian validation cohort, and n = 444 from the UK validation cohort). Following the validation, nomograms shortly became available on the Sarculator app⁵⁹. In 2022, the Sarculator was validated in n = 1326 patients from 9 US medical centres and C-indexes for 5- and 10-year OS were 0.72 and 0.73, respectively, and 0.72 for distant metastases risk⁶⁰. A separate study validated the Sarculator for the prediction of OS in n = 9738 from the National Cancer Data Base⁶¹. The C-index for the 5-year OS of the whole cohort was 0.73. Overall, the eSTS nomograms are thoroughly validated and have high predictive value. The main limitation of the above nomograms is that they were developed on a patient cohort with primary surgical excisions and, unfortunately, cannot be used in preoperative settings or for patients with metastasis.

Nomograms for STS of retroperitoneum (RPS)

One-fifth of all STS arise in the retroperitoneum, and 90% of all RPS are WDLPS, DDLPS, LMS and solitary fibrous tumours⁵⁸. The prognosis of STS patients with RPS differs from eSTS, even if the same histologies are compared, with RPS having a worse prognosis⁶², highlighting the need for a RPS-specific nomogram. Such nomogram that can predict 7-year OS and DFS for RPS patients who underwent surgery with macroscopically complete resection was developed based on n = 523 patient data from three centres (INT, the University of Texas MD Anderson Cancer Center and the University of California)⁵⁴. External validation was performed on n = 135 patients with a C-index of 0.67 and 0.68 for the OS and DFS, respectively. Gronchi *et al.* RPS nomogram is also available on the Sarcuator app⁵⁹. Consequently, the RPS nomogram was validated in n = 631 patients from 6 sarcoma centres in Europe and North America and n = 103 patients from a cancer centre in Singapore, proving its utility in different ethnic groups^{63,64}. As a result the 8th AJCC edition incorporated the Gronchi's RPS nomogram to improve individual prognosis prediction in RPS patients. Similar to the eSTS nomogram, the RPS nomogram is only accurate in postoperative settings in patients with non-metastasis resectable disease. Moreover, dynamic nomograms for eSTS and RPS were recently developed^{65–67}, which allow prognosis prediction at the baseline and at different time points during the follow-up, refining the prognosis and clinical decision-making should recurrences occur. Overall, nomograms are accurate, personalised prognostic tools, but since they do not cover all disease stages, at present they are not be able to replace the AJCC staging system.

1.3. Clinical and biological overview of select STS subtypes

1.3.1. Leiomyosarcoma

LMS is a smooth muscle malignancy that arises in the retroperitoneum, abdomen, uterus, and walls of blood vessels. Diagnosis of LMS relies on identifying spindle-shaped and elongated cells with complementary IHC

positive staining for smooth muscle markers such as alpha-smooth muscle actin, desmin and h-caldesmon⁶⁸. However, it is not uncommon for poorly differentiated LMS to lose the above markers, and LMS is occasionally misdiagnosed as UPS⁶⁹. LMS has differences in sex prevalence, which is dependent on tumour location. LMS in females is observed in the uterus, retroperitoneum, and inferior vena cava, whilst cutaneous and noncutaneous LMS is often seen in male patients^{70,71}. Different histological variants of LMS are recognised (such as epithelioid and myxoid), highlighting the histological heterogeneity of LMS. Additionally, LMS displays clinical heterogeneity. Due to differences in clinicopathological factors, treatment response and prognosis, LMS is broadly divided into extra-uterine LMS (eLMS) and uterine LMS (uLMS). uLMS more commonly occurs in the 50-59 age group, whilst the incidence peak of eLMS is at 70⁷². uLMS is considered more aggressive than eLMS, with a 50 – 70% risk of recurrence compared to 30 – 40%, respectively^{20,73–75}. Surgical resection remains the gold standard for localised disease independent of the presentation site⁷⁶. During surgery, it is essential to achieve wide negative margins, especially in retroperitoneum, where negative margins are highly prognostic of survival^{39,40}. For uLMS, a hysterectomy is performed, whilst it is advised to avoid morcellation procedures as previous reports showed that morcellation promotes the spread of malignant tissue and poorer survival outcomes⁷⁷. Metastatic disease is treated with systemic chemotherapy as palliative care⁷⁶. There is no established best first-line chemotherapy treatment, but doxorubicin and gemcitabine-based regimens are commonly used⁷⁸. LMS also rarely responds to immunotherapy^{79–81}.

1.3.1.1. Tumour biology

LMS is one of the STS subtypes with a ‘complex’ karyotype and many copy number losses, particularly in tumour suppressor regions³⁵. Mutations and deletions in *RB1*, *TP53*, and loss of *PTEN* are common events in LMS^{35,82,83}. Recently, alterations in DNA damage repair pathways have also been reported in LMS^{82,83}.

One of the notable pathways deregulated in LMS is the phosphatidylinositol 3-kinase (PI3K)/ protein kinase B (AKT)/ mechanistic target of rapamycin (mTOR) pathway, which regulates cell survival^{35,84,85}. It is believed that alterations in tumour suppressor PTEN in LMS activate the PI3K/AKT/mTOR pathway. PI3K is activated downstream of receptor tyrosine kinase (RTK) receptors that activate PI3K upon ligand binding^{86,87}. When active, PI3K phosphorylates phosphatidylinositol (4,5)-bisphosphate (PIP2) to produce phosphatidylinositol (3,4,5)-trisphosphate (PIP3), leading to the accumulation of PIP3 at the plasma membrane. This creates docking sites for AKT and phosphoinositide-dependent protein kinase 1 (PDK1). Consequently, AKT gets phosphorylated by PDK1 and mTOR complex 2 (mTORC2). Active AKT phosphorylates several targets, increasing protein synthesis and cell proliferation. PTEN acts as a negative regulator of the pathway by antagonising the activity of PI3K. In the absence of PTEN, there is no negative switch, and the pathway remains constitutively active.

Sequencing of LMS samples revealed that the *PTEN* gene can be inactivated by point mutations, deep (possibly homozygous) or shallow deletions (possibly heterozygous). The Cancer Genome Atlas Sarcoma (TCGA-SARC) study showed that mutations in *PTEN* were rare and were observed in 5% of 80 LMS patients, 13% of LMS samples had a deep deletion, and 68% had a shallow deletion of the *PTEN* gene³⁵. Similarly, Chudasama *et al.* reported <1% of 49 LMS patients with *PTEN* mutations, whilst 55% of 49 LMS patients had deleterious alterations in *PTEN*⁸². In a separate study, 30% of 165 uLMS and 20% of 120 eSTS patient samples had PI3K pathway alterations, predominantly due to deletions of *PTEN* in 21% and 14% of patients, respectively⁸³.

There is an increasing interest in investigating BRCAness in LMS research^{35,82,83,88–90}. BRCAness is a molecular characteristic observed in certain tumours resembling the biological features of breast cancer gene 1 (*BRCA1*) or *BRCA2* gene mutations. When these genes are mutated or inactivated, it increases the risk of developing certain cancers, such as breast and ovarian cancer⁹¹. *BRCA1* and *BRCA2* are tumour suppressor genes involved in DNA repair mechanisms, particularly in repairing DNA double-

strand breaks by regulation of homologous recombination repair (HRR)^{92,93}. HRR is one of the two pathways to repair double-strand breaks using a homologous DNA template to accurately reproduce the damaged DNA sequence. Interestingly, Chudasama *et al.* reported that 98% of 49 LMS patients showed an Alexandrov-COSMIC mutational signature AC3 associated with defective HRR⁸². In the same study, LMS patients showed loss of function mutations or deletions in the classic *BRCA1* (10% of patients) and *BRCA2* (53%) as well as other genes involved in HRR such as *ATM* (22%) and RAD51 recombinase (*RAD51*) (10%)⁸². Additionally, defects in HRR in LMS were prognostic for inferior progression-free survival (PFS), and non-*BRCA* gene alterations demonstrated a significant negative association with PFS and OS compared to wild-type *BRCA*⁹⁴.

Of interest, the BRCAness might be more pronounced in uLMS compared to eLMS. *BRCA1/2* loss was more common in uLMS (10% of 61 patients) compared to eLMS (1% of 109 patients)⁹⁵. A TCGA-SARC study noted that uLMS have a higher DNA damage response score, which is a reverse phase protein array (RPPA)-based pathway score, than eLMS³⁵. In a separate study, 25% of 121 uLMS showed alterations in the DNA damage response repair (DRR) pathway compared to 14% of 90 eLMS⁹⁴. Similarly, DRR gene alterations were seen in 24% of 165 uLMS and 10% of 125 eLMS in the other study, with *BRCA2* and *RAD51* mutations being the most common⁸³.

The concept of BRCAness has clinical implications as it can help identify patients who may benefit from treatments targeting DNA repair pathways. For example, certain poly ADP-ribose polymerase (PARP) inhibitors have shown effectiveness in treating tumours with BRCAness, as they exploit the underlying DNA repair deficiencies to induce synthetic lethality and tumour cell death⁹⁶. It is important to note that the BRCAness, as it relates to response to PARP inhibitors, is primarily associated with deep deletions or homozygous mutations that result in the loss of both functional *BRCA* or other HRR-related genes alleles. Heterozygous mutations, which involve one normal and one mutant allele, may not exhibit the same degree of sensitivity to PARP inhibitors. This distinction is crucial for designing targeted therapies and understanding their efficacy in different genetic contexts. Preclinical cell

models of uLMS, SK-UT-1 and SK-UT-1b harboured *BRCA2* mutation and were sensitive to PARP inhibitor olaparib⁸². In the context of clinical trials, a phase II trial investigated olaparib in combination with a chemotherapy drug (temozolomide) in n = 22 advanced uLMS⁹⁷. The study showed promising findings with an objective response rate of 27%.

1.3.2. Dedifferentiated Liposarcoma

DDLPS is one of the subtypes of liposarcoma, which has an adipocytic lineage and accounts for 20% of all LPS⁹⁸. Most DDLPS occur *de novo*, although some arise from WDLPS, another LPS subtype⁹⁹. In the case of dedifferentiation, up to 10% of WDLPS can relapse and dedifferentiate into aggressive DDLPS¹⁰⁰. DDLPS can arise in various locations within the body, with the retroperitoneum being the most common site (accounting for 64 – 75% of cases), followed by the extremities (13 – 29% of cases), the trunk (7% of cases), and other less common sites^{99,101}. The ample retroperitoneal space allows DDLPS growth without clinical symptoms, and when DDLPS is diagnosed, it typically has a very large size and has invaded the surrounding organs¹⁰². DDLPS has a dismal prognosis and is characterised by high local relapse rates and the ability to metastasise^{98,100}. The OS and metastasis rates of DDLPS are largely dependent on the grade⁵⁸. Five-year OS rate of intermediate-grade DDLPS is 70% compared to 30 – 40% in high-grade DDLPS. Intermediate and high-grade DDLPS show similar high local recurrence rates of 40% 5 years post-surgery. In contrast, high-grade DDLPS has a much higher propensity for distant metastasis rates at 30% 5 years post-surgery, compared to 10% of intermediate grade⁵⁸. A macroscopically complete surgical resection with negative microscopic margins (R0 resection) is the primary curative option for primary and locally recurrent DDLPS^{103,104}. DDLPS shows limited response to chemotherapy and radiotherapy^{105,106}. Surgery of retroperitoneal DDLPS is often challenging and requires the removal of surrounding organs (e.g. kidney and colon), which become encapsulated or invaded by DDLPS^{104,107}. Consequently, achieving macroscopically complete resection in many cases is difficult. However, when successfully accomplished, complete resection of retroperitoneal DDLPS considerably increases survival rates^{108–110}.

1.3.2.1. Tumour biology

DDLPS has the highest frequency of somatic copy number alterations compared to other most common subtypes of STS³⁵. Amplification events in DDLPS are by far more common compared to deletions. Signature genetic alterations in DDLPS are amplifications of *MDM2* (seen in up to 100% of patients) and *CDK4* (~90% of patients)^{35,111,112}. Other prominent amplifications in DDLPS involve Fibroblast Growth Factor Receptor Substrate 2 (*FRS2*) gene, an inhibitor of adipocyte differentiation Jun Proto-Oncogene (*JUN*), and amplifications of tyrosine kinases such as Discoidin Domain Receptor Tyrosine Kinase 2 (*DDR2*), human epidermal growth factor receptor 3 (*ERBB3*), Neurotrophic Receptor Tyrosine Kinase 1 (*NTRK1*) are frequently observed¹¹³.

MDM2 is an E3 ubiquitin ligase, which acts upstream of tumour suppressor p53, regulating p53 protein stability¹¹⁴. *MDM2* amplification leads to overexpression of MDM2 protein. MDM2 binds to p53 and adds ubiquitin molecules, targeting p53 for proteasomal degradation. As a result, p53 is prevented from activating its target genes involved in cell cycle arrest, DNA repair and apoptosis. Additionally, MDM2 inhibits p53 transcriptional activity by shuttling p53 out of the nucleus, preventing p53 interactions with transcriptional coactivators, and substituting them for transcriptional corepressors. Several MDM2 inhibitors have been developed and have entered clinical trials for DDLPS^{115,116}. Milademetan was tested in the phase I trial, which included n = 53 advanced DDLPS patients¹¹⁷. In phase I, a disease control rate of 59% was achieved, with 2 patients having a partial response and 34 having stable disease. The median PFS was 7.2 months. However, in the phase III clinical study milademetan did not improve PFS compared to the control trabectedin arm (NCT04979442). Another compound, brigimadlin (BI 907828), which prevents the MDM2 and p53 interaction, was investigated in a phase Ia/Ib study, which included n = 39 DDLPS patients (NCT03449381)¹¹⁸⁻¹²⁰. Among 36 evaluable patients with DDLPS, 6 (17%) achieved a partial response, and 26 (72%) had stable disease, giving a disease control rate of 89%. BI 907828 is being further investigated in a phase II/III randomised multi-

centre study comparing it with doxorubicin as a first-line treatment for advanced DDLPS (NCT05218499)¹²⁰.

CDK4 is a key regulator of cell cycle progression¹²¹. CDK4 primarily acts to inactivate Rb. CDK4 forms a complex with cyclin D1/2/3 and phosphorylates Rb, dissociating Rb from the E2F transcription factor. This upregulates the transcription of genes required for cell cycle progression and proliferation. A phase II clinical trial investigated CDK4 inhibitor abemaciclib in n = 30 advanced DDLPS patients and showed favourable results¹²². The interim analysis showed that the PFS at 12 weeks was 76%, and the median PFS was 30 weeks. Tumour reduction was observed in 11 patients, and 1 patient had a partial response. Currently, abemaciclib is in a multicentre randomised placebo-controlled phase III trial in n = 108 advanced DDLPS patients (NCT04967521).

1.3.3. Undifferentiated Pleomorphic Sarcoma

UPS is a high-grade aggressive subtype of STS, lacking any defined line of differentiation with a slight male predominance. Most UPS occurs in the extremities, followed by the trunk¹²³. Historically, UPS used to fall under the diagnosis of malignant fibrous histiocytoma (MFH), the most common STS subtype at the time. Subsequently, with advancements in molecular and histological analysis, MFH was subclassified into several new STS subtypes, including UPS^{124,125}. Current UPS diagnosis is established by ruling out other malignancies with a panel of IHC markers such as keratins (to exclude sarcomatoid carcinoma), S100 proteins (to exclude melanoma), smooth muscle actin and desmin (to exclude myogenic sarcomas), among many others¹²⁶. The management guidelines for primary UPS include surgery involving limb-sparing and imputation in certain extremity cases. Given the broad exclusion-based diagnosis, patients with UPS represent a highly heterogeneous clinical group with highly variable outcomes. The recently reported 5-year OS rate is 50 – 60%, local recurrence-free survival (LRFS) is 55 – 85%, and MFS is 50 – 70%^{123,127,128}. The benefit of systemic therapy is a subject of debate in UPS, and there are limited therapeutic options for

advanced UPS patients^{78,129}. Moreover, patients with advanced metastatic UPS have the shortest median OS, of under 12 months, compared to most common STS patients with advanced disease¹²⁹. Poor prognosis and ineffective treatment stem from a poor understanding of molecular mechanisms underlying UPS tumorigenesis.

1.3.3.1. Tumour biology

Oncogenic mechanisms in UPS are not well-explored. Genomic characterisation of UPS highlighted significant genomic instability and many copy number alterations, including genome doubling^{35,130,131}. Frequent mutations involving *TP53*, *RB1*, *ATRX* have been considered the driver mutations in UPS. The whole genome duplication (WGD) event was frequently observed in UPS patient samples and was proposed to be involved in UPS tumorigenesis¹³¹. At least one WGD was detected in 89% of 52 UPS samples, and two WGD events occurred in 19% of samples. The Hippo pathway may drive a subset of UPS tumours, as recurrent amplifications in pathway genes have been identified^{35,132}. Vestigial like family member 3 (*VGLL3*) and yes1 associated transcriptional regulator (*YAP1*), which encode cofactors of the TEAD family of transcription factors and promote proliferation, were amplified in ~10% and 5% of 44 UPS cases, respectively³⁵.

In addition to tumour cell biology, the immune microenvironment plays a considerable role in UPS pathobiology. RNA sequencing, IHC and proteomic studies observed elevated immune activity in UPS compared to other STS subtypes^{133–136}. Higher levels of infiltrating lymphocytes or higher immune infiltration scores based on gene expression signatures were correlated with improved survival outcome measures^{35,137,138}. For instance, higher tumour infiltration by T-cells was associated with improved OS and DFS in a cohort of primary UPS (n = 46)¹³⁷. Moreover, the same study classified primary UPS cases based on the expression levels of 9 immune biomarkers, revealing three groups: immune low, immune intermediate and immune high. The immune high cluster had a significantly improved OS and DFS compared to the immune intermediate and immune low clusters¹³⁷. Similarly, RNA sequencing of UPS samples (n = 25) identified an ‘immune-low’ group A and ‘immune-

high' group B¹³⁸. Group B had significant enrichment of inflammatory and interferon-gamma response pathways, as well as gene signatures corresponding to cytotoxic T cells, T-helper cells, monocytes, natural killer cells, dendritic cells, memory B cells and regulatory T cell infiltration. The same clusters were recapitulated when the classification was applied to the TCGA-SARC cohort, and the 'immune-high' group B had a significantly longer OS compared to the 'immune-low' group A¹³⁸. The above-mentioned studies highlight the heterogeneity in immune infiltrates in UPS, which might explain heterogeneity in the survival outcomes of UPS patients. Table 1.2 provides a summary of biological and clinicopathological features of LMS, DDLPS and UPS, which were discussed in Section 1.3.

Table 1.2 Summary of key biological and clinical features of LMS, DDLPS and UPS.

eLMS: extra-uterine LMS; uLMS: uterine LMS.

	LMS	DDLPS	UPS
% of STS diagnoses	~12%	~6%	~11%
Key genetic lesions	Mutations and deletions in <i>RB1</i> , <i>TP53</i> , and loss of <i>PTEN</i>	Amplifications of <i>MDM2</i> , <i>CDK4</i> and <i>FRS2</i>	Mutations in <i>TP53</i> , <i>ATRX</i> and <i>RB1</i>
Anatomical sites	Abdomen, retroperitoneum, large blood vessels, uterus and thigh region	Retroperitoneum, extremities and trunk	Extremities, trunk and head/neck
Key clinical features	The 5 year distant recurrence rate is 30-40% for eLMS and 50-70% for uLMS, 10-20% local recurrence rate	10-15% distant recurrence rate and 40-60% local recurrence rate	High-grade and aggressive, local recurrence rate is 15-45%, distant recurrence is 30-50%
Standard of care treatments	Surgery for localised disease, advanced disease is treated with palliative chemotherapy (doxorubicin, gemcitabine and ifosfamide)	Surgery, sometimes with additional radiotherapy or chemotherapy	Amputation or limb-sparing surgery, sometimes with radiotherapy
Prognostic factors	Anatomical location, age, tumour size, grade and surgical margins	Anatomical location, surgical margins, tumour grade and size	Anatomical location, age, tumour size, grade and surgical margins

1.4. Tumour microenvironment

Cancer cells cannot exist in isolation. Instead, they interact with and corrupt non-malignant surrounding cells, creating a supportive niche¹³⁹.

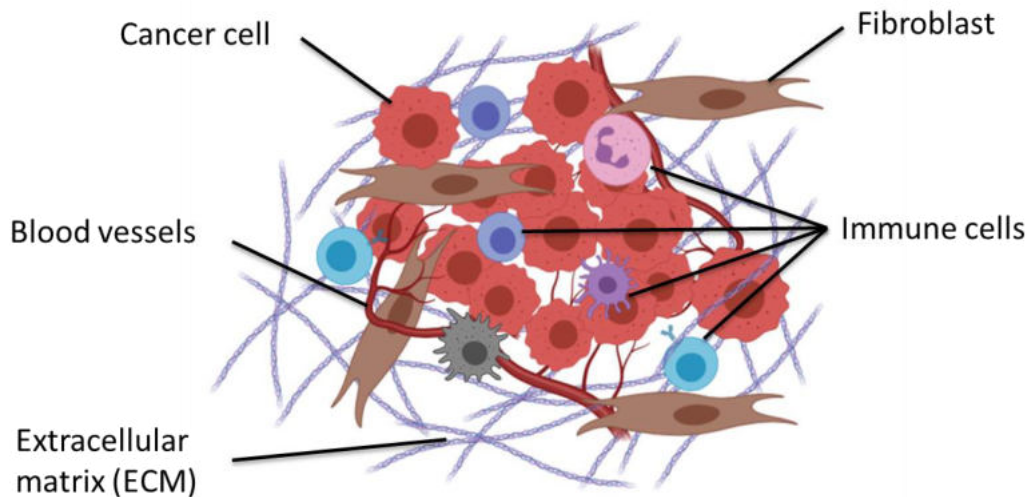


Figure 1.1 Composition of the tumour microenvironment (TME).

A tumour comprises not just cancer cells but rather represents a diverse assembly of infiltrating and resident host cells, along with secreted factors and the extracellular matrix. In the TME, cancer cells co-opt normal neighbouring cells to orchestrate angiogenesis, proliferation, invasion and metastasis, enabling cancer progression. The TME composition varies between tumour types but includes endothelial cells, fibroblasts and adipocytes. Tumours also become infiltrated with diverse adaptive and innate immune cells that can perform both pro- and anti-tumorigenic effects. Created with BioRender.com.¹⁴⁰

That niche, called the tumour microenvironment (TME), refers to the non-malignant stromal cell and acellular components within and around a tumour (Figure 1.1). Each tumour type has a unique TME composition, commonly including immune cells, blood vessels, fibroblasts, collectively referred to as stromal cells, extracellular matrix (ECM), and signalling molecules. The TME is exploited and shaped by cancer cells to facilitate tumour development, progression, and resistance to treatment¹⁴¹. Specifically, TME coordinates an ‘angiogenic switch’ to establish new blood vessels and fulfil the growing tumour’s increasing demand for oxygen and nutrients^{142,143}. Tumour cells can secrete molecules to render the TME immunosuppressive, thus evading the anti-tumour immune responses¹⁴⁴. Tumour and stromal cells can remodel and

degrade the ECM barrier, enabling local tumour invasion and metastatic dissemination¹⁴⁵. Understanding the dynamic interplay between cancer cells and their microenvironment is crucial for developing effective cancer therapies that not only target the tumour cells but also the supportive microenvironment to hinder tumour progression.

1.4.1. Immune TME in STS

Immune cells are a critical component of the TME and is comprised of infiltrating immune cells such as tumour associated neutrophils, tumour associated macrophages, dendritic cells and tumour infiltrating lymphocytes (TILs), including T cells, B cells and natural killer cells. T cells undergo a crucial developmental process, generating T cell receptors that specifically recognize peptide antigens derived from intracellularly degraded proteins¹⁴⁶. Within the TME various T cell populations can influence tumorigenesis. The CD4 and CD8 cell surface co-receptors delineate two broad classes of T cells, each characterised by distinct effector mechanisms¹⁴⁷. Cytotoxic CD8+ T cells detect abnormal tumour antigens on cancer cells and eliminate neoplastic cells. In contrast, CD4+ T cells mediate anti-tumour immunity by activating cytotoxic T cell and B cells. A specific subset of CD4+ T cells, marked by FoxP3 expression acts to dampen the immune response¹⁴⁸. Furthermore, the immune TME contains soluble immune factors such as cytokines and chemokines, along with immune molecules present on tumour and stromal cells, including the immune checkpoint proteins¹⁴⁹. Collectively, the immune TME components determine tumour immunity.

Tumours with high somatic mutation rates are thought to have elevated expression of immunogenic neoantigens, increasing the likelihood of anti-tumour immune response. STS are traditionally considered immunologically quiet/cold due to low immune infiltrate and low tumour mutational burden compared to other cancer types¹⁵⁰. Despite that a subset of STS patients show favourable responses to immunotherapy^{79,81,151}, and recent studies show that a subset of STS patients present with high TIL infiltrate, challenging the concept of all STS tumours are immunologically cold.

A recent study utilised IHC to assess T cell infiltration and immune checkpoint molecule expression within tissue microarrays (TMA) of > 1000 tumours across 22 soft tissue and bone sarcoma subtypes¹⁵². The cohort included different histologies across translocation and non-translocation-associated STS as well as reference cohorts of melanoma and carcinomas. In this study translocation-associated STS subtypes, such as synovial sarcoma, typically had fewer CD8+, CD4+, FoxP3+ and CD56+ (natural killer) TILs compared to STS with greater genomic complexity. Out of the non-translocation-associated STS, DDLPS, UPS, epithelioid sarcoma and myxofibrosarcoma (MyxFS) presented with highest number of the TILs. Notably, at least 25% of non-translocation-associated STS had TIL densities greater than the median values for the reference cohorts, which include cancer types with recognised sensitivity to immunotherapy¹⁵². Moreover, a higher levels of TILs showed an association with improved OS (HR = 4.5, p = 0.034) with a cohort of pleomorphic STS (DDLPS, LMS, MyxFS, UPS and angiosarcoma).

Similarly, another IHC-based study profiled TMAs containing 192 tumours across common and rare STS subtypes for TIL infiltrate¹⁵³. In this study ~50% of STS showed CD3+ (total T cells) infiltration and histology-based differences in CD3+, CD4+, CD8+ and FoxP3 TIL. MyxFS and UPS tumours showed higher total T CD3+, helper T CD3+CD4+, cytotoxic T CD3+CD8+ and regulatory T CD3+CD4+FoxP3+ cells compared to synovial sarcoma and liposarcoma, however not all comparisons reached significance¹⁵³. Regulatory T cells were the only TIL marker significantly associated with outcome, with a higher score being associated with poor LRFS. Other IHC-based investigation of TILs in STS showed divergent results in terms of TIL levels and outcome^{154,155}. Generally, comparability of findings between studies is limited due to variations in the IHC method and scoring, histological subtypes examined, the immune components examined, highlighting the difficulties of working with a heterogeneous and rare disease. Nevertheless, the IHC studies alluded to the presence of a subgroup of STS patients with high TIL infiltrate.

In addition to IHC-based studies, immune deconvolution strategies, which estimate immune infiltrate based on bulk transcriptomic data, were employed

to analyse the immune TME in STS. Immune cell deconvolution uses reference gene expression profiles of purified immune cell subsets to infer the relative abundance and cell type within a mixed sample¹⁵⁶. The cellular signature score is then used as a proxy for immune infiltrate. TCGA-SARC study performed RNA sequencing in a cohort of 206 STS, including UPS, DDLPS, LMS and synovial sarcoma and assessed immune infiltrate through immune deconvolution³⁵. The findings aligned with large IHC-based studies in that greater levels of immune cell infiltration was generally seen in genomically complex STS compared to histologies with a simple genotype. Specifically, higher macrophages and higher CD8 T cell scores were observed in DDLPS and MyxFS/UPS compared to synovial sarcoma samples³⁵. Moreover, scores for specific immune subsets showed an association with disease-specific survival (DSS) within histological subtypes. For instance, a high natural killer cell score was prognostic for improved DSS in LMS and MyxFS/UPS cohorts; high CD8+ score was associated with improved DSS in uLMS, whilst a high T helper 2 score was prognostic for poor DSS in DDLPS cohort.

A separate study developed an immune-based classification by performing immune deconvolution of gene expression from 608 LMS, DDLPS and UPS samples from 4 publicly available STS datasets¹³⁵. In this study LMS, DDLPS and UPS cases were assigned to one of five sarcoma immune classes (SICs) A, B, C, D and E which showed variation in estimated immune activity. From A to E, SICs were characterised by an increase in immune activity inferred from the abundance scores for 8 immune cells, endothelial cells, fibroblasts, immune checkpoints, T cell activation and T cell survival genes and the presence of tertiary lymphoid structures (TLS). TLS are lymphoid organs arising in non lymphoid tissues, commonly at the sites of chronic inflammation, including in tumour tissues¹⁵⁷. SIC A represented an immune desert population, SIC B was an immune-low group, SIC C was a highly vascularised group, SIC D a heterogeneous immune-high group and SIC E was a an immune-high group characterised by the presence of TLS¹³⁵. SIC A comprised mainly of LMS cases, SIC B of LMS and UPS, half of SIC C tumours were DDLPS, whilst SIC D and E had an equal representation of LMS, DDLPS and UPS¹³⁵. Interestingly, stratification of patients by SICs showed that patients

from the immune desert SIC A subgroup had significantly shorter OS compared to subgroup SIC D ($p = 0.048$) and SIC E ($p = 0.025$). After adjusting for established prognostic factors such as tumour grade and patient age, the more immune-high SIC D (HR = 0.4, $p = 0.011$) and SIC E (HR = 0.4, $p = 0.029$) were significantly associated with improved OS compared to SIC A.

These IHC-based and gene expression studies indicate that there is a subset of genomically complex STS of mixed histologies with an immunologically hot phenotype that are associated with more favourable survival outcomes.

1.4.2. The extracellular matrix

The ECM is a complex three-dimensional (3D) network of biomolecules comprising polysaccharides and proteins that surround and support cells within tissues¹⁵⁸ (Figure 1.2).

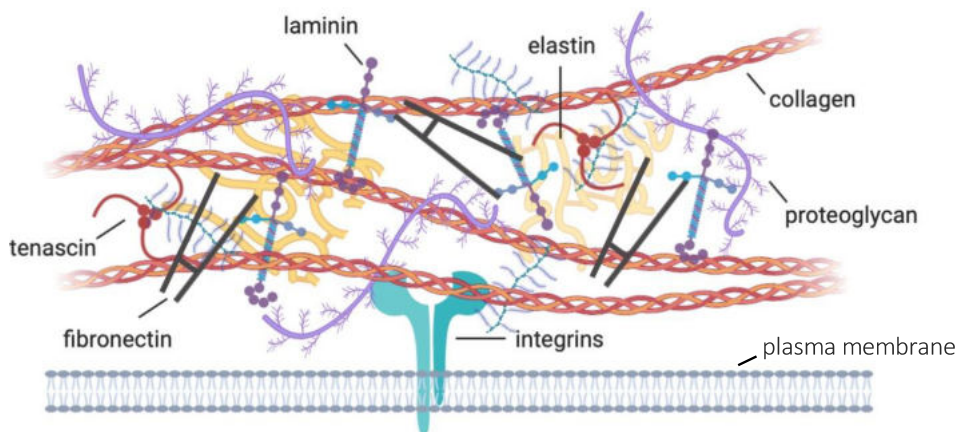


Figure 1.2 A simplified diagram of the ECM.

An interconnected network of collagen fibres, proteoglycans and glycoproteins (e.g. laminin, fibronectin, elastin and tenascin). An Integrin receptor is shown to bind to a collagen fibre. Figure created with BioRender.com.

However, the influence of the ECM on cells and tissues extends beyond the structural and mechanical support (Figure 1.3).

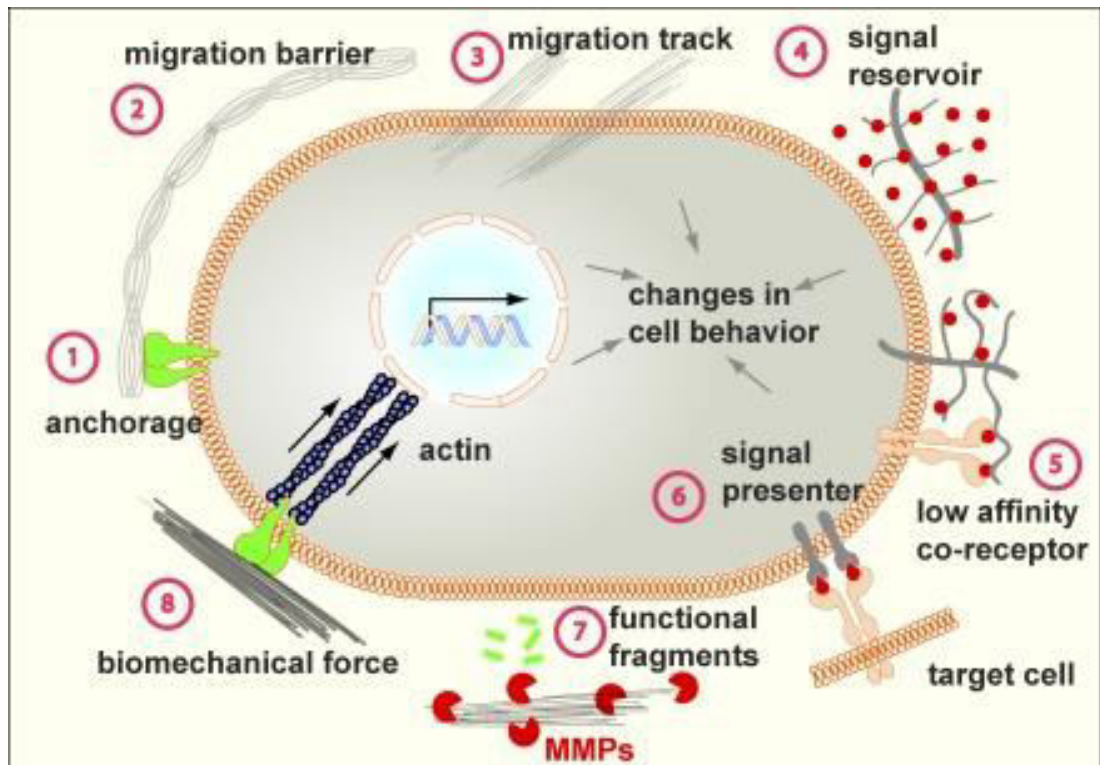


Figure 1.3 Functional roles of the ECM.

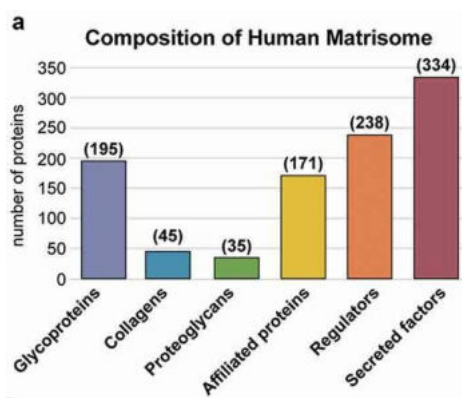
The ECM serves diverse functions and provides biochemical and biomechanical signals that induce cell behaviour changes. The versatile functions of the ECM in cancer are discussed throughout this chapter. Stage 1: the ECM is a supportive scaffold, which allows cell adhesion (anchorage) via cell adhesion molecules (in green) to maintain cell position and tissue structure. Stage 2 and 3: depending on the context, the ECM can prevent (2) or promote (3) the cell migration. Stage 4: the ECM plays a crucial role in modulating the distribution of growth factor signalling molecules by sequestering them and creating concentration gradients. Stage 5 and 6: certain ECM components, such as some proteoglycans, can act as co-receptors (5) by binding signalling molecules and cellular signalling receptors or presenting (6) signalling molecules to other cells. Stage 7: some ECM components can be processed by proteases (e.g. metalloproteinases (MMPs)) to release biologically active fragments that initiate cell signalling. Stage 8: the biomechanical characteristics of the ECM, such as its stiffness, exert significant effects on various cellular behaviours, including cell differentiation. Reproduced from Lu *et al.* (2012) under CC BY-NC-SA 4.0 licence ¹⁵⁹.

Biochemical and biomechanical cues from the ECM modulate cellular activity by activating or suppressing intracellular signalling in response to changes in the cell's environment¹⁶⁰. For instance, ECM signals can promote cell proliferation and migration, enabling tissue regeneration and wound healing^{161,162}. Likewise, stiffness and rigidity of the ECM can induce stem cells

to differentiate into specialised cell types, crucial for tissue formation and maintenance¹⁶³. The ECM is a dynamic structure that is constantly remodelled but tightly controlled during normal physiological processes¹⁶⁴. However, in the context of cancer, the ECM is dysregulated by disproportional matrix production and deposition¹⁶⁵, and changes occur to ECM biochemical and mechanical properties¹⁶⁶. For example, mammographic density, which is associated with increased collagen deposition, is a strong breast cancer prognostic factor¹⁶⁵. Mechanical properties, such as stiffness of the ECM, can be influenced by ECM crosslinking enzymes that stiffen the ECM, upregulating oncogenic signalling and driving cancer cell proliferation¹⁶⁶.

1.4.2.1. **Matrisome definition**

The ECM composition significantly differs between normal and malignant tissue¹⁶⁷, cancer types¹⁶⁸ and malignancy stages^{169–173}. The complexity and diversity of the ECM composition arise from various types of macromolecules present in the ECM and their remodelling during cancer. In the search for a better characterisation of the ECM, the Hynes laboratory has defined the ‘matrisome’, which is an *in-silico* prediction of all theoretical proteins present in the extracellular space¹⁷⁴. The ECM proteins contain repetitive and characteristic domains¹⁷⁵, and a bioinformatic sweep of the genome for the 55 specific domain sequences enabled the identification of ~ 300 candidate genes encoding ECM proteins^{174,176}. The list was further refined by excluding 20 specific domains (e.g. tyrosine kinase and phosphatase domains) and transmembrane domains incompatible with the definition of ‘ECM’ proteins¹⁷⁴.



Core matrisome Matrisome-associated

b

Class	Example
Glycoproteins	Fibronectin, Elastin, Laminins
Collagens	Collagens
Proteoglycans	Perlecan, Decorin, Prolargin
Affiliated proteins	Annexins, Galectins, Hemopexin
Regulators	Cathepsins, Serpins, MMPs
Secreted proteins	Growth factors, S100 proteins

Figure 1.4 Composition of human *in silico* matrisome.

A) Breakdown of human matrisome by 6 classes, indicating the number of components in each class. Glycoproteins, collagens and proteoglycans constitute the 'core matrisome', while affiliated proteins, regulators and secreted factors are 'matrisome-associated'. B) Selected examples of proteins from each matrisome class. MMPs: matrix metalloproteinases. Adapted from Krasny and Huang (2021) under CC BY 4.0 licence¹⁷⁷.

Knowledge-based annotation allowed the definition of subclasses of the matrisome, including 45 collagen subunits, 35 proteoglycans, and 196 glycoproteins¹⁷⁸ (Figure 1.3). This 'core matrisome' represents the foundation of the structural ECM scaffold. The structural ECM scaffold is remodelled by and can bind other ECM-associated molecules. Therefore, the Hynes lab additionally defined the list of ~ 700 'matrisome-associated' proteins: a separate list of domains was defined to identify 171 ECM-affiliated proteins, 238 ECM regulators and 334 secreted factors^{174,178} (Figure 1.3). In brief, ECM-affiliated proteins share architectural or biochemical similarities with ECM proteins, ECM regulators include ECM remodelling, ECM crosslinking enzymes, their inhibitors and regulators, whilst secreted factors include small molecules expected to bind to the ECM, such as growth factors and cytokines.

The definition of the matrisome is deliberately broad and was designed to include all potential ECM-related proteins; however, not all constituents have been experimentally validated to have ECM-related functions or detected in tissues^{179,180}. As such, it is essential to thoroughly verify the functional roles of candidate proteins identified through matrisome analysis, particularly within specific tissue contexts.

Collagens

Collagens are the most abundant proteins in the mammalian ECM. Collagens provide tensile strength and maintain tissue structure. Forty-five human genes encode 28 distinct types of collagens, which are subdivided into fibril-forming collagens (e.g. collagen I), fibril-associated collagens with interrupted triple helices (FACIT) (e.g. collagen IX), network-forming collagens (e.g. collagen IV), membrane collagens (e.g. collagen XIII) and multiplexins (e.g. collagen XV)^{174,181}.

Collagen structure

The common feature of all collagens is the presence of three polypeptide alpha chains (α chains). These α chains can be either all identical (e.g. three $\alpha 1(\text{II})$ in collagen II), two different (e.g. two $\alpha 1(\text{I})$ and one $\alpha 2(\text{I})$ in collagen I) or all different (e.g. $\alpha 1(\text{IX})$, $\alpha 2(\text{IX})$, $\alpha 3(\text{IX})$ in collagen IX). The α chains have a specific and repeating sequence of glycine (Gly)-X-Y amino acids, where X is often proline, and Y is hydroxyproline¹⁸². This repeating motif enables α chains to wind around each other and assemble into a triple helix. The extent to which the triple helix is present in different collagen types can vary significantly. For instance, triple helix contributes to 96% of the collagen type I structure but only 10% of collagen type XII¹⁸³. Collagen molecules with a higher percentage of triple helix structure tend to provide greater tensile strength and mechanical stability, while those with a smaller proportion often have more specialised functions related to the organisation of the ECM and protein-protein interactions. In addition to the signature triple helix domain, collagens can contain non-collagenous domains, which participate in structural assembly and confer biological activities to collagens^{183,184}. Fibronectin III, Kunitz, thrombospondin-1, and von Willebrand domains are the most common non-collagenous domains. The von Willebrand domain participates in protein-protein interactions, the Kunitz domain is involved in structural assembly as it is cleaved during the maturation of collagen VI and VII^{185–187}.

Fibril-forming collagens

Fibrillar collagens represent the most abundant collagen class in the body, with collagen types I, II, III, V, XI, XXIV and XXVII belonging to this class¹⁸⁸. Fibrillar collagens assemble in long, cable-like fibril structures and are the primary collagen types that provide tissue tensile strength. All fibrillar collagens are synthesised as soluble precursors, procollagens^{188,189}. Procollagens are assembled intracellularly from three α chains into the helical structure. At the end of each procollagen (N- and C-terminal ends), there are propeptides, which do not have the helical structure and maintain the structure of the procollagens. Extracellularly, procollagens are further assembled into higher-order structures called fibrils. The process of collagen maturation involves the cleavage of propeptides by proteases. Mature collagen monomers stagger

and self-assemble into fibrils. Then, the fibrils are stabilised by the covalent cross-linking of collagen's ends (telopeptides) by lysyl oxidase (LOX) enzymes. Those cross-links are essential for the mechanical properties of fibrillar collagens.

Fibril Associated Collagens with Interrupted Triple helices (FACITs)

In contrast to fibrillar collagens, FACIT collagens (collagen types IX, XII, XIV, XIX) do not form long fibrils. FACIT structure contains two or three collagen helix domains interspaced with non-collagen domains^{183,190}. Those collagens are found to be covalently bound to fibrillar collagens. For instance, FACIT collagen IX is linked to the surface of cartilage collagen II¹⁹¹. FACIT collagens are considered 'molecular bridges' important for the stability and organisation of the ECM¹⁹⁰.

Network-forming collagens

Another class of collagens is the network-forming collagens. The unique characteristic of this class of collagen is the retention of the N- and C-terminal domains of the collagen helix¹⁹², which in fibrillar collagens are removed in post-translational processing. The presence of these domains allows the head-to-head linkage of collagen fibres and their arrangement in various kinds of networks. Collagen IV is the main network-forming collagen, forming a chicken wire-type network¹⁸¹, and is localised exclusively to the basement membrane, a specialised form of the ECM providing compartmentalisation of tissues in the body¹⁹³.

Membrane collagens

Four collagens (collagen types XIII, XXIII, XXV and XVII) are transmembrane molecules composed of a cytoplasmic domain, a transmembrane domain and several extracellular triple-helical domains¹⁹⁴. Proteolytic processing of the extracellular domains results in the shedding of a soluble short version of membrane collagens. For instance, collagen type XVII is cleaved by disintegrin and metalloproteinases (ADAMs), and increased collagen XVII shedding was associated with impaired keratinocyte migration¹⁹⁵.

Multiplexins

Multiplexin collagens type XV and XVIII are multidomain collagens with a thrombospondin domain and multiple triple-helical domains, interrupted with non-collagenous domains¹⁸³. Multiplexins can be processed by proteases such as cathepsins and matrix metalloproteinases (MMPs) to generate biologically active fragments¹⁹⁶. Proteolytic fragments of collagen type XVIII, endostatin, and a fragment of collagen type XV, restin, can inhibit angiogenesis and tumour growth^{197–199}.

Collagen receptors

The biological role of collagens extends beyond the structural support and tensile strength to tissues. Collagens have several adhesion sequences in the triple helical domain and act as ligands for several cell receptors^{200,201}. Upon interaction with the cellular receptors, collagens transduce extracellular signals into the cells to regulate cell behaviour. Collagens interact with specific integrins, which are cellular receptors that lack kinase activity, and those will be further discussed in Section 1.5. Fibrillar and non-fibrillar collagens also activate discoidin domain receptors DDR1 and DDR2 and transduce signals via DDR kinase domains²⁰¹. DDRs have some shared specificity, e.g., both are activated by collagen types I–III and collagen V^{202,203}. Additionally, DDRs have differential ligand preferences, e.g., only DDR1 binds collagen type IV, while DDR2 is activated by collagen type X^{204,205}.

Glycoproteins

Glycoproteins consist of polypeptides with covalently attached oligosaccharide chains. During posttranslational modification, sugar chains are attached to specific amino acids in a process called glycosylation²⁰⁶. N-glycosylation adds carbohydrate molecules, often based on N-acetylglucosamine, to aspartic acid residue (N-linked glycoproteins). O-glycosylation involves the addition of oligosaccharides based on N-acetylgalactosamine to serine or threonine amino acids (O-linked glycoproteins). Glycoproteins have multifaceted roles, influencing the ECM assembly, organisation, and cell-matrix interactions^{176,207}. The most well-studied glycoproteins are laminins and fibronectin. Laminins are integral parts of the basement membrane, the supporting framework on which

epithelial and endothelial cells grow, and that surrounds muscle and fat cells²⁰⁸. Specific domains of glycoproteins, such as the fibronectin type-III domain, promote cell adhesion²⁰⁹. Glycoproteins can bind growth factors²⁰⁷. Bound growth factors can be stored within the ECM, acting as reservoirs that can be released when needed, often through proteolysis. Alternatively, ECM-bound growth factors can be presented to nearby cells. For instance, glycoproteins fibronectin and vitronectin can bind hepatocyte growth factor (HGF)²¹⁰. When fibronectin or vitronectin binds to HGF, it can form complexes with the HGF receptor, mesenchymal epithelial transition (Met) receptor, and integrins, and this interaction enhances cell migration.

Proteoglycans

Proteoglycans are complex molecules consisting of the core protein and glycosaminoglycan (GAG) chains^{211,212}. GAGs are long, linear polysaccharides with repeating disaccharide (two sugar units) structures. Different arrangements and types of disaccharide units give rise to several GAGs, such as heparin/heparan sulfate, chondroitin sulfate/dermatan sulfate, keratan sulfate, and hyaluronic acid²¹³. Besides hyaluronic acid, GAG chains are heavily modified with sulfate groups and covalently linked to the core protein²¹⁴. Hyaluronic acid interacts noncovalently with some proteoglycans via hyaluronan-binding motifs. Generally, sulphate groups of GAGs confer proteoglycans a negative charge, which enables proteoglycans to bind and organise water molecules and cations such as calcium²¹⁵. These properties of proteoglycans are important for space-filling and lubrication functions. In joints, for example, proteoglycans are a key component of the synovial fluid, which lubricates the surfaces of bones²¹⁶. Overall, various combinations of the core protein and the attached GAG chains give proteoglycans distinctive biological properties.

Proteoglycans as co-receptors

In contrast to signalling receptors, which transduce signals via catalytic domain or adaptors bound to their cytoplasmic domain, signalling co-receptors do not directly participate in the signal transduction²¹⁷. Instead, co-receptors facilitate signalling via ligand binding and associating with their corresponding signalling

receptors. Signalling co-receptors are integral components of major signalling pathways such as fibroblast growth factor (FGF), Hedgehog, HGF, transforming growth factor- β (TGF- β) superfamily ligands and vascular endothelial growth factor (VEGF) pathways. Several members of the proteoglycan family act as co-receptors²¹⁸. Heparan Sulfate Proteoglycan 2 (HSPG2), sometimes called perlecan, is known to bind to FGF members, which regulate angiogenesis²¹⁹. HSPG2 potentiates the high-affinity binding of FGF-2 to CHO cells and to soluble fibroblast growth factor receptors (FGFRs)²²⁰. It also strongly increases the mitogenic activity of FGF-2 and induces FGF-2-mediated angiogenesis in vivo.

Other members of the proteoglycan family, such as decorin, have been shown to act as tumour suppressors by inhibiting RTK^{221,222}. Decorin can bind to the epidermal growth factor receptor (EGFR) at a region overlapping with the native ligand's binding site, the EGF-binding domain²²³. Decorin triggers the phosphorylation of EGFR and activation of the mitogen-activated protein (MAP) kinase signal pathway^{224,225}. Consequently, this upregulates p21 expression, a potent inhibitor of cyclin-dependent kinases, resulting in the eventual arrest of the cell cycle²²⁵. Alternatively, decorin induces internalisation of EGFR, reducing the number of signalling EGFR molecules²²⁶.

ECM-affiliated factors

ECM-affiliated factors are a broad class of proteins, some of which share structural characteristics with glycoproteins and proteoglycans. Proteins affiliated to this class include glypicans, syndecans, annexin, C-type lectin domain proteins, galectins, mucins, plexins and some complement system components. While some members of this group have established physical interactions with the ECM and carry out roles directly related to ECM functions, the roles of others within the ECM context remain uncertain. For instance, galectins bind glycoproteins such as laminin and fibronectin in the ECM and can regulate integrin-mediated cell adhesion²²⁷. In contrast, C-type lectin-like domain (CTLD) proteins represent a family of transmembrane pattern recognition receptors, expressed primarily by myeloid cells. CTLD belong to

the matrisome class because their structural domains are similar to lectin-binding domain of some proteoglycans, such as versican²²⁸. It is worth noting that certain ECM-affiliated proteins have been classified as part of the matrisome due to their repeated identification in mass spectrometry analyses of insoluble ECM fractions isolated from tissues. However, the specific functions of these proteins within the ECM environment have not been definitively confirmed¹⁷⁴. As such, ECM-affiliated proteins identified through bioinformatic analyses should be interpreted cautiously and their function should be experimentally validated.

ECM regulators

ECM regulators are a sub-class of matrisome, including various ECM-remodelling (proteolytic and cross-linking) enzymes and their regulators. The class includes several families of proteases, such as MMPs and their inhibitors, tissue inhibitors of MMPs (TIMPs), ADAMs, cathepsins and their inhibitors cystatins, and serpin family of protease inhibitors. Other members include procollagen-lysine,2-oxoglutarate 5-dioxygenase (PLOD) enzymes, LOX and lysyl oxidase like-enzymes 1-4 (LOXL), which are involved in collagen maturation and crosslinking, and prolyl 4-hydroxylases (P4Hs) which are essential for collagen synthesis. ECM regulators are responsible for the breakdown, modification and synthesis of ECM components. Those processes define the ECM remodelling and alter the composition, biomechanical properties and organisation of the ECM within tissues. Many components of this class are found to be deregulated in cancer, leading to pathologic ECM remodelling^{229–232}. For instance, increased activity of LOX enzyme, which cross-links collagen type I, can induce matrix stiffness and promote cancer cell proliferation²³³ and invasiveness^{234,235}.

The main class of ECM-degrading enzymes is the MMP family. The MMP family comprises 23 enzymes classified based on substrate specificity and domain structures²³⁶. MMP-1, MMP-8, and MMP-13 are known for their ability to degrade fibrillar collagen. MMP-2 and MMP-9 cleave denatured collagens (gelatins), collagen IV, laminin and fibronectin, while MMP-3 and MMP-10 have a broader range of substrates, including proteoglycans, fibronectin, and

laminin^{236,237}. Membrane-type MMPs (MMP-14 and MMP-15) are membrane-anchored enzymes with collagen specificity. MMPs are typically synthesised as inactive zymogens or proenzymes (pro-MMPs). Pro-MMPs undergo post-translational modifications and activation steps. In some cases, other proteases, such as plasmin, MMP-3 and MMP-14, can activate pro-MMPs by cleaving off their inhibitory pro-peptide domain²³⁸. Prevention of excessive degradation is achieved through endogenous TIMPs that bind to active MMPs and inhibit their enzymatic activity. In addition to directly reorganising the ECM architecture by degrading ECM proteins, MMPs modulate cell behaviour by liberating bioactive ECM fragments, growth factors and cytokines²³⁹. For instance, MMP-2 and MMP-9 expose a cryptic epitope within collagen type IV, which promotes angiogenesis^{240,241}. MMP-7 and MMP-14 can release an anti-angiogenic fragment, endostatin, from collagen type XVIII²⁴².

Secreted factors

The ECM serves as more than a structural framework; it also functions as a reservoir for secreted factors, such as latent growth factors, cytokines and chemokines¹⁶⁴. As such, secreted factors are indirectly involved in ECM-related functions and are considered a separate matrix class.

TGF- β is a multifunctional growth factor involved in cell growth, differentiation, immune regulation, and wound healing. TGF- β is often stored in latent or inactive form within the ECM, bound to latent TGF- β -binding proteins (LTBPs)²⁴³. The LTBPs bind to other ECM proteins, such as fibronectin and fibrillins²⁰⁷. TGF- β activation depends on the proteolytic degradation of fibrillins, LTBPs or by interaction with another ECM glycoprotein, thrombospondin^{244,245}.

The release and activation of growth factors bound to the ECM molecules also require the action of proteases, such as MMPs²³⁹. Another example of a growth factor sequestered by the ECM is VEGF. VEGF is an important factor in angiogenesis because it enhances vascular permeability and promotes new vessel growth. VEGF isoforms are bound to HSPGs, fibrin and fibronectin and kept inactive within the ECM²⁴⁶. Upregulation of MMP-9 activity in a transgenic

cancer model was shown to release VEGF and induced angiogenesis and tumour progression²⁴⁷.

1.4.2.2. Basement membrane

Basement membranes are sheets of the specialised ECM and are considered the most ancient type of ECM²⁴⁸ (Figure 1.5).

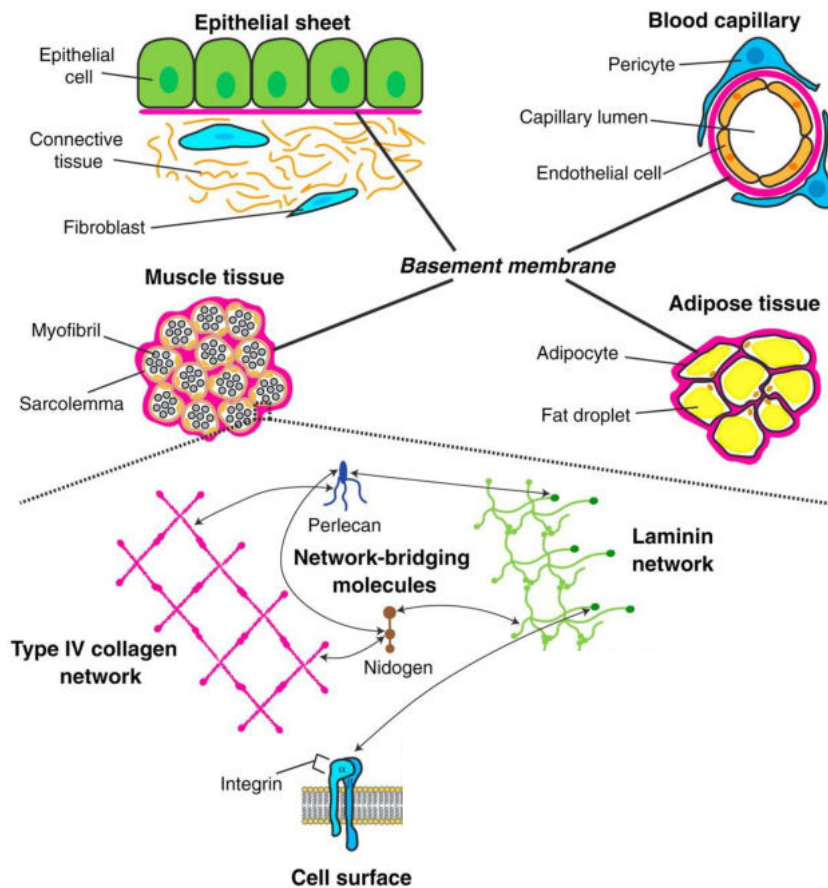


Figure 1.5 Localisation and composition of the basement membrane.

The top part of the figure shows the typical localisation of the basement membrane underlying epithelial and endothelial cells and surrounding muscle and fat cells. The bottom part shows the fundamental constituents of the basement membrane. Those include self-assembling collagen type IV and laminin networks interconnected via nidogen and perlecan (proteoglycan HSPG2) molecules. The laminin network is typically the first layer directly in contact with the cellular receptors, integrins, and the collagen type IV network is layered on top of laminin. Adapted from Jayadev and Sherwood (2017) with permission from Elsevier²⁴⁹.

Basement membranes underly epithelial cells, separating epithelial and connective tissue, and underly endothelial cells in blood vessels²⁴⁹.

Additionally, basement membranes encapsulate muscle and fat cells and form blood-brain and kidney filtration barriers^{250,251}. Fundamentally, basement membranes are composed of two independent networks, one formed by collagen IV and the other by laminin. The two networks are connected by other ECM proteins, primarily glycoprotein nidogen and proteoglycan HSPG2, also called perlecan^{252,253}.

A recent study aimed to construct a comprehensive network of genes related to the basement membrane²⁵⁴. First, 103 human “basement membrane”-annotated genes were identified from the Gene Ontology (GO) resource. GO list was expanded to 263 candidates by data curation across multiple animal species for known basement membrane components, and predicted candidates were identified through gene expression, protein interaction, and domain enrichment analyses. To verify basement membrane zone localisation for the 263 genes in the expanded network, evidence of tissue protein immunolocalisation to the basement membrane zone in humans, rodents, zebrafish, *Drosophila*, and *C. elegans* was reviewed in the published literature and protein databases, such as Human Protein Atlas. The study identified that 130 components within the extracellular space (e.g. collagens, laminins) directly localise to the basement membranes and that 40 cell-membrane bound ECM (e.g. MMP14, CSPG4) and adhesion molecules (e.g. integrins, DDR1) interact with the basement membrane²⁵⁴. Using this list of confirmed basement membrane components, the same study analysed 121 published human transcriptomic datasets from various healthy tissues. The investigation revealed a shared set of basement membrane proteins that consistently appeared across different tissues. These included COL4A1/2, COL6A1/A2/A3, laminin subunit beta-2/gamma-1 (LAMB2/C1), nidogen-1 (NID1), and HSPG2. However, it was observed that a significant portion of basement membrane proteins exhibited varying levels of abundance across different tissues. Additionally, the authors highlighted distinct patterns of basement membrane proteins specific to each human tissue, suggesting that different tissues possess unique compositions of basement membrane proteins²⁵⁴.

It is unclear if different cancer types have different basement membrane composition. Nevertheless, the basement membrane plays a crucial role in epithelial cancer invasion. Due to the compact meshwork of interconnected laminin and type IV collagen lattices, basement membranes have nanosized pores, which are too small to permit the movement of cells²⁵⁵. Consequently, cells cannot traverse this barrier without removing the basement membrane barrier. Degradation of basement membranes involves the ECM remodelling by proteases such as MMPs²⁵⁶. Alternatively, the basement membrane can be breached via force-driven invasion²⁵⁷. For instance, rapidly proliferating cancer cells can directly weaken and rupture the basement membrane by exerting forces²⁵⁸. Other cells within the TME, such as cancer-associated fibroblasts, can exert contractile forces to pull, stretch and soften the basement membrane²⁵⁶. This leads to gaps in the basement membrane, enabling cancer cells to migrate through. In mesenchymal cancers, such as STS, the basement membrane's composition or functional role has not been investigated, representing a research gap.

1.4.2.3. *Matrisome prognostic scores*

The intention behind compiling the matrisome database was to simplify annotating the ECM and ECM-associated genes and proteins within large transcriptomic and proteomic datasets²⁵⁹. Studying matrisome components in cancer research can facilitate the discovery of novel ECM proteins that could influence tumorigenesis. The identified matrisome proteins may potentially serve as targets for novel treatment development or biomarkers for various applications in the field. Over the last decade, several research groups generated prognostic signatures based on matrisome gene expression^{243–249} in carcinomas^{260–266}. Three recent matrisome prognostic signatures are discussed below, and details of the others are summarised in Table 1.3.

Table 1.3 Summary of studies describing the generation of prognostic matrix score in carcinomas.

NSCLC: non-small-cell-lung cancer; EPPI: ECM-related prognostic and predictive indicator; OS: overall survival; DFS: disease-free survival; HGSOC: High-grade serous ovarian carcinoma; GSVA: gene set variation analysis; PFS: progression-free survival; RNAseq: RNA sequencing; TCGA: The Cancer Genome Atlas.

Study	Cancer type	Matrix score construction	Composition of Matrix score	Calculation of Matrix score	Key findings
Langlois <i>et al.</i> , 2014	Derived from mouse model of pancreatic islet cell carcinoma, and applied to a cohort of human glioma and colorectal cancer datasets	RIP1-Tag2 transgenic murine model was used to study angiogenesis. Angiogenic and non-angiogenic pancreatic islets (classified by appearance under the microscope) were profiled using microarrays. Among the differentially expressed genes 110 were matrisome and were used for matrix score construction	AngioMatrix (110 matrisome genes)	The average expression level of 110 genes forming the AngioMatrix signature	In human glioma and colorectal samples high AngioMatrix score correlated with poor prognosis
Lim <i>et al.</i> , 2017	Non-small-cell-lung cancer	Based on differential expressions and significant enrichment of ECM-associated components in 1,943 primary NSCLC tumours relative to 303 normal lung tissues	29 core matrisome and matrisome-associated genes	EPPI risk score = $29 + \sum$ (expression level of gene) X (Cox regression coefficient)	A universal EPPI risk score cut off was determined, which separated NSCLC patients into two groups with significantly different OS in both discovery and validation cohorts. The score was prognostic of OS independent of traditional clinicopathological factors

Yuzhalin <i>et al.</i> , 2018	Lung, breast, ovarian, gastric, oesophageal or colorectal adenocarcinoma	Differentially upregulated genes between tumour and their corresponding normal tissue were identified in publicly available gene expression datasets. Nine matrisome genes were significantly upregulated across all cancer types studies	9 core matrisome genes	Median expression of the 9 core matrisome genes	High matrisome signature was associated with inferior DFS and OS in ovarian, gastric, lung, colorectal adenocarcinomas and prostate and bladder cancer datasets in univariable analysis. In a multivariable analysis, high matrisome signature was prognostic of DFS in gastric, lung and colorectal datasets independent of disease stage and grade
Pearce <i>et al.</i> , 2018	High-grade serous ovarian carcinoma	Transcriptomic and proteomic analysis of HGSOC samples defined matrisome genes and proteins significantly associated with tissue modulus and disease score. Twenty-two molecules were highly significant across all of the analyses	Matrix Index (22 core matrisome and matrisome-associated genes)	The ratio of the average expression levels of genes positively correlated with disease score and tissue modulus to those negatively correlated with disease score and tissue modulus	In multivariable analysis, high Matrix Index was prognostic of inferior OS in triple negative breast cancer, mesothelioma, ovarian cancer, liver hepatocellular carcinoma, lung adenocarcinoma, soft tissue sarcoma, breast invasive carcinoma, colon and colorectal adenocarcinomas, head and neck squamous cell carcinoma, kidney renal clear cell carcinoma, lung squamous cell carcinoma, glioblastoma multiform, and skin cutaneous melanoma

Yang <i>et al.</i> , 2021	Gastric cancer	Four publicly available microarray and RNA sequencing datasets of gastric cancer patients were used. GSVA scores were calculated for functional pathways obtained from pathway databases. Core matrisome gene set identified as core pathway for prognostication	ECM score (274 core matrisome genes)	The ECM score was defined as the GSVA score of the core matrisome gene set downloaded from MatrisomeDB	Gastric cancer patients were stratified by best cut off value of the ECM score. High ECM score was prognostic of inferior OS and RFS
Parker <i>et al.</i> , 2022	Squamous cell lung carcinoma	Differential gene expression analysis of RNAseq data between tumour and non-tumour samples from TCGA dataset identified 622 matrisome genes as differentially expressed	Core matrisome genes differentially expressed between tumour and non-tumour tissue	ECM high and ECM low matreotypes were identified by consensus clustering method. The consensus clustering was applied to the TCGA expression matrix of significantly differentially expressed core matrisomal genes from tumour and non-tumour tissue	ECM high matreotype was associated with inferior disease-specific survival in two independent patient cohort in an univariable analysis. In a multivariable analysis, ECM high matreotype was prognostic of OS after adjusting for age and disease stage

A study by Pearce *et al.* included matrisome analysis in a multi-omic approach to deconstruct the microenvironment of high-grade serous ovarian cancer (HGSOC) metastases²⁶⁵. HGSOC primarily metastasises to the omentum. As such, the authors studied $n = 35$ patient omental biopsies by RNA sequencing and $n = 33$ of the omental biopsies were additionally profiled with quantitative proteomics. The study identified 22 matrisome genes and proteins which were correlated with tissue stiffness and disease score. The disease score was defined as the percentage area occupied by tumour and stroma within each sample. The samples ranged from minimally-diseased with a low percentage area occupied by tumour and stroma (low disease score) omentum to biopsies with extensive disease with dense tumour and stroma areas (high disease score). The matrix index was calculated based on the expression of the 22 matrisome genes as the ratio of the average level of expression of genes positively associated to those negatively associated with disease score and tissue stiffness. Regarding the clinical significance, a high matrix index significantly correlated with shorter HGSOC patient OS in two publicly available gene expression datasets, independently of other known risk factors. Moreover, a high matrix score was an independent prognostic factor for OS in 15 gene expression datasets representing 13 major cancer types, including the STS TCGA-SARC dataset²⁶⁵.

Another study developed a prognostic matrix score for gastric cancer patients²⁶³. The authors evaluated the prognostic value of > 2,000 canonical pathways gene sets from pathway databases, including the BioCarta pathway database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database, the Pathway Interaction Database, the Reactome pathway database, and the WikiPathways pathway database, from the Molecular Signatures Database (MSigDB) in 4 publicly available gene expression datasets of gastric cancer patients. The top pathways significantly associated with OS included the core matrisome, ECM glycoproteins, ECM proteoglycans, and elastic fibre formation, indicating the key role of ECM in gastric cancer. Then, Gene Set Variation analysis was performed to calculate the ECM score for each gastric cancer patient based on the gene expression of core matrisome genes defined by Naba *et al.* group. When the gastric

patients were stratified according to the best cut-off value of the ECM score, a high ECM score was prognostic of worse OS in independent gastric cancer cohorts. Multivariable analysis in two independent cohorts additionally showed that the high ECM score was prognostic of OS after adjusting for age, stage and grade²⁶³.

A study by Parker *et al.* analysed the ECM landscape in squamous cell carcinoma (SqCC), a subtype of non-small cell lung cancer²⁶⁴. The authors evaluated matrisome gene expression in n = 223 tumours and matched, adjacent n = 17 non-tumour tissue from a publicly available lung SqCC TCGA cohort. The analysis identified 170 core matrisome and 488 matrisome-associated genes as differentially expressed in tumours compared to non-tumour tissues. By focusing on the core matrisome gene expression, the authors investigated if distinct ECM molecular subtypes (matreotypes) are present in SqCC. Monte Carlo reference-based consensus clustering of core matrisome genes identified two matreotypes with high (ECM-high) and low (ECM-low) expression of core matrisome genes. ECM-high matreotype was significantly associated with shorter DSS in the TCGA cohort used to generate the matreotypes as well as in an independent SqCC cohort. The ECM-high matreotype was an independent prognostic factor in both cohorts after adjusting for disease stage and age. The prognostic value of matreotypes was confirmed in other solid tumours. However, the SqCC ECM-high matreotype was not associated with survival in the STS TCGA-SARC dataset, implying that SqCC has distinct ECM biology²⁶⁴.

The above-mentioned studies highlight that cancer ECM is a source of prognostic biomarkers and provide a rationale for developing matrisome-related biomarkers for prognostication of various types of cancer. While one study demonstrated the potential of the carcinoma-derived matrix score for risk-stratifying STS patients, further validation and global characterisation of matrisome components across different STS subtypes are required.

Gene expression analysis techniques like microarray and RNA sequencing are valuable tools for quantifying matrisome components at the transcript level. These methods provide high coverage of the matrisome components and

serve as surrogates for matrisome protein expression. However, it is important to note that changes at the mRNA level are not always reflected at the protein level. In fact, the correlation between RNA expression and protein expression is low^{267–269}. This is apparent for ECM proteins, and many studies have shown a poor correlation between ECM-transcript and ECM-protein level^{270–272}. The ECM proteins are subject to post-translational modifications²⁷³, such as glycosylations and cross-linking, and some ECM proteins, such as collagens, may exhibit a low turnover rate²⁷⁴. Ultimately, the characteristics and levels of matrisome proteins can greatly deviate from the initial transcript level. Moreover, ECM proteins mediate cellular functions at the protein level, not the transcript. Thus, comprehensive proteomic analysis of matrisome is required for an accurate picture of its role in cancers, including STS.

1.5. Integrin signalling

Cells perceive and respond to cues from the ECM through adhesion proteins, with integrins serving as a prominent family of cell-matrix receptors. Integrins facilitate essential cellular processes by engaging with the ECM, including cell survival, proliferation, differentiation, shape, polarity, and motility²⁷⁵. Their aberrant function in cancer underscores the importance of understanding integrin-mediated ECM signalling²⁷⁶. Integrins function as heterodimers composed of α and β subunits (Figure 1.6). In mammals, 24 distinct integrin pairs form from the combination of 18 α and 8 β subunits. Integrins are transmembrane molecules containing a short cytoplasmic and an extracellular domain. The terminal domains of extracellular regions of α and β subunits assemble to form a 'head' that acts as a ligand-binding site. Remarkably, integrins can recognise numerous ECM ligands, and each heterodimer exhibits a distinct specificity for ligand binding. Integrin heterodimers can be classified into subfamilies based on their ligand specificity or cellular distribution. The subfamilies include integrins ($\alpha 5$, αV , $\beta 1$, $\beta 5$) recognising the arginine-glycine-aspartate (RGD) sequence in molecules such as fibronectin and vitronectin, collagen-specific ($\alpha 1$, $\alpha 2$, $\alpha 10$, $\beta 1$), laminin-specific ($\alpha 3$, $\alpha 6$, $\alpha 7$, $\beta 1$), and those primarily found on leukocytes (αL , αM , αX , $\beta 2$).

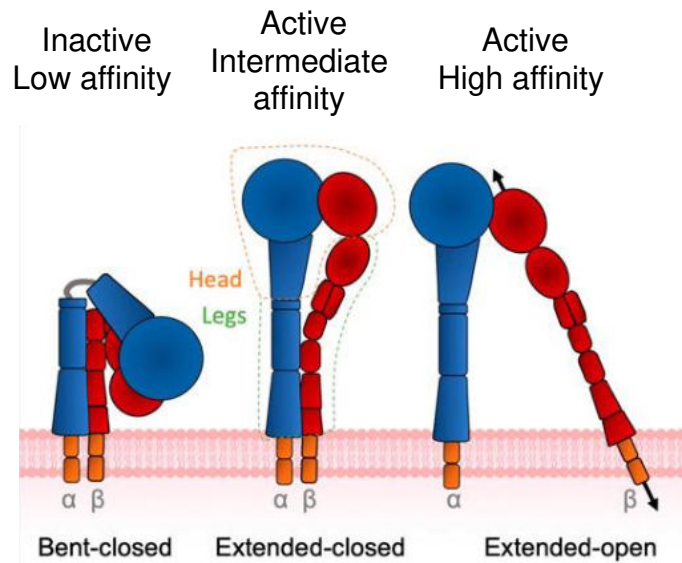


Figure 1.6 Three conformational states of an integrin receptor.

The conformation of alpha (α) and beta (β) integrin subunits in the three activation states indicates an extracellular ligand's affinity (low, intermediate and high). Adapted from Brown and Marshall (2019) under CC BY 4.0 licence²⁷⁷.

Integrins exhibit three conformational states with distinct affinities for their extracellular ligands^{278,279} (Figure 1.6). Integrins remain inactive in their bent or closed conformation, characterised by a low affinity for ECM ligands. Conversely, integrins become active in the extended-closed conformation, primed for binding to ECM ligands with increased affinity. The third conformation state is extended-open (ligand-occupied) when the integrin heterodimer is fully extended with an open headpiece and intracellular signalling is initiated.

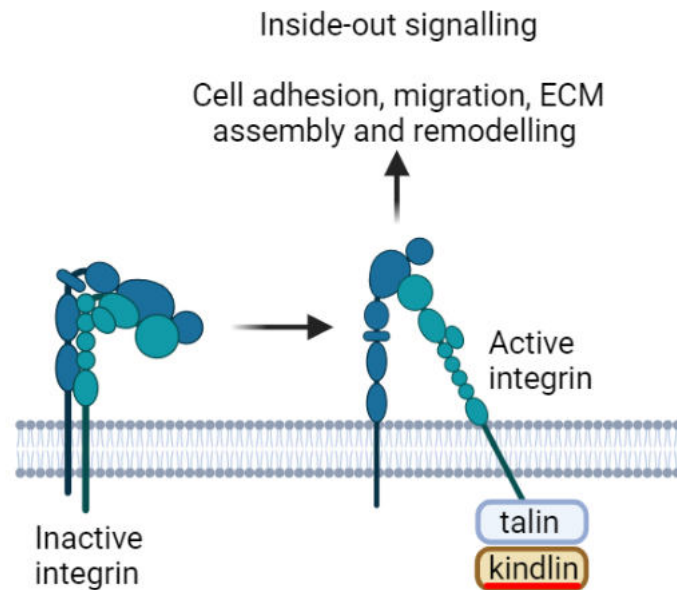


Figure 1.7 Integrin activation and ‘inside-out’ signalling.

Inside-out integrin signalling is initiated by cellular events leading to receptor activation. Talin binds to the cytoplasmic tail of the β subunit of integrin, causing conformational changes that shift the receptor from a closed to an open, high-affinity state. Kindlin further reinforces this conformational change, stabilising the active integrin conformation. This activation increases integrin affinity for ECM ligands and facilitates various cellular processes such as cell adhesion, migration, ECM assembly and remodelling. Created using BioRender.com.²⁸⁰

The process of integrin activation to adopt the fully extended conformation is regulated by the binding of intracellular adaptor proteins, such as talin and kindlins, to the cytoplasmic region of β integrin subunit (Figure 1.7). The talin and kindlin binding initiates ‘inside-out’ signalling, which involves conformational changes and separates the cytoplasmic and transmembrane regions of α and β integrin subunits^{281,282}. This unbends the ligand-binding headpiece, and conformational changes increase ligand-binding affinity^{283,284}. The inside-out signalling plays a pivotal role in regulating the strength of cellular adhesion^{285,286}. Moreover, it empowers integrins to transmit mechanical forces essential for critical cellular processes such as migration, ECM assembly, and ECM remodelling^{287–289}.

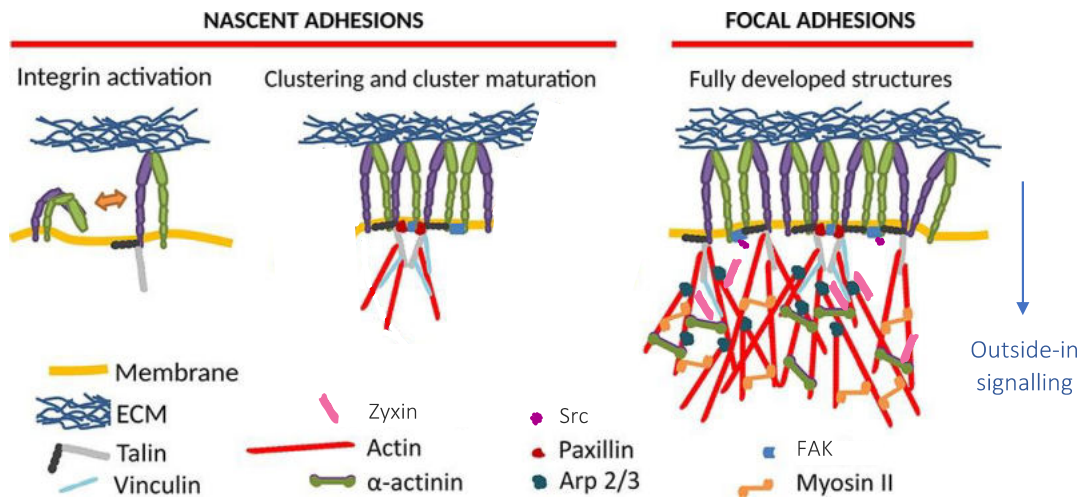


Figure 1.8 Formation of integrin-mediated nascent and focal adhesion and the ‘outside-in signalling’.

Integrin activation and binding to the ECM induces the clustering of integrins and the formation of nascent adhesions, which connect to the actin cytoskeleton via talin. Recruitment of various intracellular adaptors and signalling proteins to the nascent adhesions and cytoskeletal reorganization enhances the strength and stability of cell-ECM adhesions. This maturation of the nascent adhesions into focal adhesions further recruits intracellular adaptors (paxillin, vinculin, zyxin), kinases (focal adhesion kinase (FAK) and proto-oncogene c-Src (Src)) and actin regulator proteins (α -actinin, arp2/3) and assembles multicomponent intracellular adhesion structures. The multicomponent adhesion structures can propagate signals in an ‘outside-in’ signalling. The signalling is context-dependent and can include changes in gene expression, cytoskeletal rearrangements, cell migration, proliferation, differentiation, and apoptosis. Adapted from Stumpf *et al.* (2020) under CC BY 4.0 licence²⁹⁰.

Integrins are unique receptors that signal bidirectionally²⁹¹. In addition to the ‘inside-out’ signalling, integrins participate in the classical ligand-dependent or the so-called “outside-in” signalling to mediate cellular responses to ECM adhesion (Figure 1.8). Integrin heterodimers lack inherent enzymatic activity and rely on association with intracellular kinases and adaptors for signal transmission. The binding of integrins to their ECM ligands initiates the clustering of integrin complexes and the formation of nascent adhesions (Figure 1.8). These adhesions serve as molecular bridges, connecting ECM molecules to the cytoskeleton through interaction with the protein talin²⁹². As these adhesions mature to form focal adhesions, many scaffolding and signalling proteins come into play. Vinculin, focal adhesion kinase (FAK), proto-oncogene c-Src (Src), paxillin, zyxin and others are recruited to

assemble intracellular adhesion complexes²⁹³. A remarkable feature of focal adhesion components is their multidomain nature, allowing them to engage with multiple partner molecules. A prime example of this is observed with molecules like vinculin, FAK, Src kinases, and paxillin, each of which can bind to over ten distinct partner molecules^{293,294}. Consequently, the potential number of diverse combinations of molecular interactions that may participate in bridging integrins to the actin cytoskeleton is vast.

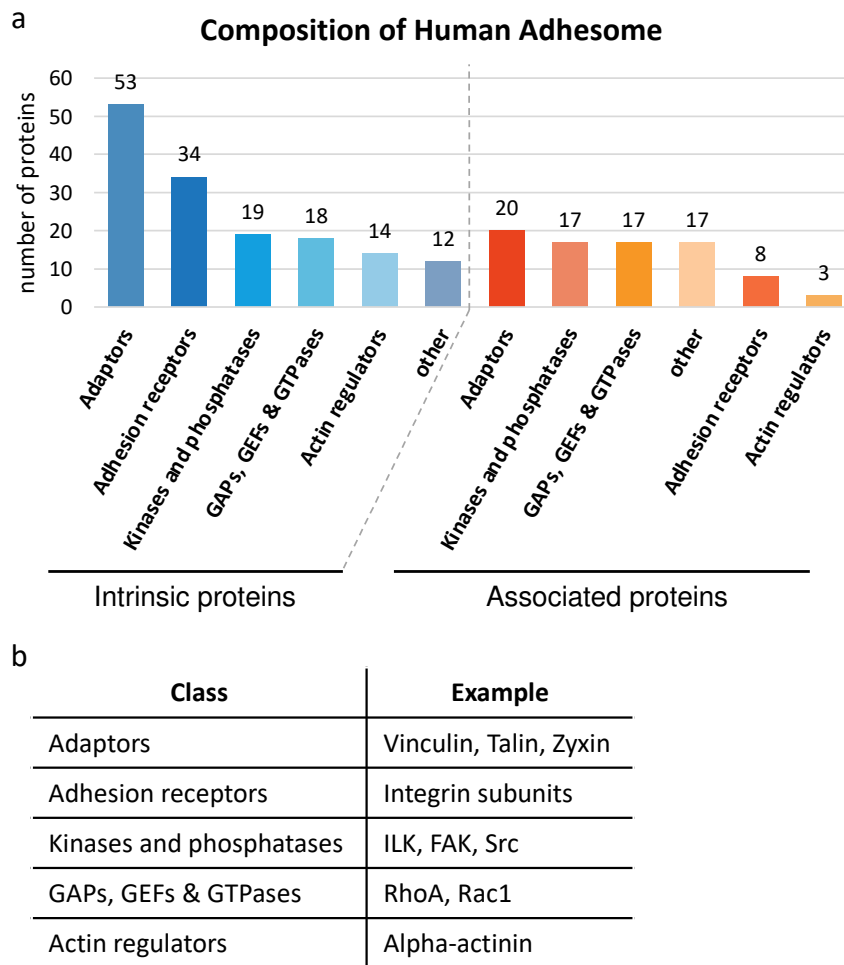


Figure 1.9 Composition of literature-curated human adhesome.

A) Breakdown of adhesome by intrinsic and associated proteins and by functional classes (adaptors, adhesion receptors, actin regulators and others). The numbers of components in each class are shown. B) Selected examples of proteins from adhesome classes. ILK: integrin-linked kinase, FAK: focal adhesion kinase, Src: proto-oncogene c-Src, RhoA: Ras homolog family member A, Rac1: Ras-related C3 botulinum toxin substrate 1. Created in Excel.

1.5.1. Adhesome definition

In search of a comprehensive list encompassing all the possible molecules involved in integrin adhesion complexes, the Geiger laboratory performed an extensive literature search to identify 156 components (151 proteins plus four lipids and calcium ions) constituting the integrin 'adhesome'²⁹⁵. Furthermore, the list was updated due to the isolation and identification of new integrin adhesion-related proteins by proteomic studies^{270,296,297}, and the most recent adhesome version includes 232 proteins²⁹⁸ (Figure 1.9). The components of integrin adhesion complexes can be broadly classified as 'intrinsic', and 'associated'. The 'intrinsic' components directly link the intracellular adhesion complexes to the actin cytoskeleton, forming an integral part of the complex²⁹⁵. On the other hand, 'associated' proteins are considered effectors that interact with and regulate the intrinsic molecules to contribute to the overall function of the complex. The intrinsic and associated proteins can be further subclassified according to the functional role or localisation into adhesion receptors (e.g., integrin subunits), adaptor proteins (e.g., vinculin, talin, zyxin), actin regulators (e.g., alpha-actinin), serine/threonine (e.g., integrin-linked kinase (ILK)) kinases, tyrosine (e.g., focal adhesion kinase (FAK)) kinases, GTPases (e.g., Ras homolog family member A (RhoA), Ras-related C3 botulinum toxin substrate 1 (Rac1)) and other classes. Notably, adhesome is an *in-silico* definition and, therefore, theoretical and context-dependent. Experimental determination of the actual cell-specific adhesion partners is required.

The cell's specific integrin expression patterns dictate which ECM substrate the cell can bind²⁹⁹. Additionally, the composition of integrin adhesome governs subsequent signalling cascades, ultimately shaping the cell's behaviour and fate. The complex interplay between the surrounding ECM and cells underscores the need for integrative analysis of matrisome and adhesome¹⁷⁷. Such approaches can help to define the dynamic and bidirectional alterations in ECM remodelling and integrin-based adhesion signalling that occur in physiological and disease processes.

1.6. Matrisome and adhesome components in STS

1.6.1. ECM in STS

Limited research has been conducted to comprehensively characterise the ECM in STS³⁰⁰ (Table 1.4 summarises the current literature). Specifically, our global understanding is largely shaped by a single proteomics study that employed mass spectrometry to analyse intramuscular myxoma and grade IMyxFS³⁰¹. In addition, two transcriptomic studies have contributed to our knowledge, one focusing on rhabdomyosarcoma and the other comparing desmoid tumours (DES) with benign solitary fibrous tumours^{302,303}. Most ECM studies in STS have relied on IHC and immunofluorescence (IF) analyses, which have probed for the expression of collagen types I, III, IV, and glycoproteins, such as fibronectin, laminin, vitronectin and tenascin in various STS subtypes^{304–312}.

Table 1.4 Summary of the studies on ECM components in STS.

IF: immunofluorescence; IHC: immunohistochemistry; STS: soft tissue sarcoma; ASPS: alveolar soft part sarcoma; LMS: LMS; SS: synovial sarcoma; WDLPS: well-differentiated liposarcoma; LPS: liposarcoma; MFH: malignant fibrous histiocytoma; MPNST: malignant peripheral nerve sheath tumour; SS: synovial sarcoma; DES: desmoid tumours; SFT: solitary fibrous tumours; ADAM: a disintegrin and metalloproteinase; MMP: matrix metalloproteinase; MyFS: myxofibrosarcoma; aRMS: alveolar rhabdomyosarcoma; eRMS: embryonic rhabdomyosarcoma; IM: intramuscular myxoma; STS: soft tissue sarcoma; UPS: undifferentiated pleomorphic sarcoma; DFS: disease-free survival; CSPG4: chondroitin sulfate proteoglycan 4.

Study	ECM components	Biochemical characterisation	Number of samples	Subtype breakdown	Key findings
Stenman and Vaheri, 1981	Fibronectin	IF	18	4 LMS, 1 fibrosarcoma, 2 neurofibrosarcoma, 2 giant-cell tumour of bone, 1 Kaposi's sarcoma, 1 LPS, 1 malignant hemangiopericytoma, 2 leiomyoma, 1 desmoid tumour, 2 neurilemoma, 1 cavernous hemangioma	Dense network of fibronectin in all 12 sarcomas and 6 benign soft-tissue tumours
d'Ardenne <i>et al.</i> , 1984	Collagen III, laminin, fibronectin	IHC	60	36 STS, 24 benign tumours	Collagen III has the most variable and irregular distribution, fibronectin is abundant in nearly all cases. Laminin was strongly expressed in ASPS, malignant and benign schwannomas, neurofibroma, leiomyoma and some LMS, but not the rest of STS

Ogawa <i>et al.</i> , 1986	Collagen IV	IHC	103	Tumours of peripheral nerve, muscle origin, tumours of fibrous, adipose and synovial tissue, tumours of blood vessels and of uncertain histogenesis	Strong staining in STS of peripheral nerve origin and blood vessel tumours, weak to moderate staining is seen in leiomyomas, angiomyomas and LMS. SS, fibroblastic and fibrohistiocytic tumours are negative for collagen IV. In ASPS strong staining was around nests of cells
Persson <i>et al.</i> , 1988	Laminin	IHC	10	10 ASPS	Strong laminin staining around 'alveoli' of tumour cells
Ordonez <i>et al.</i> , 1990	Collagen IV and laminin	IHC	39	15 biphasic, 24 monophasic SS	6/15 biphasic SS show continuous staining for laminin and collagen IV around epithelioid areas. In monophasic SS and spindle-cell component of biphasic SS the staining of laminin and collagen IV is weak and focal
Guarino and Christense, 1994	Fibronectin, collagen I, III, IV, laminin and tenascin	IHC	4	2 biphasic SS and 2 monophasic SS	Strong collagen I, III and fibronectin staining in the mesenchymal but not epithelioid areas. Strong staining for laminin, collagen IV and tenascin around epithelioid areas in biphasic SS.

Haraida <i>et al.</i> , 1996	Collagen IV, laminin, heparan sulphate proteoglycan and fibronectin	IHC	50	20 normal adipose tissue, 13 benign lipomatous (lipomas, hibernomas and lipoblastomas) and 17 malignant (WDLPS, myxoid, spindle cell and pleomorphic LPS) tumours	Lipomas and WDLPS show predominant expression of collagen IV staining, myxoid LPS mainly stain for laminin and pleomorphic LPS mainly express fibronectin
Benassi <i>et al.</i> , 1998	Fibronectin, laminin, collagen IV, vitronectin	IHC	58	20 MFH, 17 MPNST, 21 SS	Fibronectin, laminin and vitronectin are detected in MFH, MPNST and in spindle cell component of SS. Collagen IV detected in MPNST and epithelioid areas of SS
Fukuda and Tsuneyoshi, 2000	Fibronectin	IF	13	13 myxoid LPS	Fibronectin staining is randomly present in immature tumour cells in 5/13 samples
West <i>et al.</i> , 2005	Non-targeted, various genes	cDNA microarray	23	13 SFT, 10 DES	SFT had higher expression of basement membrane genes (COL4A5, COL17A1). DES overexpressed fibrillar collagen genes (COL1A1, COL5A1, COL3A1) and ECM remodelling enzymes (ADAM12/19, MMP11/19/23b)
Willems <i>et al.</i> , 2009	Collagens and proteoglycans	Mass spectrometry, IHC and qPCR	20	10 IM and 10 grade I MyxFS	Decorin and α 1 chain of collagen VI were significantly overexpressed at mRNA and protein levels in grade I MFS versus IM

Lian <i>et al.</i> , 2021	Collagen XVIII, collagen IV	RNA sequencing, IHC	42	Human: 1 aRMS and 3 eRMS Murine: 15 aRMS, 6 eRMS, 17 undifferentiated sarcomas	Collagen XVIII was expressed in 69% of all murine RMS, with slight predominance in aRMS compared to eRMS, α 1 chain of collagen IV was expressed in 79% of RMS and α 2 chain of collagen IV was mainly expressed in 67% eRMS compared to 14% of aRMS.
Boudin <i>et al.</i> , 2022	Proteoglycan chondroitin sulfate proteoglycan 4	Gene expression from 15 public STS databases	610	LMS, LPS, UPS, MyxFS and other	The 'CSPG4-high' patients had a significantly shorter DFS, and CSPG4 score was independent of STS subtype, tumour size and FNCLCC grade (HR=3.47, 95% CI 1.73–6.95, p=4.59E-04)

1.6.1.1. Mass spectrometry-based analysis

Intramuscular myxoma is a benign tumour with no potential for recurrence or metastasis. Conversely, MyxFS exhibits a significantly higher recurrence rate and an increased tendency to metastasise, particularly after recurrence. A mass spectrometry analysis was conducted on tumour lysates from $n = 10$ intramuscular myxoma and $n = 10$ grade I MyxFS³⁰¹. Grade I MyxFS exhibited unique ECM protein expression patterns compared to intramuscular myxoma. Specifically, MyxFS samples displayed elevated expression of FACIT collagen type XII chain $\alpha 1$ and collagen type XIV chain $\alpha 1$, as well as five proteoglycans (lumican, proteoglycan 4, prolargin, decorin, and biglycan). Both intramuscular myxoma and MyxFS shared collagen type VI. Further investigation confirmed increased mRNA and protein levels of collagen VI chain $\alpha 1$ and decorin in grade I MyxFS compared to intramuscular myxoma. These findings suggest that the ECM in intramuscular myxoma and grade I MyxFS comprises distinct proteoglycan and collagen components. However, whether these differences are linked to the observed variations in recurrence and metastasis rates between these tumours remains to be established.

1.6.1.2. Transcriptomic studies

A study by Lian *et al.* investigated the role of multiplexin collagen type XVIII in rhabdomyosarcoma, a paediatric STS³⁰³. Type XVIII collagen levels were investigated through RNA sequencing and IHC in a series of specimens from human alveolar rhabdomyosarcoma and embryonic rhabdomyosarcoma and genetically engineered mouse models of alveolar rhabdomyosarcoma, embryonic rhabdomyosarcoma and pleomorphic rhabdomyosarcoma. The results indicated that collagen XVIII mRNA expression was notably higher in human and murine alveolar and embryonic rhabdomyosarcomas when compared to normal muscle tissue. To determine the clinical significance of collagen XVIII expression, RNA expression data from human alveolar rhabdomyosarcoma and embryonic rhabdomyosarcoma biopsies ($n = 65$) obtained from the Intergroup Rhabdomyosarcoma Study-IV was correlated with OS outcome. Kaplan-Meier survival analysis showed that the 9-year OS rate was significantly lower for patients with high collagen XVIII expression

than those with low collagen XVIII ($p = 0.01$). Importantly, the prognostic value of high collagen XVIII expression persisted even after considering other clinical factors. Nevertheless, further validation using independent cohorts is necessary to confirm the potential of collagen XVIII as a prognostic biomarker in rhabdomyosarcoma. These findings collectively suggest that collagen XVIII potentially plays a role in rhabdomyosarcoma growth and progression.

A transmembrane proteoglycan chondroitin sulfate proteoglycan 4 (CSPG4), has previously been investigated in STS. Bertucci *et al.* collated gene expression data of clinical STS samples from 15 public data sets through the National Centre for Biotechnology Information (NCBI)/Genbank GEO, ArrayExpress databases, and authors' websites³¹³. CSPG4 mRNA expression was associated with DFS in $n = 610$ STS patient samples across LPS, LMS, UPS, MyxFS and other STS subtypes. 'The CSPG4-high' patient group had a significantly shorter DFS than the 'CSPG4-low' group (HR = 1.49, 95% CI 1.14 – 1.94; p -value = 0.0034). The stratification based on the CSPG4 expression remained prognostic for DFS after adjusting for known risk factors such as histological subtypes, tumour size and FNCLCC grade (HR = 3.47, 95% CI 1.73 – 6.95, p -value = 0.0005). This study provides another compelling example of the value of ECM analysis in STS.

1.6.1.3. IHC and IF studies

IHC and IF studies were restricted to investigations of collagen type I, III, IV, laminin and fibronectin expression in pan-soft tissue tumours and STS-subtype specific studies (Table 1.4).

Collagen III

Expression of fibrillar collagen type III was investigated in formalin-fixed paraffin-embedded (FFPE) specimens of 36 malignant and 24 benign soft tissue tumours by IHC³¹². Collagen type III was present in the stroma of most analysed tumours, with varying staining intensity across different histological subtypes. Notably, all 13 LMS cases exhibited strong pericellular collagen III staining, resembling the staining pattern found in benign leiomyomas. Conversely, tumours of peripheral nerve origin (6 neurofibromas, 6

schwannomas, and 2 malignant schwannomas) displayed weak and irregular collagen III staining. The differential staining of collagen III may be attributed to the distinct cellular origins (smooth muscle vs. nerve) of these subtypes, although further studies are required to confirm this.

Collagen IV

IHC was used investigated the expression of collagen type IV on FFPE samples in pan-STS and studies focused on synovial sarcoma^{305,306,309-311} (Table 1.4). Ogawa *et al.* analysed 103 soft tissue tumour samples, including tumours of peripheral nerve (schwannomas, neurofibromas, malignant schwannomas), tumours of smooth muscle origin (leiomyomas, angiomyomas, and LMSs), tumour of fibrous origins (desmoid-type fibromatosis, fibrosarcoma, dermatofibroma, and malignant fibrous histiocytoma), tumours of adipose origin, synovial sarcoma and alveolar soft part sarcoma³¹¹. They found strong collagen IV staining around tumour cells in peripheral nerve tumours and alveolar soft part sarcoma. Leiomyomas showed weak to moderate staining, while angiomyomas exhibited moderate staining around tumour cells. LMSs were weakly positive in eight out of ten cases. In contrast, fibrous origin tumours were negative for collagen IV³¹¹. Ogawa *et al.* and Haraida *et al.* studies found that normal adipose, benign and well-differentiated LPS cases showed strong collagen IV expression, while malignant myxoid and pleomorphic LPS were negative for collagen IV^{309,311}. Collagen IV production in synovial sarcoma appeared to depend on cellular differentiation, with biphasic synovial sarcomas often showing positivity around epithelioid areas, while spindle cell areas and monophasic synovial sarcomas were negative^{310,311}. In summary, collagen IV distribution varies among soft tissue tumours, with peripheral nerve, smooth muscle origin, and benign adipocytic tumours typically showing abundant collagen IV. In contrast, fibrous origin tumours and some LPS subtypes tend to lack collagen IV expression.

Fibronectin

Fibronectin is an ECM glycoprotein that undergoes cell-mediated assembly into an insoluble fibrillar matrix. This process involves secreted FN molecules

binding to the specific adhesion receptor integrin $\alpha 5\beta 1$ ³¹⁴. Integrin binding facilitates interactions between fibronectin molecules, forming fibronectin fibrils that cluster into a network of thicker fibril bundles. Fibronectin is pivotal in incorporating other ECM components like collagens, fibrillins, and fibulin^{315,316}. Beyond its role in matrix assembly, fibronectin also participates in various biological processes, including adhesion, growth, cell migration, and differentiation. The expression and distribution of fibronectin have been examined through IF and IHC in various histological subtypes of soft tissue tumours^{305,307,308,312,317}. These studies found that fibronectin is abundant in soft tissue tumours, with moderate to strong fibronectin staining in various tumours of smooth muscle, peripheral nerve, adipose and fibrous origin.

Laminin

Laminins, high-molecular-weight glycoproteins, constitute the primary component of the basement membrane. These heterotrimeric proteins consist of α , β , and γ chains, forming a cross-shaped structure that enables interaction with other ECM molecules³¹⁸. In both the mouse and human genomes, there are currently identified 5 α -chain genes, 3 β -chain genes, and 3 γ -chain genes, which combine to create a total of 16 unique laminin proteins³¹⁸. The presence of laminin has been examined in various soft tissue tumour types^{304,306,309,310,312}. Strong laminin staining was observed in benign tumours, such as leiomyomas, neurofibromas, schwannomas³¹², lipomas and hibernomas³⁰⁹. In malignant STS, laminin was abundant in LMS³¹², alveolar soft part sarcoma³⁰⁴, and synovial sarcoma³¹⁰, while in myxoid LPS and pleomorphic LPS, laminin staining was weak or absent³⁰⁹. These studies collectively demonstrate that laminin is a prevalent ECM component in benign and malignant soft tissue tumours originating from muscle, peripheral nerve, and benign adipose tissue tumours. However, it should be noted that the antibodies used in these studies were generated against the laminin isolated from Engelbreth-Holm-Swarm mouse sarcoma, which is predominantly composed of laminin-111 isoform^{319,320}. Therefore, it remains unclear if other specific laminin isoforms are expressed in soft tissue tumours.

In summary, in the last decades, there have been efforts to understand the composition of the ECM in STS. However, while the ECM consists of hundreds of proteins and polysaccharides, the exploration of the ECM in STS has so far focused on the most common ECM proteins and lacked mechanistic follow-up studies. Moving forward, it would be crucial to apply proteome-wide approaches to a diverse range of STS subtypes to gain a deeper understanding of the composition and function of the ECM in STS progression.

1.6.2. Integrins in STS

Given the rarity and heterogeneous nature of STS, research on integrins in these diseases is limited, and the studies lack independent validation. Nevertheless, efforts to characterise integrin expression patterns by IHC have been undertaken in rhabdomyosarcoma, MFH, malignant peripheral nerve sheath tumour and synovial sarcoma patient samples^{317,321} (Table 1.5). Additionally, some studies investigated the mechanistic role of integrin subunits in MyxFS and rhabdomyosarcoma tumorigenesis^{322–324} (Table 1.5).

Table 1.5 Summary of the studies on integrins in STS.

IHC: immunohistochemistry; MFH: malignant fibrous histiocyoma; MPNST: malignant peripheral nerve sheath tumour; SS: synovial sarcoma; MyFS: myxofibrosarcoma; DSS: disease-specific survival; DFR: disease recurrence-free; TRIO: triple functional domain protein; RICTOR: rapamycin-insensitive companion of mammalian target of rapamycin.

Study	Integrin(s)	Biochemical characterisation	Number of samples	Subtype breakdown	Key findings
Barth <i>et al.</i> , 1995	α 1/2/3/5/6, β 1/3/4	IHC	22	7 rhabdomyosarcomas, 8 primitive peripheral neuroectodermal tumours and 7 Ewing sarcomas	All subtypes were positive for β 1. Rhabdomyosarcoma tumours were mostly α 1 and α 3 negative and showed heterogenous expression of α 5 and α 6 subunits. Ewing sarcomas and primitive peripheral neuroectodermal tumours shared a similar integrin subunit profile and were only positive for α 5 and β 1
Benassi <i>et al.</i> , 1998	α 2/5/6, α v, β 3	IHC	58	20 MFH, 17 MPNST, 21 SS	Patients with \geq 35% integrin α 6-positive cells have significantly worse 2-year DFS
Roma <i>et al.</i> , 2011	α 9	Rhabdomyosarcoma cell lines	3	RH30, CW9019 and HTB82	Notch signalling pathway upregulated integrin α 9 expression and is associated with invasive phenotype in rhabdomyosarcoma
Okada <i>et al.</i> , 2016	α 10	cDNA microarray	64	MyxFS	DSS (HR=2.5; p-value < 0.01), DRF survival (HR=3.8; p-value = 0.001). α 10 subunit may control growth and metastasis development in MyxFS via integrin α 10/TRIO/RICTOR pathway

In the Barth *et al.* study, the distribution of integrin subunits was investigated using IHC analysis on fresh frozen samples from $n = 32$ different cancers, which included 7 cases of rhabdomyosarcoma, 8 cases of primitive peripheral neuroectodermal tumours, and 7 cases of Ewing sarcomas³²¹. The results revealed distinct integrin expression patterns in these tumour subtypes. Rhabdomyosarcomas displayed a unique integrin profile, being mostly negative for $\alpha 1$ and $\alpha 3$ subunits and showing heterogeneous expression of $\alpha 5$ and $\alpha 6$ subunits. Ewing sarcomas and primitive peripheral neuroectodermal tumours shared similar integrin subunit profiles, being only positive for $\alpha 5$ and $\beta 1$ subunits. The shared integrin expression patterns between Ewing sarcomas and primitive peripheral neuroectodermal tumours suggested that these two subtypes likely engage similar ECM components and activate comparable intracellular signalling networks. This observation is in keeping with these two subtypes now being known to represent the same entity³²⁵. Conversely, the differential expression of $\alpha 1$, $\alpha 3$, $\alpha 5$, and $\alpha 6$ integrin subunits in rhabdomyosarcoma in comparison to primitive peripheral neuroectodermal tumours and Ewing sarcoma hinted at potentially distinct integrin adhesion complexes and biological functions in rhabdomyosarcoma. However, it is important to note that this study had a relatively small sample size, and the findings should be validated in independent patient cohorts to establish their robustness and clinical relevance.

Two studies investigated the mechanistic role of integrin subunits in MyxFS³²² and rhabdomyosarcoma³²³. Gene expression microarray analysis of $n = 64$ of primary untreated high-grade MyxFS identified integrin $\alpha 10$ subunit, high expression of which was associated with inferior DSS (HR = 2.46; p -value < 0.01) and inferior distant recurrence-free survival (HR = 3.75; p -value = 0.001)³²². The integrin $\alpha 10$ subunit associates with $\beta 1$ integrin subunit to form an $\alpha 10\beta 1$ collagen receptor, which preferentially binds collagen IV and VI. Integrin $\alpha 10$ short hairpin RNA (shRNA) knockdown reduced cell growth in 4 MyxFS cell lines. The authors reported that integrin $\alpha 10$ subunit controls MyxFS growth via downstream effectors, the nucleotide exchange factor, triple functional domain protein (TRIO), and the subunit of the mTORC2 complex, rapamycin-insensitive companion of mammalian target of rapamycin

(RICTOR). TRIO and RICTOR activated the ras-related C3 botulinum toxin substrate 1 (RAC1)/RAC1 activated kinase 1 (PAK) and AKT/mTOR pathways in MyxFS cell lines. Blocking these proteins with Ehop-016 (a RAC inhibitor) and INK128 (an mTOR inhibitor) demonstrated anti-tumour effects in both MyxFS cell lines and mouse xenograft models³²².

The role of the integrin α 9 subunit in invasion was investigated in three rhabdomyosarcoma cell lines (RH30, CW9019, and HTB82)^{323,324}. It was established that the Notch signalling pathway regulated integrin α 9 subunit levels³²³. Genetically manipulating Notch signalling in rhabdomyosarcoma cells by overexpressing the Notch ligand Delta-like protein 1 (Delta 1) to constitutively activate the Notch pathway, increased integrin α 9 subunit levels at both mRNA and protein levels. Transfecting rhabdomyosarcoma cells with the dominant negative form of the Notch receptor coactivator, mastermind-like protein 1 (MAML1), reduced integrin α 9 subunit expression. Delta1-overexpressing rhabdomyosarcoma cells exhibited higher invasiveness compared to mock-transfected cells³²³. Moreover, when Delta1-overexpressing cells were treated with an integrin α 9 blocking antibody, their invasiveness was significantly impaired compared to controls³²⁴. These studies showed that the integrin α 9 subunit in rhabdomyosarcoma might have a pro-invasive role and that cooperation of Notch signalling with integrin α 9 might be an underlying mechanism of rhabdomyosarcoma cell invasion.

Overall, our understanding of adhesion receptor biology in STS is currently in its early stages. The emerging data presented above suggest that integrins play crucial functional roles in STS tumour progression and metastasis. It is worth noting that these studies have predominantly focused on a limited number of STS subtypes, and future research should expand to other histological subtypes and investigate the role of other adhesion-related components.

1.7. Preclinical models mimicking tumour ECM

The ECM is a complex milieu of proteins, glycoproteins and proteoglycans, and assembled ECM molecules activate multiple intracellular signalling pathways^{160,176,291}. The ECM within tumours is a dynamic and complex entity produced and shaped by cancer cells and stromal cells, leading to a varied composition across cancer types^{167,326,327}. As such, to understand the collective contribution of the intricate ECM network in cancer, it is crucial to replicate the complexity of the ECM in preclinical cancer models.

1.7.1. Matrix material commonly used in cancer research

To mimic the ECM environment, cells can be grown in two-dimension (2D) on layers of different ECM substrates, the most common being fibronectin, laminin and collagen. Alternatively, cells can be mixed with ECM-mimicking hydrogels and cultured in 3D.

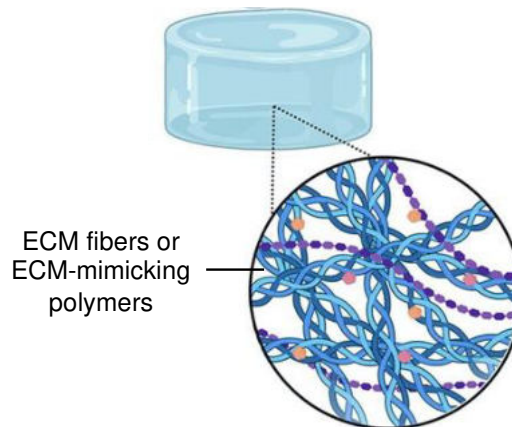


Figure 1.10 Schematic of a typical hydrogel.

A hydrogel is a three-dimensional, cross-linked network of ECM fibres or ECM-mimicking polymers that can absorb and retain a significant amount of water or other aqueous solutions while maintaining its structural integrity. Created with BioRender.com.

Hydrogel is a semi-solid 3D network of natural ECM or synthetic polymer materials that absorbs and retains water³²⁸ (Figure 1.10). Matrigel and collagen I are the natural ECM hydrogels commonly employed in tumour models. Matrigel is a commercially available product and is a basement

membrane extract derived from murine Engelbreth-Holm-Swarm sarcoma³¹⁹. Matrigel stands as the primary matrix of choice in research and finds application in various areas³²⁹, including angiogenesis³³⁰, stem cell maintenance³³¹, tumour spheroid formation³³², invasion assays³³³, and as a material for implanting cancer cells in *in vivo* mouse models³³⁴. Collagen gel, a type I collagen commonly isolated from rat tail tendons³³⁵, is commonly used in tumour spheroid formation and invasion assays³³⁶. However, being composed of one major ECM component, collagen I gels are not able to recapitulate the biochemical complexity of the ECM.

Despite its ability to support cancer research applications, Matrigel has important limitations³³⁷. Being of animal origin, it includes xenogenic components that are not representative of human ECM. Additionally, Matrigel displays considerable inconsistency from one batch to another, which can unpredictably impact the cell culture and has a poorly defined composition. The primary components of Matrigel are 4 major basement membrane ECM proteins: glycoproteins, laminin, collagen IV, nidogen and the heparin sulfate proteoglycan 2³³⁸. A recent study reported that glycoproteins account for ~ 96%, proteoglycans for 1% and collagens for 0.4% of the total matrisome in Matrigel³²⁰. As such, Matrigel is not representative of tissues and TMEs with high collagen and proteoglycan content. Moreover, Matrigel is not a pure ECM extract; it contains ~ 2,000 different proteins, which include other cellular proteins, and those could affect cell phenotypes^{338,339}. Finally, Matrigel preparation suffers from low batch-to-batch reproducibility. Growth-factor-reduced versions of Matrigel, share approximately a 53% similarity in protein content from batch to batch³³⁸, raising questions to the reproducibility of experiments using different Matrigel batches. Matrigel continues to be extensively employed in cancer research, including studies involving sarcomas; nevertheless, there is a growing interest in exploring alternative ECM-mimicking preclinical models that can address Matrigel's limitations.

1.7.2. Decellularised ECM scaffolds and tissue-derived ECM hydrogels for cancer research

In the last decade, tissue decellularisation methods were developed to obtain reconstituted ECM *in vitro* from multiple organ and tissue types^{340,341}. The method was developed for tissue engineering and regenerative medicine and is now also used in cancer research.

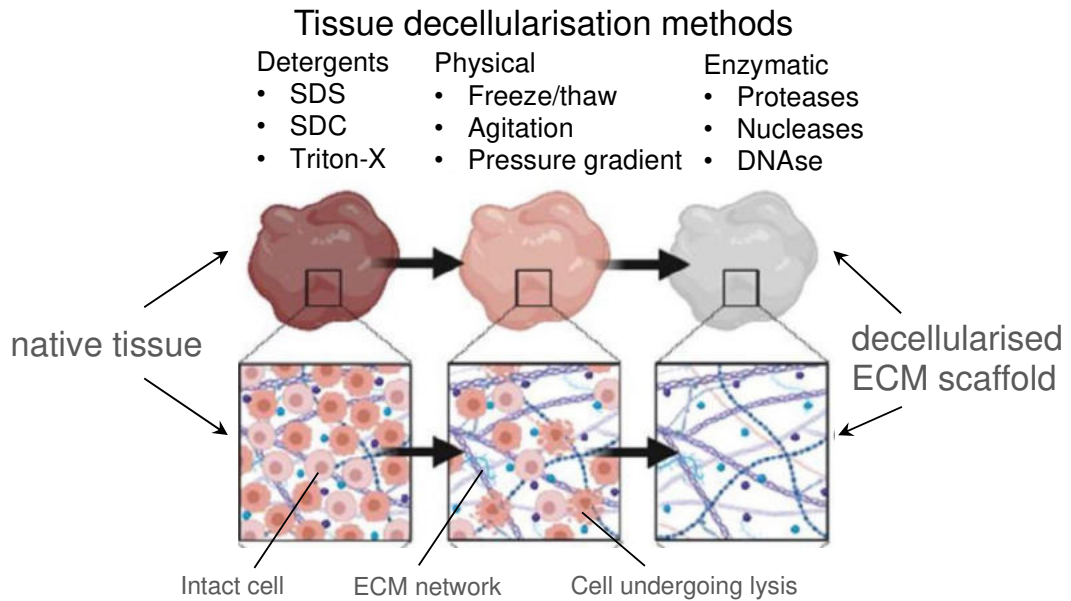


Figure 1.11 Tissue decellularisation.

Schematic showing progressive tissue decellularisation. In this process, intact cells are lysed and removed from native tissue by incubations with chemical detergents, enzymes or physical methods. The tissue ECM network remains intact after the decellularisation. Adapted from Moffat *et al.* (2022) under CC BY NC 4.0 licence³⁴².

This technique involves the removal of nuclear and cellular components, leaving behind a decellularised ECM scaffold (Figure 1.11). Following decellularisation, the native composition and structural organisation of the tissue's ECM remain intact. Consequently, many inherent cell binding sites and protease degradation sites are conserved within the decellularised scaffold, enabling cell adhesion, proliferation, and differentiation^{343–345}.

The discovery that decellularised ECM scaffolds can be solubilised and then moulded into hydrogels has broadened their scope for both *in vitro* and *in vivo*

applications^{346,347}. ECM hydrogels can serve as 2D culture coating substrates^{348–350} or 3D gels³⁴⁸ similar to commercially available collagen type I or Matrigel^{319,351}. Additionally, ECM hydrogels can be used as injectable materials for *in vivo* implantation³⁵².

Decellularised ECM scaffolds and ECM hydrogels can be derived from patient tumour samples, which ensures that the extracted models represent ECM from a clinically relevant TME. Both models serve as valuable platforms for studying the roles of ECM in various cancer processes³⁵³. Breast and colon cancer patient-derived decellularised ECM scaffolds can successfully support tumour growth and proliferation^{354,355}, can be used to study the influence of ECM on migration and angiogenesis³⁵⁶, cancer cell invasion³⁵⁷, to mimic metastatic niche³⁵⁸, and show utility as a platform for evaluating therapeutic drugs^{355,359}. Decellularised ECM scaffolds and ECM hydrogels derived from patients are increasingly being integrated into cancer drug testing platforms^{360–362}. This integration enhances disease modelling and enables drug testing within an environment that closely mimics the native tumour TME. Those testing platforms can be leveraged for high-throughput screening of candidate therapeutics to advance personalised medicine.

1.7.2.1. Decellularisation methods overview

Tissue decellularisation methods include physical, chemical and enzymatic approaches.

Physical methods

Physical methods for tissue decellularisation encompass various techniques, including freezing, direct pressure, sonication, and agitation^{340,341}. Freeze-thaw cycles involve rapid freezing to induce intracellular ice crystal formation, disrupting cell membranes and leading to cell lysis. Direct pressure applies mechanical force to lyse cells. Mechanical agitation and sonication are generally combined with chemical treatment to aid cell lysis and cellular debris removal. Methods like magnetic stir plates, orbital shakers, or low-profile rollers are used for mechanical agitation. Physical decellularisation methods have the advantage of avoiding chemical agents or enzymes that can

sometimes alter the ECM's biochemical properties. Secondly, physical methods do not introduce chemical residues or detergents into the ECM, reducing the risk of potential interference with subsequent cell seeding or biological assays. However, physical techniques have limitations in achieving complete cell removal.

Chemical methods

Non-ionic detergents

Non-ionic detergents like Triton X-100 disrupt DNA-protein, lipid-lipid and lipid-protein interactions while preserving protein-protein interactions³⁶³. As a result, when tissues or organs are treated with non-ionic detergents, the proteins within them are typically left in a functional conformation. This approach allows for cell removal while minimising alterations to the structural and biochemical integrity of the tissue's ECM protein components. However, the studies show that Triton X-100 is ineffective at completely removing cellular material^{364–366}. Regarding the ECM preservation, Triton X-100 can lead to a loss of GAGs and decreased laminin and fibronectin content³⁶⁷.

Ionic detergents

Ionic detergents act by solubilising cytoplasmic and nuclear membranes but tend to denature proteins by disrupting protein–protein interactions³⁶³. Commonly used ionic detergents are sodium dodecyl sulfate (SDS) and sodium deoxycholate. SDS appears to be more effective than non-ionic Triton X-100 for removing nuclei and cellular proteins from dense tissues and organs while preserving tissue mechanics^{368,369}. However, there are reports that SDS alters ECM ultrastructure, including reduced collagen, elastin, GAG and growth factor content^{370–372}. Nevertheless, the effect of residual SDS upon cell repopulation has been considered insignificant^{373,374}.

Enzymatic methods

Decellularisation protocols can include nucleases, trypsin, collagenase, lipase and other enzymes. Enzymes offer specificity in removing cell residues or undesirable ECM components. Nucleases (e.g., DNases and RNases) cleave nucleic acid sequences, aiding nucleotide removal post-cell lysis. Lipase aids in delipidation from tissues with high lipid content, such as adipose tissue³⁷⁵.

A serine protease, trypsin, is superior at preserving GAGs but more disruptive to elastin and collagens than detergents^{367,376}. Nevertheless, complete cell removal via enzymatic treatment alone can be challenging, and enzymes are typically combined with other decellularisation methods.

Protease inhibitors

During decellularisation protocols, the disruption of cells releases intracellular proteases. In the case of extended chemical treatments, these proteases can disrupt the native ECM ultrastructure. To mitigate this, it is advisable to introduce protease inhibitors like aprotinin and leupeptin into the solutions used for tissue immersion^{340,341}.

1.7.3. Tissue-derived ECM hydrogels

Tissue-derived hydrogels have been successfully isolated from diverse tissue sources, encompassing healthy specimens³⁷⁷ and disease-affected tissues, including tumour samples^{378,379}. The generation of tissue-derived ECM hydrogels involves two fundamental steps³⁸⁰. Firstly, the ECM from the starting material (decellularised ECM scaffold) is homogenised and solubilised to ECM protein monomers. After the solubilisation step, the next phase induces the spontaneous reformation of intramolecular bonds among the monomeric ECM components. The solubilised ECM also called the pre-gel ECM solution, is neutralised to physiologic pH, salt concentration, and temperature in vitro to form a hydrogel. The gelation of liquid pre-gel solution to a semi-solid hydrogel is driven by entropy and is primarily governed by collagen kinetics. During this process, collagen monomers lose water, aggregate together, and bury hydrophobic residues within the fibril, leading to the self-assembly of the hydrogel^{381,382}.

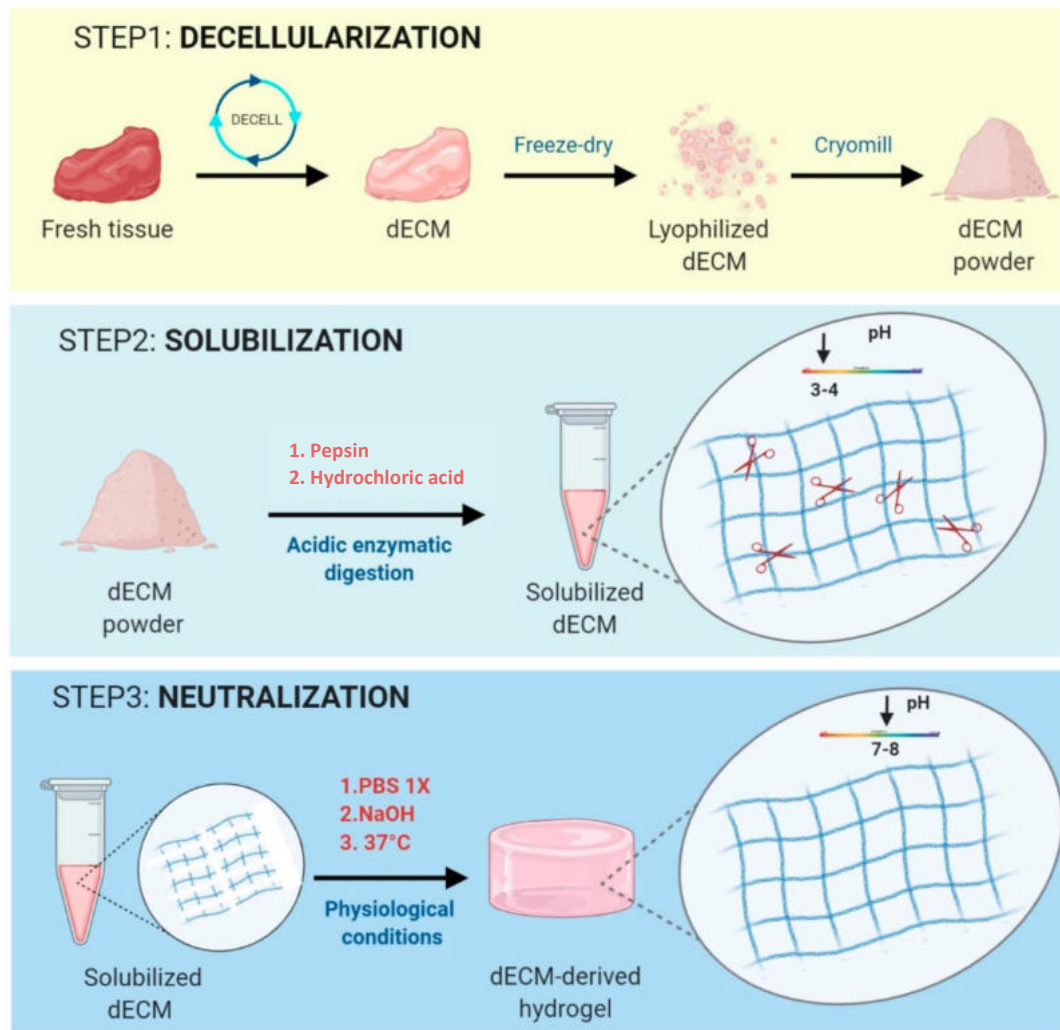


Figure 1.12 Post-processing of decellularised tissue to generate tissue-derived ECM hydrogel.

Tissue-derived hydrogels can be generated through pepsin-mediated digestion involving three sequential steps. Step 1 entails the decellularisation of fresh or frozen tissues (DECELL). Subsequently, the decellularised ECM (dECM) scaffold undergoes freeze-drying and cryomilling to yield a fine decellularised ECM powder. In Step 2, dECM powder is solubilised using pepsin under acidic conditions (pH 3–4), where pepsin cleaves collagen molecules (depicted by red scissors). Finally, Step 3 involves neutralising the acidic environment to achieve physiological osmotic conditions (pH 7–8). This neutralisation process prompts the spontaneous reformation of intramolecular bonds among the digested collagen fragments, resulting in the formation of a homogeneous dECM-derived hydrogel. Adapted from Boso *et al.* (2020) under CC BY 4.0 licence³⁸³.

1.7.3.1. Pepsin-based preparation of ECM hydrogel

Preparing ECM hydrogel involves pepsin-mediated solubilisation of a finely powdered ECM material³⁴⁷ (Figure 1.12). Pepsin is an endopeptidase isolated from porcine gastric mucosa and can solubilise collagen by cleaving the telopeptide bonds of the collagen triple helix³⁸⁴. This enzymatic cleavage leads to the unravelling of collagen fibrils to separate collagen α chains³⁸⁵, increasing collagen solubility. Freytes *et al.* hydrogel preparation protocol involves tissue decellularisation followed by lyophilisation and cryo-milling to obtain a fine ECM powder³⁴⁷. The ECM powder is stirred into pepsin (at a 10:1 mass ratio) diluted in 0.01 M hydrochloric acid and left to digest for 48 h at room temperature. An alternative method involves substituting 0.5 M acetic acid for 0.01 M HCl as the base medium for the pepsin enzyme³⁸⁶. The solubilised ECM (pre-gel ECM) is polymerised to form a hydrogel when the liquid is neutralised to pH 7.0, physiological salt concentration and 37 °C.

A potential limitation of decellularised ECM scaffolds and ECM hydrogels is that interpatient heterogeneity is not captured with a model from a single patient, and experiments should include several patient-derived models. Although neither decellularised ECM scaffolds nor tissue-derived hydrogels have been reported in STS research, both are attractive models to improve our understanding of the role of ECM in STS pathobiology, drug response and resistance. As decellularised ECM scaffolds and tissue-derived hydrogels capture the complexity of TME, incorporating them into developing STS preclinical models may improve drug screening and enable the discovery of more effective therapies.

1.8. Hypothesis and aims

The hypothesis of my thesis is that a comprehensive understanding of the ECM and corresponding integrin adhesion signalling in STS would enable the identification of novel biomarkers and potential therapeutic targets. The hypothesis will be addressed via the following aims:

Aim 1: To perform a comprehensive analysis of matrisome and integrin adhesome components across multiple STS subtypes.

Aim 2: To understand matrix signalling heterogeneity in LMS, DDLPS and UPS.

Aim 3: To generate and characterise STS-specific ECM models suitable for investigating the role of ECM in preclinical settings.

Aim 4: To identify candidate metastasis drivers in LMS and validate the role of zyxin in cell growth and migration in a panel of LMS cell lines.

Chapter 2 Materials and methods

2.1. Bioinformatics and statistical methods

All data was analysed using custom R scripts in R v4.1.1 or later.

2.1.1. Patient cohort and clinical data

The cohort consisted of $n = 321$ STS patients spanning 11 histological subtypes. Retrospective collection of FFPE tissue and clinical data was approved as part of the Royal Marsden Hospital (RMH) PROgnoStic and PrEdiCTive ImmUnoprofiling of Sarcomas (PROSPECTUS) study (RMH Committee for Clinical Research reference 4371, NHS Research Ethic Committee reference 16/EE/0213), National Taiwan University Hospital (Research Ethics Committee Reference 201912226RINB), and as part of Children's Cancer and Leukaemia Group (CCLG) Biological Study 2012 BS 05 (Research Ethics Committee Reference 8/EM/0134). Baseline clinicopathological characteristics and survival data were collected by retrospective review of medical records as part of a previous study by our laboratory¹³⁶. Pseudonymised clinicopathological data were password-protected and securely stored. Analyses were conducted without access to personally identifiable information. Each FFPE block underwent histological evaluation using haematoxylin and eosin (H&E) stained sections. Tumour blocks with $> 75\%$ tumour content were sectioned ($20 \mu\text{m}$) and were used for protein extraction. Samples with $< 75\%$ tumour content were macrodissected prior to protein extraction to enrich for tumour content.

2.1.2. Proteomic data filtering and renormalisation

Previously acquired and processed proteomic data was used for the $n = 321$ cohort analysis¹³⁶. For the sub-cohort and subtype-specific analyses, the raw proteomic data was renormalised to include only samples of interest. The proteomics data was processed in the following way: 1) proteins identified in less than 75% of samples were removed, 2) to address the issue of missing

data, the remaining missing data was imputed using k-nearest neighbours (k-NN) algorithm in the 'impute' R package, 3) each sample was divided by corresponding reference sample within the same tandem mass tag (TMT) set, 4) to remove batch effects, normalisation was performed by \log_2 transformation median centring across samples and z-scoring of proteins within each sample. The resultant dataset was filtered for matrisome and adhesome proteins. The list of matrisome and adhesome proteins was obtained from matrisome database¹⁷⁴ developed by Naba lab, and adhesome database^{295,298}, developed by Geiger lab. Three proteins, ADAM12, SDC4 and MMP14, were present in both matrisome and adhesome databases. ADAM12 was not detected in our proteomic data. SDC4 was considered an adhesome component, and MMP14 was considered a matrisome component for all analyses in this thesis.

2.1.3. Clustering

The processed matrisome and adhesome dataset was visualised using two-way unsupervised clustering based on Pearson's correlation coefficient and using a dimension reduction by uniform manifold approximation and projection (UMAP)³⁸⁷. The clustering was visualised with an annotated heatmap plotted using the 'ComplexHeatmap' package. The UMAP analysis was performed using the default settings and the optimal number of 6 neighbours was chosen. Consensus clustering using the 'ConsensusClusterPlus' package in R was performed to generate stable and robust clustering in LMS, DDLPS and UPS-specific analyses of matrisome and adhesome³⁸⁸. Consensus clustering was conducted using an agglomerative hierarchical clustering algorithm, with Spearman correlation distance and average linking. Protein and sample resampling were set at 80%, and consensus clustering was performed for up to 10 clusters (k). The optimal value of k was determined by inspecting consensus matrices, the cluster tracking plot and the consensus cumulative distribution function (CDF) plot. The statistical significance of the clusters was confirmed using the 'SigClust' package in R with hard thresholding and 1000 sample simulations ($p < 0.05$)³⁸⁹.

2.1.4. Differential expression analysis

Differentially expressed proteins (DEPs) were identified using significance analysis of microarrays (SAM) with the 'samr' package in R³⁹⁰. Normalised and imputed datasets were analysed with two-class unpaired tests based on Student's t-test statistic with 100 permutations for comparisons between two sets of samples. In each comparison, the delta value was selected as the threshold for significance at which the median false discovery rate (FDR) was less than 0.01 and fold change ≥ 2 for upregulated DEPs and ≤ 0.5 for downregulated DEPs.

2.1.5. Overrepresentation analysis

Overrepresentation analyses were performed using the online tool g: Profiler (v.e110_eg57_p18_4b54a898), and the Benjamini–Hochberg FDR method was applied for multiple testing correction with 0.01 FDR threshold³⁹¹. The query lists were compared against custom protein backgrounds specified in each chapter.

2.1.6. Survival analyses

To evaluate whether the survival rates within my STS cohort align with those observed in the general STS population and to identify prognostic biomarkers, three clinical outcome measures were assessed in survival analyses. 1) Local recurrence-free survival (LRFS) was defined as the time from primary disease surgery to radiologically confirmed local recurrence or death. 2) Metastasis-free survival (MFS) was defined as the time from primary disease surgery to radiologically confirmed metastatic disease or death. 3) Overall survival (OS) was defined as the time from primary disease surgery to death from any cause. Clinical data was censored at 5 years to ensure consistency, and patients who had not experienced a survival event were censored at their last follow-up. To evaluate matrixome and adhesome proteins associated with LRFS, MFS and OS, SAM was performed in the 'samr' package in R using the 'survival' response type, with 100 permutations³⁹⁰. In each comparison, the delta value

was selected as the threshold for significance at which the median FDR was less than 0.01. Kaplan-Meier curves were plotted using 'surv' package to visualise the association of protein of interest with the clinical outcomes over time. Median normalised protein expression was used as a cut-off to stratify patients into high and low groups. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test. Multivariable Cox regression analysis was performed to adjust HR and p-values for known prognostic clinicopathological variables. Variable categories were grouped where necessary, as detailed in the thesis chapters.

2.1.7. Matrix scores

To develop matrix scores, the single sample GSEA (ssGSEA) was performed using ssGSEA (v10.1.0) on the GenePattern public server^{392,393}. Rank normalisation and a weighting exponent of 0.75 were used to assess the enrichment of different matrisome classes with gene sets containing at least 10 genes. All human matrisome gene sets were obtained from the c2.cgp.v2023 within the Molecular Signatures Database (MSigDB) v2023.1.Hs^{174,394}. The median value of enrichment scores for each separate gene set was used as a cut-off to stratify patients in high and low groups for survival analyses.

2.1.8. Matrisome and adhesome networks correlation analysis

To assess which matrisome and adhesome proteins are co-regulated in STS, Pearson's correlation coefficients were calculated in R for all possible pairwise combinations of matrisome and adhesome proteins. Distinct clusters of co-regulated matrisome and adhesome proteins were identified by the 'ConsensusClusterPlus' and 'SigClust' packages in R. Hierarchical clustering of the resultant similarity matrix was performed and visualised as a heatmap with the 'ComplexHeatmap' package in R.

2.1.9. Protein-protein interaction network analysis

Protein-protein interaction (PPI) networks were built by querying matrixome and adhesome proteins of interest against the Search Tool for the Retrieval of Interacting Genes/Proteins database (STRINGdb)^{395,396}. Only textmining, experiments and curated databases were used as active interaction sources. A default medium confidence cut-off score of 0.4 in the network was applied; the nodes represent proteins with at least one interaction, and the edges represent interactions between the proteins. The resulting networks were visualised and annotated in Cytoscape v3.10.0³⁹⁷.

2.1.10. Statistical analyses

Statistical analyses were performed using GraphPad Prism software v.9.3.1 for Windows (GraphPad Software, Boston, Massachusetts, USA). As applicable, data distribution was evaluated for normality using Shapiro-Wilk tests, and in cases where $p < 0.05$, non-parametric tests that do not rely on the assumption of a normal distribution were applied. The statistical tests used in this thesis included Kruskal-Wallis one-way analysis of variance (ANOVA) tests, Dunn's tests, one-way ANOVA tests, Tukey's honestly significant difference (HSD) tests, Mann-Whitney U tests and chi-squared tests of independence. Additional information regarding the specific statistical tests conducted can be found in the figure legends.

2.2. Immunohistochemistry for tumour infiltrating lymphocytes

Tissue microarrays (TMA) containing a subset of the mass-spectrometry profiled DDLPS, UPS and LMS cohort were used for IHC. The TMA were constructed, stained and scored by Dr Alex Lee, Dr Cornelia Szecei, Ms Nafia Guljar, with guidance from sarcoma histopathologist Dr Khin Thway. At least two replicate cores per patient were stained for tumour infiltrating lymphocytes (TILs) markers CD3, CD4 and CD8. Serial 4 μm TMA sections were cut and mounted on slides. DAKO link automated stainer (Agilent, CA, USA) with

EnVision FLEX kit (K8002; Agilent) was used for all IHC processing. The slides were deparaffinised with xylene and rehydrated in a series of ethanol washes with decreasing concentration. Antigen retrieval was performed by either pressure cooking in citrate (pH6) for 2 min (CD3) or incubating with pH9 pre-treatment module (PTM) buffer (Agilent, CA, USA) for 20 min at 97 °C (CD4 and CD8). The slides were stained with relevant primary antibodies (CD3 DAKO M0452 at 1:600 dilution; CD4 DAKO 4B12 at 1:80 dilution; CD8 DAKO C8/144B at 1:100 dilution) for 60 minutes at room temperature. Secondary antibody staining was performed using mouse horseradish peroxidase (HRP)-linked antibodies from the DAKO FlexEnvision (Mouse) Kit and applying 3,3' diaminobenzidine (DAB) substrate. The slides were then counterstained with haematoxylin. The CD3/4/8+ TILs counts were conducted under direct brightfield microscopy at a magnification of x400 by Dr Alex Lee and Dr Cornelia Szecsei. Cell counts for cores with section preservation ranging from 50% to 100% were adjusted to represent 100% of the area. Data from cases with section preservation below 50% were excluded from the analysis. Replicate scores were averaged and multiplied by 1.274 ($\pi^*(r^2)$, where $r = 0.5$) to adjust TIL counts to 1 mm². Digital microscopy images for all stained TMA sections were captured at a resolution of x40 using the Nanozoomer-XR microscope (Hamamatsu Photonics, Japan). The relevant IHC data was available for 35 DDLPS, 50 UPS and 63 LMS patients, and the breakdown by TIL score and STS subtype is in Table 2.1.

Table 2.1 Summary of the number of DDLPS, UPS and LMS patients with available IHC data for CD3/4/8+ TILs.

DDLPS: dedifferentiated liposarcoma; UPS: undifferentiated pleomorphic sarcoma and LMS: leiomyosarcoma.

	CD3+	CD4+	CD8+
DDLPS	32	32	35
UPS	50	47	50
LMS	63	63	63
Total	145	142	148

2.3. Generation of preclinical ECM models

2.3.1. Generation of decellularised ECM scaffolds

Patient samples were selected for inclusion based on availability of fresh frozen tissue from primary excision extremity samples with histopathologically confirmed LMS diagnosis. Seven LMS tumour specimens were acquired from the RMH Biobank as part of the CCR 4371 PROSPECTUS. Six LMS tumours occurred in the leg region (5 were from the thigh region and the exact location of the 6th tumour was not specified) and one tumour occurred in the elbow region. Firstly, tumour samples were subjected to three freeze-thaw cycles. Next, tumours were cut into 2-3 mm² cubes and washed in phosphate-buffered saline (PBS) solution containing 1% penicillin-streptomycin. Decellularisation was performed with a standard protocol which uses sodium dodecyl sulphate (SDS) detergent for decellularisation of dense tissues, such as tumour samples³⁹⁸³⁷⁸. Tumours are submerged in PBS solution for 24 h at 4 °C. Tumours were then moved into a detergent solution (0.1% w/v of SDS (Sigma-Aldrich), 10mM Tris buffer pH 8.0 and 0.1% v/v ethylenediaminetetraacetic acid (EDTA)) (Sigma-Aldrich) and incubated for at 4 °C for 4 days until the tissue chunks looked white and semi-transparent. During all the washes and decellularisation steps, the solutions were agitated and changed every 8 – 16 h. All detergent solutions contained 1x Halt Protease and Phosphatase Inhibitor Cocktail (ThermoFisher Scientific). Tumours were rinsed with PBS to remove the detergent and incubated in PBS for 24 h at 4 °C. Tissue was then submerged in DNase solution (30 µg/ml) (BD) in 1.3 mM MgSO₄ and 2 mM CaCl₂ for 1h at room temperature. At the end of the decellularisation, the tissue was washed in ultra-filtrated water for 24 h to remove as much decellularisation agent and cell debris as possible. The resulting ECM decellularised scaffolds were dried in a SpeedVac concentrator (Thermo Scientific).

2.3.2. ECM hydrogel generation

In the absence of a commonly used cryo-milling equipment for homogenising decellularised scaffolds, the protocol for preparing ECM hydrogel was adapted from Nehrenheim *et al.* protocol^{347,399}. Every 10 mg of dry decellularised ECM scaffolds were digested with 1 mg/ml pepsin (Sigma-Aldrich) in 0.01 M hydrochloric acid (HCL). Dry scaffolds in pepsin-HCL solution were mechanically homogenised in 0.5 ml tubes with ceramic beads (1.4 mm diameter, Bertin) at 6800 rpm (3 cycles of 20 sec, 30 sec break in between cycles) using commercial homogeniser system Precellys Evolution (Bertin). The homogenised solution was left at room temperature for 48 h to digest. The digested pre-gel solution was chilled on ice and neutralised with cold 10% (v/v) 0.1M NaOH and 11.1% (v/v) 10x PBS to pH 7-7.5. The gel was formed by incubating the neutralised ECM solution for 1 h at 37 °C. Protein concentration was measured with a Pierce bicinchoninic acid (BCA) assay kit (ThermoFisher Scientific) per the manufacturer's recommendations.

2.4. Characterisation of preclinical ECM models

2.4.1. Protein digestion and sample preparation for mass spectrometry (frozen tumours, scaffolds and hydrogels)

For protein digestion, 1 mg of either a frozen tumour specimen or a dry decellularised ECM scaffold was used. For ECM hydrogel characterisation, 100 µl of digested pre-gel solution was solidified. Then, all samples were homogenised in 8 M urea (Sigma-Aldrich) and 100 mM ammonium bicarbonate (ABC, Sigma-Aldrich) in Precellys Evolution at 6800 rpm (3 cycles of 20 sec, 30 sec break in between cycles). Protein homogenate was diluted 1:4 with water, and the concentration of each sample was measured with a BCA assay. The protein samples were reduced with 10 mM dithiothreitol (Sigma-Aldrich) at 56 °C for 40 min and alkylated with 25 mM iodoacetamide (Sigma-Aldrich) at room temperature for 30 min in the dark. Samples were diluted with 100 mM ABC to 2 M urea concentration. A total of 20 µg of protein sample was digested overnight at 37 °C with 0.8 µg trypsin. The trypsin was

neutralised with 10% trifluoroacetic acid (TFA, Sigma-Aldrich). The neutralised, digested samples were desalted using Pierce C18 spin columns (Thermo Fisher Scientific), dried by SpeedVac and dissolved in 2% acetonitrile (ACN, Fisher Scientific), 0.1% formic acid (FA, Honeywell) (ICR stores).

2.4.2. Liquid chromatography and mass spectrometry (frozen tumours, scaffolds and hydrogels)

For liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, samples were dissolved in Buffer A (2% ACN, 0.1% FA), spiked with iRT calibration mix (Biognosys AG) and analysed on an Agilent 1260 HPLC system coupled to a TripleTOF 5600+ mass spectrometer with NanoSource III (AB SCIEX). Sequential window acquisition of all theoretical mass spectra (SWATH) mass spectrometry data was acquired using two μg of peptides for each sample which was loaded onto a ZORBAX C18 (Agilent Technologies) trap column and separated with an integrated manually pulled tip packed with Reprosil Pur C18AQ beads (3 μm , 120 \AA particles, Dr. Maisch) with a linear gradient of 2 – 40% of Buffer B (98% ACN, 0.1% FA). Full-profile MS scans were acquired in the mass range of m/z 340–1400 in positive ion mode. Precursor isolation windows had a fixed size of 25 Da across the mass range of m/z 350–1250 with 1 Da overlap. MS/MS scans were acquired in the m/z 100–1500 mass range. All SWATH data were analysed against a publicly available pan-human library using DIA-NN (version 1.8)⁴⁰⁰, and trypsin was specified as the cleavage enzyme. The spectral library was refined using the dataset with 0.01 FDR to generate an *in-silico* library, which was used to reanalyse the data. The subsequent report was filtered at a q-value of 0.01 for both precursor and proteins. Each sample was analysed with two technical replicates, and a protein was considered expressed if it was detected in at least one of the technical replicates. The average of raw values was calculated if a protein was detected in two technical replicates. Proteins were annotated as matrisome and non-matrisome according to MatrisomeDB¹⁷⁴. Raw protein intensities were used to identify proteins in the samples and calculate the percentage of relative abundance of the matrisome ((sum of all matrisome proteins intensities/sum of all proteins intensities in a sample)*100%).

2.5. Cell line maintenance

Human uterine leiomyosarcoma SK-UT-1 and SK-UT-1b cell lines (obtained from Dr Priya Chudasama, German Cancer Research Centre, Heidelberg, Germany) were cultured in Minimum Essential Media (MEM) (Gibco) supplemented with 0.5% penicillin/streptomycin and 10% fetal bovine serum (FBS)^{401,402}. Human uterine leiomyosarcoma SHEF-LMS w1 and SHEF-LMS ws cells (obtained from Dr Karen Sisley, The University of Sheffield, Sheffield, UK) were cultured in Roswell Park Memorial Institute (RPMI) 1640 supplemented with 0.5% penicillin/streptomycin, 10% FBS, 1% v/v of 240mM L-glutamine and 0.4% v/v of 1M D-glucose⁴⁰³. Human uterine leiomyosarcoma ICR-LMS-1 were previously established and characterised in our lab and were grown in Dulbecco's modified Eagle medium (DMEM):Ham's F12 1:1 + 15 mM HEPES, supplemented with 0.5% penicillin/streptomycin, 10% FBS, 1% v/v of 240mM L-glutamine, 5 µg/ml of bovine insulin (Sigma), 0.4 µg/ml hydrocortisone (Sigma), 10 ng/ml epidermal growth factor (Peprotech), 250 ng/ml amphotericin B (ThermoFisher Scientific), 9.6 ng/ml cholera toxin (Sigma), 5 µM Y-27632 dihydrochloride (LC labs). All cells were cultured in 5% CO₂ at 37 °C in humidified incubators. After reaching 70-80% confluency, the cells were rinsed with PBS, and detached using 0.05% Trypsin in 0.02% EDTA. Trypsin was neutralised by the addition of fresh FBS-supplemented growth media; the cells were diluted in the media to the appropriate concentration and transferred to a new flask. The procedure was repeated approximately every 3-4 days.

2.6. Molecular biology techniques

2.6.1. Short hairpin RNA (shRNA) oligo design and cloning

The ZYX and scramble control shRNA sequences were obtained from the RNAi Consortium of the Broad Institute. The oligos were ordered from Sigma and cloned into the lentiviral Tet-pLKO-puro vector (Addgene: #21915), according to Wiederschain *et al.* protocol⁴⁰⁴. Briefly, complementary oligo

sequences (Table 2.3) are annealed using 10X annealing buffer (1M NaCl, 100 mM Tris-HCl, pH = 7.4) with a thermocycler program (5 min at 95, cool down at 95 °C – 1 °C/min rate to 25 °C). Then, 1 µl of the oligo mixture was diluted 1:400 in 0.5X annealing buffer. The tet-pLKO-puro vector was digested with restriction enzymes AgeI (New England Biolabs) and EcoRI (New England Biolabs), and the digested vector was gel-purified. The ligation reaction was set up with 1 µl of diluted oligo mixture, 40 ng Gel-purified digested Tet-pLKO-puro, 10 µl of Quick ligase buffer (New England Biolabs), 1 µl of Quick ligase (New England Biolabs) and nucleases-free water was used to make up the solution to 10 µl. The reaction was incubated at RT for 15 min, and bacteria transformation with the ligation reaction product was performed on the same day.

Table 2.2 Short hairpin RNA (shRNA) oligo sequences used for cloning.

	Forward Primer (5' - 3')	Reverse Primer (5' - 3')
Scramble	CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTTTCATCTTGTGTTTTG	AATTCAAAAACAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTTTCATCTTGTG
ZYX1	CCGGCTTCCACATGAAGTGTACAACCTCGAGTTGTAACACTTCATGTGGAAGTTTTG	AATTCAAAAACTTCCACATGAAGTGTACAACCTCGAGTTGTAACACTTCATGTGGAAG
ZYX2	CCGGCTGGGTCACAACCAATCAAACCTCGAGTTTGATTTGGTTGTGACCCAGTTTTG	AATTCAAAAACTGGGTCACAACCAATCAAACCTCGAGTTTGATTTGGTTGTGACCCAG
ZYX5	CCGGGAAGGTGAGCAGTATTGATTTCTCGAGAAATCAATACTGCTCACCTTCTTTTTG	AATTCAAAAAGAAGGTGAGCAGTATTGATTTCTCGAGAAATCAATACTGCTCACCTTC

2.6.2. Bacterial transformation

To generate greater quantities of plasmid DNA, 33 µl aliquots of XL-10-Gold ultracompetent bacteria (Agilent) were mixed with 1.3 µl β-mercaptoethanol (Agilent) and incubated for 10 min on ice. Next, 3 µl of the ligation reaction was added, and bacteria were incubated for 30 min on ice. Bacteria were then transformed by heat shock for 30 sec at 42 °C and incubated on ice for 2 min to recover. Following this, 250 µl of Luria-Bertani broth (LB broth; 5 g/L NaCl, 10 g/L bacto tryptone, 5 g/L yeast extract) was added to the transformation mix, and samples were incubated for 1 h at 37 °C with shaking at 225 rpm. Working next to the flame, 100 µl of transformed bacteria were spread onto LB-agar (LB broth, 15 g/L agar) plates containing the selection antibiotic, ampicillin (100 µg/ml). Plates were incubated at 37°C for 16 h until individual colonies could be observed.

2.6.3. Plasmid DNA preparation

A starter culture of 5 ml of LB broth containing 100 µg/ml ampicillin was inoculated with single bacterial colonies of interest and incubated at 37°C with shaking at 225 rpm overnight. For small-scale plasmid preparation, 4 ml of bacterial culture was centrifuged at 12,000 x g for 10 minutes to pellet bacteria and plasmid DNA was isolated using QIAprep Miniprep kit following manufacturer's instructions (QIAGEN). After elution, DNA concentration was assessed by measuring absorbance using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific).

For long-term storage of transformed bacteria, liquid bacterial culture was mixed 1:1 with a 50% glycerol solution and stored at -80°C. All plasmids were sequenced using Eurofins Genomics TubeSeq service to verify their coding sequences, and sequence chromatograms were visualised and aligned using Benchling (<https://benchling.com>). The primer used for sequencing (GGCAGGGATATTCACCATTATCGTTTCAGA).

2.6.4. Lentiviral plasmid preparation, transfection and transduction

Lentiviruses were produced by co-transfecting HEK-293T cells with the Tet-pLKO-puro vector containing the 8 µg of ZYX or scramble control shRNA sequences and with 4 µg pMD2.G (Addgene: #12259) and 4 µg psPAX2 (Addgene: #12260) vectors⁴⁰⁵. For the GFP transduction, HEK-293T cells were co-transfected with 8 µg of PGK-H2B GFP (Addgene: #21210), 4 µg pMD2.G and 4 µg psPAX2. After 96 h, the viral supernatant was harvested and filtered using a 0.45 µm pore filter. SK-UT-1 (150,000 cells/well), SHEF-LMS w1 (100,000 cells/well) and SHEF-LMS ws (100,000 cells/well) were plated in 6 well plates. After 24 h, the medium was aspirated and replaced with twofold serial dilutions of medium containing lentivirus (ranging from 1:2 to 1:8) and a final concentration of 8 µg/ml polybrene (Sigma-Aldrich), a cationic polymer used to enhance transduction. Transduced cells were selected using 2 µg/ml of puromycin (Sigma-Aldrich) for 72 h.

2.7. Phenotypic assays

2.7.1. ECM-coated plate preparation

Immediately prior to the assay, plates were coated with collagen IV (C5533, Sigma LOT109K3801), patient-derived LMS ECM pre-gel solution (LMS ECM) diluted in ice-cold 0.1 M acetic acid or fibronectin (F1141, Sigma LOT SLCD2908), laminin (CC095, Merck, LOT 3660319) diluted in ice-cold in PBS, according to the manufacturer's instruction. The plastic control condition was coated with 0.1 M acetic acid. After 2 h of incubation at 37 °C and 5% CO₂, the coating solution was aspirated, washed with PBS, and the cells were seeded (density depended on cell types). If the plates were not used on the day, they were paraffin-sealed and stored at 4 °C for no longer than a week.

2.7.2. Cell adhesion assay

The wells of a black-walled PhenoPlate 96-well microplate (Perkin Elmer) were pre-coated with patient-derived LMS ECM or 0.1 M acetic acid (for plastic control). SK-UT-1 cells (10,000 cells/well) were seeded. Media was removed after either 1 h or 24 h, and cells were fixed in 10% neutral-buffered formalin solution (Sigma Aldrich) for 15 mins at room temperature. To enable visualisation of cell nuclei, the cells were incubated with 0.05% (v/v) Hoechst 33342 (in PBS) (Tocris) for 15 mins at 37 °C and washed with PBS. Afterwards, the cells were counted using Celigo Image Cytometer (Nexcelom Biosciences).

2.7.3. Growth curve assay

Cells (250 cells/well) were seeded into black-walled PhenoPlate 96-well microplates (Perkin Elmer) pre-coated with patient-derived LMS ECM for the experiment in Chapter 5 or seeded directly on plastic for the experiments in Chapter 6. After 24 hours, one plate was fixed with 10% neutral-buffered formalin solution (Sigma Aldrich) and stored at 4 °C. Formalin fixing was repeated daily on a single plate for 7 days. Media was replenished every 72 h.

After 7 days, cells were fixed in 10% neutral-buffered formalin solution (Sigma Aldrich) for 15 mins at room temperature and stained with 0.05% (v/v) Hoechst 33342 (in PBS) (Tocris) for 15 mins at 37 °C and washed with PBS. Direct cell count was undertaken using Celigo Image Cytometer (Nexcelcom BioScience). Direct cell count was plotted against time, the Malthusian exponential growth curve was fitted to calculate the doubling time using GraphPad Prism software v.9.3.1 for Windows (GraphPad Software, Boston, Massachusetts, USA).

2.7.4. Colony formation assay

ZYX-knockdown cells were seeded into 6 well plates at a density of 1,000 cells/well (SK-UT-1) and 2,000 cells/well (SHEF-LMS w1 and SHEF-LMS ws). After 24 hours, either 2 µg/ml doxycycline-supplemented media or media without doxycycline were added, and the cultures were maintained for 2 weeks. Media and doxycycline were replenished every 72 hours. To assess colony formation of SK-UT-1 cells under dasatinib treatment, the cells were seeded into 12 well plates at a density of 100 cells/well either on plastic or plastic pre-coated with 10 µg/ml LMS ECM. After 24 hours, cells were treated with dasatinib at the indicated concentration or media with dimethyl sulfoxide (DMSO), and the cultures were maintained for 2 weeks. Media with DMSO or dasatinib were replenished every 72 hours. After 2 weeks, cells were fixed using the Carnoy's fixative (3:1 methanol:acetic acid) and stained with 1% crystal violet solution (Sigma Aldrich). Plates were digitally imaged using G-Box Chemi-XX6 (Syngene) imagers. Cell density was quantified using the ImageJ software program (National Institutes of Health).

2.7.5. Dose response proliferation assay

Cells (1,500 cells/well) were seeded in triplicate into black-walled PhenoPlate 96-well microplates (Perkin Elmer) previously pre-coated with 10 µg/ml of LMS ECM or 0.1 M acetic acid. After 24 h, the media was replaced with fresh media containing serial dilutions of dasatinib (0-50,000 nM) or DMSO. After 72 h, media was removed, and cells were fixed in a 10% neutral-buffered formalin

solution (Sigma Aldrich) for 15 mins at room temperature. To enable visualisation of cell nuclei, the cells were incubated with 0.05% (v/v) Hoechst 33342 (in PBS) (Tocris) for 15 mins at 37 °C and washed with PBS. Afterwards, the cells were counted with Celigo Image Cytometer (Nexcelom Biosciences). Raw cell count values were normalised against the vehicle control (plated on its respective ECM coating or plastic), and four-parameter non-linear regression curve-fitting was performed using GraphPad Prism v.9.3.1 for Windows (GraphPad Software, Boston, Massachusetts, USA) to calculate the IC₅₀ values. A minimum of three biological replicates were used to calculate the average IC₅₀ value.

2.7.6. Small molecule inhibitor screen

Cells (1,500/well) were seeded into black-walled PhenoPlate 96-well microplates (Perkin Elmer) previously pre-coated with 10 µg/ml of patient-derived LMS ECM or 0.1 M acetic acid. After 24 hours, cells were treated with inhibitors at 500nM and incubated for 72 hours before the cell viability measurement. Direct cell count was undertaken using a Celigo Image Cytometer (Nexcelcom BioScience). For LMS ECM coating and plastic, raw cell count values were normalised against their respective vehicle control (plated on LMS ECM coating or plastic). A list of the small molecule inhibitors utilised, along with their primary target and supplier, is presented in Table 2.2. Data was clustered two-way based on Pearson distance and visualised as a heatmap with the 'ComplexHeatmap' package in R.

Table 2.3 Small molecule inhibitor screen components.

List of small molecule inhibitors, targets and suppliers. ALK; Anaplastic lymphoma kinase, Bcl-2/x; B-cell lymphoma 2/extra large protein; Bcl-w; Bcl-2-like protein 2, Bcr; Breakpoint cluster region protein, CDK4/6; Cyclin-dependent kinase 4/6; Chk1; Checkpoint kinase 1, CK2; Casein kinase 2, EGFR; Epidermal growth factor, FAK; Focal adhesion kinase, HER2; Human epidermal growth factor receptor 2, Hsp90; Heat shock protein 90, IGF1R; Insulin-like growth factor 1 receptor, IKK(1/2); IκB kinase (1/2), InsR; Insulin receptor, JAK(1/2); Janus kinase (1/2), JNK(1/2/3); c-Jun N-terminal kinase (1/2/3), MAPK; Mitogen-activated protein kinase, MEK; Mitogen-activated protein kinase kinase, mTOR; Mechanistic target of rapamycin, N-terminal; Amino-terminal, NTRK(1/2/3); Neurotrophic tyrosine kinase receptor (1/2/3) PARP; Poly (ADP-ribose) polymerase, PI3K; Phosphoinositide 3-kinase, PLK1; Polo-like kinase 1,

(B/C)-Raf; Rapidly accelerated fibrosarcoma, RTK; Receptor tyrosine kinase, STAT(3); Signal transducer and activator of transcription (3), TGFβR1; Transforming growth factor β receptor

1.

Small molecule inhibitor	Primary target(s)	Supplier
Adezmapimod (SB203580)	p38 MAPK	Selleck Chemicals
Alisertib	Aurora A	Selleck Chemicals
BI-2536	PLK1	Selleck Chemicals
Binimetinib	MEK1/2	LC Laboratories
BMS345541	IKK1/2	Sigma Aldrich
Bosutinib	Src, Abl1	LC Laboratories
Capivasertib (AZD-5363)	Akt1/2/3	Selleck Chemicals
Cediranib	Broad spectrum: RTKs	LC Laboratories
Ceritinib	ALK	Selleck Chemicals
Cilengitide trifluoroacetate	Integrins αβ3, αβ5	Selleck Chemicals
Crizotinib	ALK, MET	LC Laboratories
Dabrafenib	B-Raf V600E	Selleck Chemicals
Dactolisib (BEZ235)	PI3K, mTOR	LC Laboratories
Dasatinib	Src, Abl1, Broad spectrum: RTKs	LC Laboratories
Erlotinib	EGFR	LC Laboratories
Foretinib	Broad spectrum: RTKs	LC Laboratories
Galunisertib	TGFβR1	Selleck Chemicals
Gefitinib	EGFR	LC Laboratories
GW441756	NTRK1	Selleck Chemicals
Imatinib	Broad spectrum: RTKs, Abl1	LC Laboratories
Lapatinib	EGFR, HER2	LC Laboratories
Lenvatinib	Broad spectrum: RTKs	LC Laboratories
Linsitinib	IGF1R	LC Laboratories
Luminespib (NVP-AUY922)	Hsp90	LC Laboratories
MK2206	Akt1/2/3	Selleck Chemicals
Momelotinib	JAK1/2	Selleck Chemicals
Navitoclax	Bcl-2, Bcl-w, Bcl-xL	Selleck Chemicals
Neratinib	EGFR, HER2	LC Laboratories
Nicosamide	STAT3	Selleck Chemicals
NVP-AEW541	IGF1R, InsR	Selleck Chemicals
NVP-TAE684	ALK	Selleck Chemicals
Osimertinib (AZD-9291)	EGFR	Selleck Chemicals
Palbociclib	CDK4/6	Selleck Chemicals
Pazopanib	Broad spectrum: RTKs	LC Laboratories
PF562271	FAK	Selleck Chemicals
Ponatinib	Broad spectrum: RTKs, Abl1	LC Laboratories
Rabusertib (LY2603618)	Chk1	Selleck Chemicals
Rapamycin (Sirolimus)	mTOR	LC Laboratories
Regorafenib	Broad spectrum: RTKs	LC Laboratories
Rucaparib	PARP	Selleck Chemicals
Saracatinib	Src	Selleck Chemicals
Silmitasertib	CK2	Selleck Chemicals
Sorafenib	Broad spectrum: RTKs, C-Raf, B-Raf	LC Laboratories
SP600125	JNK1/2/3	Selleck Chemicals
Sunitinib	Broad spectrum: RTKs	LC Laboratories
Talazoparib	PARP	Selleck Chemicals
Trametinib	MEK1/2	LC Laboratories
Vandetanib	Broad spectrum: RTKs	LC Laboratories

2.7.7. Live single-cell imaging

GFP+ SK-UT-1, ICR-LMS-1, SHEF-LMS w1 and SHEF-LMS ws (3,000 cells/well), and SK-UT-1b (8,000 cells/well) were seeded into plastic black-walled PhenoPlate 96-well plastic microplates (Perkin Elmer) or plastic previously pre-coated with 10 µg/ml of collagen IV, fibronectin, laminin or

patient-derived LMS ECM. Those densities ensured single cell tracking during migration experiments. To assess the effect of ZYX knockdown on LMS cell migration, GFP+ SK-UT-1, SHEF-LMS w1 and SHEF-LMS ws cells transduced with three independent ZYX or scramble control shRNA constructs were seeded on plastic. The ZYX knockdown was induced with 2 µg/ml doxycycline for 72 h before the migration experiment, and the doxycycline was withdrawn for the duration of the migration experiment. After allowing the cells to attach for 5 h, the cells were washed twice with complete growth media to remove non-attached cells. Live-cell imaging was performed with a confocal laser scanning microscope ZEISS LSM 980 (Carl Zeiss AG, Oberkochen, Germany) equipped with a thermostatic humid chamber with 5% CO₂ and 37 °C. All images were acquired with a 10x 0.3-NA air objective (model EC Plan-Neofluar 10x/0.30 M27), inverted microscope Axio Observer Z1/7 and LSM980 Airyscan detector 1.3x zoom, laser 488 nm at 0.7% power. Images were acquired every 30 min for 18 h for 5 LMS cell lines plated on ECM coatings or for 12 h for the ZYX knockdown experiment across 20 fields of view per condition. Raw files were processed in Zen 2.1 Blue software (Carl Zeiss) with Airyscan-Processing, the fields of view were stitched, and files were exported with OME-TIFF export. The processed images were analysed with the TrackMate v7 plugin in Fiji/ImageJ (National Institutes of Health)^{406,407}. The TrackMate parameters were optimised on fibronectin conditions separately for each cell line (Figure 2.1). The optimised parameters were used for batch analysis with TrackMate Batch v1.2.3 for each cell line. Information about migration speed and directionality index were extracted from the track's features calculated in TrackMate. The directionality index is calculated as the net distance (Euclidean distance) between the start and finish divided by the total distance travelled⁴⁰⁸.

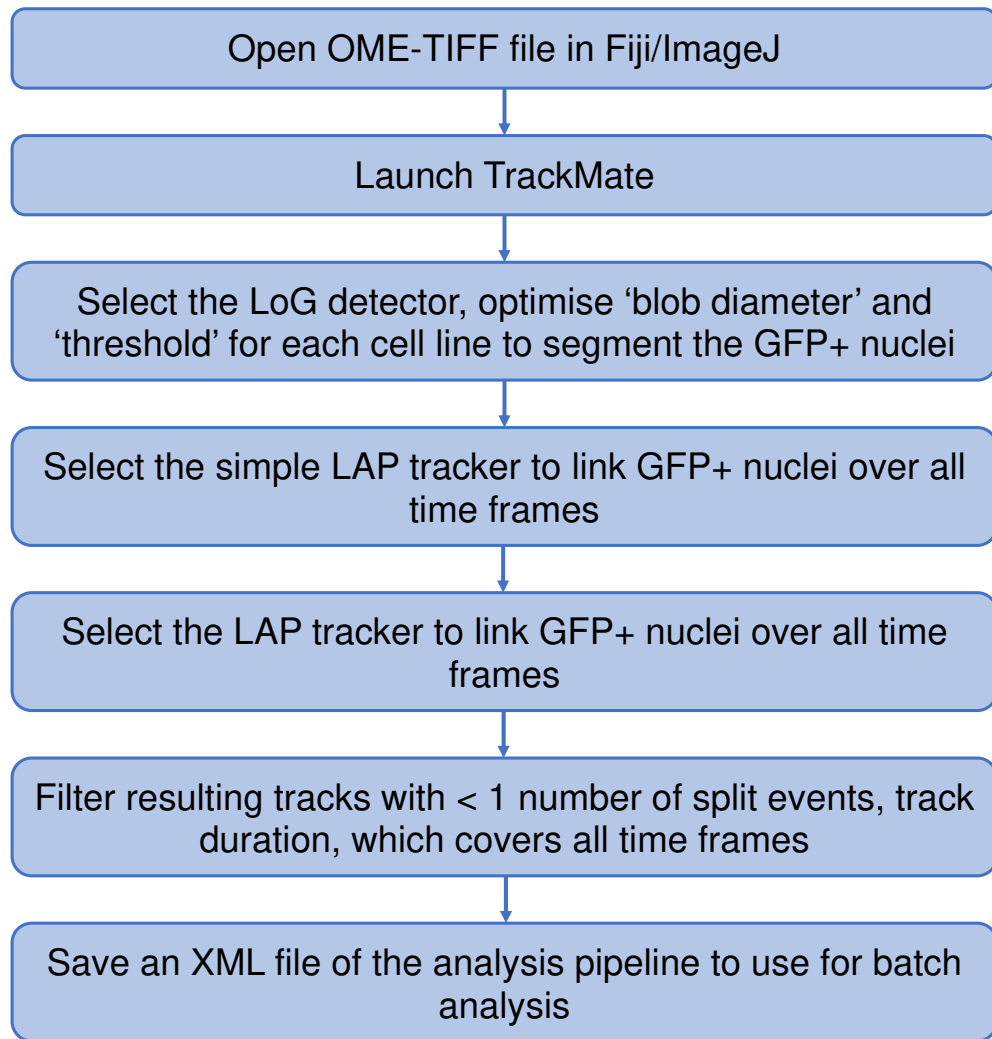


Figure 2.1 A flowchart showing the step-by-step procedure to track GFP+ cells using the TrackMate plugin in Fiji/ImageJ.

Given that LMS cells showed one of the fastest migrations on fibronectin, the OME-TIFF files from this condition were used to optimise the cell tracking, separately for each LMS cell line. The steps were the following. Open the OME-TIFF file in ImageJ using the hyperstack option. Within the ImageJ software, launch the TrackMate plugin. To optimise the segmentation parameters select the LoG detection algorithm, optimise 'blob diameter' and 'threshold' for each cell line to make sure all the GFP+ nuclei are correctly segmented. Next, the Linear Assignment Problem (LAP) particle-linking algorithm is selected and the track i.e. cell's path are generated for each GFP+ nuclei. Set track filters: 1) < 1 number of split events to ignore the tracks generated due to cell division, 2) set the maximum track duration to track all cells from the beginning till the end of the assay. Save the analysis pipeline as an XML file, to be used to run batch analysis for the rest of the conditions.

2.8. Protein analysis

2.8.1. Immunoblotting

To confirm ZYX knockdown SK-UT-1, SHEF-LMS w1 and SHEF-LMS ws cells (50,000/well) were seeded into 6 well plates. After 24 hours, cells were treated with 2 µg/ml doxycycline for 72 h to induce the ZYX knockdown. After the indicated time post-treatment, the cells were lysed in radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% IGEPAL CA-630 (NP-40) (Sigma Aldrich), 0.1% SDS (Sigma Aldrich), and 0.5% sodium deoxycholate (Sigma Aldrich)), supplemented with the Halt protease and phosphatase inhibitors and EDTA (ThermoFisher Scientific) at 4 °C. To measure ZYX expression in a panel of LMS cell lines, the cell lysates (SK-UT-1, SK-UT-1b, ICR-LMS-1, SHEF-LMS w1, SHEF-LMS ws, SARC-393, ICR-LMS-4 and ICR-LMS-6) collected by a previous lab member were used. Protein quantification was undertaken using a BCA assay kit (ThermoFisher Scientific) per the manufacturer's recommendations. Cell lysates (20-25 ug) were loaded onto NuPAGE Novex 4-12% Bis-Tris gels (Invitrogen), followed by blotting onto iBlot polyvinylidene fluoride (PVDF) membranes (Invitrogen). Membranes were blocked for 1-2 hours at RT with 5% dried skimmed milk (Marvel) in 1x tris-buffered saline (TBS) with 0.1% Tween 20 (TBST) (Sigma Aldrich). Blots were incubated with primary antibodies (diluted in 5% dried skimmed milk in TBST) at 4 °C overnight. Membranes were washed at RT with 3 x 10 minutes of TBST incubation. Following this, membranes were incubated at RT with HRP-conjugated secondary antibodies (diluted in 5% dried skimmed milk in TBST) for 1 hour. Membranes were washed with 3 x 10 minutes of TBST incubation at RT. Primary and secondary antibodies utilised within this thesis are outlined in Table 2.4 with associated dilutions. Immunoreactive bands were visualised by SuperSignal West Pico PLUS chemiluminescence substrate (ThermoFisher Scientific), and the blots were digitally imaged using ChemiDoc Touch Imaging System (Bio-Rad) or G-Box Chemi-XX6 (Syngene) imagers. Blots were analysed using Image Lab (Bio-Rad) and GNU image manipulation program (GIMP).

Table 2.4 Primary and secondary antibodies used for immunoblotting with associated dilutions and supplier information.

GAPDH: glyceraldehyde-3-phosphate dehydrogenase HRP: horse radish peroxidase.

	Antibody (clone)	Dilution	Supplier
Primary	Anti-Zyxin (polyclonal)	1:2000	Atlas Antibodies (HPA073497)
	Anti- α -tubulin (B512)	1:10000	Sigma-Aldrich (T5168)
	Anti-GAPDH (D4C6R)	1:10000	Cell Signaling (97166S)
Secondary	Anti-rabbit HRP	1:10000	Cell Signaling (7074)
	Anti-mouse HRP	1:10000	Strattech (G32-62G-SGC)

Chapter 3 Overview of STS matrisome and adhesome

3.1. Background and objectives

Multiple proteins within the ECM act together in a coordinated manner to initiate integrin-mediated adhesome signalling pathways^{176,291,409}. Studying the interactions of these components is essential for understanding tumour growth, metastasis, therapy resistance and the development of targeted therapies⁴¹⁰. Moreover, specific matrisome and adhesome proteins have demonstrated potential as biomarkers in other cancer types⁴¹¹. Identifying and validating matrisome and adhesome biomarkers in STS can improve prognostication and personalised treatment approaches. At present, we have a limited understanding of the matrisome constituents in STS through immunohistochemistry-based studies and a handful of mass spectrometry-based analyses (discussed in Section 1.6.1). Notably, none of these studies was specifically dedicated to characterising the adhesome in STS. Chapter 3 aims to bridge this gap in the knowledge of STS pathobiology by conducting a comprehensive analysis of the matrisome and adhesome sub-proteomes. This analysis leverages previously collected mass spectrometry proteomic data encompassing 11 histological subtypes of STS¹³⁶. Furthermore, it is worth noting that components of the tumour matrisome play a role in shaping the tumour's immune microenvironment⁴¹². For instance, dense collagen matrix was shown to restrict T cell infiltration in lung cancer⁴¹³. Additionally, high density tumour ECM impaired T cell proliferation, downregulated genes involved in T cell cytotoxicity and impaired ability of T cells to kill breast cancer⁴¹⁴. Therefore, the relationship between matrisome components and TILs was evaluated. Accordingly, the objectives of this chapter were twofold. Firstly, I set out to comprehensively characterise components of the matrisome and adhesome and identify networks of co-regulated matrisome and adhesome proteins that operate in STS. The second objective was to assess the prognostic utility of matrisome proteins and TILs in STS.

3.2. Results

3.2.1. Cohort characteristics

Proteomic profiling and clinical data collation was done by previous lab members. Proteomic and clinical data were available for 321 primary tumours across 11 histological subtypes. Among these subtypes, the most common were LMS (25%), UPS (17%), synovial sarcoma (SS 13%), DDLPS (12%) and DES (12%) (Table 3.1 and Supplementary Table 3.1). The cohort also contained some of the ultra-rare STS subtypes such as alveolar soft part sarcoma (1%), desmoplastic small round cell tumours (DSRCT 1%) and clear cell sarcoma (CCS 1%). Rhabdoid tumours (RT), a paediatric STS subtype, comprised 5% of the cohort. DES are locally aggressive tumours lacking metastatic potential, and together with RT, those were used in descriptive analysis but were not included in survival analyses. For the survival analyses, the clinical data was censored at 5 years (60 months) post-surgical resection.

In terms of clinicopathological features of the cohort, the median age at diagnosis was 58 years (range: 0.1 – 90). Across the subtypes, the highest median age was for UPS (at 74 years) and the lowest for RT (at 1.1 years). UPS had the highest percentage (93%) of high-grade (grade 3) tumours, and ES had the highest percentage (63%) of intermediate (grade 2) tumours. None of the tumours were of low grade (grade 1) in the cohort. Most of the tumours across the cohort arose in the extremities (39%), and for each subtype, the extremities were the most common location. The exceptions were DDLPS, where 82% of tumours arose in the retroperitoneum, and AS and DES, where most tumours (70% and 60%, respectively) were found in the trunk. Seventy-eight percent of tumours in the cohort were located in deep anatomical regions. Median tumour size (maximum tumour diameter) was 90 mm across the cohort (range: 4 – 1090 mm), and DDLPS tumours were the largest, followed by DSRCT (median size of 190 and 132.5 mm, respectively).

Table 3.1 Clinicopathological characteristics of n=321 soft tissue sarcoma (STS) cases

Summary of features of the whole cohort. For continuous variables, median, minimum (min) and maximum (max) values are indicated. For categorical variables, count (n) and percentages (%) are shown. F = female; M = male.

Characteristic	Category	Whole cohort
Number of patients	n (%)	321 (100)
Histological subtype	Angiosarcoma	30 (9)
	Alveolar Soft Part Sarcoma	4 (1)
	Clear Cell Sarcoma	3 (1)
	Dedifferentiated Liposarcoma	39 (12)
	Desmoid Tumour	37 (12)
	Desmoplastic Small Round Cell Tumour	4 (1)
	Epithelioid Sarcoma	16 (5)
	Leiomyosarcoma	80 (25)
	Rhabdoid Tumour	12 (4)
	Synovial Sarcoma	43 (13)
	Undifferentiated Pleomorphic Sarcoma	53 (17)
Age at excision (years)	median	58.4
	min	0.1
	max	90
Anatomical site [n (%)]	Extremity	125 (38.9)
	Head/neck	13 (4.0)
	Intra-abdominal	28 (8.7)
	Retroperitoneal	57 (17.8)
	Trunk	65 (20.2)
	Pelvic	24 (7.5)
	Uterine	9 (2.8)
Grade [n (%)]	2	115 (35.8)
	3	139 (43.3)
	unknown	67 (20.9)
Tumour depth [n (%)]	Deep	250 (77.9)
	Superficial	54 (16.8)
	unknown	17 (5.3)
Tumour size (mm)	median	90
	min	4
	max	1090
Tumour margins [n (%)]	R0	133 (41.4)
	R1	151 (47.0)
	R2	4 (1.2)
	unknown	33 (10.3)
Pre-op treatment [n (%)]	Chemo	19 (5.9)
	Radio	8 (2.5)
	Chemo & radio	13 (4.0)
	None	267 (83.2)
	unknown	14 (4.4)
Performance status [n (%)]	0	158 (49.2)
	1	82 (25.5)
	2	16 (5.0)
	3	5 (1.6)
	unknown	60 (18.7)
Sex [n (%)]	F	201 (62.6)
	M	119 (37.1)
	unknown	1 (0.3)
Status at excision [n (%)]	Local	301 (93.8)
	Metastatic	15 (4.7)
	Locally Metastatic	3 (0.9)
	Multifocal	1 (0.3)
	unknown	1 (0.3)

Collectively, 88% of tumours had R0 or R1 surgical margins. Notably, DDLPS had the highest percentage of R1 tumours, likely due to the predominance of deep and hard-to-access tumours in the retroperitoneum, where a compromise between a sufficient resection and limiting the potential surgical morbidity needs to be achieved. Patients in the cohort were mainly treatment-naïve (83%); however, all DSRCT received either pre-operative chemotherapy or radiotherapy. Additionally, a subset of SS patients received pre-operative chemo-, radiotherapy or both prior to the surgical excision (16, 14 and 28%, respectively), and 5 AS patients (17%) received neoadjuvant chemotherapy. In terms of the survival outcome measures, the median OS for the cohort was 50 months (95% CI 41 – not achieved (NA)), MFS was 48 months (95% CI 33 - NA) and the median survival for LRFS was not reached (Supplementary Figure 3.1).

3.2.2. Matrisome and adhesome overview in STS

To characterise the components of the matrisome and adhesome in the STS cohort, the proteomic dataset (n = 3,290) was queried against publicly available databases. The Matrisome database covers n = 1,027 proteins, of which n = 193 (19%) were present in our dataset (Figure 3.1A)¹⁷⁴. Of these, n = 99 were core matrisome and n = 94 were matrisome-associated proteins (Figure 3.1B). Within the core matrisome, n = 18 collagen chains, n = 64 glycoproteins and n = 12 proteoglycan were detected. The breakdown of matrisome-associated proteins was as follows: n = 60 ECM regulator proteins, n = 23 ECM-affiliated proteins and n = 15 secreted factors. The STS adhesome was assessed using a functional atlas of the integrin adhesome^{295,415}. The integrin adhesome database contains n = 232 proteins, of which n = 109 proteins (47%) were present in our dataset (Figure 3.1C). According to the functional categories of adhesome proteins, there were n = 40 adaptors, n = 15 actin regulators, n = 20 adhesion receptors and n = 34 others (Figure 3.1D). The most well-represented matrisome classes in our dataset were glycoproteins and ECM regulators and for the adhesome, actin regulation, serine/threonine kinase, adhesion receptor and adaptor classes.

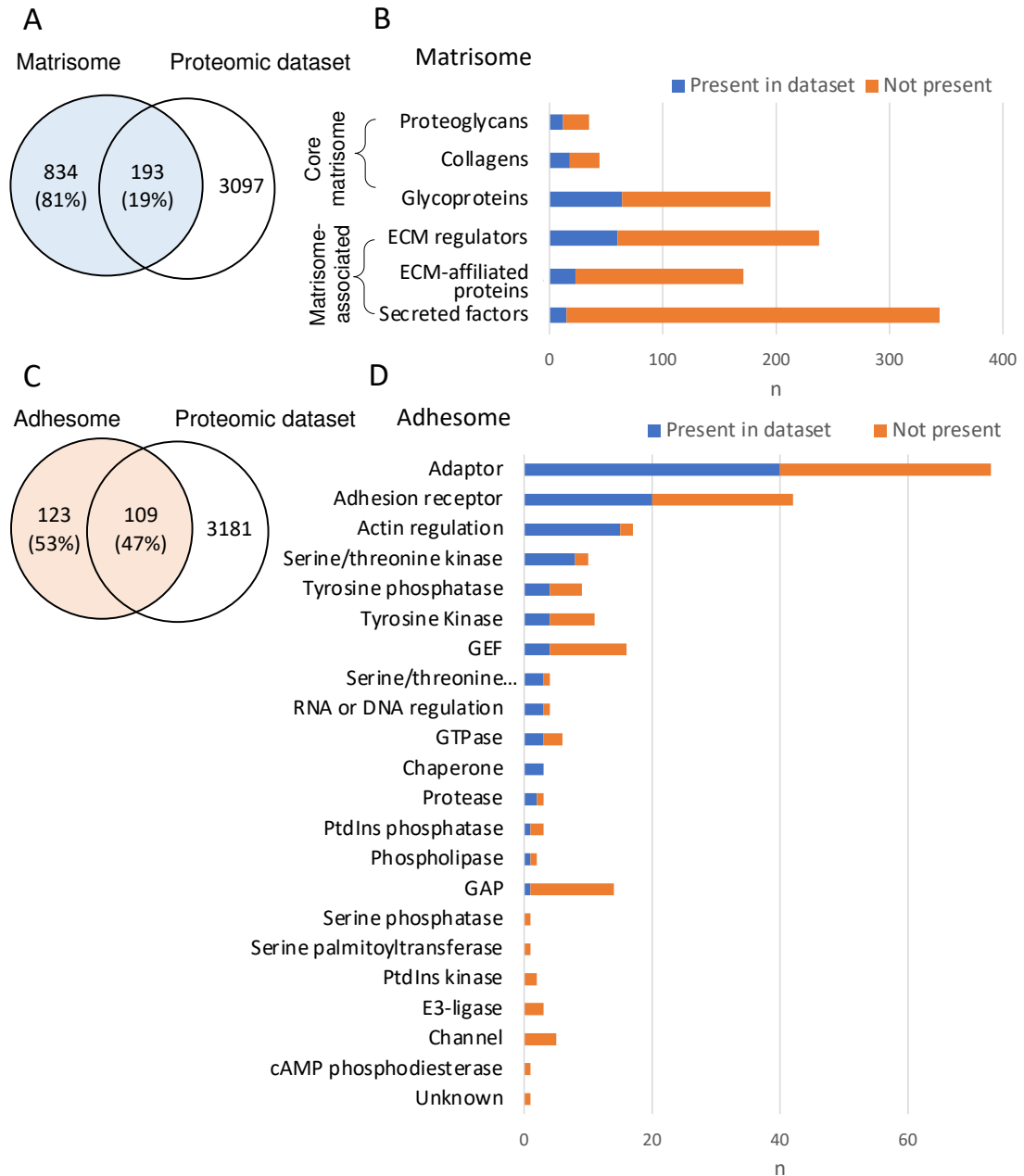


Figure 3.1 Coverage of matrisome and adhesome within the soft tissue sarcoma proteomic dataset.

A) Venn diagram to show the overlap and percentage coverage between the proteomic dataset and matrisome database¹⁷⁴. B) Stacked bar chart showing a representation of each matrisome class in the proteomic dataset. C) Venn diagram to show the overlap between the proteomic dataset and adhesome database^{295,298}. D) Stacked bar chart showing a representation of each functional category of adhesome in the proteomic dataset.

To explore the patterns of matrisome and adhesome expression in relation to clinicopathological factors, unsupervised clustering of proteomic profiles was performed and the resulting heatmap was annotated with the clinical features of each patient (Figure 3.2).

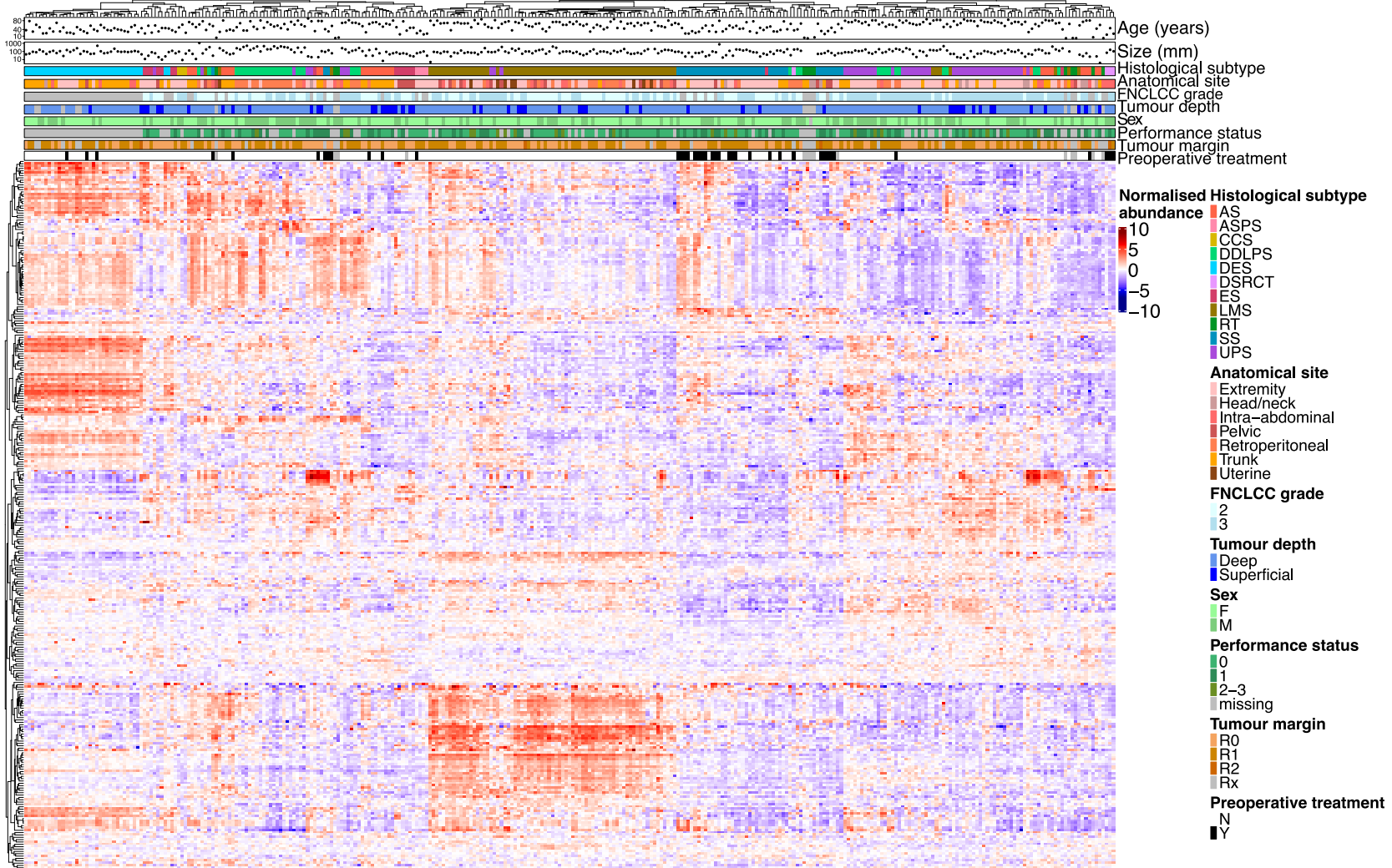


Figure 3.2 The matrisome and adhesome profiles in soft tissue sarcoma (STS).

Annotated heatmap illustrating the unsupervised clustering (Pearson's correlation distance) of 302 matrisome and adhesome components in the STS cohort. Annotation panels at the top correspond to patient age, tumour size, histological subtype, anatomical site, tumour grade, patient sex, performance status, surgical margin and preoperative treatment status. AS = angiosarcoma; ASPS = alveolar soft part sarcoma; CCS = clear cell sarcoma; DDLPS = dedifferentiated liposarcoma; DES = desmoid tumour; DSRCT = desmoplastic small round cell tumour; ES = epithelioid sarcoma; LMS = leiomyosarcoma; RT = rhabdoid tumour; SS = synovial sarcoma; UPS = undifferentiated pleomorphic sarcoma; R = residual tumour; Y = yes; N = no.

Prominently, clustering (groupings of samples) by histological subtype was observed. This implies that the expression patterns of matrisome and adhesome proteins differed between the STS subtypes, separating them into distinct clusters in the analysis. Meanwhile, samples of the same histological subtype had similar patterns of matrisome and adhesome protein expression. The unsupervised clustering of matrisome and adhesome proteomics profiles revealed the presence of distinct clusters corresponding to three histological subtypes, DES, LMS and SS. Although LMS cases grouped closely together, some variability in the expression patterns of matrisome and adhesome proteins was observed among LMS patients. This intra-subtype heterogeneity in LMS will be discussed in more detail in Chapter 4. UPS cases were mostly clustered together. Approximately half of DDLPS cases were present as one distinct cluster, indicating relatively homogeneous expression of matrisome and adhesome proteins within the subset of DDLPS cases. The other half of DDLPS cases were found spread across the heatmap, frequently close to UPS cases. This suggests that a subset of DDLPS cases exhibited expression patterns more similar to UPS cases than to the distinct cluster of DDLPS cases. Consequently, this indicates potential molecular overlap or similarities of matrisome and adhesome profiles between some DDLPS and UPS samples. This is consistent with The Cancer Genome Atlas Sarcoma (TCGA-SARC) study, which showed that some UPS and DDLPS cases had similar mRNA expression profiles³⁵. AS did not form a separate cluster according to the histological subtype, indicating substantial heterogeneity in this subtype.

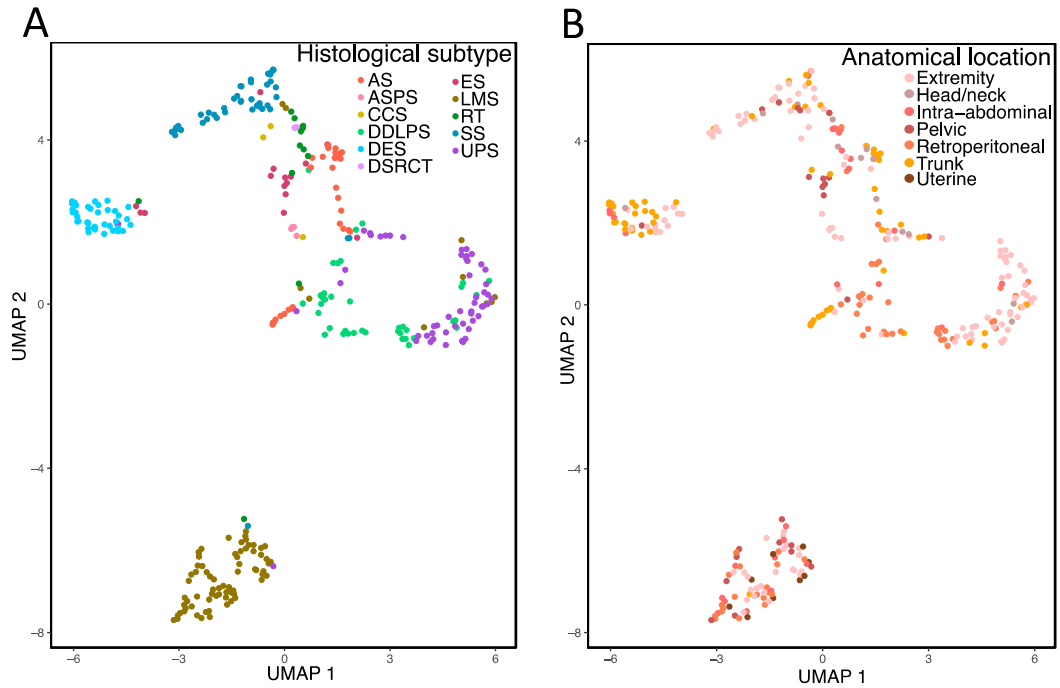


Figure 3.3 Uniform Manifold Approximation and Projection (UMAP) of matrisome and adhesome profiles.

A) Uniform Manifold Approximation and Projection (UMAP) plot showing matrisome and adhesome proteomic expression of individual cases coloured by histological subtype. B) UMAP coloured by anatomical location. AS = angiosarcoma; ASPS = alveolar soft part sarcoma; CCS = clear cell sarcoma; DDLPS = dedifferentiated liposarcoma; DES = desmoid tumour; DSRCT = desmoplastic small round cell tumour; ES = epithelioid sarcoma; LMS = leiomyosarcoma; RT = rhabdoid tumour; SS = synovial sarcoma; UPS = undifferentiated pleomorphic sarcoma.

To complement the unsupervised clustering, uniform manifold approximation and projection (UMAP) was used for data visualisation (Figure 3.3). UMAP is a dimension reduction technique that transforms a high-dimensional dataset (such as the proteomics dataset presented here) into a 2D plot³⁸⁷. UMAP was designed to preserve the overall structure of the dataset and relationship between data points. Applying UMAP to the matrisome and adhesome proteomics dataset and colouring individual cases by STS subtype, showed that the majority of cases with the same histological subtypes clustered closely together (Figure 3.3A). Of note, LMS cases clustered separately from the rest of the cases, suggesting LMS has the most distinct matrisome and adhesome profile. Different anatomical sites can have site-specific tumour

microenvironments, which can shape tumour progression⁴¹⁶. Interestingly, my analysis showed that although arising from a range of anatomical locations, the STS cases did not cluster by anatomical location, i.e. cases of the same anatomical location were not found in separate clusters (Figure 3.3B). This suggests that matrisome and adhesome composition within the subtypes is more likely due to intrinsic tumour biology rather than anatomical location.

Supervised comparisons were performed to explore which matrisome and adhesome proteins are unique and consistently expressed in different STS subtypes within the dataset. Only the subtypes with ≥ 20 cases (LMS, DDLPS, DES, UPS, AS and SS) were considered for this analysis to increase the robustness of results. Histological subtype-specific matrisome and adhesome profiles were derived by Significance Analysis of Microarrays (SAM) two-class unpaired tests that compared each subtype with the rest of the cohort³⁹⁰. Differentially expressed proteins (DEPs) were considered significant if they had a false discovery rate (FDR) < 0.01 and a fold change ≥ 2 . Upregulated DEPs, which were not overlapping between multiple subtypes, were considered unique. Unique upregulated DEPs for LMS, DDLPS, DES, UPS, AS and SS are depicted in Figure 3.4.

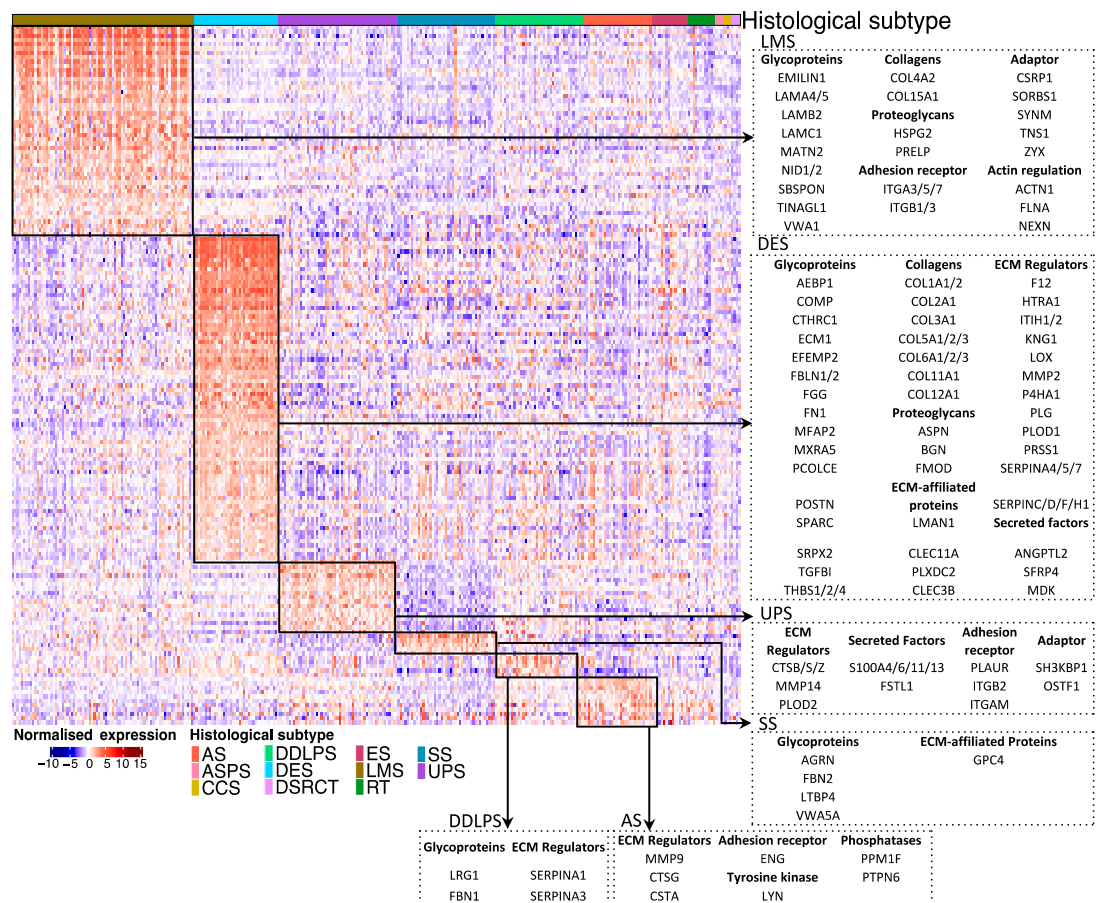


Figure 3.4 Unique matrixome and adhesome components in the six most common subtypes of soft tissue sarcoma (STS).

Heatmap showing matrixome and adhesome proteins uniquely upregulated (indicated by black boxes) in histological subtypes (false discovery rate < 0.01, fold change \geq 2), arranged by histological subtype. A selection of proteins which are upregulated in each histological subtypes is shown. AS = angiosarcoma; ASPS = alveolar soft part sarcoma; CCS = clear cell sarcoma; DDLPS = dedifferentiated liposarcoma; DES = desmoid tumour; DSRCT = desmoplastic small round cell tumour; ES = epithelioid sarcoma; LMS = leiomyosarcoma; RT = rhabdoid tumour; SS = synovial sarcoma; UPS = undifferentiated pleomorphic sarcoma.

In LMS, 56 DEPs were upregulated and 10 downregulated, and of those, 49 were uniquely enriched in LMS (Supplementary Figure 3.2A). Upregulated unique proteins included structural muscle constituents (CSR1P1, SYNM and NEXN), consistent with the LMS smooth muscle lineage of origin (Figure 3.4). Other enriched proteins included components of the basement membrane such as type-IV collagen chain (COL4A2), glycoproteins, nidogens (NID1/2), and laminins (LAMA4/A5/B2/C1). DES had the highest number of DEPs,

having 93 upregulated DEPs, and 21 downregulated DEPs compared to all other samples (Supplementary Figure 3.2B). Of the upregulated DEPs, 75 proteins were uniquely enriched in DES, including fibrillar collagen chains (COL1A1/2, COL2A1, COL3A1, COL5A1/2/3, COL11A1), collagen synthesis and remodelling molecules (P4HA1, LOX and PLOD1), various glycoproteins (CTHRC1, FBLN1/2) and proteoglycans (ASPN, BGN) (Figure 3.4). The detection of a high number of collagen chains in the analysis corroborates with pathological reports, which describe DES presenting with dense collagenous stroma^{417,418}. In UPS, 28 DEPs were upregulated and 29 downregulated, with 17 unique upregulated proteins (Supplementary Figure 3.2C). Those included integrin receptors predominantly expressed by leukocytes (ITGB2 and ITGAM), cysteine cathepsins (CTSB/S/Z) and S100 family proteins (S100A4/6/11/13) found to play a role in regulating immune system^{232,419,420} (Figure 3.4). This is consistent with previous reports of high immune infiltration in UPS tumours¹³⁷.

AS, SS and DDLPS had only a handful of uniquely upregulated proteins (Supplementary Figure 3.2D-F). In SS, 9 DEPs were upregulated and 62 downregulated, with 5 proteins unique to SS, mainly glycoproteins (AGRN, FBN2, LTBP4 and VWA5A) and ECM-affiliated proteins (GPC4) (Figure 3.4). In DDLPS, 15 DEPs were upregulated and 22 DEPs downregulated, with 5 proteins unique to DDLPS, which comprise glycoproteins (FBN1 and LRG1), ECM regulators (SERPINA1/3) and ECM-affiliated proteins (GPC6) (Figure 3.4). In AS, 16 DEPs were upregulated and 42 DEPs downregulated, compared to the rest of the cohort. Of those, 11 were unique to AS and included cell surface protein endoglin (ENG) found on endothelial cells, which aligns with AS's vascular origin⁴²¹ (Figure 3.4). Other upregulated proteins were ECM regulators (MMP9, CTSG and CSTA) and intracellular adhesion signalling molecules such as the tyrosine kinase LYN and phosphatases PPM1F and PTPN6.

3.2.3. Matrisome and adhesome networks in STS

Matrisome and adhesome molecules rarely function in isolation, and instead, they are part of a cooperative network that drives tumorigenesis. The matrisome represents multiple proteins that provide coordinated signals to activate integrin-mediated adhesome signalling^{176,415}. Concurrent changes in protein expression, whether upregulation or downregulation, may indicate coordinated signalling. To determine if there are sets of co-regulated matrisome and adhesome components, Pearson's correlation analysis was performed on all possible pairwise combinations of the 302 matrisome and adhesome proteins in my dataset. The resulting similarity matrix was analysed with consensus clustering to determine clusters of co-regulated matrisome and adhesome proteins (Figure 3.5). Consensus clustering provides quantitative evidence for determining the optimal number and composition of different clusters (k) within a dataset³⁸⁸. The consensus clustering analysis results were summarised in graphical outputs such as consensus matrices heatmaps, consensus cumulative distribution function (CDF) plot and tracking plot and those were examined to make an informed choice on the optimal number of clusters (Supplementary Figure 3.3). Consensus matrices indicated the 'cleanest' cluster transition (from white to dark blue) for $k = 3$ and $k = 4$ (Supplementary Figure 3.3A). CDF depicts consensus distribution for each k , and $k = 3$ showed a sharp increase in CDF between 0 and 0.3, a horizontal line between 0.3 and 0.9, and a sharp increase close to 1, indicating optimal clustering at $k = 3$ (Supplementary Figure 3.3B). The tracking plots, where samples are in columns and k are in rows, allow one to assess the history of a sample's assignment to different clusters for each k value. The tracking plot did not have samples that frequently changed from one cluster to another, indicating stable clustering (Supplementary Figure 3.3C). Consensus matrices, CDF and tracking plots suggest that $k = 3$ and $k = 4$ show optimal clustering within the correlation matrix. However as $k = 4$ had only a handful of samples, to ensure the statistical robustness of the results, $k = 3$ was chosen. Visual inspection using the above outputs can be complemented with a statistical Significance of Clustering (SigClust) test to determine if clusters at

a putative k are significantly distinct⁴²². At $k = 3$, all pairwise comparisons between the clusters were significant at $p < 0.0001$.

After assessing consensus clustering and SigClust results (Supplementary Figure 3.3A-C), $k = 3$ was chosen, meaning three distinct clusters of co-regulated matrisome and adhesome proteins were present in the dataset. Hierarchical clustering was used to visualise the dataset grouped according to the three consensus clusters (Figure 3.5).

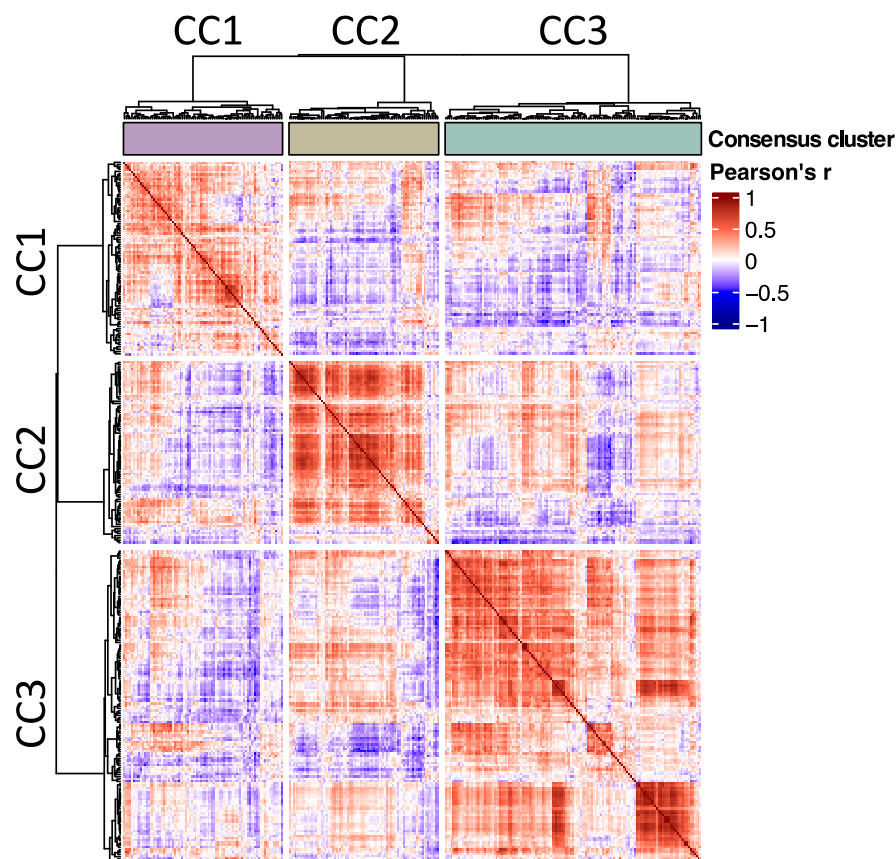


Figure 3.5 Matrisome and adhesome networks in soft tissue sarcoma (STS)

Heatmap showing similarity matrix of Pearson's correlation coefficients for all pairwise comparisons of matrisome and adhesome proteins. Heatmap is split into three clusters (CC1, CC2, CC3) of co-regulated matrisome and adhesome proteins identified by consensus clustering analysis.

The composition of each cluster was assessed against database membership (adhesome, core matrisome, matrisome-associated), division into matrisome classes (collagens, glycoproteins, proteoglycan, ECM-affiliated proteins, ECM

regulators and secreted factors) and adhesome functional categories (actin regulation, adaptor protein, adhesion receptor, and all the other classes were grouped into 'other') (Figure 3.6A). CC1 (n = 85) had 41 adhesome and 41 matrisome-associated proteins (both at 48%), while only 3 proteins (4%) were core matrisome. CC2 (n = 80) was predominantly adhesome proteins with 46 proteins (58%), followed by core matrisome with 28 proteins (35%) in that category and 6 matrisome-associated proteins (7%). CC3 (n = 137) was mainly made up of 63 core matrisome (46%) and 52 matrisome-associated proteins (38%), with 22 adhesome proteins (16%). Furthermore, clusters had differences in matrisome class composition. CC1 consisted mainly of ECM

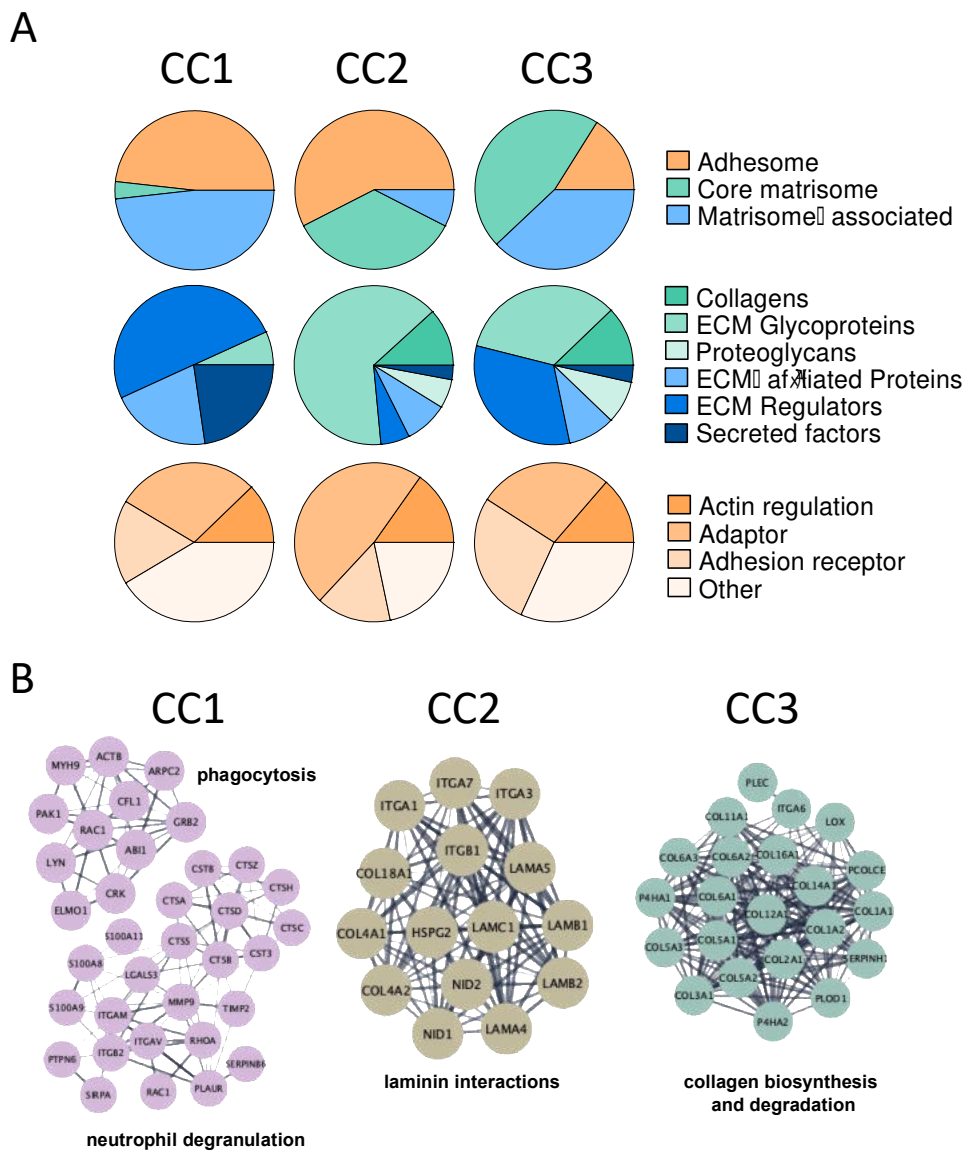


Figure 3.6 Biological characterisation of clusters of co-regulated matrisome and adhesome proteins in soft tissue sarcoma (STS).

A) Pie charts showing breakdown of clusters by into adhesome, core matrisome or matrisome-associated proteins (top panel), breakdown by matrisome class (middle panel) and by functional annotation of adhesome (bottom panel). B) Selected protein-protein interaction networks (coloured by CC) constructed using STRINGdb are shown for each cluster, including Reactome pathway annotations.

regulators (50%); in contrast, in CC2 glycoproteins (65%) were the most predominant class, and CC3 was a mix of glycoproteins (34%) and ECM

regulators (32%). Interestingly, CC3 contained 14 collagen chains out of 18 detected across the whole cohort. The most notable difference in adhesive composition was an expansion of the adaptor category in CC2 (48%) compared to CC1 (29%) and CC3 (27%). To identify biological functions that are enriched in each cluster, overrepresentation analysis was performed using the Reactome database^{423,424}. Supplementary Table 3.2 summarises biological terms significantly enriched in each cluster, highlighting the top 10 hits. In the CC1, the top hits included biological processes associated with the innate immune system, phagocytosis, and haemostasis. Subnetworks for 'neutrophil degranulation' and 'Fc gamma receptor-dependent phagocytosis' are presented in Figure 3.6B. The CC2 was enriched for 'laminin interactions' (subnetwork in Figure 3.6B) and 'cell-extracellular matrix interactions'. The CC3 included general 'ECM organisation and degradation' and specific 'collagen biosynthesis and degradation'. The network for collagen biosynthesis and modifying enzymes is shown in Figure 3.6B.

To investigate if the three consensus clusters were related to clinicopathological factors (histological subtype, anatomical location and tumour grade), median expression values for proteins in clusters CC1, CC2 and CC3 were calculated for each patient so that each patient had three values corresponding to CC1, CC2 and CC3. In relation to the histological subtypes, subtypes with < 20 samples were grouped together in the 'other' group (Figure 3.7). CC1 was significantly enriched in UPS (all pairwise comparisons with other histological subtypes were at $p < 0.01$) and, to a lesser extent, in DDLPS. In the multiple comparisons, DDLPS had a significantly higher score than LMS ($p = 0.002$), other ($p = 0.03$), and SS ($p < 0.0001$) but not compared to AS and DES. SS was significantly less represented in CC1 compared to other subtypes with all pairwise comparisons at $p < 0.0001$ (Supplementary Table 3.3). CC2 significantly enriched LMS histology (all pairwise comparisons with other histological subtypes were at $p < 0.0001$).

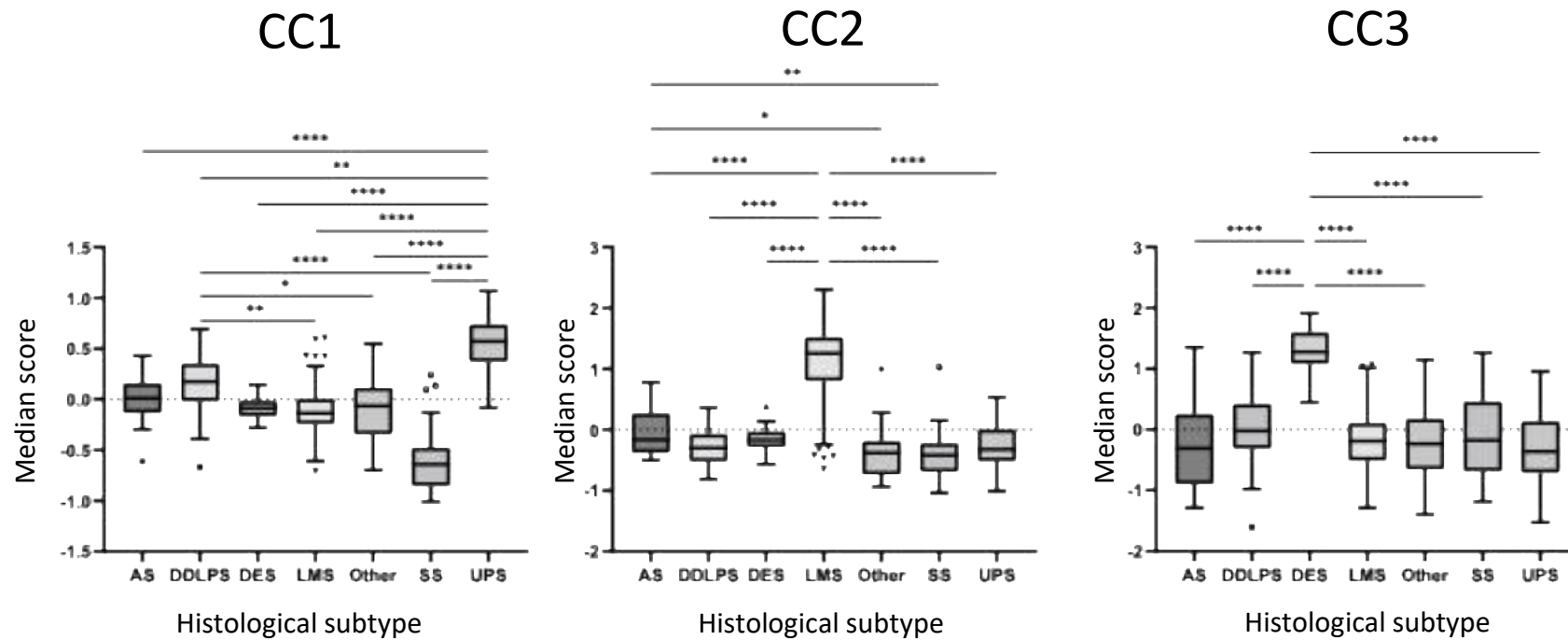


Figure 3.7 Histological subtypes associated with matrisome and adhesome networks in soft tissue sarcoma (STS).

Box plots showing distributions of median score of CC1, CC2 and CC3 proteins in histological subtypes. Alveolar soft part sarcoma, clear cell sarcoma, desmoplastic small round cell tumour, epithelioid sarcoma and rhabdoid tumour were grouped into 'Other'. Boxes indicate 25th and 75th percentile, with median line in the middle, whiskers extending from 25th percentile-(1.5*interquartile range (IQR)) to 75th percentile+(1.5*IQR), and outliers plotted as points. Significance determined by Kruskal-Wallis tests with Dunn's multiple corrections tests. * = $p < 0.05$, ** = $p < 0.01$, **** = $p < 0.0001$. AS = angiosarcoma; DDLPS = dedifferentiated liposarcoma; DES = desmoid tumour; LMS = leiomyosarcoma; SS = synovial sarcoma; UPS = undifferentiated pleomorphic sarcoma.

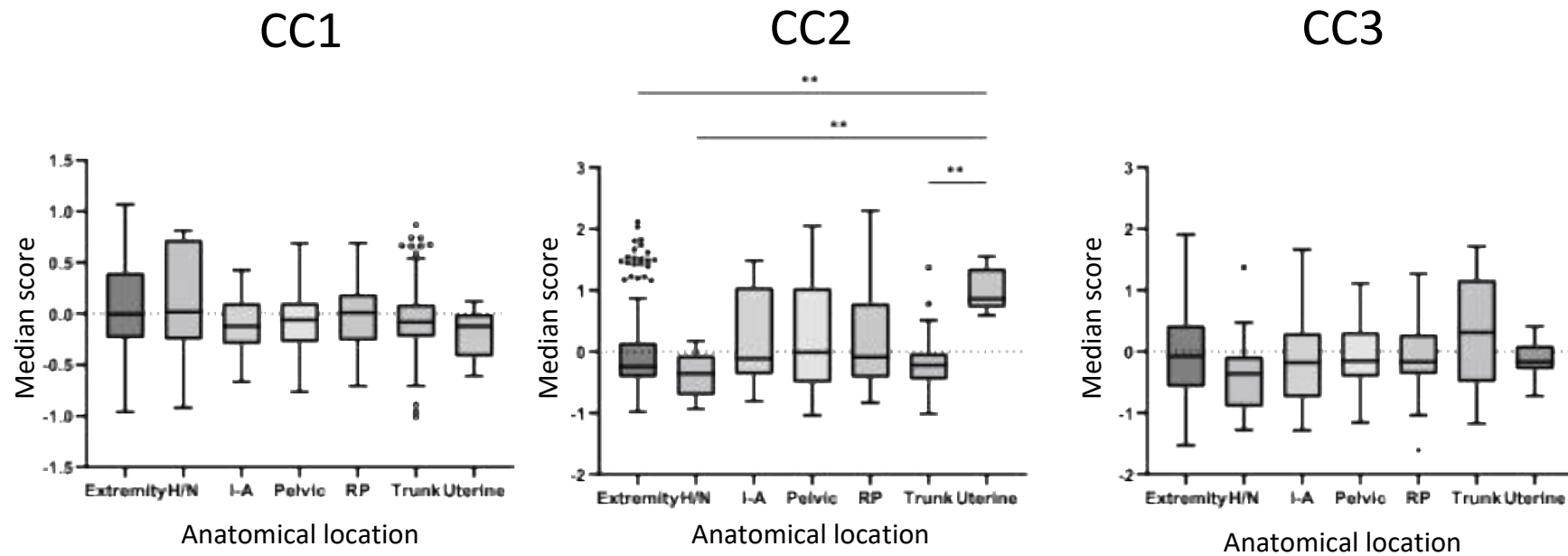


Figure 3.8 Anatomical location associated with matrisome and adhesome networks in soft tissue sarcoma (STS).

Box plots showing distributions of median score of CC1, CC2 and CC3 proteins across anatomical locations of the tumour. Boxes indicate 25th and 75th percentile, with median line in the middle, whiskers extending from 25th percentile-(1.5* interquartile range (IQR)) to 75th percentile+(1.5*IQR), and outliers plotted as points. A-B) Significance determined by Kruskal-Wallis tests with Dunn's multiple corrections tests. ** = $p < 0.01$. H/N = head and neck; I-A = intra-abdominal; RP = retroperitoneum.

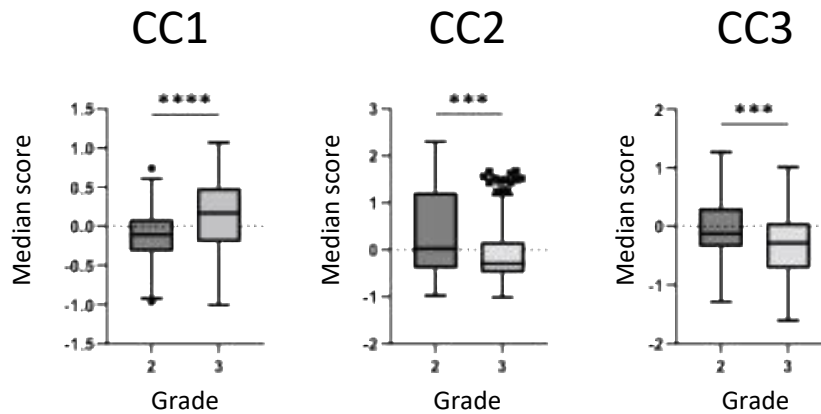


Figure 3.9 Histological grade associated with matrisome and adhesome networks in soft tissue sarcoma (STS).

Box plots showing distributions of median score of CC1, CC2 and CC3 proteins across tumour grades. Boxes indicate 25th and 75th percentile, with median line in the middle, whiskers extending from 25th percentile-(1.5* interquartile range (IQR)) to 75th percentile+(1.5*IQR), and outliers plotted as points. Significance determined by Mann-Whitney *U* tests, *** = $p < 0.001$, **** = $p < 0.0001$.

Additionally, the median score for AS was significantly higher than 'other' ($p = 0.049$) and SS ($p = 0.009$) in CC2. In CC3, DES was the dominant histology (all pairwise comparisons with other histological subtypes were at $p < 0.0001$), and there was no enrichment of other subtypes (other comparisons were ns).

In terms of anatomical location, there was no enrichment of any particular locations in CC1 and CC3 (Figure 3.8). In CC2, uterine tumours had the highest score of all anatomical locations, with significance against extremities ($p = 0.007$), head and neck ($p = 0.002$) and trunk ($p = 0.002$). All uterine tumours were LMS, consistent with the enrichment of LMS in CC2. The significance to the rest of the locations was not reached, potentially due to a small number of cases in the uterine group ($n = 9$). In regards to the grade, CC1 had a significant enrichment in higher grade (grade 3) tumours ($p < 0.0001$) (Figure 3.9). In CC2, tumours with grade 2 had a significantly higher score than grade 3 ($p < 0.0006$). In CC3, the grade 2 was more prevalent ($p = 0.0004$).

3.2.4. Identification of a prognostic matrix score for STS

Recently, different groups have developed various prognostic matrix scores based on gene expression of selected matrix components^{260,261,265}. These matrix signatures were shown to be prognostic of survival outcomes in carcinomas. Notably, a matrix signature originally defined through ovarian cancer analysis was applied to the TCGA-SARC dataset and was able to stratify STS patients into high and low-risk groups²⁶⁵. Tumour matrix remodelling contributes to the establishment of inflammation and modulates the immune TME⁴¹². Consequently, I aimed to identify a protein expression-based matrix score applicable for prognostication in STS and investigate its relationship with TILs.

I focused on a subset of STS patients with sufficient tissue to generate TMA to evaluate TILs counts by IHC. The resultant cohort comprised n = 148 patients of the most common STS subtypes (LMS, DDLPS and UPS). The baseline clinicopathological characteristics for this subset of patients are summarised in Supplementary Table 3.4. In comparison to the full cohort, the clinical characteristics of this subset of patients, (in Section 3.2.1 and Table 3.1) with the exception of age and tumour grade, were broadly similar. In the reduced sub-cohort, the age at excision was slightly higher than in the main cohort, with a median of 68 years (range: 28-90), given that no paediatric cases were present. Additionally, this cohort had a slightly higher proportion of high-grade (grade 3) tumours, as UPS were mainly high-grade (grade 3) tumours.

A single-sample gene set expression analysis (ssGSEA) was performed to define a prognostic matrix score. ssGSEA allows for assessing the relative activity or enrichment of predefined gene sets within individual samples, providing a valuable tool for understanding biological processes^{392,393}. A flowchart outlining the steps for identifying prognostic matrix scores is depicted in Figure 3.10A. Ten matrix-related gene sets¹⁷⁴ from The Molecular Signatures Database (MSigDB)³⁹⁴ were evaluated and are listed in Table 3.2.

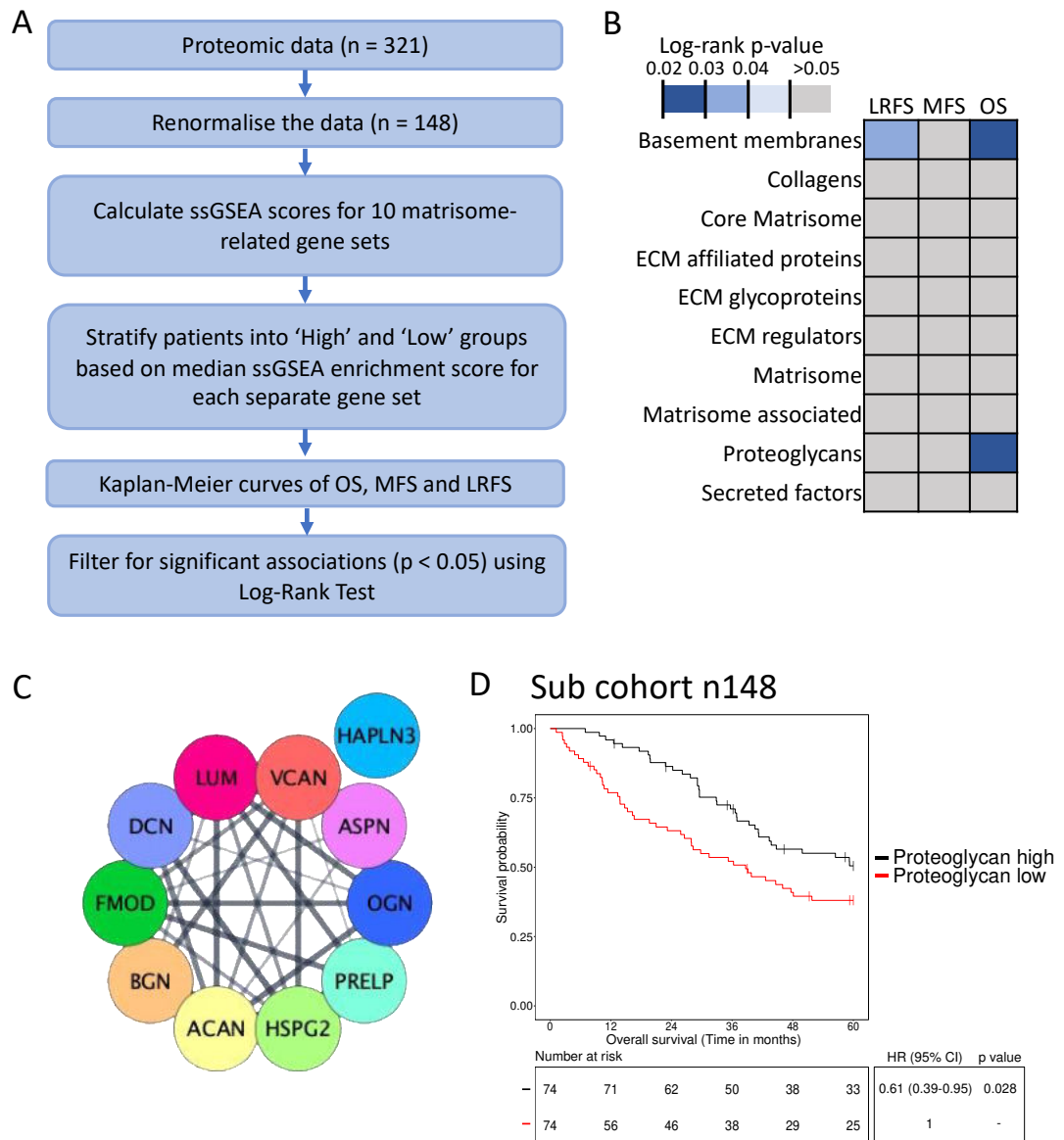


Figure 3.10 Proteoglycan score is prognostic of overall survival in a sub-cohort of soft tissue sarcoma (STS).

A) A flowchart summarising the steps to identify matrix prognostic score (n indicates number of patients). B) Summary of log-rank tests used to assess significant associations of matrisome-related gene sets with local recurrence-free survival (LRFS), metastasis-free survival (MFS) and overall survival (OS). C) A network of 11 proteoglycans based on STRINGdb interaction score. The figure was created using Cytoscape v3.10.0. D) Kaplan-Meier plot of OS with stratification by proteoglycan score using the median score as a cut-off. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test. ssGSEA: single-sample gene set expression analysis. ACAN: aggrecan; BGN: biglycan; FMOD: fibromodulin; DCN: decorin; LUM: lumican; VCAN: versican; ASPN: asporin; HAPLN3: Hyaluronan And Proteoglycan Link Protein 3; OGN: osteoglycin; PRELP: prolargin; HSPG2: Heparan Sulfate Proteoglycan 2/perlecan.

Table 3.2 Matrisome gene sets from The Molecular Signatures Database (MSigDB) evaluated for prognostic score.

Breakdown by gene sets indicating the total number of genes belonging to a gene set and the number of overlapping proteins between the proteomic dataset and a defined gene set, which were used for single-sample gene set expression analysis (ssGSEA).

Gene sets from MSigDB	Number of proteins used for ssGSEA score	Total number of genes in a gene set
Basement membranes	16	40
Collagens	17	44
Core Matrisome	93	275
ECM affiliated proteins	26	170
ECM glycoproteins	65	196
ECM regulators	65	238
Matrisome	200	1027
Matrisome associated	107	751
Proteoglycans	11	35
Secreted factors	16	343

The ssGSEA enrichment score was calculated using the protein expression data for gene sets of interest. To evaluate the association of gene sets with survival outcomes, patients were stratified into high and low groups based on the median value of the ssGSEA enrichment score for each separate gene set. Kaplan-Meier curves were used to visualise the association of a particular gene set with OS, MFS and LRFs, and a log-rank test was used to filter for significant ($p < 0.05$) associations. From the analysis, the basement membrane gene set was prognostic for LRFs (log-rank test $p = 0.034$) and OS (log-rank test $p = 0.025$) (Figure 3.10B; Supplementary Figure 3.4A, B). The proteoglycan gene set was prognostic for OS (log-rank test $p = 0.026$) (Figure 3.10B; Supplementary Figure 3.4A). An association of the glycoprotein gene set with OS was borderline (log-rank test $p = 0.054$) and was not investigated further (Kaplan-Meier curve is in Supplementary Figure 3.4C). The protein-protein interaction network of 11 proteoglycan proteins used to generate the ssGSEA score is shown in Figure 3.10C. Kaplan-Meier curve (Figure 3.6D) and univariate Cox regression analyses (Table 3.3) revealed that patients with low proteoglycan score had a significantly worse OS compared to patients with high proteoglycan score (HR = 0.61, 95% CI = 0.39 – 0.95 and $p = 0.028$). The proteoglycan score was not prognostic of LRFs (Supplementary Figure 3.4D) and MFS (Supplementary Figure 3.4E).

Table 3.3 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing association of clinicopathological factors and proteoglycan score with overall survival (OS).

Clinicopathological factors which were significantly associated with OS in UVA analysis were included in the MVA model. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test. I-A = intra-abdominal; RP = retroperitoneal, LMS = leiomyosarcoma, DDLPS = dedifferentiated liposarcoma and UPS = undifferentiated pleomorphic sarcoma.

Variable		Groups	n	Univariable analysis (OS)		Multivariable analysis (OS)	
				HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	-	1.03 (1.01-1.05)	0.003	1.01 (0.99-1.04)	0.219
Grade	3 (reference)	95	-	-	-	-	-
	2	52	0.56 (0.32-0.86)	0.01	0.67 (0.37-1.24)	0.202	
STS subtype	LMS (reference)	63	-	-	-	-	
	UPS	50	1.63 (0.98-2.71)	0.058	1.1 (0.59-2.07)	0.757	
	DDLPS	35	1.4 (0.79-2.46)	0.246	0.93 (0.49-1.75)	0.819	
Anatomical location	Extremity/Trunk/Head-neck (reference)	78	-	-	-	-	
	I-A/RP/Pelvic	62	1.13 (0.72-1.78)	0.59	-	-	
	Uterine	8	1.43 (0.56-3.63)	0.451	-	-	
Log [tumour size] (mm)	4-5 (reference)	77	-	-	-	-	
	>5	46	1.34 (0.84-2.14)	0.224	1.8 (1.03-3.15)	0.04	
	<4	24	0.49 (0.23-1.05)	0.068	0.4 (0.17-0.93)	0.034	
Tumour depth	Deep (reference)	124	-	-	-	-	
	Superficial	24	0.70 (0.36-1.37)	0.3	-	-	
Tumour margin	R1&2 (reference) (R1-	77	-	-	-	-	
	R0	64	0.84 (0.54-1.31)	0.451	-	-	
	Rx	7	0.20 (0.03-1.45)	0.111	-	-	
Sex	F (reference)	83	-	-	-	-	
	M	65	1.57 (1.01-2.44)	0.044	1.52 (0.91-2.55)	0.11	
Performance status	0 (reference)	66	-	-	-	-	
	1	38	2.53 (1.47-4.36)	0.001	2.91 (1.63-5.18)	0.0003	
	2-3 (2-12cases; 3-5cases)	17	5.21 (2.72-9.97)	0.000001	4.04 (1.98-8.24)	0.0001	
	unknown	27	1.53 (0.786-2.98)	0.21	1.11 (0.53-2.32)	0.783	
Proteoglycan score	Proteoglycan low	74	-	-	-	-	
	Proteoglycan high	74	0.61 (0.39-0.95)	0.028	0.551 (0.33-0.97)	0.022	

Additionally, Chi-squared and Kruskal-Wallis tests were performed to test for the association of clinicopathological factors with the proteoglycan set. Tumour grade and histological subtype were found to be associated with the

low proteoglycan score group (Supplementary Table 3.5). Briefly, patients in the low proteoglycan group had a higher proportion of grade 3 tumours (Chi-squared test: $X^2 = 16.6$, $p = 0.0001$), a higher proportion of UPS and a lower proportion of LMS tumours (Chi-squared test: $X^2 = 12.2$, $p = 0.0022$). A multivariable analysis was performed to assess if the proteoglycan score was independent of known baseline prognostic clinicopathological variables. After adjusting for age, tumour grade, STS subtype, tumour size, sex and performance status, the proteoglycan score remained an independent prognostic factor for OS (Table 3.3). As expected, tumour size and performance status were significant factors associated with OS in the multivariable analysis. Although the basement membrane score was prognostic for LRFS and OS in the univariable analysis, after adjusting for known prognostic variables in the multivariable analysis, the basement membrane score was not prognostic for LRFS and OS (Supplementary Table 3.6).

3.2.4.1. Combination of proteoglycan score with TILs counts

DDLPS and UPS are associated with relatively high levels of TILs, compared to other STS subtypes, whilst LMS has a wide range of TILs^{133,135,425}. Moreover, high levels TILs were shown to be associated with improved survival in DDLPS and UPS^{137,154,426}. For this study, TILs levels (CD3+, CD4+ and CD8+) were assessed by IHC on TMA and analysed by previous lab members for a subset of LMS, DDLPS and UPS cases in my cohort. CD3+ is considered a marker for all mature T cells, CD4+ is predominantly expressed on helper T cells, while CD8+ is a marker for cytotoxic T cells. To factor in intra-patient heterogeneity, only cases with data with a minimum of two cores for each stain were included in the analysis. Briefly, $n = 148$ patients (combination of DDLPS, UPS and LMS) had relevant IHC data. Of those, $n = 145$ patients stains for CD3+, $n = 142$ for CD4+ and $n = 148$ for CD8+ TILs.

CD3+ TILs

The CD3+ TIL/mm² numbers ranged from 0 – 1239 across the cohort of $n = 145$ patients (median = 95 TIL/mm²). For survival analyses, CD3+ TILs counts

were dichotomised, and patients were classified into high and low groups using the median value as a cut-off. Representative images of high and low CD3+ are shown in Figure 3.11A. Dichotomous scores were associated with LRFs, MFS and OS to assess the prognostic value of CD3+ TILs in DDLPS, UPS and LMS patients. The univariable analysis did not find associations between the CD3+ score and LRFs, MFS or OS outcomes (Supplementary Table 3.7).

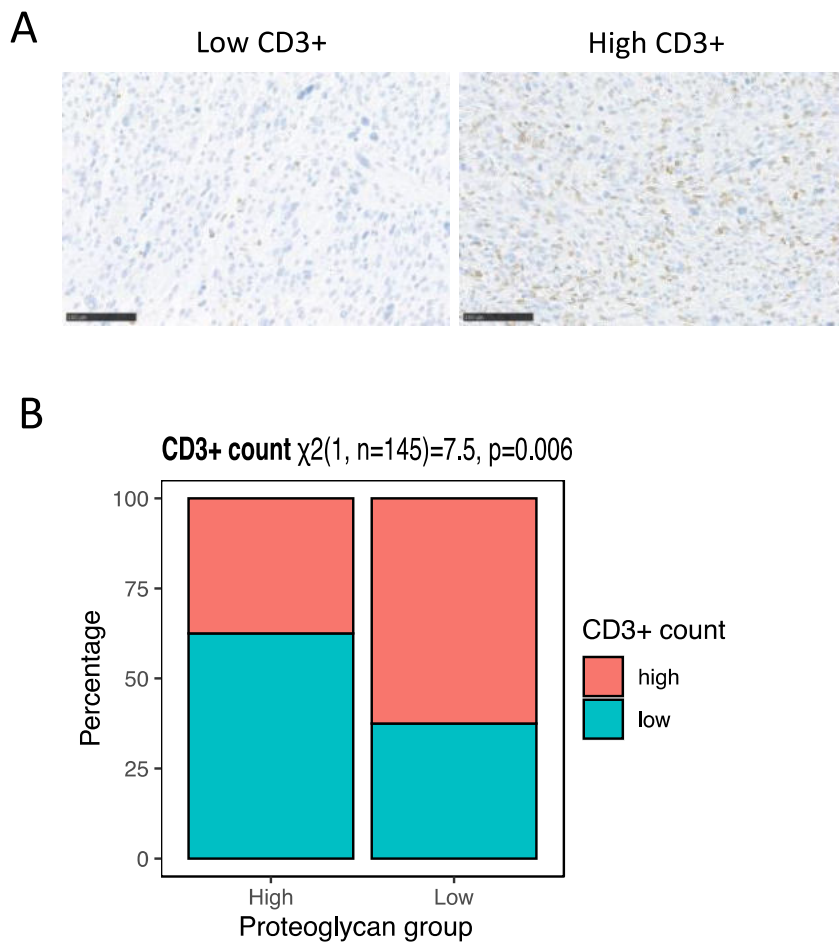


Figure 3.11 CD3+ tumour infiltrating lymphocytes (TILs) are inversely associated with proteoglycan groups.

A) Representative images of low and high CD3+ TILs staining. B) Stacked bar charts showing percentages of CD3+ TILs in proteoglycan low and high groups. The patient cases were divided into high and low categories according to the median CD3+ TILs score. The chi-square test results are presented at the top of the plot.

To investigate the potential association between the CD3+ infiltrate and risk groups based on the proteoglycan score (as derived in Section 3.2.4), dichotomous CD3+ scores were assessed in relation to the proteoglycan low and proteoglycan high groups. A significantly higher percentage of high CD3+ was observed in the low proteoglycan group (high risk) compared to the high proteoglycan group (Chi-squared test: $X^2 = 7.5$, $p = 0.006$) (Figure 3.11B). To investigate if the CD3+ TIL score can add further prognostic information to the proteoglycan score, the cohort was divided into four strata based on the median proteoglycan score and median TILs score (proteoglycan-high/CD3+-high, proteoglycan-high/CD3+-low, proteoglycan-low/CD3+-high and proteoglycan-low/CD3+-low). Given that $n = 145$ patients had CD3+ counts, the protein expression data was renormalised for 145 cases, and the ssGSEA median cut-off was derived for the proteoglycan gene set. The CD3+ score did not add additional prognostic value to the proteoglycan score, and the separation of the curves was primarily driven by the proteoglycan score (Supplementary Figure 3.5; Supplementary Table 3.8).

CD4+ TILs

CD4+ TILs counts ranged from 0 – 869 (median = 24 TIL/mm²) across the cohort of $n = 142$ patients. Representative images of high and low CD4+ are shown in Figure 3.12A. Similarly to the CD3+ TILs analysis, the univariable analysis did not find associations between the dichotomous CD4+ score and LRFs, MFS or OS outcomes (Supplementary Table 3.9).

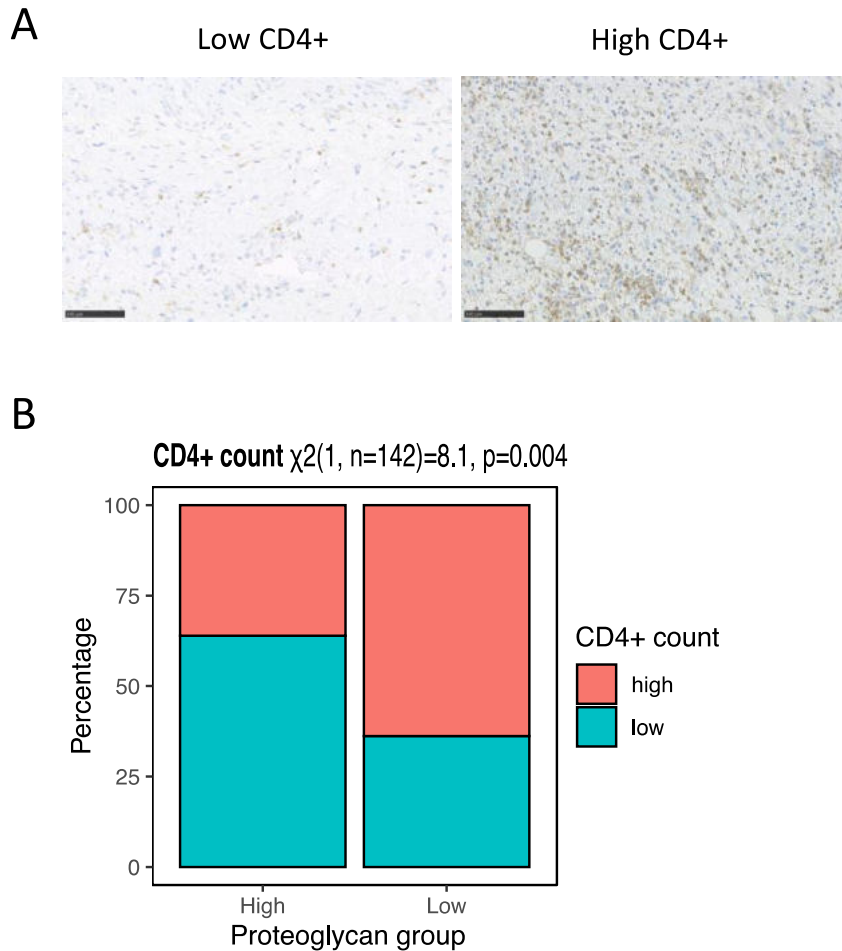


Figure 3.12 CD4+ tumour infiltrating lymphocytes (TILs) are inversely associated with proteoglycan groups.

A) Representative images of low and high CD4+ TILs staining. B) Stacked bar charts showing percentages of CD4+ TILs in proteoglycan low and high groups. The patient cases were divided into high and low categories according to the median CD4+ TILs score. The chi-square test results are presented at the top of the plot.

To investigate the potential association between the CD4+ infiltrate and risk groups based on the proteoglycan score (as derived in Section 3.2.4), dichotomous CD4+ scores were assessed in relation to the proteoglycan low and proteoglycan high groups. A significantly higher percentage of high CD4+ was observed in the low proteoglycan group (high risk) compared to the high proteoglycan group (Chi-squared test: $X^2 = 8.1, p = 0.004$) (Figure 3.8B). To determine if the CD4+ TIL score can add further prognostic information to the proteoglycan score, the cohort was divided into four strata based on the

median proteoglycan score and median TILs score. Given that n = 142 patients had CD4+ counts, the protein expression data was renormalised for 142 cases, and the ssGSEA median cut-off was derived for the proteoglycan gene set. There was no added prognostic value based on proteoglycan and CD4+ score (Supplementary Figure 3.6; Supplementary Table 3.10).

CD8+ TILs

CD8+ TILs counts ranged from 0 – 1736 (median = 68 TIL/mm²). Representative images of high and low CD8+ are shown in Figure 3.13A. Dichotomous CD8+ score was not prognostic for OS, LRFS and MFS (Supplementary Table 3.11).

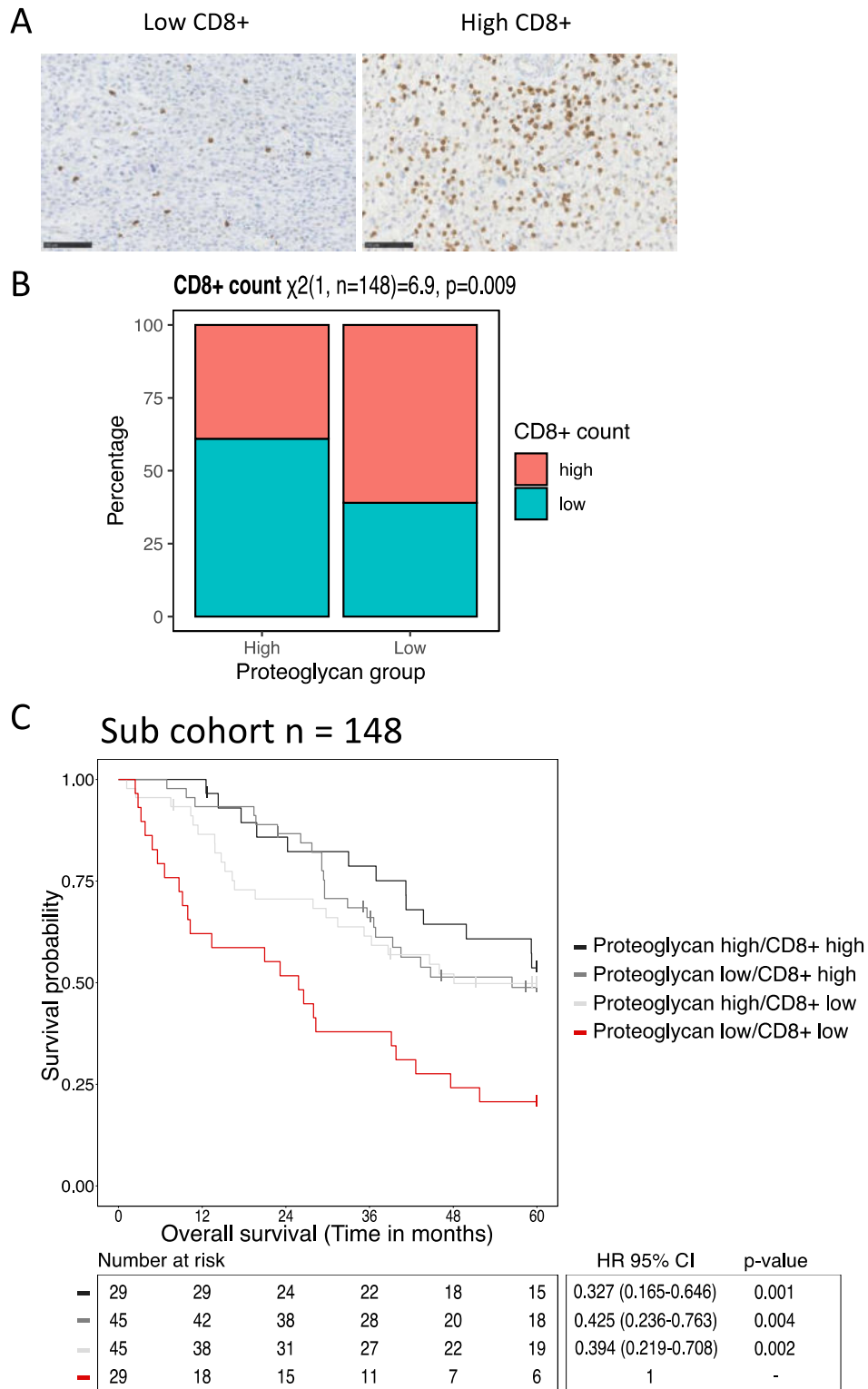


Figure 3.13 CD8+ tumour infiltrating lymphocytes (TILs) are associated with proteoglycan groups, and proteoglycan/CD8+ score identifies a high-risk STS group.

A) Representative images of low and high CD8+ TILs staining. B) Stacked bar charts showing percentages of CD8+ TILs in proteoglycan low and high groups. The patient cases were divided into high and low categories according to the median CD8+ TILs score. The chi-square

test results are presented at the top of the plot. C) Kaplan-Meier plot of OS with stratification by proteoglycan and CD8+ TILs score. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression analysis with a two-sided Wald test.

To investigate the potential association between the CD8+ immune infiltrate and risk groups based on the proteoglycan score (as derived in Section 3.2.4), dichotomous CD8+ TILs scores were assessed in relation to the proteoglycan low and high groups (Figure 3.13B). A significantly higher percentage of high CD8+ (Chi-squared test: $X^2 = 6.9$, $p = 0.009$) TILs were found in the low proteoglycan group (high risk), compared to the high proteoglycan group. To investigate if the CD8+ TILs score can add further prognostic information to the proteoglycan score, the cohort was divided into four strata based on the median proteoglycan score and median CD8+ TILs score. The Kaplan-Meier analysis of LRFS, MFS and OS with the four strata, showed that the CD8+ score contributed additional OS prognostic information to the proteoglycan score (Figure 3.13C). There was no added improvement in LRFS and MFS (Supplementary Figure 3.7; Supplementary Table 3.12). The analysis identified a proteoglycan-low/CD8+-low group with a significantly worse OS compared to proteoglycan-low/CD8+-high (HR = 0.43, 95% CI = 0.24 – 0.76, p-value = 0.004), proteoglycan-high/CD8+-low (HR = 0.39, 95% CI = 0.22 – 0.71, p-value = 0.002) and proteoglycan-high/CD8+-high (HR = 0.33, 95% CI = 0.17 – 0.65, p-value = 0.001) (Figure 3.13C). After adjusting for age, tumour grade, STS subtype, tumour size, sex and performance status, the proteoglycan-low/CD8+-low score remained an independent prognostic factor for OS (Table 3.4).

Table 3.4 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing association of clinicopathological factors and proteoglycan/CD8+ tumour infiltrating lymphocytes (TILs) score with overall survival (OS).

Clinicopathological factors which were significantly associated with OS in UVA analysis were included in the MVA model. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test. I-A = intra-abdominal; RP = retroperitoneal, LMS = leiomyosarcoma, DDLPS = dedifferentiated liposarcoma and UPS = undifferentiated pleomorphic sarcoma.

Variable	Groups	n	Univariable analysis (OS)		Multivariable analysis (OS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
Age	-		1.03 (1.01-1.05)	0.003	1.01 (0.99-1.04)	0.298
Grade	3 (reference)	95	-	-	-	-
	2	52	0.53 (0.32-0.86)	0.010	0.69 (0.37-1.27)	0.232
STS subtype	LMS (reference)	63	-	-	-	-
	UPS	50	1.63 (0.98-2.71)	0.058	1.22 (0.64-2.33)	0.543
	DDLPS	35	1.4 (0.79-2.46)	0.246	0.89 (0.45-1.68)	0.712
Anatomical location	Extremity/Trunk/Head-neck (reference)	78	-	-	-	-
	I-A/RP/Pelvic	62	1.13 (0.72-1.78)	0.59	-	-
	Uterine	8	1.43 (0.56-3.63)	0.451	-	-
Log [tumour size] (mm)	4-5 (reference)	77	-	-	-	-
	>5	46	1.34 (0.84-2.14)	0.224	1.76 (1-3.08)	0.048
	<4	24	0.49 (0.23-1.05)	0.068	0.37 (0.16-0.86)	0.021
Tumour depth	Deep (reference)	124	-	-	-	-
	Superficial	24	0.70 (0.36-1.37)	0.3	-	-
Tumour margin	R1&2 (reference) (R1-76 cases; R2-1case)	77	-	-	-	-
	R0	64	0.84 (0.54-1.31)	0.451	-	-
	Rx	7	0.20 (0.03-1.45)	0.111	-	-
Sex	F (reference)	83	-	-	-	-
	M	65	1.57 (1.01-2.44)	0.044	1.55 (0.91-2.64)	0.103
Performance status	0 (reference)	66	-	-	-	-
	1	38	2.53 (1.47-4.36)	0.001	2.46 (1.36-4.48)	0.003
	2-3 (2-12cases; 3-5cases)	17	5.21 (2.72-9.97)	0.000001	3.52 (1.69-7.33)	0.001
	unknown	27	1.53 (0.79-2.98)	0.21	1.2 (0.58-2.49)	0.629
Proteoglycan/CD8+	Proteoglycan low/CD8+ low (reference)	29	-	-	-	-
	Proteoglycan high/CD8+ low	45	0.39 (0.22-0.71)	0.002	0.43 (0.23-0.80)	0.007
	Proteoglycan low/CD8+ high	45	0.43 (0.24-0.76)	0.004	0.50 (0.26-0.96)	0.037
	Proteoglycan high/CD8+ high	29	0.33 (0.17-0.65)	0.001	0.33 (0.16-0.70)	0.004

3.3. Discussion

Chapter 3 provided an in-depth characterisation of matrisome and adhesome components across multiple histological subtypes of STS. To our knowledge, it is the most comprehensive integrative analysis of matrisome and adhesome in STS. Additionally, the chapter demonstrated the potential clinical value of profiling the matrisome and adhesome for prognostication of the most common STS subtypes.

3.3.1. Matrisome and adhesome overview in STS

It has been historically challenging to extract and analyse the matrisome by mass spectrometry due to the biochemical properties of some matrisome components³⁶⁸. Structural ECM are large, highly crosslinked molecules which undergo extensive posttranslational modifications, making ECM highly insoluble and difficult for digestion. Poor solubility and incomplete digestion can result in lower peptide yields, reducing the sensitivity of mass spectrometry detection. As such, matrisome proteins are commonly underrepresented in global proteomic datasets⁴²⁷. Therefore, it is not surprising that it was the case in our dataset, with $n = 193$ matrisome proteins detected, which corresponded to 19% of *in-silico*-predicted matrisome. In contrast, adhesome components are more soluble, less cross-linked, and smaller molecules compared to matrisome components. Consequently, adhesome components were more represented in our dataset, with $n = 109$, representing 47% coverage of the proteins defined by the adhesome database. Given that only a portion of the proteins from the matrisome and adhesome databases could be profiled, it could have potentially introduced bias in the downstream data analyses.

The majority of recent studies on matrisome characterisation analysed fresh or fresh frozen specimens, specifically focusing on adding an ECM enrichment step to sample processing protocols to increase the number of detectable matrisome proteins⁴²⁷. A limitation of our study was that no ECM enrichment step was included during the FFPE sample preparation for our cohort.

Potentially, the ECM enrichment in our samples prior to the mass spectrometry analysis can result in detection of a higher number of matrisome proteins. At present, ECM enrichment protocols specific to FFPE tissue have not been reported.

The composition of the ECM differs between tissue types, organs and cancer types^{168,428,429}. Chapter 3 focused on characterising the matrisome and adhesome composition in different histological subtypes of STS. Indeed, we found that different histological subtypes, especially LMS and DES, had strikingly distinct expression profiles of matrisome and adhesome proteins. Interestingly, matrisome and adhesome profiles clustered by histological subtype but not by anatomical location. One of the limitations of my analysis is the lack of matching normal tissues, which made it difficult to rule out the possibility that the differences between histological subtype composition of matrisome and adhesome were due to tissue of origin (e.g. we cannot rule out that differences between LMS and DES were due to intrinsic differences of matrisome and adhesome components which were present in smooth muscle and fibroblastic cells). Integration of our proteomic dataset with normal tissue gene and protein expression databases would address this limitation, although only for subtypes where the cell origin is known. Unfortunately, for most STS subtypes, the cell of origin was not identified, and therefore, selecting a suitable normal control tissue remains challenging.

The integrin adhesome historically has not received as much experimental characterisation as the matrisome. Such studies were almost exclusively performed *in vitro* by seeding cells on the purified substrate (e.g. fibronectin) and isolation of integrin adhesion complexes (IACs)^{430–433}. Such studies lacked the complexity of tissue ECM and may not represent biologically relevant interactions of matrisome and adhesome components. The present study is one of the first analyses to specifically evaluate adhesome components in patient tumour samples. Furthermore, by combining this assessment with the evaluation of matrisome components, this present study provides a holistic picture of the role of cell-matrix interactions in STS tumours. Herein, three major matrisome-adhesome signalling networks operating in

STS were identified. These include networks related to the innate immune system, laminin biology and integrin interactions, and collagen turnover. UPS and DDLPS were the subtypes associated with the innate immune system axis, consistent with reports of higher immune activity in UPS and DDLPS compared to other STS subtypes^{133,134}. In contrast, the innate immune proteins were misrepresented in SS, aligning with previous findings of low immune infiltrate in this histological subtype^{133,134,434}. Overall, the results suggest that the ECM and integrin adhesion signalling in UPS and DDLPS subtypes may play a role in modulating the immune microenvironment.

Collagen synthesis and collagen turnover axis were exclusively associated with DES. In our dataset, DES were found to contain a high number of collagen chains, collagen biosynthesis, cross-linking and degradation enzymes, which is expected based on pathology reports of DES being surrounded by collagenous matrix^{417,418}. Several specific matrisome components identified as significantly upregulated in DES (compared to the rest of the cohort) in our dataset were also detected in the study by West *et al.* at a transcriptome level in $n = 10$ DES³⁰². Both our and West *et al.* studies detected a high number of fibrillar collagen chains (COL1A1, COL3A1, COL5A1, COL6A1 and others), glycoproteins (AEBP1, CTHRC1, DPT, FBN1 and others) and collagen remodelling enzymes (LOX and P4HA2). Although several members of ECM remodelling MMP proteins were present in DES samples in our dataset, we did not detect any members of ADAM protease family that were seen in the study by West *et al.* In fact, across the whole cohort of 321 STS samples, only one member of this family was detected, suggesting poor recovery of proteins within this class from FFPE samples. In contrast, at the protein level, several members of protease inhibitors from the Serpins family were detected in our DES cohort, which have not been reported previously in DES.

Our analysis demonstrated that the laminin and integrin signalling axis was significantly associated with LMS. Although there are no studies of comprehensive characterisation of matrisome and adhesome in LMS, a handful of IHC-based studies reported on the presence of basement membrane proteins^{308,311,312}. Those studies detected strong to moderate

staining for laminin, collagen III and IV, and fibronectin in LMS samples. This was in line with our data, where the basement membrane components, apart from collagen III, were uniquely upregulated in LMS samples. Several research groups have identified three molecular subtypes of LMS^{82,136,435–439}. Beck *et al.* reported differential enrichment of adhesome and matrisome components between the molecular subtypes of LMS⁴³⁵. Beck's group I (muscle-enriched) showed upregulation of actin cytoskeleton genes compared to the other groups and included some of the components we identified in our LMS dataset (FERMT2, NEXN, LPP and SORBS1). Beck's group III was enriched in ECM genes, including collagen chains, integrin subunits and ECM remodelling enzymes, consistent with the components identified in our dataset. The molecular heterogeneity of matrisome and adhesome expression profiles in LMS will be examined in Chapter 4 to determine whether the molecular-level biological heterogeneity correlates with the clinical heterogeneity observed in LMS patients.

My analysis has uncovered novel biology related to the ECM and integrin adhesion signalling in different subtypes of STS. In this study, entire tumour sections were analysed, providing an overall proteomic profile within the section, but the spatial resolution was not captured. Given the inherent limitations of bulk proteomic analysis, at present, it is not possible to determine the specific cellular origin or spatial localisation of the ECM and adhesome components in this study. This information would further contribute to understanding the complex interplay between cancer cells, stromal cells, immune cells, and the ECM. It could reveal the role of different cell types in shaping the ECM TME in STS. Recently, matrix-assisted laser desorption/ionisation imaging mass spectrometry (MALDI-IMS), which combines principles of mass spectrometry with 2D spatial information to generate a spatially resolved map of profiled peptides, was used to analyse collagen peptides on the breast cancer TMA⁴⁴⁰ and several glycoproteins, proteoglycans and collagens in aortic valve⁴⁴¹. The authors reported the heterogeneous distribution of different ECM proteins in both studies. As such, MALDI-IMS is an attractive complementary method to bulk proteomic profiling.

MALD-IMS could enable spatial ECM profiling in STS and studying the implications of potential intrapatient ECM heterogeneity.

3.3.2. Proteoglycan score in STS

The analysis in Section 3.2.2 identified subtype-specific matrisome and adhesome components, which might contribute to tumorigenesis in specific STS subtypes. Section 3.2.4 showed that the characterisation of matrisome components has a potential utility as prognostic biomarkers in STS. In particular, proteoglycan score correlated with OS in a sub-cohort of LMS, DDLPS and UPS patients. Moreover, the CD8+ score added further prognostic value to the proteoglycan score. This suggests that despite subtype-specific matrisome and adhesome profiles, common biology involving proteoglycans and CD8+ T cells might contribute to the aggressiveness and progression of the most common STS subtypes.

This chapter described the generation of the combined score of 11 proteoglycans which was comprised of different proteoglycan classes, including the small leucine-rich proteoglycan Class I (BGN, DCN, ASPN), Class II (FMOD, LUM, PRELP) and Class III (OGN) proteins, basement membrane zone proteoglycan (HSPG2) and hyalactinas (ACAN, VCAN)²¹⁵. Low proteoglycan score correlated with shorter OS in our LMS, DDLPS and UPS patients cohort. This result suggests that proteoglycans in STS might have a tumour-suppressive effect. Previous studies reported the anti-tumorigenic role of some of the proteoglycan proteins that constitute our proteoglycan score. Proteoglycans such as LUM^{442,443}, HSPG2⁴⁴⁴, DCN^{222,445–447} and BGN^{448,449} were shown to be tumour-suppressive in carcinomas. The tumour-suppressive effects are believed to be facilitated by the capacity of proteoglycans to sequester growth factors and act as co-receptors to their corresponding receptors^{215,450,451}. This process can lead to the suppression of downstream oncogenic and angiogenic signalling pathways. For instance, the mechanism of DCN-mediated tumour suppression involves binding to and down-regulation multiple RTK, including the EGFR, the HGF receptor Met, the IGF-I receptor (IGF-IR), the VEGF receptor (VEGFR) and the PDGF receptor

(PDGFR)²²¹. A transmembrane proteoglycan CSPG4, which was not found in our prognostic score, was investigated in STS^{313,452,453}. In a genetically engineered autochthonous mouse model of UPS, genetic deletion of CSPG4 at the time of tumour initiation promoted tumour growth⁴⁵². This supports the tumour-suppressive role of proteoglycans in STS. However, the studies investigating the association of CSPG4 mRNA levels and survival outcome measures in STS found the opposite effect. Higher CSPG4 levels correlated with reduced MFS in n = 55 mixed histological subtype STS cohort⁴⁵³. In a larger cohort of n = 610 STS patients, a higher CSPG4 mRNA expression was associated with shorter DFS³¹³. This discrepancy could be due to a poor correlation between mRNA expression and protein expression, cohort-specific differences, or the fact that CSPG4 has a distinct role from the 11 proteoglycans in our signature.

3.3.3. Tumour infiltrating lymphocytes in STS

Low proteoglycan score was associated with higher levels of infiltrating CD3+, CD4+ and CD8+ TILs in our cohort. Proteoglycans have multifaceted functions that promote or suppress the immune response^{221,454–456}. Depending on the tumour type and proteoglycan molecule, proteoglycan can recruit or prevent tumour immune infiltration. Proteoglycans comprise protein cores that can be integrated within a cell membrane or exist in a soluble form in the extracellular space. These protein cores are accompanied by side chains of sulphated GAG sugars. GAG chains bind chemokines, facilitating the leukocyte recruitment^{456,457}. The immune-suppressive function of proteoglycans is associated with the glycocalyx⁴⁵⁸. Proteoglycans line endothelial vessels, where proteoglycans contribute to the glycocalyx structure due to the ability of GAG chains to trap water. The glycocalyx is a hydrated gel-like structure that forms a thick barrier, preventing the interactions between adhesion receptors on leukocytes and endothelial lining, thus impairing leukocyte recruitment into tissues^{459–461}.

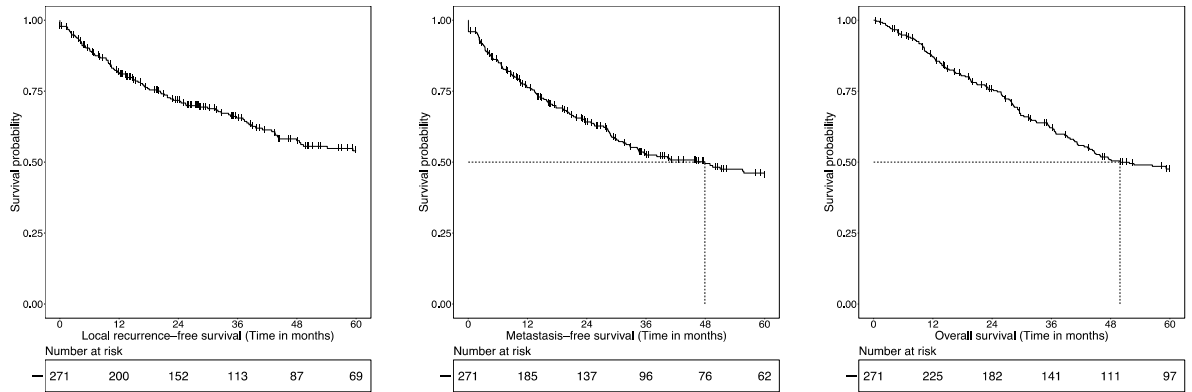
At present, the role of proteoglycan in regulating immune infiltration in STS is unclear. Our findings of inverse association between low proteoglycan

expression and TILs count suggest that proteoglycans might have immunosuppressive function in STS. Based on gene expression data, Boudin *et al.* quantified the infiltration level of lymphoid and myeloid cells in STS tumour tissues classified on CSPG4 proteoglycan expression³¹³. 'CSPG4-low' STS displayed higher infiltration score of lymphoid and myeloid cells than 'CSPG4-high' STS, consistent with higher infiltration of TILs in our STS cohort with low proteoglycan score.

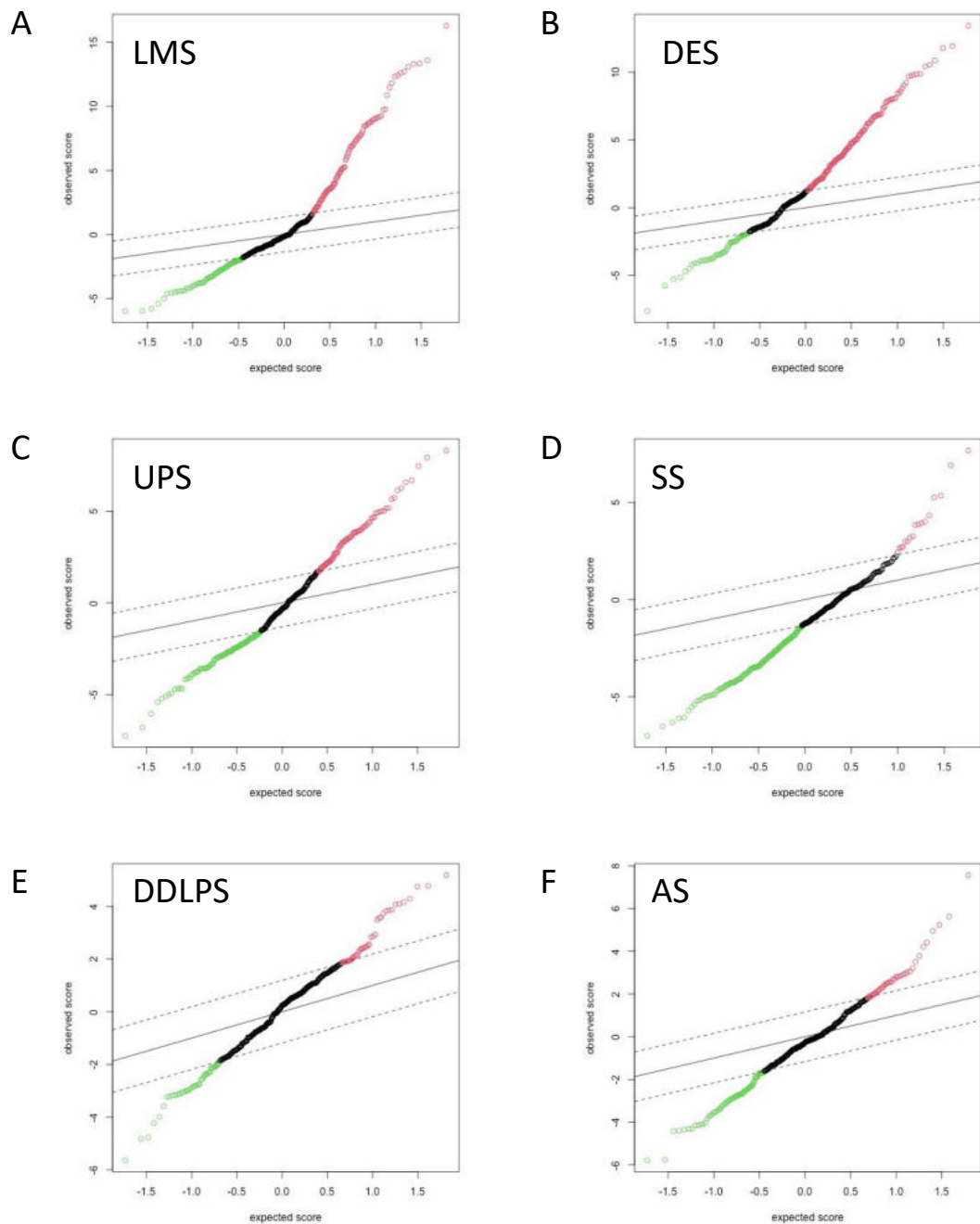
Given that the low proteoglycan/CD8+ score emerged as an independent prognostic factor for OS, it is a potential biomarker in patients with LMS, DDLPS, and UPS. However, it is essential to note that establishing a reliable biomarker would require further validation of the score in an independent cohort. Moreover, to validate the immunosuppressive role of proteoglycans in STS, it will be important to experimentally study the interactions between proteoglycans and TILs. Exploring approaches to modulate these interactions to enhance immune infiltration in STS tumours may offer valuable insights on how to enhance the action of immunotherapy in STS.

3.4. Supplementary Material

3.4.1. Supplementary Figures

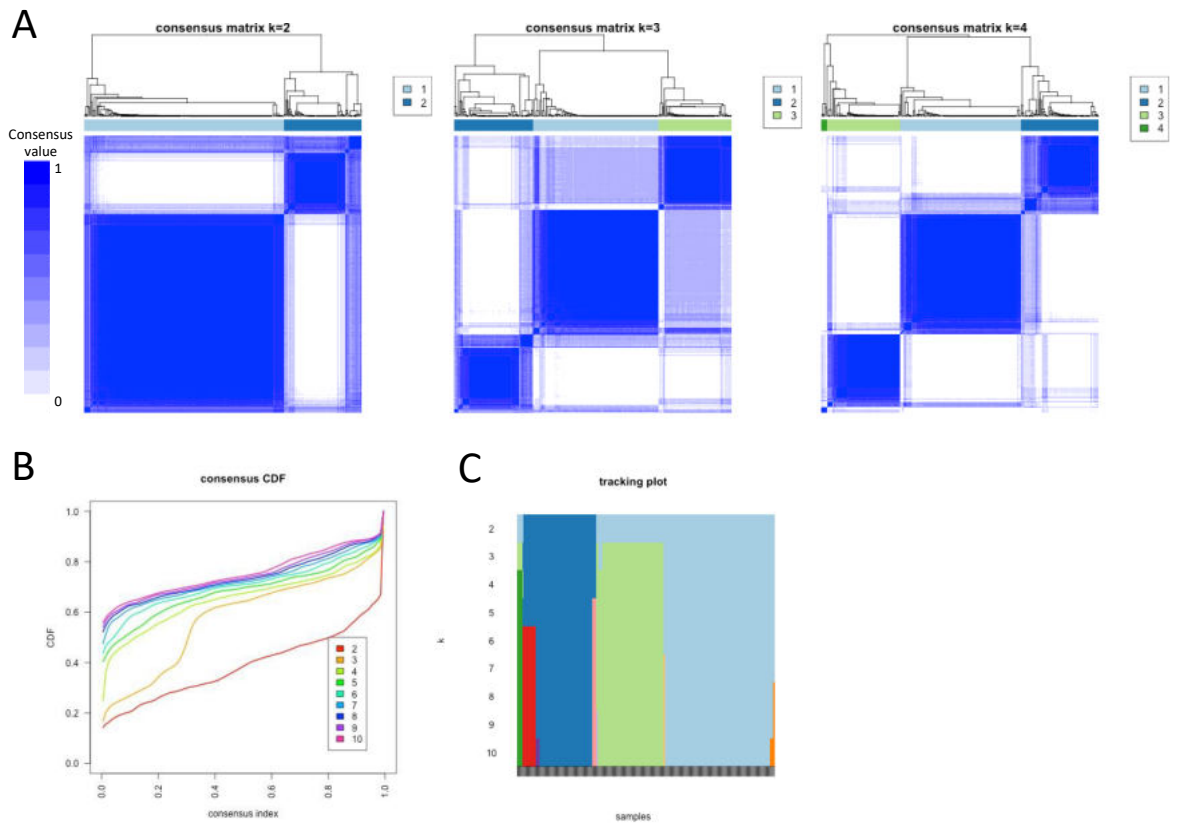


Supplementary Figure 3.1 Kaplan-Meier plots of local recurrence-free survival (LRFS), metastasis-free survival (MFS) and overall survival (OS) in a cohort of n=321 soft tissue sarcoma (STS). Horizontal and vertical lines indicate the median MFS and OS.



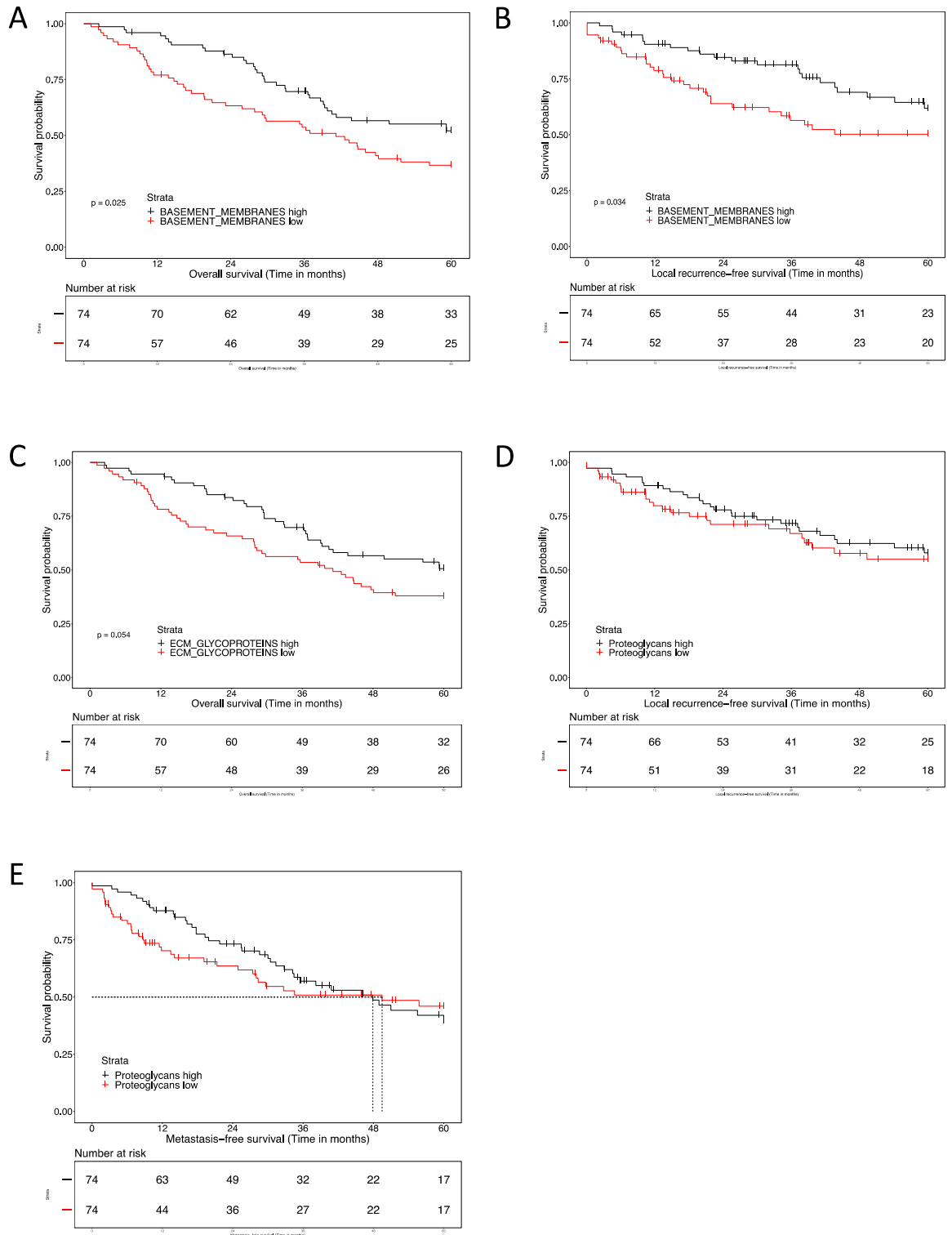
Supplementary Figure 3.2 Significantly differentially expressed matrisome and adhesome proteins in histological subtypes of soft tissue sarcoma (STS).

A-F Significant analysis of microarray (SAM) 2-class unpaired plots for STS subtypes compared to the rest of the cohort A) leiomyosarcoma (LMS), B) desmoid tumour (DES), C) undifferentiated pleomorphic sarcoma (UPS), D) synovial sarcoma (SS), E) dedifferentiated liposarcoma (DDLPS), F) angiosarcoma (AS). Each point is a protein. Proteins within the dashed lines have an FDR ≥ 0.01 and therefore are not significantly differentially expressed proteins (DEPs). Proteins in red are significantly upregulated DEPs (fold change ≥ 2) in the subtype, and proteins in green are significantly downregulated DEPs (fold change < 0.5) in the subtypes.



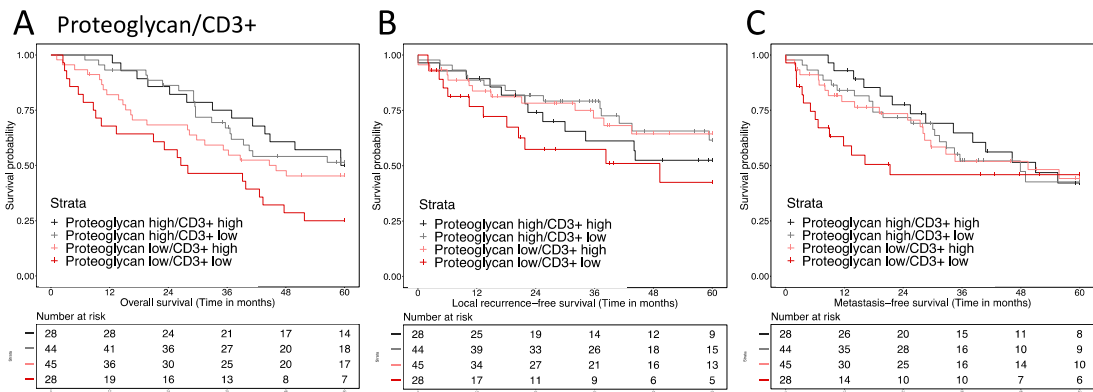
Supplementary Figure 3.3 Identification of matrisome and adhesome networks in soft tissue sarcoma (STS).

A) Consensus matrices for $k = 2, 3$ and 4 , B) Consensus CDF plot for $k = 2-10$, C) Tracking plot for $k = 2-10$.



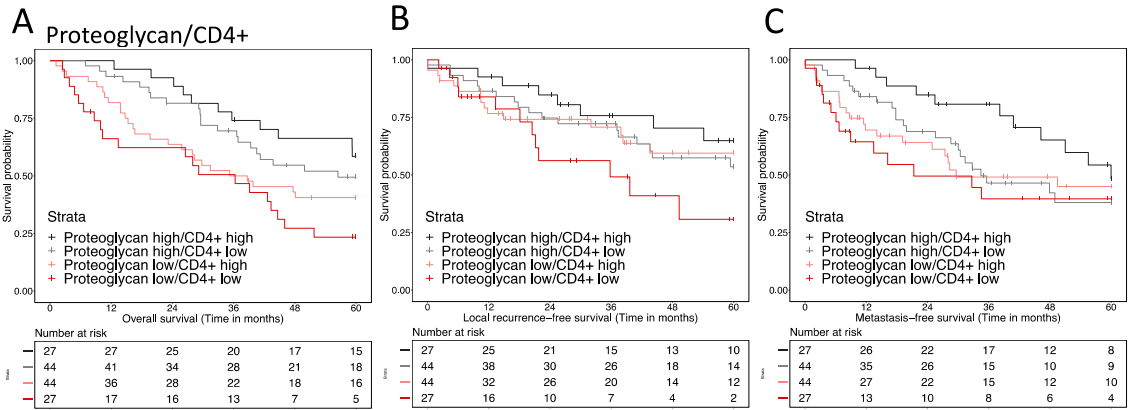
Supplementary Figure 3.4 Association of Naba-defined matrisome related gene sets with survival outcomes in a sub-cohort of soft tissue sarcoma (STS)

Kaplan-Meier plots of A) overall survival, B) local recurrence-free survival with stratification by basement membranes gene set, C) overall survival with stratification by extracellular matrix (ECM) glycoproteins, D) local recurrence-free survival, E) metastasis-free survival with stratification by proteoglycans gene set. Significance is tested with log-rank tests.



Supplementary Figure 3.5 Associations of a proteoglycan/ CD3+ tumour infiltrating lymphocytes (TILs) score with survival outcomes in a sub cohort of soft tissue sarcoma (STS).

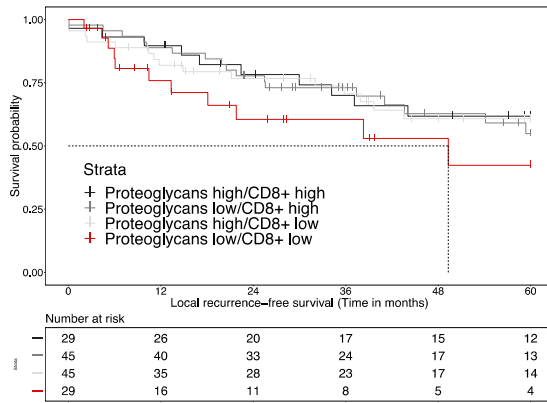
A-C) Kaplan-Meier plots of A) overall survival, B) local recurrence-free and C) metastasis-free survival outcomes with stratification by proteoglycan/CD3+ score. Corresponding univariable Cox regression analysis results are detailed in Supplementary Table 3.5



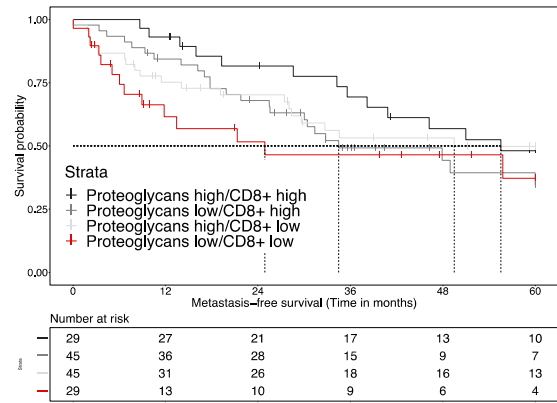
Supplementary Figure 3.6 Associations of a proteoglycan/ CD4+ tumour infiltrating lymphocytes (TILs) score with survival outcomes in a sub-cohort of soft tissue sarcoma (STS).

A-C) Kaplan-Meier plots of A) overall survival, B) local recurrence-free and C) metastasis-free survival outcomes with stratification by proteoglycan/CD4+ score. Corresponding univariable Cox regression analysis results are detailed in Supplementary Table 3.5.

A Proteoglycan/CD8+



B



Supplementary Figure 3.7 Associations of a proteoglycan/ CD8+ tumour infiltrating lymphocytes (TILs) score with survival outcomes in a sub-cohort of soft tissue sarcoma (STS).

Kaplan-Meier plots of A) local recurrence-free, B) metastasis-free survival outcomes with stratification by proteoglycan/CD8+ score. Corresponding univariable Cox regression analysis results are detailed in Supplementary Table 3.5.

3.4.2. Supplementary Tables

Supplementary Table 3.1 Clinicopathological characteristics of n=321 soft tissue sarcoma (STS) cases.

Summary of features of the whole cohort and different histological subtypes of STS. For continuous variables median, minimum (min) and maximum (max) values are indicated. For categorical variables count (n) and percentages (%) are shown. F = female; M = male. AS = angiosarcoma; ASPS = alveolar soft part sarcoma; CCS = clear cell sarcoma; DDLPS = dedifferentiated liposarcoma; DES = desmoid tumour; DSRCT = desmoplastic small round cell tumour; ES = epithelioid sarcoma; LMS = leiomyosarcoma; RT = rhabdoid tumour; SS = synovial sarcoma; UPS = undifferentiated pleomorphic sarcoma.

Characteristic	Category	Whole cohort	AS	ASPC	CCS	DDLPS	DES	DSRCT	ES	LMS	RT	SS	UPS
Number of patients	n (%)	321 (100)	30 (9)	4 (1)	3 (1)	39 (12)	37 (12)	4 (1)	16 (5)	80 (25)	12 (4)	43 (13)	53 (17)
Age at excision (years)	median	58.4	68.8	22.3	49.1	63	39.3	28.7	38.5	65.3	1.1	42.3	73.5
	min	0.1	27.3	18.1	25.2	35.1	21.2	16.6	18.3	29.3	0.1	19.6	28.2
	max	90	82.7	33.9	61.9	81.3	78.3	46.1	76.8	86.9	4.7	79.4	90
Anatomical site [n (%)]	Extremity	125 (38.9)	2 (6.7)	4 (100.0)	3 (100.0)	2 (5.1)	9 (24.3)	-	9 (56.2)	31 (38.8)	1 (8.3)	26 (60.5)	38 (71.7)
	Head/neck	13 (4.0)	4 (13.3)	-	-	-	1 (2.7)	-	-	-	2 (16.7)	2 (4.7)	4 (7.5)
	Intra-abdominal	28 (8.7)	2 (6.7)	-	-	3 (7.7)	4 (10.8)	3 (75.0)	-	10 (12.5)	3 (25.0)	2 (4.7)	1 (1.9)
	Retroperitoneal	57 (17.8)	1 (3.3)	-	-	32 (82.1)	-	1 (25.0)	-	19 (23.8)	2 (16.7)	2 (4.7)	-
	Trunk	65 (20.2)	21 (70.0)	-	-	2 (5.1)	22 (59.5)	-	1 (6.2)	2 (2.5)	2 (16.7)	7 (16.3)	8 (15.1)
	Pelvic	24 (7.5)	-	-	-	-	1 (2.7)	-	6 (37.5)	9 (11.2)	2 (16.7)	4 (9.3)	2 (3.8)
Uterine	9 (2.8)	-	-	-	-	-	-	-	9 (11.2)	-	-	-	
Grade [n (%)]	2	115 (35.8)	12 (40.0)	1 (25.0)	-	19 (48.7)	-	-	10 (62.5)	47 (58.8)	-	23 (53.5)	3 (5.7)
	3	139 (43.3)	13 (43.3)	-	3 (100.0)	20 (51.3)	-	3 (75.0)	5 (31.2)	33 (41.2)	-	13 (30.2)	49 (92.5)
	unknown	67 (20.9)	5 (16.7)	3 (75.0)	-	-	37 (100.0)	1 (25.0)	1 (6.2)	-	12 (100.0)	7 (16.3)	1 (1.9)
Tumour depth [n (%)]	Deep	250 (77.9)	15 (50.0)	4 (100.0)	3 (100.0)	38 (97.4)	30 (81.1)	4 (100.0)	8 (50.0)	66 (82.5)	-	39 (90.7)	43 (81.1)
	Superficial	54 (16.8)	15 (50.0)	-	-	1 (2.6)	2 (5.4)	-	8 (50.0)	14 (17.5)	-	4 (9.3)	10 (18.9)
	unknown	17 (5.3)	-	-	-	-	5 (13.5)	-	-	-	12 (100.0)	-	-
Tumour size (mm)	median	90	58	68	55	190	90	132.5	50	92.5	-	71	80
	min	4	4	30	9	35	25	70	10	5	-	18	15
	max	1090	400	100	95	1090	500	175	240	400	-	760	360
Tumour margins [n (%)]	R0	133 (41.4)	17 (56.7)	1 (25.0)	1 (33.3)	9 (23.1)	10 (27.0)	-	11 (68.8)	42 (52.5)	-	18 (41.9)	26 (49.1)
	R1	151 (47.0)	11 (36.7)	3 (75.0)	2 (66.7)	25 (64.1)	20 (54.1)	1 (25.0)	5 (31.2)	35 (43.8)	-	22 (51.2)	27 (50.9)
	R2	4 (1.2)	-	-	-	-	1 (2.7)	1 (25.0)	-	1 (1.2)	-	1 (2.3)	-
	unknown	33 (10.3)	2 (6.7)	-	-	5 (12.8)	6 (16.2)	2 (50.0)	-	2 (2.5)	12 (100.0)	2 (4.7)	-
Pre-op treatment [n (%)]	Chemo	19 (5.9)	5 (16.7)	-	-	1 (2.6)	3 (8.1)	3 (75.0)	-	-	-	7 (16.3)	-
	Radio	8 (2.5)	-	1 (25.0)	-	-	-	-	-	1 (1.2)	-	6 (14.0)	-
	Chemo & radio	13 (4.0)	-	-	-	-	-	1 (25.0)	-	-	-	12 (27.9)	-
	None	267 (83.2)	25 (83.3)	3 (75.0)	3 (100.0)	38 (97.4)	34 (91.9)	-	15 (93.8)	79 (98.8)	-	17 (39.5)	53 (100.0)
unknown	14 (4.4)	-	-	-	-	-	-	1 (6.2)	-	12 (100.0)	1 (2.3)	-	
Performance status [n (%)]	0	158 (49.2)	15 (50.0)	4 (100.0)	2 (66.7)	17 (43.6)	28 (75.7)	3 (75.0)	7 (43.8)	40 (50.0)	-	20 (46.5)	22 (41.5)
	1	82 (25.5)	12 (40.0)	-	-	12 (30.8)	4 (10.8)	1 (25.0)	5 (31.2)	16 (20.0)	-	17 (39.5)	15 (28.3)
	2	16 (5.0)	-	-	-	2 (5.1)	-	-	-	7 (8.8)	-	3 (7.0)	4 (7.5)
	3	5 (1.6)	-	-	-	1 (2.6)	-	-	-	1 (1.2)	-	-	3 (5.7)
	unknown	60 (18.7)	3 (10.0)	-	1 (33.3)	7 (17.9)	5 (13.5)	-	4 (25.0)	16 (20.0)	12 (100.0)	3 (7.0)	9 (17.0)
Sex [n (%)]	F	201 (62.6)	26 (86.7)	3 (75.0)	1 (33.3)	15 (38.5)	29 (78.4)	1 (25.0)	8 (50.0)	56 (70.0)	7 (58.3)	27 (62.8)	28 (52.8)
	M	119 (37.1)	4 (13.3)	1 (25.0)	2 (66.7)	24 (61.5)	8 (21.6)	3 (75.0)	8 (50.0)	24 (30.0)	4 (33.3)	16 (37.2)	25 (47.2)
	unknown	1 (0.3)	-	-	-	-	-	-	-	-	1 (8.3)	-	-
Status at excision [n (%)]	Local	301 (93.8)	29 (96.7)	3 (75.0)	2 (66.7)	36 (92.3)	37 (100.0)	2 (50.0)	13 (81.2)	78 (97.5)	6 (50.0)	42 (97.7)	53 (100.0)
	Metastatic	15 (4.7)	1 (3.3)	1 (25.0)	1 (33.3)	2 (5.1)	-	2 (50.0)	-	2 (2.5)	5 (41.7)	1 (2.3)	-
	Locally Metastatic	3 (0.9)	-	-	-	-	-	-	3 (18.8)	-	-	-	-
	Multifocal	1 (0.3)	-	-	-	1 (2.6)	-	-	-	-	-	-	-
	unknown	1 (0.3)	-	-	-	-	-	-	-	-	1 (8.3)	-	-

Supplementary Table 3.2 Reactome enriched pathways in matrisome and adhesome networks of soft tissue sarcoma (STS).

Tables for each consensus cluster with enriched term description and false discovery rate (FDR). Top 10 hits with FDR < 0.001 are highlighted for each cluster.

CC1

Term description	FDR
Immune System	1.53E-11
Innate Immune System	1.49E-09
Neutrophil degranulation	6.40E-07
Fcgamma receptor (FCGR) dependent phagocytosis	1.65E-05
Hemostasis	1.88E-05
Signaling by Receptor Tyrosine Kinases	2.88E-05
Regulation of actin dynamics for phagocytic cup formation	3.11E-05
Extracellular matrix organization	6.65E-05
EPH-Ephrin signaling	0.00013
FCGR3A-mediated phagocytosis	0.00014
EPHB-mediated forward signaling	0.00051
VEGFA-VEGFR2 Pathway	0.00051
Signaling by Interleukins	0.0012
PTK6 Regulates RHO GTPases, RAS GTPase and MAP kinases	0.0013
Degradation of the extracellular matrix	0.0027
Cell surface interactions at the vascular wall	0.0027
Signal Transduction	0.0028
Adaptive Immune System	0.003
Interleukin-3, Interleukin-5 and GM-CSF signaling	0.0032
Platelet activation, signaling and aggregation	0.0033
Cytokine Signaling in Immune system	0.0039
Signaling by SCF-KIT	0.0039
Signaling by PTK6	0.0039
Regulation of signaling by CBL	0.0039
Costimulation by the CD28 family	0.004

CC2

Term description	FDR
Laminin interactions	1.09E-13
Extracellular matrix organization	1.12E-12
Non-integrin membrane-ECM interactions	9.23E-08
Cell-extracellular matrix interactions	5.49E-07
ECM proteoglycans	6.27E-07
Cell-Cell communication	1.05E-06
Integrin cell surface interactions	1.05E-06
Signaling by Receptor Tyrosine Kinases	7.68E-06
MET activates PTK2 signaling	1.80E-05
Hemostasis	0.00023
Signaling by MET	0.00023
Cell surface interactions at the vascular wall	0.00029
Degradation of the extracellular matrix	0.002
Post-translational protein phosphorylation	0.002
Axon guidance	0.0023
Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs)	0.0029
Elastic fibre formation	0.0063
L1CAM interactions	0.0092
A tetrasaccharide linker sequence is required for GAG synthesis	0.0182
Defective B4GALT7 causes EDS, progeroid type	0.0182
Defective B3GAT3 causes JDSSDHD	0.0182
Defective B3GALT6 causes EDSP2 and SEMD11	0.0182
Diseases of glycosylation	0.025
Signal Transduction	0.0286
Molecules associated with elastic fibres	0.0286

CC3

Term description	FDR
Extracellular matrix organization	1.55E-39
Degradation of the extracellular matrix	6.80E-16
ECM proteoglycans	2.06E-15
Integrin cell surface interactions	5.44E-15
Collagen formation	9.67E-14
Collagen biosynthesis and modifying enzymes	3.43E-12
Assembly of collagen fibrils and other multimeric structures	3.83E-11
Collagen degradation	2.37E-10
Hemostasis	7.42E-10
Collagen chain trimerization	8.39E-10
Elastic fibre formation	1.20E-09
Signaling by PDGF	1.28E-09
Formation of Fibrin Clot (Clotting Cascade)	1.78E-09
Molecules associated with elastic fibres	4.20E-09
Platelet activation, signaling and aggregation	1.52E-08
Syndecan interactions	7.11E-08
Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs)	7.11E-08
Non-integrin membrane-ECM interactions	1.05E-07
Intrinsic Pathway of Fibrin Clot Formation	1.60E-07
Platelet degranulation	1.75E-07
MET activates PTK2 signaling	5.64E-06
Platelet Aggregation (Plug Formation)	1.15E-05
Diseases of glycosylation	1.22E-05
Signaling by Receptor Tyrosine Kinases	2.30E-05
Common Pathway of Fibrin Clot Formation	2.45E-05

Supplementary Table 3.3 Statistical comparisons between histological subtypes in consensus cluster 1 (CC1).

Significant results are in bold. Kruskal-Wallis test was performed and adjusted p values were obtained with Dunn's multiple correction tests. AS = angiosarcoma; DDLPS = dedifferentiated liposarcoma; DES = desmoid tumour; LMS = leiomyosarcoma; SS = synovial sarcoma; UPS = undifferentiated pleomorphic sarcoma.

Dunn's multiple comparisons test	Summary	Adjusted P Value
AS vs. SS	****	<0.0001
AS vs. UPS	****	<0.0001
DDLPS vs. LMS	**	0.0016
DDLPS vs. Other	*	0.0334
DDLPS vs. SS	****	<0.0001
DDLPS vs. UPS	**	0.0041
DES vs. SS	****	<0.0001
DES vs. UPS	****	<0.0001
LMS vs. SS	****	<0.0001
LMS vs. UPS	****	<0.0001
Other vs. SS	****	<0.0001
Other vs. UPS	****	<0.0001
SS vs. UPS	****	<0.0001
AS vs. DDLPS	ns	>0.9999
AS vs. DES	ns	>0.9999
AS vs. LMS	ns	>0.9999
AS vs. Other	ns	>0.9999
DDLPS vs. DES	ns	0.0654
DES vs. LMS	ns	>0.9999
DES vs. Other	ns	>0.9999
LMS vs. Other	ns	>0.9999

Supplementary Table 3.4 Clinicopathological characteristics of n = 148 soft tissue sarcoma (STS) cases.

Summary of features of the whole cohort. For continuous variables median, minimum (min) and maximum (max) values are indicated. For categorical variables count (n) and percentages (%) are shown. F = female; M = male.

Characteristic	Category	Total
Number of patients	n	148
Histological subtype	Leiomyosarcoma	63 (42.6)
	Dedifferentiated Liposarcoma	35 (23.6)
	Undifferentiated Pleomorphic Sarcoma	50 (33.8)
Age at excision (years)	Median	67.7
	Min	28.2
	Max	90
Anatomical site [n (%)]	Extremity	64 (43.2)
	Intra-abdominal	11 (7.4)
	Retroperitoneal	43 (29.1)
	Trunk	10 (6.8)
	Pelvic	8 (5.4)
	Uterine	8 (5.4)
Grade [n (%)]	Head/neck	4 (2.7)
	2	52 (35.1)
	3	95 (64.2)
Tumour depth [n (%)]	missing	1 (0.7)
	Deep	124 (83.8)
Tumour size (mm)	Superficial	24 (16.2)
	Median	100
	Min	5
Tumour margins [n (%)]	Max	500
	R0	64 (43.2)
	R1	76 (51.4)
	Rx	7 (4.7)
Pre-op treatment [n (%)]	R2	1 (0.7)
	Chemo	1 (0.7)
Performance status [n (%)]	None	147 (99.3)
	0	66 (44.6)
	1	38 (25.7)
	2	12 (8.1)
	3	5 (3.4)
Sex [n (%)]	missing	27 (18.2)
	F	83 (56.1)
Status at excision [n (%)]	M	65 (43.9)
	Local	144 (97.3)
	Metastatic	3 (2)
	Multifocal	1 (0.7)

Supplementary Table 3.5 Statistical association of clinicopathological features with proteoglycan groups.

Abbreviations: F = female; M = male; LMS = leiomyosarcoma; UPS = undifferentiated pleomorphic sarcoma; DDLPS = dedifferentiated liposarcoma.

Variable		Proteoglycan group		Test results			
		Low	High	Test performed	χ -squared	Degrees of freedom	p-value
Age at excision (years)	median	68.6	67.5	Kruskal Wallis	0.38	1	0.537
	min	28.2	30.5				
	max	86.3	90.0				
Tumour size (mm)	median	110	87.5	Kruskal Wallis	0.71	1	0.398
	min	15	5				
	max	450	500				
Anatomical site	Extremity	30	34	chi-square	7.58	6	0.271
	Head/neck	4	-				
	Intra-abdominal	6	5				
	Pelvic	4	4				
	Retroperitoneal	18	25				
	Trunk	7	3				
Uterine	5	3					
Grade	2	14	38	chi-square	16.64	1	0.00005
	3	59	36				
Performance status	0	28	38	chi-square	6.95	4	0.138
	1	17	21				
	2	7	5				
	3	3	2				
	unknown	19	8				
Sex	F	45	38	chi-square	1.34	1	0.246
	M	29	36				
Subtype	LMS	22	41	chi-square	12.23	2	0.002
	DDLPS	18	17				
	UPS	34	16				
Tumour margins	R0	33	31	chi-square	0.25	1	0.621
	R1	36	40				
Tumour depth	Deep	62	62	chi-square	0	1	1
	Superficial	12	12				

Supplementary Table 3.6 Summary of multivariable (MVA) Cox regression analysis assessing basement membrane score with local recurrence-free survival (LRFS) and overall survival (OS).

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by multivariable Cox regression with Wald tests. I-A = intra-abdominal; RP = retroperitoneal, LMS = leiomyosarcoma, DDLPS = dedifferentiated liposarcoma and UPS = undifferentiated pleomorphic sarcoma.

Variable	Groups	n	Multivariable analysis (LRFS)		Multivariable analysis (OS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
Age	-	-	1 (0.98-1.03)	0.989	1.02 (0.99-1.04)	0.181
Grade	3 (reference)	95	-	-	-	-
	2	52	0.97 (0.49-1.92)	0.927	0.58 (0.31-1.07)	0.079
STS subtype	LMS (reference)	63	-	-	-	-
	UPS	50	0.97 (0.39-2.44)	0.946	1.25 (0.62-2.52)	0.529
	DDLPS	35	2.1 (0.91-4.87)	0.083	1.04 (0.52-2.08)	0.918
Anatomical location	Extremity/Trunk/Head-neck (reference)	78	-	-	-	-
	I-A/RP/Pelvic	62	1.13 (0.72-1.78)	0.590	-	-
	Uterine	8	1.43 (0.56-3.63)	0.451	-	-
Log [tumour size] (mm)	4-5 (reference)	77	-	-	-	-
	>5	46	1.7 (0.82-3.5)	0.152	1.85 (1.05-3.26)	0.035
	<4	24	0.76 (0.30-1.96)	0.570	0.429 (0.19-0.99)	0.049
Tumour depth	Deep (reference)	124	-	-	-	-
	Superficial	24	0.70 (0.36-1.37)	0.300	-	-
Tumour margin	R1&2 (reference) (R1-76 cases; R2-1case)	77	-	-	-	-
	R0	64	0.84 (0.54-1.31)	0.451	-	-
	Rx	7	0.20 (0.03-1.45)	0.111	-	-
Sex	F (reference)	83	-	-	-	-
	M	65	1.21 (0.65-2.24)	0.549	1.31 (0.79-2.19)	0.297
Performance status	0 (reference)	66	-	-	-	-
	1	38	2.34 (1.17-4.68)	0.016	2.67 (1.5-4.75)	0.001
	2-3 (2-12cases; 3-5cases)	17	1.09 (0.36-3.31)	0.885	3.89 (1.89-7.99)	0.0002
	unknown	27	0.97 (0.40-2.31)	0.937	1.32 (0.63-2.76)	0.465
Basement membrane score	Basement membrane low (reference)	74	-	-	-	-
	Basement membrane high	74	0.70 (0.34-1.47)	0.351	1.02 (0.56-1.88)	0.937

Supplementary Table 3.7 Summary of univariable (UVA) Cox regression analysis assessing overall survival (OS), local recurrence-free survival (LRFS) and metastasis-free survival (MFS) in n = 145 STS patients with CD3+ tumour infiltrating lymphocytes (TILs) counts.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression analysis with Wald tests.

Variable	OS		LRFS		MFS	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
CD3+ high (reference)	-	-	-	-	-	-
CD3+ low	1.2 (0.77-1.86)	0.426	1.09 (0.63-1.89)	0.749	1.16 (0.72-1.86)	0.534

Supplementary Table 3.8 Summary of univariable (UVA) Cox regression analyses assessing overall survival (OS), local recurrence-free survival (LRFS) and metastasis-free survival (MFS) in n = 145 STS patients stratified by proteoglycan and CD3+ tumour infiltrating lymphocytes score.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression analysis with Wald tests.

Variable	Groups	n	Univariable analysis (OS)		Univariable analysis (LRFS)		Univariable analysis (MFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Proteoglycan/CD3+	Proteoglycan low/CD3+ low (reference)	28	-	-	-	-	-	-
	Proteoglycan high/CD3+ low	44	0.44 (0.24-0.82)	0.009	0.50 (0.23-1.08)	0.076	0.71 (0.36-1.41)	0.331
	Proteoglycan low/CD3+ high	45	0.59 (0.33-1.05)	0.073	0.52 (0.24-1.14)	0.105	0.73 (0.37-1.45)	0.367
	Proteoglycan high/CD3+ high	28	0.43 (0.22-0.85)	0.016	0.69 (0.31-1.54)	0.367	0.65 (0.31-1.36)	0.255

Supplementary Table 3.9 Summary of univariable (UVA) Cox regression analysis assessing overall survival (OS), local recurrence-free survival (LRFS) and metastasis-free survival (MFS) in n = 142 STS patients with CD4+ tumour infiltrating lymphocytes (TILs) counts.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression analysis with Wald tests.

Variable	OS		LRFS		MFS	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
CD4+ low (reference)	-	-	-	-	-	-
CD4+ high	1.19 (0.76-1.86)	0.443	1.39 (0.81-2.4)	0.232	1.31 (0.81-2.1)	0.269

Supplementary Table 3.10 Summary of univariable (UVA) Cox regression analyses assessing overall survival (OS), local recurrence-free survival (LRFS), and metastasis-free survival (MFS) in n = 142 STS patients stratified by proteoglycan and CD4+ tumour infiltrating lymphocytes score.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression with Wald tests.

Variable	Groups	n	Univariable analysis (OS)		Univariable analysis (LRFS)		Univariable analysis (MFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Proteoglycan/CD4+	Proteoglycan low/CD4+ low (reference)	27	-	-	-	-	-	-
	Proteoglycan high/CD4+ low	44	0.46 (0.25-0.85)	0.013	0.58 (0.28-1.21)	0.146	0.75 (0.38-1.45)	0.391
	Proteoglycan low/CD4+ high	44	0.67 (0.38-1.21)	0.184	0.57 (0.27-1.23)	0.153	0.79 (0.40-1.55)	0.492
	Proteoglycan high/CD4+ high	27	0.34 (0.16-0.72)	0.004	0.40 (0.16-0.99)	0.047	0.46 (0.21-1.01)	0.053

Supplementary Table 3.11 Summary of univariable (UVA) Cox regression analysis assessing overall survival (OS), local recurrence-free survival (LRFS) and metastasis-free survival (MFS) in n = 148 STS patients with CD8+ tumour infiltrating lymphocytes (TILs) counts.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression analysis with Wald tests.

Variable	OS		LRFS		MFS	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
CD8+ low (reference)	-	-	-	-	-	-
CD8+ high	0.67 (0.43-1.04)	0.0764	0.791 (0.461-1.36)	0.396	0.706 (0.442-1.13)	0.146

Supplementary Table 3.12 Summary of univariable (UVA) Cox regression analyses assessing overall survival (OS), local recurrence-free survival (LRFS), and metastasis-free survival (MFS) in n = 148 STS patients stratified by proteoglycan and CD8+ tumour infiltrating lymphocytes score.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression analysis with Wald tests.

Variable	Groups	n	Univariable analysis (OS)		Univariable analysis (LRFS)		Univariable analysis (MFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Proteoglycan/CD8+	Proteoglycan low/CD8+ low (reference)	29	-	-	-	-	-	-
	Proteoglycan high/CD8+ low	45	0.39 (0.22-0.71)	0.002	0.62 (0.29-1.33)	0.223	0.79 (0.41-1.54)	0.491
	Proteoglycan low/CD8+ high	45	0.43 (0.24-0.76)	0.004	0.60 (0.27-1.3)	0.195	0.66 (0.34-1.32)	0.241
	Proteoglycan high/CD8+ high	29	0.33 (0.17-0.65)	0.001	0.55 (0.23-1.3)	0.175	0.53 (0.25-1.14)	0.105

Chapter 4 Understanding heterogeneity in matrix signalling in LMS, DDLPS and UPS

4.1. Background and objectives

Chapter 3 comprehensively analysed the protein expression of matrisome and adhesome components across 11 histological STS subtypes. These findings revealed that most histological subtypes exhibited unique ECM and integrin adhesion signalling profiles. Additionally, clear intra-subtype heterogeneity was observed in the protein expression profiles of matrisome and adhesome components in LMS, DDLPS and UPS. As such, the first objective of Chapter 4 was to investigate the molecular heterogeneity of matrix signalling in LMS, DDLPS and UPS.

Intra-subtype molecular heterogeneity in cancers is often reflected in clinical heterogeneity, where patients within the same subtype may exhibit different treatment responses, metastasis rates, and overall outcomes. LMS, DDLPS and UPS are known for clinical heterogeneity^{20,462,463}. An underlying factor potentially contributing to this clinical diversity is the dysregulation of ECM remodelling. ECM remodelling involves altering the composition and arrangement of ECM components, thereby impacting tumour cell behaviour. In the context of cancer, ECM dysregulation has been implicated as a driving force behind disease progression^{410,464}, potentially underscoring the observed clinical heterogeneity. The second objective of this chapter was to determine whether the matrix signalling heterogeneity is associated with the clinicopathological features and survival outcome measures in LMS, DDLPS and UPS patients, thus establishing a link between matrix signalling heterogeneity and clinical heterogeneity observed in LMS, DDLPS and UPS patients.

ECM remodelling is known to induce aggressive tumour phenotypes and modulate the immune microenvironment in other cancer types^{145,412,465}. Matrisome and adhesome network analysis in Chapter 3 revealed the enrichment of immune-related pathways in DDLPS and UPS. Moreover, given

the notable T lymphocyte infiltration observed in DDLPS and UPS patients^{133,134,466}, Chapter 4 additionally sought to characterise the relationship between ECM remodelling and TILs in DDLPS and UPS.

4.2. Results

4.2.1. Intra-subtype heterogeneity in LMS

4.2.1.1. *Clinicopathological features of LMS cohort*

Proteomic and clinical data were available for 80 LMS patients. The clinicopathological features of the cohort are summarised in Table 4.1. Briefly, the median age at excision was 65 years (range: 29-87). The cohort had a high proportion of females compared to males (70% compared to 30%). The most common site of tumour presentation was the extremities (39%), followed by the retroperitoneum (24%). Eleven percent of the cohort had uterine LMS tumours. There was a slight prevalence for intermediate grade (grade 2) LMS tumours, compared to the high grade (59% and 41%, respectively). The majority of tumours were deep-seated (83%). Patients presented with relatively large tumours (median size – 95 mm, range: 5 – 400 mm). Tumour margins were mainly R0 and R1 (53% and 44%, respectively). The cohort consisted of treatment naïve patients and one LMS patient (1.2%) received radiotherapy prior to surgery. Half of the patients had a performance status score of 0, and 20% had a score of 1. Performance status information was missing for 20% of LMS patients. Metastatic disease at resection was present in 2.5% of LMS patients, and 97.5% had localised primary disease. All the survival analyses were censored at the 5-year (60 months) mark. The median MFS was 36 months (95% CI 30 – 55), while the median LRFS and OS were not reached (Supplementary Figure 4.1). LMS is characterised by high rates of metastasis and low rates of local recurrence²⁰, and in line with that, 56% of patients had developed metastasis, and 26% had a local recurrence event at 60 months after surgery (Supplementary Figure 4.1).

Table 4.1 Clinicopathological characteristics of n=80 leiomyosarcoma (LMS) cases.

For continuous variables, median, minimum (min) and maximum (max) values are indicated.

For categorical variables, count (n) and percentages (%) are shown. F = female; M = male.

Characteristic	Category	Leiomyosarcoma
Number of patients	n (%)	80 (100)
Age at excision (years)	Median	65.3
	Min	29.3
	Max	86.9
Anatomical site [n (%)]	Extremity	31 (38.8)
	Intra-abdominal	10 (12.5)
	Pelvic	9 (11.2)
	Retroperitoneal	19 (23.8)
	Trunk	2 (2.5)
Grade [n (%)]	2	47 (58.8)
	3	33 (41.2)
Tumour depth [n (%)]	Deep	66 (82.5)
	Superficial	14 (17.5)
Tumour size (mm)	Median	95
	Min	5
	Max	400
Tumour margins [n (%)]	R0	42 (52.5)
	R1	35 (43.8)
	R2	1 (1.2)
	Rx	2 (2.5)
Pre-op treatment [n (%)]	None	79 (98.8)
	Radio	1 (1.2)
Performance status [n (%)]	0	40 (50)
	1	16 (20)
	2	7 (8.8)
	3	1 (1.2)
	unknown	16 (20)
Sex [n (%)]	F	56 (70)
	M	24 (30)
Status at excision [n (%)]	Local	78 (97.5)
	Metastatic	2 (2.5)

4.2.1.2. Identification of LMS subtypes

In Section 3.2.2, variability in the protein expression patterns of matrisome and adhesome was observed within the LMS cohort. To assess whether this molecular heterogeneity contributes to diverse clinical characteristics observed in LMS, I first identified subgroups of LMS patients grouped based

on similarities of matrisome and adhesome protein expression profiles. Consensus clustering was applied to matrisome and adhesome proteomic profiles in our LMS cohort of 80 patients. The optimal number of clusters was assessed as described in Section 3.2.3. Briefly, the consensus matrices for $k = 3, 4$ and 5 showed a clear separation of clusters (Supplementary Figure 4.2A). CDF depicts consensus distribution for each k , and the $k = 3$ showed a steady increase in CDF between 0 and 1 , and there was a minimum improvement in CDF beyond $k = 3$, indicating optimal clustering at $k = 3$ (Supplementary Figure 4.2B). As indicated by the tracking plot, starting from $k = 5$, cases showed a frequent cluster reassignment, indicating unstable clustering beyond $k = 5$ (Supplementary Figure 4.2C). Moreover, the SigClust test indicated the presence of three significantly different consensus clusters, with all pair-wise comparisons being significant at $p < 0.05$ (Supplementary Figure 4.2D). After visual inspection of the consensus clustering results and SigClust, $k = 3$ was chosen, indicating that in our cohort, there were three objectively distinct subgroups of LMS patients (LMS1, LMS2 and LMS3).

4.2.1.3. Biological characterisation of LMS subgroups

Characterisation of ECM and integrin adhesion networks in LMS subgroups

To identify matrisome and adhesome proteins specific to each LMS subgroup, two-class unpaired SAM tests were performed (LMS1 vs the rest, LMS2 vs the rest and LMS3 vs the rest). Only the upregulated DEPs (with $FDR < 0.01$ and fold change ≥ 2) were considered, and DEPs which were found in more than 1 subgroup were removed. Figure 4.1 illustrates the expression profiles of 112 matrisome and adhesome DEPs in the three LMS subgroups. LMS1 had 30 upregulated DEPs, LMS2 had only three upregulated DEPs (F13A1, PTPN6 and CTSS), and LMS3 had 79 upregulated DEPs (Supplementary Table 4.1). Notably, 20 out of 30 upregulated DEPs in LMS1 were adhesome proteins. In contrast, 1 out of 79 upregulated DEPs in LMS3 was an adhesome protein (chaperone HSPA2).

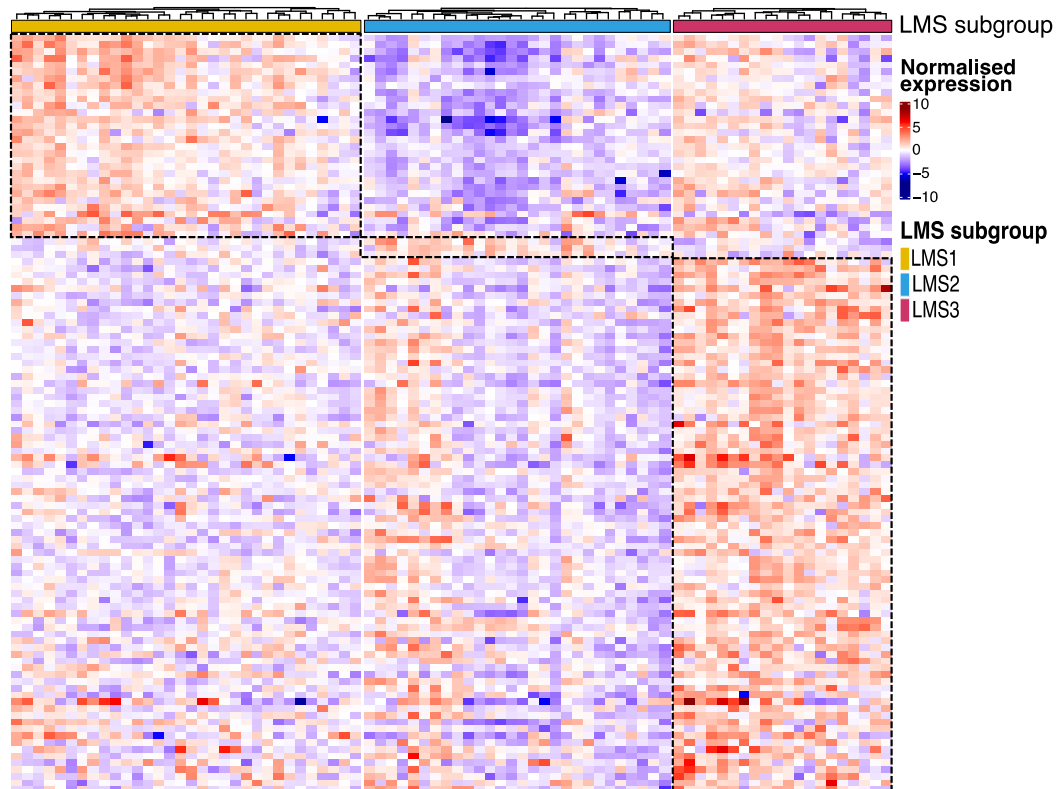


Figure 4.1 Matrisome and adhesome components specific for each leiomyosarcoma (LMS) subgroup.

Heatmap showing the supervised clustering of 112 differentially expressed proteins (DEPs) uniquely upregulated in each LMS subgroup defined by matrisome and adhesome protein expression. Black boxes indicate unique upregulated matrisome and adhesome DEPs in each of the subgroups.

The upregulated DEPs were used to construct PPI interaction networks to evaluate molecular interaction between matrisome and adhesome components in the LMS subgroups (Figure 4.2A-E). STRINGdb was used to construct the PPI based on interaction scores derived from a curated database, experimental evidence and literature searches³⁹⁶. Overrepresentation analysis against the Reactome pathway database was performed for each set of upregulated DEPs, and the 3,262 LMS proteins were used as a background instead of the whole genome. The enrichment analysis for LMS1 DEPs identified the 'ECM-receptor interaction' pathway. Within this pathway, 12 proteins localise to the basement membrane (Figure 4.2A)²⁵⁴. These proteins were glycoproteins (LAMA4/B2, NID1, TINAGL1 and VWA1), proteoglycan HSPG2 and collagen

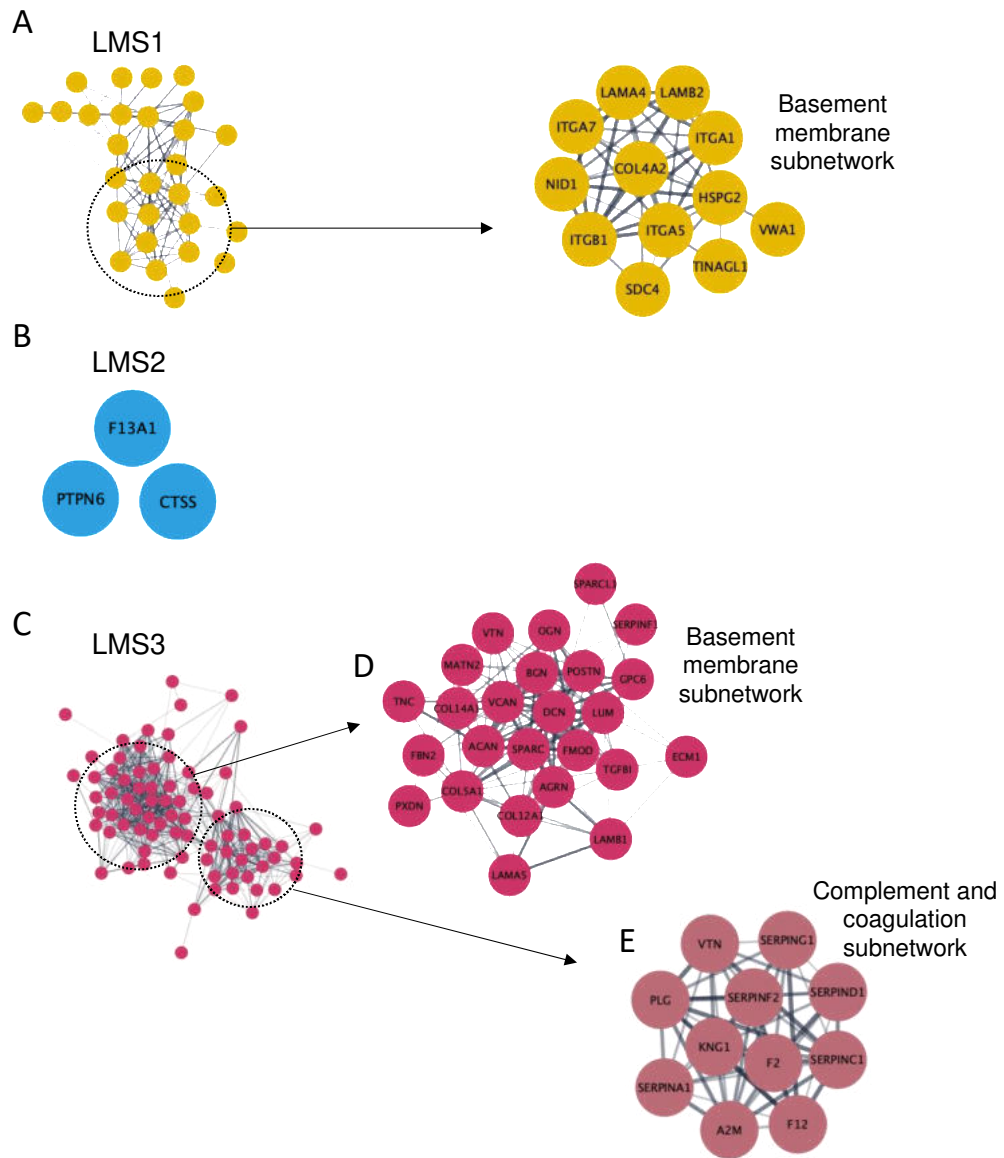


Figure 4.2 Extracellular matrix and integrin adhesion networks in leiomyosarcoma (LMS) subtypes.

A) Protein-protein interaction (PPI) networks of all upregulated DEPs in LMS1. The basement membrane subnetwork is shown on the right. B) Upregulated DEPs in LMS2. C) PPI network of all upregulated DEPs in LMS3. The basement membrane (D) and complement and coagulation (E) subnetworks are shown on the right.

chain COL4A2. The subnetwork additionally included a plasma membrane proteoglycan syndecan-4 (SDC4) and adhesion receptors integrins (ITGA1/A5/A7/B1). Integrins function as heterodimers (e.g., ITGA7/ITGB1 and ITGA1/ITGB1) and each pair interact with specific ECM molecules⁴⁶⁷. The integrins identified in LMS1 are known to bind to the glycoproteins (laminin and collagen IV) detected in this LMS subgroup and to syndecan SDC4.

SDC4, in turn, is known to bind to heparan sulfate chains of HSPG2 and can modulate integrin binding to the ECM⁴⁶⁸. Additionally, LMS1 included several smooth muscle proteins such as muscle and smooth muscle cell actin cytoskeleton components SYNM, SVIL and PARVA^{469–471}, as well as glycoproteins EMILIN1 and TINAGL1^{472,473}, which are found in vascular smooth muscle cells. This suggests that matrisome and adhesome proteins in LMS1 are involved in smooth muscle biology. Overall, LMS1 showed the presence of basement membrane proteins and their respective adhesion receptors.

The three upregulated DEPs from LMS2 did not form a network (Figure 4.2B). The number of upregulated DEPs in LMS2 was too small to perform the overrepresentation analysis. Additionally, the LMS2 subgroup was characterised by the downregulation of basement membrane proteins identified in the LMS1 subgroup (Figure 4.1). As such, LMS2 was considered a subgroup characterised by the lack of basement membrane features.

The enrichment analysis for LMS3 DEPs identified the 'ECM-receptor interaction' pathway and the 'complement and coagulation cascades' pathway. The basement membrane sub-network of LMS3 differed from the one in LMS1 (Figure 4.2D, Supplementary Table 4.1). It included 22 proteins which were all matrisome components, mainly glycoproteins (LAMA5/B1, TNC), collagen chains (COL5A1, COL14A1 and COL12A1) and proteoglycans (ACAN, BGN, DCN, OGN) (Figure 4.2D). The 'complement and coagulation cascade' network consisted of coagulation factors F2 and F12, serine proteases PLG and serine protease inhibitors A2M and SERPINA1/C1/F2 (Figure 4.2E). Additionally, the network contained a complement component inhibitor SERPING1. The glycoprotein vitronectin (VTN) was identified in both 'basement membranes' and 'complement and coagulation cascades' networks in LMS3 (Figure 4.2D, E). Although not identified in the enrichment analysis, fibrillar collagen chains (COL1A1/A2 and COL3A1), collagen biosynthesis (P4HA1 and P4HA2) and collagen remodelling enzymes (LOX, PLOD1 and PLOD2) were significantly upregulated in LMS3, compared to other LMS subgroups (Supplementary Table 4.1).

Characterisation of functional proteomic networks in LMS subgroups

To investigate what other functional pathways outside of matrisome and adhesome operate in the three LMS subgroups, ssGSEA was performed on the full dataset of 3,262 proteins using the Hallmarks gene set³⁹⁴. The average ssGSEA enrichment scores between LMS subgroups were compared to identify differentially expressed biological features in each LMS subgroup. ANOVA followed by post-hoc Tukey's multiple comparisons tests were used to assess for such differences (Figure 4.3).

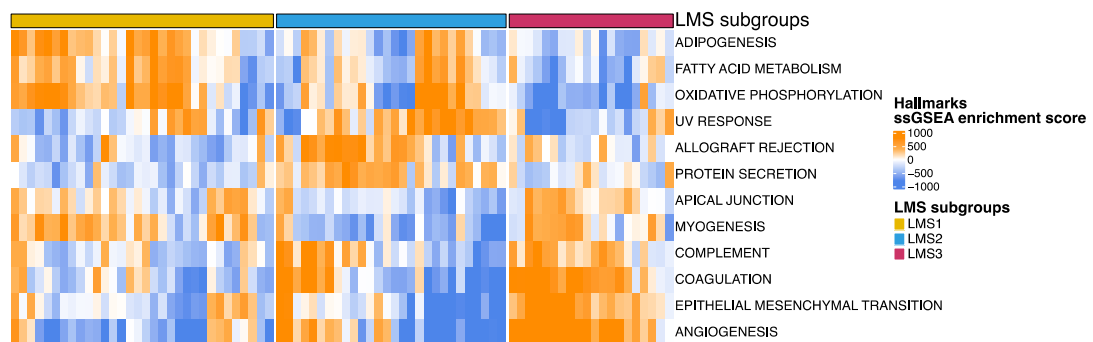


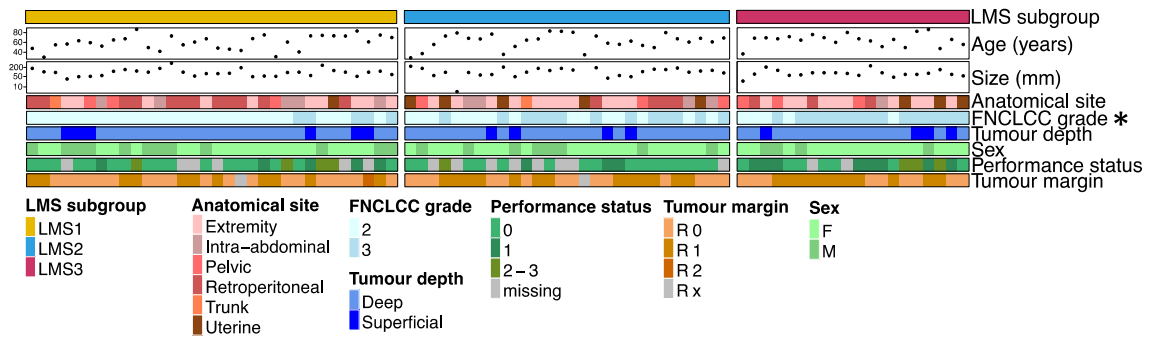
Figure 4.3 Hallmarks in leiomyosarcoma (LMS) subtypes.

Heatmap of significant biological features obtained from single sample Gene Set Enrichment Analysis (ssGSEA) of the MSigDB Hallmark gene sets, arranged by LMS subgroup (top annotation). Significance was tested using one-way ANOVA with Tukey's honestly significant difference (HSD) test; FDR < 0.05.

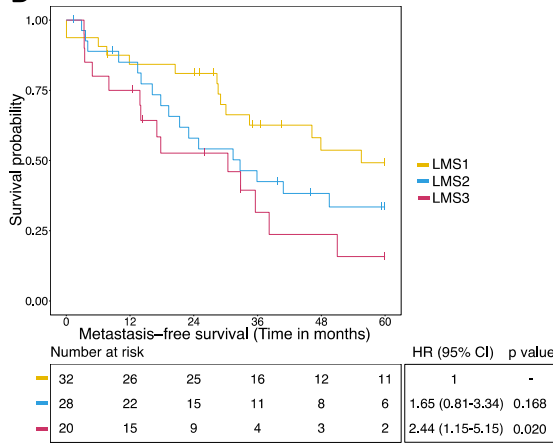
LMS1 showed significant enrichment of 'adipogenesis' compared to LMS2 ($p = 0.0030$) and LMS3 ($p < 0.0001$). The DNA damage pathway 'UV response' (LMS1 $p = 0.0058$; LMS3 $p < 0.0001$), immune pathway 'allograft rejection' (LMS1 $p < 0.0001$; LMS3 $p = 0.023$), and 'protein secretion' (LMS1 $p < 0.0001$; LMS3 $p < 0.0001$) were uniquely enriched in LMS2 compared to both LMS1 and LMS3. Compared to LMS1 and LMS2, 'apical junction' (LMS1 $p = 0.0004$; LMS3 $p < 0.0001$) and 'myogenesis' (LMS1 $p < 0.0001$; LMS3 $p < 0.0001$) were significantly downregulated in LMS2. LMS3 had significant enrichment of immune and development pathways such as 'complement' (LMS1 $p = 0.0008$; LMS2 $p = 0.0490$), 'coagulation' (LMS1 $p < 0.0001$; LMS2 $p < 0.0001$), 'angiogenesis' (LMS1 $p < 0.0001$; LMS2 $p < 0.0001$), and 'epithelial to

mesenchymal transition' pathway (LMS1 $p < 0.0001$; LMS2 $p < 0.0001$) compared to LMS1 and LMS2. In comparison to LMS1 and LMS2, metabolic pathways 'fatty acid metabolism' (LMS1 $p = 0.0005$; LMS2 $p = 0.0496$), 'oxidative phosphorylation' (LMS1 $p < 0.0001$; LMS2 $p = 0.0013$) were significantly downregulated in LMS3.

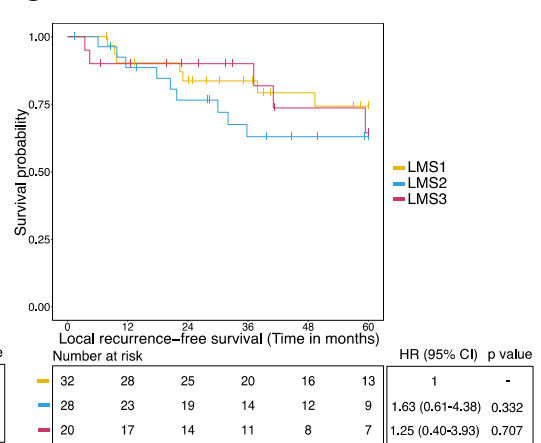
A



B



C



D

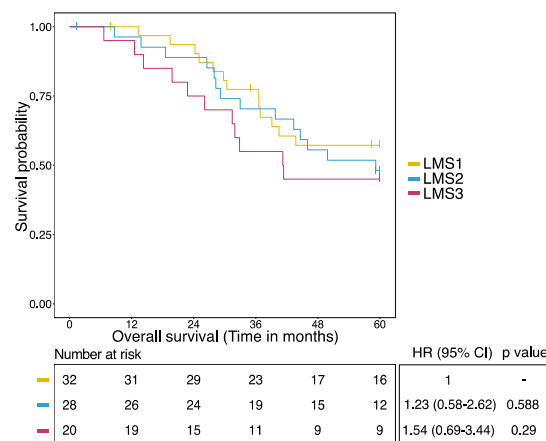


Figure 4.4 Clinical characterisation of leiomyosarcoma (LMS) subgroups

A) Patient and tumour characteristics are summarised and arranged by LMS subgroups.

‘**’ Indicates that a clinical feature is significantly associated with LMS subgroups. Statistical tests are detailed in Supplementary Table 4.2. B) Kaplan-Meier plots of B) metastasis-free survival, C) local recurrence-free survival and D) overall survival with stratification by LMS subgroups. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

Clinical characterisation of LMS subgroups

Matrisome components have been shown to be associated with the tumour stage in other cancer types^{265,474}. Chi-square and Kruskal Wallis tests were performed to assess whether LMS subgroups were associated with any clinical features of LMS. Clinicopathological features for each LMS subgroup were assessed, and the tumour grade was the only clinical feature showing a significant association with the subgroups (Figure 4.4A and Supplementary Table 4.2). LMS1 had a higher proportion of intermediate grade tumours (27 out of 32), LMS2 had an almost equal number of intermediate and high-grade tumours (15 and 13, respectively), whilst LMS3 had a higher proportion of high-grade tumours (15 out of 20) (Chi-squared test: $X^2 = 18.4$, $p = 0.0001$) (Supplementary Table 4.2).

In order to assess the clinical significance of LMS subgroups, survival outcome measures (LRFS, MFS and OS) were statistically tested for associations with the LMS1, LMS2 and LMS3. In univariable analysis, LMS3 was significantly associated with shorter MFS compared to the reference LMS1 (HR = 2.4, 95% CI = 1.15 – 5.15, p-value = 0.02) (Figure 4.4B). To evaluate if LMS3 was an independent prognostic factor in LMS, multivariable Cox regression was undertaken to adjust for grade, anatomical location, tumour size and tumour depth. The association of LMS3 with MFS was not independent of the other prognostic clinical features (HR = 2.2, 95% CI = 0.88 – 5.64, p-value = 0.091) (Supplementary Table 4.3). As expected, grade 3 and extremity/trunk location of the tumour were independent prognostic factors for MFS in the multivariable analysis (Supplementary Table 4.3). No statistically significant association was observed between the LMS subgroups and LRFS or OS in univariable and multivariable analyses (Figure 4.4C-D; Supplementary Table 4.4).

4.2.2. Intra-subtype heterogeneity in DDLPS

4.2.2.1. *Clinicopathological features of DDLPS cohort*

Proteomic and clinical data were available for 39 patients with DDLPS. The clinicopathological features of the cohort are summarised in Table 4.2. The median age at excision was 63 years (range: 35 – 81). Almost all of the tumours occurred in the retroperitoneum (82%). There was an equal split of intermediate (grade 2) and high-grade (grade 3) tumours (49% and 51%, respectively). Ninety-seven percent of tumours were deep-seated and had a large size at presentation (median – 190 mm, range: 35 – 1090 mm). It is challenging to achieve clear margins for large retroperitoneal tumours, and as such, 23% of tumours had R0 margins, whilst 64% had R1 margins. A higher proportion of patients were males (62%), and 74% of patients had a performance status of 0 or 1. The median OS of the cohort was 52 months (95% CI 39 – NA); the median LRFS was 37 months (95% CI 20 – 49), and the median MFS was not reached (Supplementary Figure 4.3). In contrast to LMS, DDLPS are characterised by a high risk for local recurrence and fewer distant metastasis⁵⁸. In our cohort, 62% of patients recurred locally, and 23% had a metastasis event at 60 months post-surgery (Supplementary Figure 4.3).

Table 4.2 Clinicopathological characteristics of n = 39 dedifferentiated liposarcoma (DDLPS) cases.

For continuous variables, median, minimum (min) and maximum (max) values are indicated. For categorical variables, count (n) and percentages (%) are shown. F = female; M = male.

Characteristic	Category	Dedifferentiated Liposarcoma
Number of patients	n (%)	39 (100)
Age at excision (years)	Median	63
	Min	35.1
	Max	81.3
Anatomical site [n (%)]	Extremity	2 (5.1)
	Intra-abdominal	3 (7.7)
	Retroperitoneal	32 (82.1)
	Trunk	2 (5.1)
Grade [n (%)]	2	19 (48.7)
	3	20 (51.3)
Tumour depth [n (%)]	Deep	38 (97.4)
	Superficial	1 (2.6)
Tumour size (mm)	Median	190
	Min	35
	Max	1090
Tumour margins [n (%)]	R0	9 (23.1)
	R1	25 (64.1)
	Rx	5 (12.8)
Pre-op treatment [n (%)]	Chemo	1 (2.6)
	None	38 (97.4)
Performance status [n (%)]	0	17 (43.6)
	1	12 (30.8)
	2	2 (5.1)
	3	1 (2.6)
	unknown	7 (17.9)
Sex [n (%)]	F	15 (38.5)
	M	24 (61.5)
Status at excision [n (%)]	Local	36 (92.3)
	Metastatic	2 (5.1)
	Multifocal	1 (2.6)

4.2.2.2. Identification DDLPS subgroups

In previous studies, molecular heterogeneity of DDLPS was explored by investigating changes in DNA copy number and the patterns of DNA methylation^{35,103}. These approaches identified subgroups of DDLPS patients with different survival outcomes. To identify if matrisome/adhesome proteomic subtypes of DDLPS are present, consensus clustering of the matrisome and adhesome proteomic profiles in our cohort of 39 DDLPS cases was performed.

Consensus matrices at $k = 3$ and 4 showed a clear separation of clusters (Supplementary Figure 4.4A). Consensus CDF for $k = 3$ had a sharp increase around the consensus index of 0 , a horizontal line between 0.3 and 0.7 , ending with a sharp rise close to 1 , indicating stable clustering at $k = 3$ (Supplementary Figure 4.4B). The tracking plot indicated good cluster stability as the samples did not switch between groups frequently (Supplementary Figure 4.4C). SigClust test comparisons between clusters showed that cluster 1 versus cluster 3 and cluster 2 versus cluster 3 comparisons were significant; however, cluster 1 was not significantly different from cluster 2 (Supplementary Figure 4.4D). Based on visual inspection of consensus clustering results and the SigClust, $k = 3$ was chosen as the optimal number of clusters, identifying three DDLPS subgroups DDLPS1, DDLPS2 and DDLPS3.

4.2.2.3. *Biological characterisation of DDLPS subgroups*

Characterisation of ECM and integrin adhesion networks in DDLPS subgroups

To understand the ECM-integrin signalling networks in each DDLPS subgroup, matrisome and adhesome components were inspected for each DDLPS subgroup. Two-class unpaired SAM tests were undertaken to compare each DDLPS subgroup with the rest of the cohort. Only the upregulated matrisome and adhesome DEPs (FDR < 0.01 and fold change ≥ 2) were considered (Figure 4.5A). No upregulated DEPs overlapped between the DDLPS subgroups. DDLPS1 had 9 upregulated DEPs, DDLPS2 had 39, and DDLPS3 had 9 upregulated DEPs. DDLPS1 included glycoproteins (MFAP2/5, FBLN1/2), ECM-affiliated proteins (GPC6, SDC2 and ANXA4), a proteoglycan PRELP and a secreted factor FSTL1 (Figure 4.5B). DDLPS2 had a high number of ECM regulators (SERPIN proteins and coagulation factors), followed by glycoproteins (FGA/B/G, VTN and VWA5A) and ECM-affiliated proteins (Figure 4.5C). DDLPS3 was the only subgroup which contained both adhesome (adhesion receptors ITGB2, PLAUR, SLC3A2 and actin regulator ACTN1) and matrisome (ECM regulators CTSA/B/Z and a glycoprotein TNC) components (Figure 4.5D).

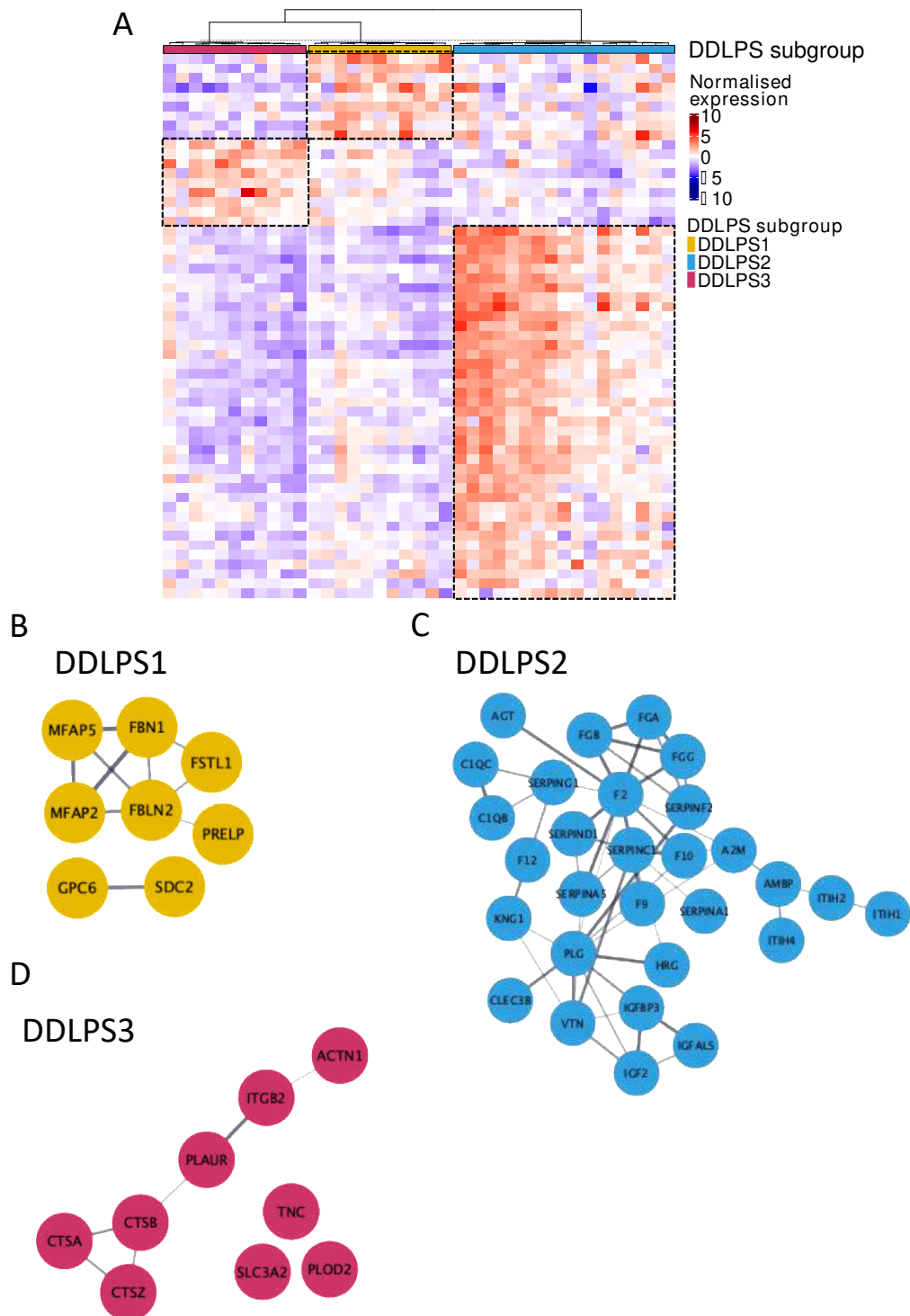


Figure 4.5 Biological characterisation of dedifferentiated liposarcoma (DDLPS) subgroups.

A) Heatmap showing the supervised clustering of 57 differentially expressed matrisome and adhesome proteins (DEPs) uniquely upregulated in each DDLPS subgroup. Black boxes indicate unique upregulated matrisome and adhesome DEPs in each of the subgroups. Protein-protein interaction (PPI) networks of all upregulated DEPs in B) DDLPS1, C) DDLPS2 and D) DDLPS3.

Overrepresentation analysis against the Reactome pathway database was performed for each set of upregulated DEPs, and the 3,304 DDLPS proteins were used as a background instead of the whole genome. The overrepresentation analysis identified enrichment of ‘elastic fibre formation’ in DDLPS1. In DDLPS2 ‘common pathway of fibrin clot formation’ and ‘regulation of complement cascade’ pathways were enriched. DDLPS3 showed enrichment of ‘ECM organisation’ pathway.

Characterisation of functional proteomic networks in DDLPS subgroups

To understand what other functional pathways operate in the three DDLPS subgroups, ssGSEA was performed on the full dataset of 3,304 DDLPS proteins using the Hallmarks gene set³⁹⁴. The average ssGSEA enrichment scores between DDLPS subgroups were compared to identify differentially expressed biological features in DDLPS subgroups. ANOVA followed by post-hoc Tukey’s multiple comparisons tests were used to assess for such differences (Figure 4.6).

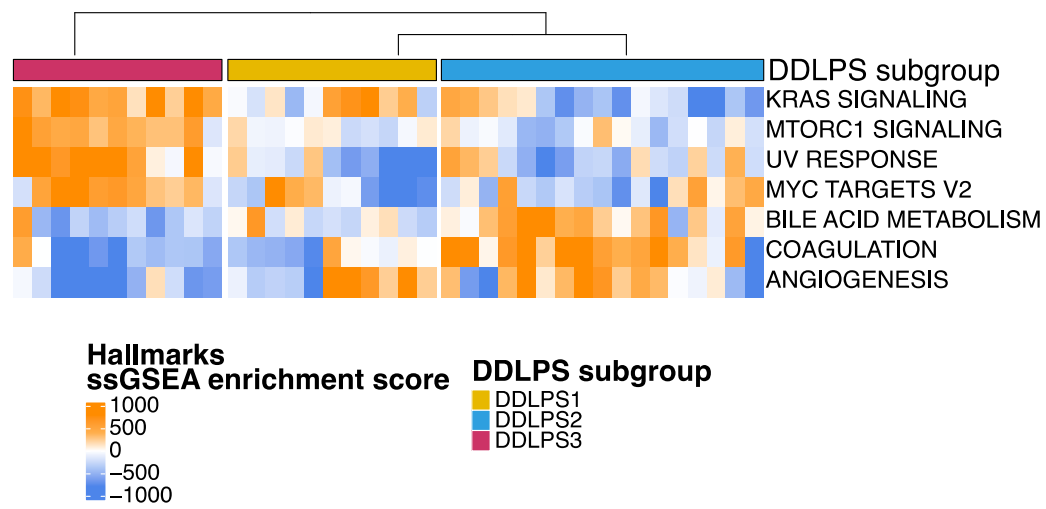


Figure 4.6 Hallmarks in dedifferentiated liposarcoma (DDLPS) subgroups.

Heatmap of significant biological features obtained from single sample Gene Set Enrichment Analysis (ssGSEA) of the MSigDB Hallmark gene sets, arranged by DDLPS subgroups (top annotation). Significance was tested using one-way ANOVA with Tukey’s honestly significant difference (HSD) test; FDR < 0.05.

DDLPS1 did not display significantly upregulated pathways compared to DDLPS2 and DDLPS3. The ‘coagulation pathway’ (DDLPS1 $p = 0.0004$;

DDLPS3 $p < 0.0001$) and 'bile acid metabolism' (DDLPS1 $p = 0.0488$; DDLPS3 $p = 0.0002$) were significantly enriched in DDLPS2, compared to other subgroups. DDLPS3 showed significant upregulation of oncogenic signalling pathways such as 'KRAS signalling' (DDLPS1 $p = 0.049$; DDLPS2 $p < 0.0001$), 'MTORC1 signalling' (DDLPS1 $p = 0.0003$; DDLPS2 $p < 0.0001$) and 'MYC targets v2' (DDLPS1 $p = 0.034$; DDLPS2 $p = 0.013$) compared to both DDLPS1 and DDLPS2. Additionally, the DNA damage pathway 'UV response pathway' was enriched in DDLPS3 compared to both DDLPS1 ($p < 0.0001$) and DDLPS2 ($p < 0.0001$). 'Angiogenesis' was significantly downregulated in DDLPS3, compared to both DDLPS1 ($p = 0.047$) and DDLPS2 ($p = 0.014$).

Overview of tumour infiltrating lymphocytes in DDLPS subgroups

DDLPS is known to be associated with relatively high immune infiltrate compared to other STS^{35,134,425,475}. In our DDLPS cohort, out of 39 cases, IHC staining for CD3+, CD4+ and CD8+ TILs with at least two cores per patient was available for $n = 32$, $n = 32$ and $n = 35$ patients, respectively. DDLPS2 displayed the lowest median TILs count for all tested immune cell markers. In terms of the CD3+ TILs, DDLPS1 had a median count of 112 TILs/mm² (range: 82.8 – 404.7), DDLPS2 had 61.5 TILs/mm² (range: 1.7 – 251.1), and the median in DDLPS3 was 106.8 TILs/mm² (range: 27 – 229.8). All three subgroups had a lower density of the CD4+ TILs compared to the CD3+ and CD8+ TILs. DDLPS1 median CD4+ was 29.7 TILs/mm² (range: 19.1 – 75.4), DDLPS2 had a median of 17.1 TILs/mm² (range: 2.5 – 88), and DDLPS3 median CD4+ was 27.1 TILs/mm² (range: 8.5 – 869). The CD8+ TILs median count for DDLPS1 was 117.4 TILs/mm² (range: 60.8 – 183); for DDLPS2 median count was 43.8 TILs/mm² (range: 60.8 – 183) and for DDLPS3 was 65.6 TILs/mm² (range: 42.4 – 1451.5).

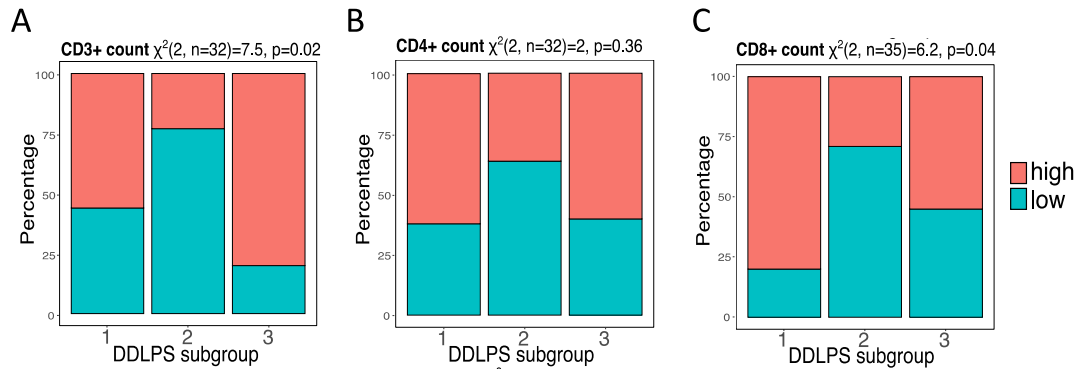


Figure 4.7 Tumour infiltrating lymphocytes (TILs) in dedifferentiated liposarcoma (DDLPS) subgroups.

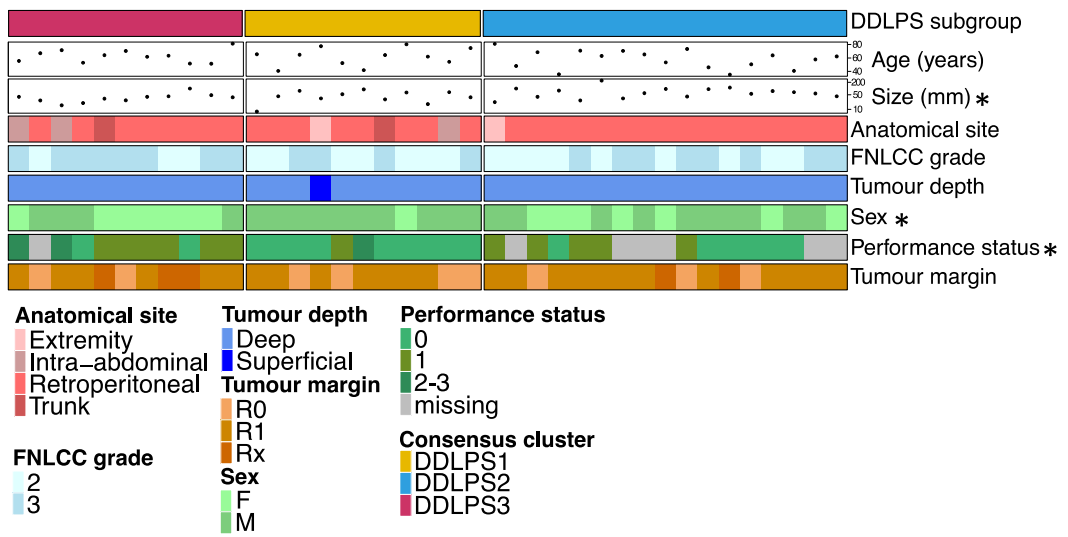
Stacked bar charts showing the percentages of A) CD3+, B) CD4+ and C) CD8+ tumour infiltrating lymphocytes (TILs) in each DDLPS subgroup. DDLPS cases were divided into high and low categories according to the median TILs score, separately for each stain. The chi-squared test results are presented at the top of each plot.

Chi-square tests were performed to explore the association between DDLPS subgroups and TILs. CD3+/4+/8+ scores were dichotomised for statistical analysis, and DDLPS cases were characterised as low and high groups (Figure 4.7A-C). Associations of TILs high and low groups with DDLPS subgroups showed that CD3+ and CD8+ were significantly associated but not the CD4+ TILs. DDLPS2 had a higher percentage of cases with low CD3+ (Chi-squared tests: $X^2 = 7.5$, $p = 0.02$) (Figure 4.7A) and low CD8+ (Chi-squared tests: $X^2 = 6.2$, $p = 0.04$) (Figure 4.7C) than the other subgroups. Similarly, DDLPS2 had a higher percentage of low CD4+ cases, but the comparison was not significant (Figure 4.7B). Overall, DDLPS2 was considered to be an ‘immune-cold’ subgroup.

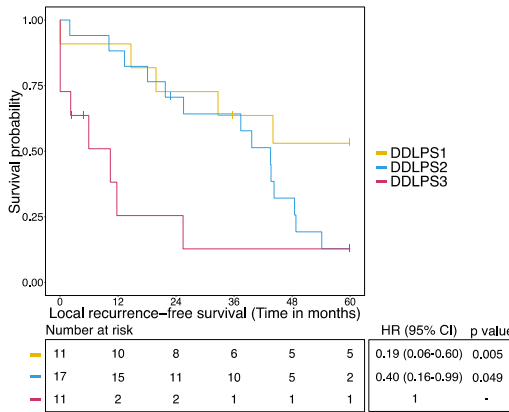
4.2.2.4. *Clinical characterisation of DDLPS subgroups*

Further to the observed biological differences between DDLPS subgroups, the associations with clinicopathological features and survival outcome measures were undertaken. A summary of clinical annotations arranged by DDLPS subgroups is depicted in Figure 4.8A. Chi-square and Kruskal Wallis statistical tests were performed to associate DDLPS subgroups with clinicopathological features (Supplementary Table 4.5).

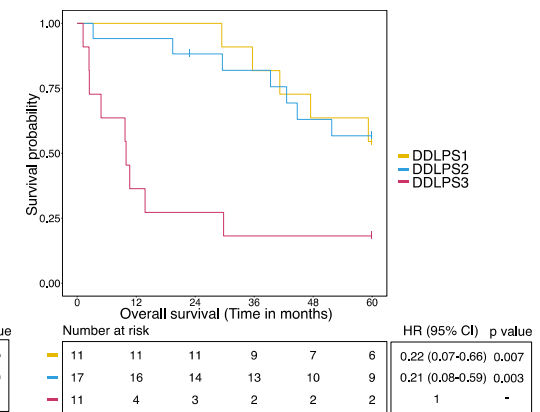
A



B



C



D

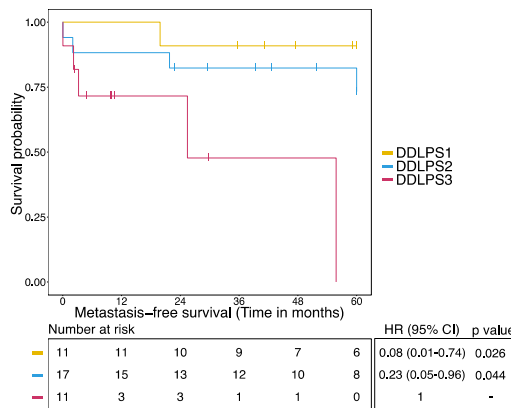


Figure 4.8 Clinicopathological characteristics specific to each dedifferentiated liposarcoma (DDLPS) subgroups and association of the DDLPS subgroups with survival outcomes.

A) Patient and tumour characteristics are summarised and arranged by DDLPS subgroups.

‘**’ Indicates that a clinical feature is significantly associated with DDLPS subgroups.

Statistical tests are detailed in Supplementary Table 4.8. B) Kaplan-Meier plots of B) local recurrence-free survival, C) overall survival and D) metastasis-free survival with stratification

by DDLPS matrixome and adhesome clusters. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

DDLPS subgroups were significantly associated with tumour size (Kruskal Wallis: $X^2 = 6.5$, $p = 0.038$), with DDLPS2 having larger tumours than DDLPS1 and DDLPS3. Moreover, performance status (Chi-squared tests: $X^2 = 17.5$, $p = 0.008$) was significantly associated with DDLPS subgroups, with most patients in DDLPS1 having a performance status 0. Additionally, almost all patients in DDLPS1 were male (10/11), and therefore patient sex (Chi-squared tests: $X^2 = 7$, $p = 0.03$) was significantly associated with the subgroups. In contrast to the LMS cohort, tumour grade was not associated with the DDLPS subgroups.

Five-year survival outcomes (LRFS, OS and MFS) were evaluated in DDLPS subgroups. The DDLPS3 subgroup was consistently the worst-performing group for all the mentioned survival outcome measures (Figure 4.8B-D). DDLPS3 had shorter LRFS compared to DDLPS1 (HR = 0.19, 95% CI = 0.06 – 0.60, p-value = 0.005) and DDLPS2 (HR = 0.40, 95% CI = 0.16 – 0.99, p-value = 0.049) (Figure 4.8B). Shorter LRFS translated into a significantly shorter OS in DDLPS3 compared to DDLPS1 (HR = 0.22, 95% CI = 0.07 – 0.66, p-value = 0.007) and DDLPS2 (HR = 0.21, 95% CI = 0.08 – 0.59, p-value = 0.003) (Figure 4.8C). Additionally, DDLPS3 described the group with a significantly shorter MFS compared to DDLPS1 (HR = 0.08, 95% CI = 0.01 – 0.74, p-value = 0.026) and DDLPS2 (HR = 0.23, 95% CI = 0.05 – 0.96, p-value = 0.044) (Figure 4.8D). Moreover, after adjusting for age, tumour grade, tumour size, patient sex and performance status, DDLPS3 remained an independent prognostic factor for LRFS compared to both DDLPS1 (HR = 0.11, 95% CI = 0.02 – 0.56, p-value = 0.008) and DDLPS2 (HR = 0.28, 95% CI = 0.08 – 0.98, p-value = 0.046) in multivariable analysis (Supplementary Tables 4.6). In addition, in the multivariable analysis, DDLPS3 remained prognostic for OS compared to DDLPS1 (HR = 0.15, 95% CI = 0.03 – 0.69, p-value = 0.015) and DDLPS2 (HR = 0.13, 95% CI = 0.03 – 0.53, p-value = 0.005) (Supplementary Tables 4.7). Following multivariable Cox analysis,

DDLPS3 lost prognostic significance for MFS compared to DDLPS2 and DDLPS1 (Supplementary Tables 4.8).

4.2.3. Intra-subtype heterogeneity in UPS

4.2.3.1. Cohort characteristics

Proteomic and clinical data were available for 53 UPS patients. The median age of the cohort was 73.5 years (range: 28.2 – 90), and both sexes were equally represented (Table 4.3). Most tumours occurred in the extremities (72%), were deep-seated (81%), of large size (median – 80mm; range: 15-360 mm), and almost all tumours were of high grade (93%), which is consistent with other UPS clinical reports^{127,128,476}. The median LRFS was not achieved in the cohort. The median MFS and OS were 34 months (95% CI 19 – NA) and 37 months (95% CI 26 – NA), respectively (Supplementary Figure 4.5).

Table 4.3 Clinicopathological characteristics of n = 53 undifferentiated pleomorphic sarcoma (UPS) cases.

For continuous variables, median, minimum (min) and maximum (max) values are indicated. Count (n) and percentages (%) are shown for categorical variables. F = female; M = male.

Characteristic	Category	Undifferentiated Pleomorphic Sarcoma
Number of patients	n (%)	53 (100)
Age at excision (years)	Median	73.5
	Min	28.2
	Max	90
Anatomical site [n (%)]	Extremity	38 (71.7)
	Head/neck	4 (7.5)
	Intra-abdominal	1 (1.9)
	Pelvic	2 (3.8)
	Trunk	8 (15.1)
Grade [n (%)]	2	3 (5.7)
	3	49 (92.5)
	unknown	1 (1.9)
Tumour depth [n (%)]	Deep	43 (81.1)
	Superficial	10 (18.9)
Tumour size (mm)	Median	80
	Min	15
	Max	360
Tumour margins [n (%)]	R0	26 (49.1)
	R1	27 (50.9)
Performance status [n (%)]	0	22 (41.5)
	1	15 (28.3)
	2	4 (7.5)
	3	3 (5.7)
	unknown	9 (17)
Sex [n (%)]	F	28 (52.8)
	M	25 (47.2)

4.2.3.2. Identification of UPS subgroups

Consensus clustering was applied to matrixome and adhesome proteomic profiles in our cohort of 53 UPS. Consensus matrices showed clear separation of clusters up to $k = 5$ (Supplementary Figure 4.6A). Consensus CDF for $k = 2$ had a sharp increase between the consensus index of 0 and 0.4, a horizontal line between 0.4 and 0.6, ending with a sharp increase close to 1. Consensus CDF for $k = 3$ showed an increase in the area under the curve compared to $k = 2$, and followed a similar trend as the CDF at $k = 2$, indicating stable clustering at $k = 2$ and $k = 3$ (Supplementary Figure 4.6B). All the pair-wise comparisons between the clusters using SigClust tests were significant for $k = 2$ and $k = 3$. At $k = 4$, cluster 1 versus cluster 3 pair was the only comparison which failed to achieve significance (Supplementary Figure 4.6D). Although $k = 2$ or $k = 3$ are optimal clusters based on CC analysis, the downstream

analyses did not show an association of clusters at $k = 2$ and $k = 3$ with survival outcome measures. As we specifically set out to characterise matrisome and adhesome in the context of clinical applications, $k = 4$ was selected.

4.2.3.3. *Biological characterisation of UPS subgroups*

Characterisation of ECM and integrin adhesion networks in the UPS subgroups

Two-class unpaired SAM tests were performed to characterise the matrix signalling in the 4 identified UPS subgroups (UPS1, UPS2, UPS3 and UPS4). Comparisons of matrisome and adhesome proteins within each UPS subgroup with the rest of the cohort were undertaken to identify significant DEPs. Only the upregulated matrisome and adhesome DEPs with $FDR < 0.01$ and fold change ≥ 2 were considered, and DEPs overlapping between the subgroups were removed from the analysis. Across the 4 UPS subgroups, 95 upregulated DEPs were identified and arranged by the UPS subgroup (Figure 4.9). The UPS1 subgroup had 33 uniquely upregulated DEPs, which were fibrillar collagen chains (COL1A1/A2, COL3A1, COL5A1 and COL11A1), FACIT collagen chains (COL14A1 and COL16A1), which associate with fibrillar collagens. Additionally, upregulated DEPs in this UPS subgroup included 11 glycoproteins (SPARC, FBLN1, MATN2 and others), 7 ECM regulators (MMP2/14, PLOD1/2, TIMP2, P4HA1 and SERPINH1), three proteoglycans (ASPN, OGN and ACAN), two integrin chains (ITGAV and ITGB5), an intracellular adaptor PALLD and an ECM-affiliated protein PLXDC2. The UPS2 and UPS3 subgroups lacked any DEPs which were significantly enriched. The UPS4 subgroup contained 62 uniquely upregulated

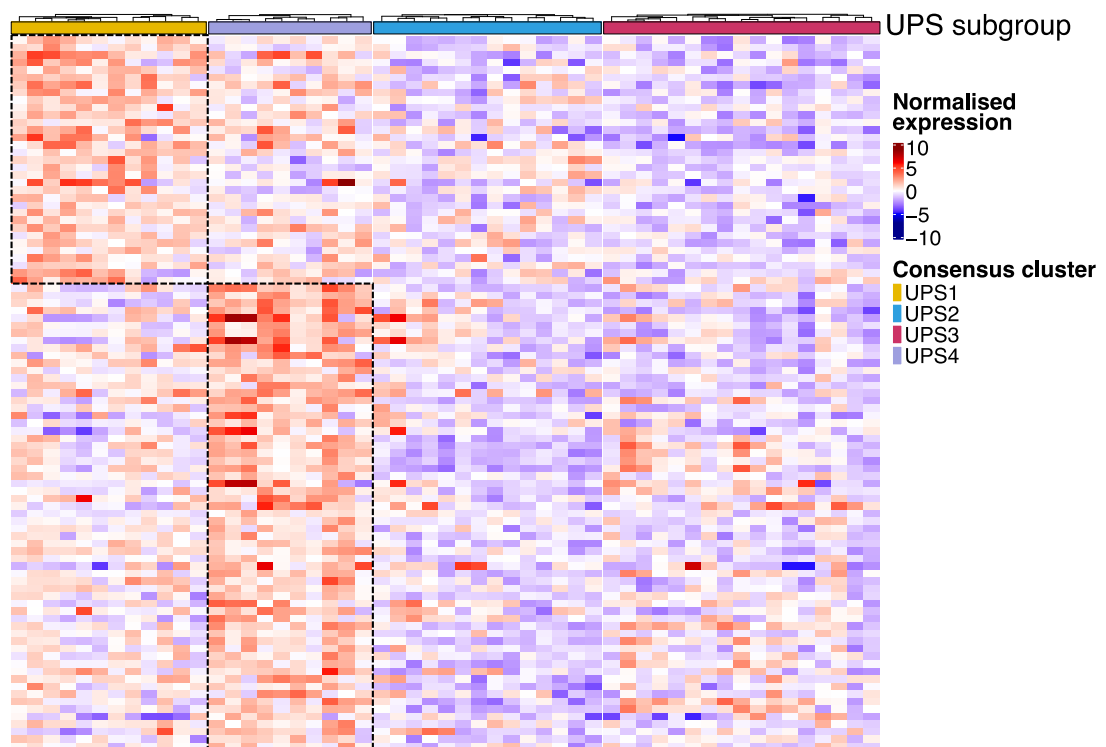


Figure 4.9 Identification of undifferentiated pleomorphic sarcoma (UPS) subgroups based on matrisome and adhesome proteins expression.

Heatmap showing the supervised clustering of 95 differentially expressed proteins (DEPs) uniquely upregulated in each UPS subgroup. Black boxes indicate unique upregulated matrisome and adhesome DEPs in UPS1 and UPS4.

DEPs. Those included non-fibrillar collagen chains (COL4A1/A2 and COL6A1/A2/A3), which constitute the basement membrane, 25 glycoproteins (including the basement membrane components LAMA4, LAMC1, NID1/2), 16 ECM regulators (serpin family serine proteases inhibitors), five proteoglycans (PRELP, LUM, FMOD, HAPLN3 and DCN), three ECM-affiliated proteins (CLEC11A, CLEC3B and ANXA3) and two secreted factors S100A8/9. The overrepresentation analysis against the Reactome database was performed for each set of upregulated DEPs, using 3,355 UPS proteins as a background.

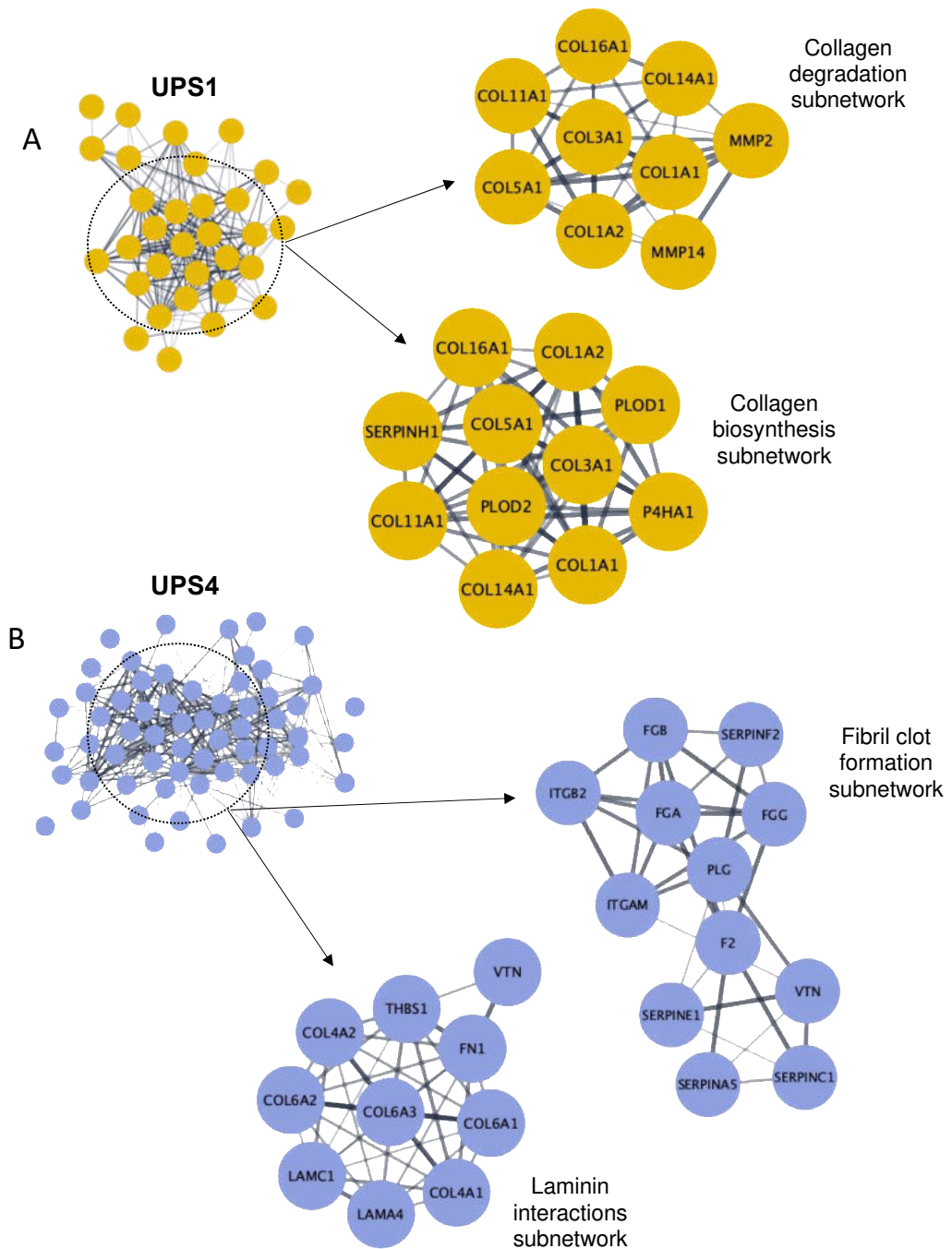


Figure 4.10 Matrisome and adhesome components specific for each consensus clusters in undifferentiated pleomorphic sarcoma (UPS).

A) Protein-protein interaction (PPI) network of all upregulated DEPs in the UPS1 subgroup, with collagen degradation subnetwork (top) and collagen biosynthesis (bottom) subnetwork.

B) PPI network of all upregulated DEPs in the UPS4 subgroup, with fibrin clot formation subnetwork (top) and laminin interactions (bottom) subnetwork.

The overrepresentation analysis showed the upregulation of pathways related to 'collagen chain trimerisation', 'collagen degradation' and 'collagen biosynthesis and modifying enzymes' as the top significantly enriched pathways in the UPS1 subgroup. Figure 4.10A shows the 'collagen degradation' subnetwork with collagen-degrading proteinases (MMP2 and MMP14) and the 'collagen biosynthesis and modifying enzymes' subnetwork with collagen biosynthesis enzymes P4HA1, PLOD1 and PLOD2, identified in the UPS1 patient subgroup. Enriched pathways in the UPS4 subgroup included the 'common pathway of fibrin clot formation', 'laminin interactions' and 'integrin cell surface interactions' (Figure 4.10B).

Overview of tumour infiltrating lymphocytes in the UPS subgroups

UPS is considered one of the STS subtypes with the highest immune infiltrate^{133,466}. Moreover, high TILs count correlated with improved survival and a higher response rate to immunotherapy with checkpoint inhibitors than sarcoma patients with low TILs^{108,137,477}. In our UPS cohort, the IHC staining for T cell markers (CD3+, CD4+ and CD8+) was performed on at least two cores per patient. Of the 53 UPS cases, IHC data for CD3+, CD4+ and CD8+ stains were available for n = 50, n = 47 and n = 50 patients, respectively. To determine the association between UPS patient subgroups and TILs, CD3+/4+/8+ TIL scores were dichotomised, and all UPS cases were characterised as either low or high for each TILs group. Then the chi-square tests assessed the relationship between UPS patient subgroups and TILs groups (Figure 4.11).

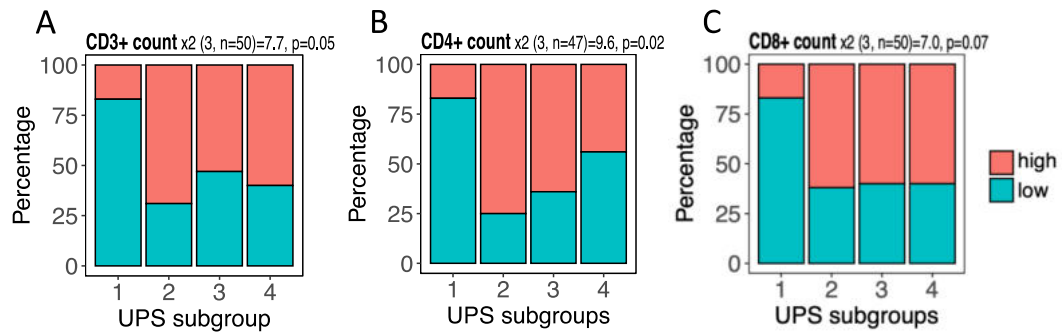


Figure 4.11 Tumour infiltrating lymphocytes (TILs) in undifferentiated pleomorphic sarcoma (UPS) subgroups.

Stacked bar charts showing the percentages of A) CD3+, B) CD4+ and C) CD8+ tumour infiltrating lymphocytes (TILs) in each UPS subgroup. UPS cases were divided into high and low categories according to the median TILs score, separately for each stain. The chi-squared test results are presented at the top of each plot.

A relationship between TILs groups and UPS subgroups showed that CD3+ and CD4+ were significantly associated with the subgroups but not the CD8+ TILs. The UPS1 subgroup had a lower percentage of cases with high CD3+ (Chi-squared tests: $X^2 = 7.7$, $p = 0.05$) and a lower percentage of high CD4+ (Chi-squared tests: $X^2 = 9.6$, $p = 0.02$) than the other subgroups (Figure 4.11A, B). Additionally, the UPS1 subgroup had the lowest percentage of high CD8+ cases (not significant) (Figure 4.11C). Overall, the UPS1 subgroup was considered an ‘immune cold’.

To investigate what other functional pathways operate in the four UPS subgroups, ssGSEA was performed on the full dataset of 3,355 UPS proteins using the Hallmarks gene set. The average ssGSEA enrichment scores between UPS subgroups were compared to identify differentially expressed biological features in UPS subgroups. ANOVA followed by post-hoc Tukey’s multiple comparisons tests were used to assess for such differences.

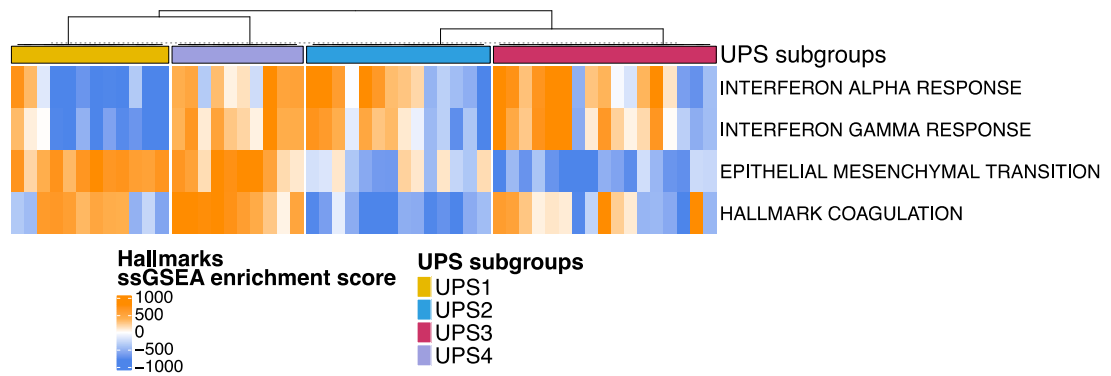


Figure 4.12 Hallmarks in undifferentiated pleomorphic sarcoma (UPS) subgroups.

Heatmap of significant biological features obtained from single sample Gene Set Enrichment Analysis (ssGSEA) of the MSigDB Hallmark gene sets, arranged by UPS subgroups (top annotation). Significance was tested using one-way ANOVA with Tukey's honestly significant difference (HSD) test; FDR < 0.05.

Consistent with being 'immune-cold', the UPS1 subgroup showed downregulation of immune-related pathways (Figure 4.12). 'Interferon alpha response' was downregulated in the UPS1 compared to UPS2 ($p = 0.037$), UPS3 ($p = 0.006$) and UPS4 ($p = 0.015$). Similarly, 'interferon gamma response' was downregulated in the UPS1 compared to UPS2 ($p = 0.039$), UPS3 ($p = 0.001$) and UPS4 ($p = 0.0004$).

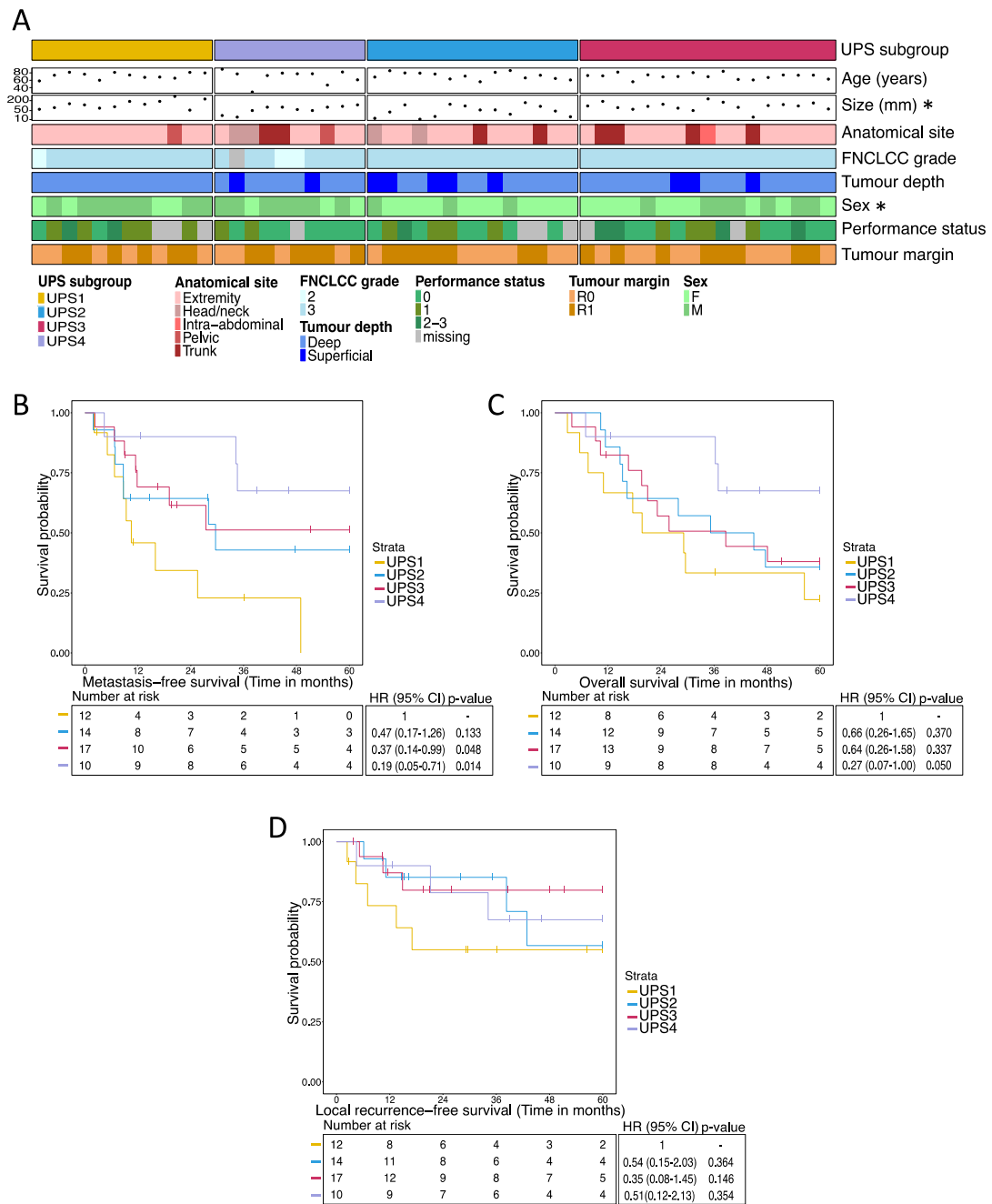


Figure 4.13 Clinicopathological characteristics specific to each undifferentiated pleomorphic sarcoma (UPS) subgroup and association of the subgroups with survival outcomes.

A) Patient and tumour characteristics are summarised and arranged by UPS subgroups.

“**” Indicates that a clinical feature is significantly associated with UPS subgroups. Statistical tests are detailed in Supplementary Table 4.1. B) Kaplan-Meier plots of B) metastasis-free survival, C) overall survival and D) local recurrence-free survival with stratification by UPS subgroups. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

Clinical characterisation of the UPS subgroups

Chi-square and Kruskal Wallis tests were performed to test for the association of UPS subgroups with clinicopathological features (Supplementary Table 4.9, Figure 4.13A). Sex was found to be significantly associated (Chi-square: $X^2 = 7.9$, $p = 0.048$), and the UPS2 and UPS3 subgroups had a higher proportion of females, whilst the UPS1 and UPS4 had a higher proportion of male patients. Tumour size was another clinicopathological feature significantly associated with the UPS subgroups (Kruskal Wallis: $X^2 = 11.4$, $p = 0.01$), with the UPS1 subgroup having the largest tumours.

In order to assess the clinical significance of UPS subgroups, survival outcome measures (LRFS, MFS and OS) were statistically tested for associations with the UPS subgroups (Figure 4.13B-D). In univariable analysis, UPS1 had a significantly worse outcome in terms of MFS compared to UPS3 (HR = 0.37, 95% CI = 0.14 – 0.99, p -value = 0.048) and UPS4 (HR = 0.19, 95% CI = 0.05 – 0.71, p -value = 0.014) (Figure 4.13B). Additionally, the UPS1 subgroup displayed shorter OS compared to UPS4 (HR = 0.27, 95% CI = 0.07 – 1, p -value = 0.05) (Figure 4.13C). To test if the UPS1 subgroup is an independent prognostic factor for survival, multivariable Cox regression analysis was performed, adjusting for age, anatomical site, tumour size, patient sex and performance status (Supplementary Tables 4.10-4.12). The UPS1 remained prognostic of MFS compared to UPS4 (HR = 0.18, 95% CI = 0.04 – 0.84, p -value = 0.029).

4.3. Discussion

Chapter 4 investigated the intra-subtype heterogeneity of matrix signalling in LMS, DDLPS and UPS. In each subtype, robust patient subgroups were identified, which showed distinct matrix signalling, and some were correlated with clinicopathological features and survival outcomes (Table 4.4).

Table 4.4 Summary of Chapter 4 analysis.

Table shows the number of subgroups identified in LMS, DDLPS and UPS and significant associations with clinicopathological features and survival outcome measures. Y: yes (statistically significant association present); UVA: univariable analysis; MVA: multivariable analysis; LRFS: local recurrence-free survival; MFS: metastasis-free survival; OS: overall survival; TILs: tumour infiltrating lymphocytes.

	LMS subgroups	DDLPS subgroups	UPS subgroups
Number of subgroups	3	3	4
Grade	Y	-	-
Size	-	Y	Y
Sex	-	Y	Y
Performance status	-	Y	-
Significant UVA for LRFS	-	DDLPS3 worse than DDLPS1 and DDLPS2	-
Significant UVA for MFS	LMS3 worse than LMS1	DDLPS3 worse than DDLPS1 and DDLPS2	UPS1 worse than UPS3 and UPS4
Significant UVA for OS	-	DDLPS3 worse than DDLPS1 and DDLPS2	UPS1 worse than UPS4
Significant MVA for LRFS	-	DDLPS3 worse than DDLPS1 and DDLPS2	-
Significant MVA for MFS	-	-	UPS1 worse than UPS4
Significant MVA for OS	-	DDLPS3 worse than DDLPS1 and DDLPS2	-
TILs infiltrate	-	DDLPS2 had lowest TILs levels Higher CD8+ TILs in DDLPS1	UPS1 had lowest TILs levels

4.3.1. Matrix signalling heterogeneity in LMS

Consensus clustering analysis of the matrixome and adhesome proteomic profiles classified the LMS cohort into three robust subgroups. The LMS1 subgroup showed overexpression of smooth muscle-specific intracellular adaptors, which regulate the actin cytoskeleton. Additionally, LMS1 had enrichment in the myogenesis pathway compared to LMS2. This is in line with the Beck *et al.* study, which reported upregulated expression of actin cytoskeleton genes in a smooth-muscle enriched LMS molecular group¹⁴³⁵. Additionally, fundamental basement membrane components such as nidogen, laminins, nonfibrillar collagen type IV chains, HSPG2²⁵⁵ and their corresponding integrin subunits (ITGA1/7/5 and ITGB1) were upregulated in LMS1 compared to other subgroups. In my analysis, the LMS1 subgroup was enriched in adipogenesis and metabolism pathways ‘fatty acid metabolism’ and ‘oxidative phosphorylation’, consistent with studies reporting transcriptomic LMS subtypes enriched in metabolic activity^{82,435,436}. Additionally, LMS1 was enriched in lower-grade tumours compared to other LMS subgroups, and LMS1 patients had the longest MFS in our LMS cohort. It would be interesting to further investigate whether the enrichment of

fundamental basement membrane components and metabolism-related pathways underpins the biology of tumour grade in LMS.

LMS3 showed higher expression of glycoproteins, ECM remodelling enzymes, proteoglycans and fibrillar collagens compared to LMS1 and LMS2 subgroups and a loss of adhesome components. Some of these upregulated proteins agree with the previously reported Beck *et al.* LMS group III, which showed upregulated expression of collagen chains and ECM remodelling enzymes⁴³⁵. An ECM proteoglycan versican (VCAN) was enriched in the LMS3 subgroup compared to LMS1 and LMS2 in my analysis. A previously published study investigated the expression and the role of VCAN in LMS⁴⁷⁸. Clinical samples of n = 80 LMS showed higher VCAN mRNA expression levels than n = 24 benign leiomyomas. The same study showed that the knockdown of VCAN using small interfering RNA (siRNA) in LMS cell line SK-LMS-1 reduced cell growth and migration *in vitro* and tumour formation *in vivo*⁴⁷⁸. It highlights the importance of VCAN in the biology of LMS and suggests it may serve as a potential therapeutic target. In ovarian biopsy samples, a higher disease score was associated with the increased mRNA and protein expression of glycoproteins, proteoglycans and ECM remodelling enzymes compared to biopsies with low disease score²⁶⁵. This is in line with the findings in this thesis since the LMS3 subgroup was significantly associated with a higher grade and also showed increased number and protein expression of glycoproteins, proteoglycans, ECM regulators and collagen remodelling enzymes compared to LMS1 and LMS2 groups. Moreover, the LMS3 subgroup overexpressed the collagen biosynthesis enzymes (P4HA1 and P4HA2) and collagen maturation and crosslinking enzymes (PLOD1, PLOD2, LOX and PXDN), suggesting increased collagen deposition and stabilisation in this patient subgroup. PLOD enzymes are responsible for catalysing the hydroxylation of specific lysine residues within the collagen protein chains⁴⁷⁹. Hydroxylysine residues are essential for forming collagen crosslinks by the LOX enzyme⁴⁸⁰. The crosslinks stabilise collagen fibrils and contribute to increased stiffness of the matrix⁴⁸¹. Elevated PLOD1, PLOD2 and LOX expression are associated with increased cancer cell migration, invasion and metastasis in various cancers, including sarcoma^{482–484}. LOX family enzymes have been proposed as promising

targets for drug intervention to prevent breast cancer metastasis^{485,486}. One of the promising LOX inhibitors is a copper chelator tetrathiomolybdate (TM)⁴⁸⁷. Copper is an essential cofactor for the enzymatic activity of LOX enzymes, and sequestering copper by TM indirectly inhibits LOX activity. In mouse xenograft models of triple-negative breast cancer, it was observed that TM treatment inhibited the development of lung metastases⁴⁸⁸. Lung tissue from these models had reduced LOX levels and collagen cross-linking and fibrillar length, suggesting a role of collagen remodelling in TM-mediated metastases prevention⁴⁸⁸. A phase II study of TM is currently ongoing in patients with breast cancer at a moderate to high risk of recurrence (NCT00195091). Given that the LMS3 subgroup was significantly associated with shorter MFS in the univariable analysis compared to the reference LMS1, increased collagen deposition and crosslinking might contribute to the enhanced propensity for metastasis in this subgroup of patients. Taken together, increased expression of glycoproteins, proteoglycans, fibrillar collagens and collagen remodelling enzymes could contribute to an aggressive phenotype in LMS3. Inhibiting collagen crosslinking might be a promising treatment strategy for the LMS3 subgroup.

In addition to differential matrix signalling, the LMS3 subgroup had the enrichment of several hallmark pathways (complement, coagulation, EMT and angiogenesis) compared to LMS1 and LMS2, which might explain the poor MFS in the LMS3 subgroup. The majority of STS, including LMS, primarily spread through the vascular system^{489,490}. Considering this, enhanced angiogenesis at the primary site (observed in LMS3 patients) may contribute to the increased propensity for metastatic spread via newly formed blood vessels. The LMS3 subgroup displayed upregulation of the EMT pathway. Upregulation of EMT is associated with a more aggressive phenotype and an increased propensity for metastasis in epithelial cancers⁴⁹¹. Nevertheless, the EMT involvement in STS is unclear, considering the paradoxical nature of EMT in a mesenchymal tumour⁴⁹². EMT is characterised by specific molecular changes that drive the transition from an epithelial to a mesenchymal cell phenotype. Those involve loss of the epithelial cell-cell adhesion molecule E-cadherin and the upregulation of mesenchymal markers such as vimentin,

Twist-related protein 1 (TWIST-1), Snail Family Transcriptional Repressor 1 (SNAIL) and Zinc finger protein SNAI2 (SLUG)^{493,494}. Sarcomas are mesenchymal malignancies and, by extension, are considered to lack EMT ability. Nevertheless, certain STS subtypes, such as biphasic sarcoma, a subset of LMS and epithelioid sarcoma, were shown to express epithelial markers such as E-cadherin^{495–497}. Recently, it has been proposed that EMT is not a binary process where cells switch abruptly from epithelial to mesenchymal states⁴⁹⁸. Rather, EMT can be activated to various extents and is a reversible process, which can result in the existence of a spectrum of intermediate states between fully epithelial and fully mesenchymal phenotypes. In line with the reversible notion of EMT, Mesenchymal to Epithelial Reverting Transition (MET) has been defined in LMS with the loss of vimentin and increased expression of E-cadherin via SLUG regulation⁴⁹⁶. This resulted in reduced cell proliferation, migration and invasion in SK-LMS-1 LMS cell line. Moreover, high IHC staining of EMT mesenchymal marker vimentin was shown to be an independent prognostic factor for worse OS in LMS and strong vimentin staining correlated with stage III-IV and recurrence or metastasis occurrence⁴⁹⁹. Therefore, whilst LMS might not undergo the full classical EMT, it is possible that EMT/MET-like processes might play a role in the aggressive nature of LMS and other STS⁴⁹².

In the multivariable analysis, after adjusting for known prognostic factors in LMS, the LMS3 subgroup was not prognostic for MFS. As expected, higher tumour grade was the primary factor prognostic for MFS after the multivariable analysis. Although the molecular classification of LMS, based on matrisome and adhesome protein expression profiles, was not prognostic in LMS, Chapter 4 uncovered novel LMS pathobiology related to the ECM and integrin adhesion signalling. The identified LMS subgroups had differential expression of basement membrane components and the subgroups were associated with tumour grade, suggesting that remodelling of the basement membrane might be may be linked to the biology underpinning LMS tumour grade. As such, the basement membrane composition in LMS might be capturing a different biological aspect of the tumour grade beyond the traditional FNCLLC grading system. Further work could investigate if the assessment of basement

membrane proteins in LMS could refine the accuracy of the FNCLLC grading for LMS.

4.3.2. Matrix signalling heterogeneity in DDLPS

Three robust patient subgroups were identified in DDLPS based on matrixome and adhesome protein expression profiles. The subgroups showed differences in matrix signalling components, biological pathways, and T cell infiltrate. Additionally, the subgroups showed a significant association with survival outcomes independent of other known clinicopathological prognostic factors in DDLPS.

The DDLPS1 subgroup had the longest LRFS, MFS and OS compared to other subgroups. The DDLPS1 subgroup showed upregulation of specific glycoproteins (FBN1, FBLN2, MFAP2 and MFAP5), which modulate the assembly of elastic fibres. Elastic fibres provide resilience and elasticity to tissues subjected to repetitive stretching, such as blood vessels, lungs, and skin⁵⁰⁰. Elastic fibres are present in adipose tissue, where they help the tissue to expand and contract as fat storage and release occur^{501,502}. Fibrillin-1 (FBN1) is one of the glycoproteins upregulated in the DDLPS1 subgroup that forms a microfibrillar scaffold around elastin fibres⁵⁰³. Recently, FBN1 was shown to maintain adipose tissue homeostasis by inhibiting excessive adipogenic differentiation⁵⁰⁴. FBN1 acted by sequestering insulin and negatively regulating the PI3K/AKT/mTOR cascade. This aligns with the downregulation of mTORC1 and other oncogenic pathways in the DDLPS1 patient subgroup. Considering the positive outlook of the DDLPS1 patients in our cohort and the absence of other upregulated cancer hallmark processes, elastic fibre glycoproteins might act to restrain DDLPS progression.

Additionally, the DDLPS1 subgroup had elevated CD8+ TILs levels compared to DDLPS2 and DDLPS3. Compared to other liposarcoma subtypes and some other STS, DDLPS often harbour elevated levels of T cells^{134,425,475}. In DDLPS, high CD4+ infiltrate has previously been shown to be associated with improved survival⁴²⁶. However, studies investigating the association of CD8+ TILs with

survival in DDLPS are lacking. Oike *et al.* found no association of CD8+ TILs levels with PFS in DDLPS⁵⁰⁵; however, the cohort was small (n = 17) and further studies are required to confirm if CD8+ TILs are associated with survival outcome measures in DDLPS.

The DDLPS2 subgroup showed upregulation of matrixome proteins involved in coagulation and complement pathways. As part of the innate immune system, the complement system provides a nonspecific immune response⁵⁰⁶. Notably, a significant cross-talk exists between coagulation and complement pathways⁵⁰⁷. Although the complement system is part of the immune response, it has been shown to play an immunosuppressive role in cancer⁵⁰⁸. In melanoma and lung cancer, activation of complement was reported to suppress CD4+ and CD8+ TILs^{509–511}. In our cohort, the DDLPS2 subgroup was 'immune-cold' due to low infiltrate of CD3+, CD4+ and CD8+ TILs, implying that the complement system might play an immunosuppressive in this group of patients.

The main clinical feature of DDLPS is the high rate of local relapse, which significantly contributes to the dismal outcome of DDLPS^{98,100}. The identified DDLPS3 subgroup had a significantly shorter LRFs, MFS and OS. The median LRFs and OS for DDLPS3 patients were both under 12 months, which is considerably lower than the median LRFs of 37 months and OS of 52 months in our whole DDLPS cohort and in other reported cohorts^{512,513}. The matrix signalling network in DDLPS3 contained cathepsin enzymes (CTSA, CTSB and CTSZ), which are lysosomal proteases involved in protein degradation, protein and lipid metabolism, autophagy, and antigen presentation²³². These enzymes can also be secreted in the extracellular space, where they turnover and degrade the ECM. Increased expression of cathepsins is associated with a more aggressive cancer phenotype and reduced survival in multiple cancer types, including STS such as MyFS and UPS^{232,514,515}. Genetic loss of cathepsin B (CTSB) reduced invasion and metastasis of pancreatic epithelial cells in the pancreatic ductal adenocarcinoma mouse model and in neuroendocrine cells in the RIP1-Tag2 mouse model^{516,517}. Moreover, a selective CTSB inhibitor CA-074 reduced

bone metastasis *in vivo* in a breast cancer mouse model⁵¹⁸. The DDLPS3 subgroup was associated with worse LRFS and MFS and had elevated levels of cathepsins A, B and Z, implicating them as candidate proteins involved in local and distant metastases development in this subgroup of DDLPS patients.

The DDLPS3 subgroup also presented with an upregulated expression of an adhesome components PLAUR. *PLAUR*, which encodes for urokinase-type plasminogen activator receptor (uPAR), was shown to promote cancer cell proliferation and metastasis in multiple cancer types, including paediatric rhabdomyosarcoma and osteosarcoma^{519–522}. Recent progress in immunotherapy has led to the development of strategies to target uPAR. Notably, anti-uPAR chimeric antigen receptor (CAR) T-cells and antibody recruiting molecules (ARMs) have demonstrated strong and enduring responses in preclinical studies^{523,524}. While further research and clinical trials are essential to validate their effectiveness and safety in human patients, the anti-uPAR CAR T-cells and ARMs could be a potential therapeutic option for DDLPS patients with high uPAR expression.

When considering the global proteomic profile, DDLPS3 patients had an enriched activity in oncogenic KRAS, MTORC1 and MYC pathways and DNA damage pathways. These signalling pathways are interconnected and often dysregulated in cancer, where they collectively promote the hallmark features of uncontrolled cell division and survival^{525–528}. Aggressive phenotype induced through KRAS, MTORC1 and MYC pathways aligns with poor LRFS, MFS and OS in the DDLPS3 patients.

In multivariable analysis, after adjusting for known prognostic factors in the DDLPS, DDLPS3 remained an independent prognostic factor for LRFS and OS, supporting its clinical applicability as a valuable biomarker. DDLPS3 had a short list of uniquely upregulated matrisome and adhesome proteins, which can be clinically used as candidate markers to identify patients at high risk of local relapse. Identifying a high-risk subpopulation would allow for the stratification of patients, enabling the implementation of more aggressive treatment strategies upfront. Additionally, it would facilitate more frequent

monitoring of these high-risk individuals to ensure timely intervention and improved disease management. The matrix signalling profiles identified in this thesis could serve as the foundation for the prognostic classifier in DDLPS. Complementing the proteomics-based approach with orthogonal validation techniques, such as IHC staining, would aid in the classifier development, paving the way for its integration into clinical practice.

4.3.3. Matrix signalling heterogeneity in UPS

Four UPS subgroups were identified based on the protein expression of matrisome and adhesome. The UPS subgroups displayed differential ECM integrin adhesion networks, were associated with differential T cell infiltrate and stratified UPS patients for the risk of metastasis development.

The UPS1 matrisome and adhesome biology were related to fibrillar collagens. Specifically, enriched PPI networks in UPS1 included collagen biosynthesis, modification, and degradation processes. These findings imply that patients within this subgroup could have a relatively high rate of collagen turnover and cross-linking. Furthermore, in univariable analysis, UPS1 was significantly associated with a higher risk of metastasis compared to the UPS3 and UPS4 subgroups. This association highlights the potential importance of collagen turnover and cross-linking in UPS metastatic dissemination and disease progression. In a genetically engineered mouse model of UPS, shRNA-mediated knockdown of collagen-crosslinking enzyme PLOD2 reduced collagen density and changed the collagen organisation in the primary tumour⁴⁸². The PLOD2 knockdown reduced UPS cell migration and reduced lung metastasis compared to the control shRNA. Moreover, overexpression of collagen-crosslinking enzyme PLOD2 promoted UPS cell migration and lung metastasis by deposition of disorganised collagen. Furthermore, the same study showed that the inhibition of PLOD2 with minoxidil decreased pulmonary metastasis *in vivo*⁴⁸². However, minoxidil is not a PLOD2-specific inhibitor and other targets of minoxidil were reported, including Kir6/SUR potassium channel complex⁵²⁹. As such, further investigations are needed to validate and fully understand the mechanistic role of collagen turnover and cross-linking in

UPS metastasis. Nevertheless, in my analysis, UPS1 demonstrated an association with poor MFS even after adjustment for known prognostic factors in the multivariable Cox model. This emphasises the added prognostic value that the molecular classification based on expression profiles of matrisome and adhesome can offer. This classification has the potential to become a valuable tool in predicting UPS patient metastasis risk.

A collagen-degrading proteinase MMP14 was upregulated in the UPS1 subgroup compared to other UPS subgroups. Recently, MMP14 was investigated as a target for a bicyclic peptide drug conjugate BT1718 and is currently being evaluated in a phase I/II clinical trial for solid cancers, including sarcoma (NCT03486730)⁵³⁰. BT1718 is a peptide with a high affinity for MMP14, and it acts to deliver a toxin mertansine, a potent anti-tubulin agent, to tumours with high MMP14 expression⁵³¹. Upon tumour-localised cleavage of the linker, connecting mertansine to BT1718, the liberated active mertansine inhibits the assembly of microtubules, hindering the division of tumour cells^{532,533}. This cascade of events ultimately results in cell death and a consequent reduction in tumour size. This offers a therapeutic opportunity for the UPS1 subgroup, suggesting that patients within this group are more likely to respond to BT1718 compared to other subgroups.

In my analysis, UPS1 had the lowest T cell infiltrate of the four subgroups. Collagens have been shown to negatively modulate T cell infiltration in tumours^{413,414,534,535}. Collagens exert an immunosuppressive effect by creating physical barriers and trapping T lymphocytes, preventing intra-tumoral infiltration^{413,534–536}. In human lung tumour slices, high-density collagen areas were associated with lower infiltrating CD3+ T cells⁴¹³. In breast cancer, collagen-dense areas had lower CD8+ T cells, and those CD8+ T cells displayed reduced proliferation and cytotoxicity compared to CD8+ T cells in low-density collagen areas⁴¹⁴. In addition to limiting the intra-tumoral infiltration, higher fibrillar collagen deposition (collagen I and collagen III) was associated with increased expression of CD8+ T cell exhaustion markers and resistance to PD-1/PD-L1 blockage in lung cancer⁵³⁷. Further counterstaining using multiplexed IHC for fibrillar collagens (collagen I, III or IV) and T cell

markers in our UPS cohort would strengthen the observed inverse association between T cell infiltrate and fibrillar collagen expression in UPS1 patients.

A subset of UPS patients show favourable responses to immune-checkpoint inhibitors (ICIs) in phase II trials, however, there is a pressing need to identify biomarkers to improve patient selection and develop future combination strategies^{81,151,538}. STS tumours with increased T cell infiltration (i.e., ‘immune-hot’) generally have higher response rates to ICI^{108,137,477}. As such, patients in the UPS1 subgroup (immune cold) might have a poor response to ICIs monotherapy. Recently, a combination of inhibitors of collagen-crosslinking enzymes was shown to improve response to ICIs in preclinical models^{537,539}. In a study by Peng *et al.* inhibition of intra-tumoral collagen deposition through suppression of collagen-crosslinking enzyme increased CD8+ T cell infiltrate and resensitised resistant lung tumours to PD-L1 blockade⁵³⁷. In a mouse model of pancreatic ductal adenocarcinoma, inhibition of collagen-crosslinking enzyme improved response for PD-1 blockade⁵³⁹. This could offer a therapeutic opportunity for combining treatment strategies to reduce collagen deposition and cross-linking with ICIs in UPS1 patients.

4.3.4. Limitations

There are some limitations associated with the analysis in this chapter. Firstly, LMS, DDLPS and UPS cohorts were obtained from a single institution. As such, the findings might be specific to these cohorts. In addition, the treatment and management of STS depend on the expertise of the treatment centres. For instance, the OS rate was significantly higher for retroperitoneal sarcoma patients receiving surgery at a specialist centre (such as RMH) compared to surgeries performed in a non-specialist centre⁵⁴⁰. As such, findings from a single treatment centre might not be reproducible in a wider population, as not all STS patients receive treatment at the specialist centres. To avoid the overfitting of the results, external validation is crucial. Secondly, the cohort in this thesis exclusively focused on primary tumours. Whilst this enabled insights into the baseline biology of the ECM integrin signalling in LMS, DDLPS and UPS, it may have restricted the relevance of the findings. It is known that ECM

undergoes remodelling during carcinoma progression, and the ECM composition of primary tumours and tumours from recurrent or metastatic sites is different^{171,541,542}. It remains to be determined if the ECM composition and associated signalling change during STS progression. Consequently, the valuable biological insights gained from primary tumours may not necessarily apply to patients with advanced LMS, DDLPS and UPS. Additional limitations relate to the underrepresentation of certain anatomical locations in LMS and DDLPS cohorts and small cohort size, which are discussed below.

LMS-specific limitations

Some of the LMS cohort characteristics in this study deviate from other LMS studies. In our LMS cohort, the percentage of uterine cases (11%) was lower than the reported 20-50% of uLMS in other LMS studies^{82,435–437,439}, and 25% incidence of uLMS in the general LMS population⁵⁴³. Additionally, our LMS cohort lacked low-grade (grade 1) tumours, likely reflecting the patient cases typically encountered at RMH. It is well-established that uLMS compared to eLMS and low-grade versus high-grade LMS have different clinical behaviour. Patients with uLMS have higher rates of metastasis than eLMS^{20,73–75}. Similarly, high-grade LMS is more aggressive with higher propensity for metastasis compared to low-grade LMS⁴¹. As such, the cohort with the underrepresentation of uLMS and low-grade LMS may not be adequately capture whole spectrum of LMS biology, potentially leading to a skewed understanding of the disease. These deviations introduce selection bias, making it challenging to extend the findings to the general LMS population.

DDLPS-specific limitations

Due to the rarity of DDLPS, the limitations of the DDLPS analysis are related to a relatively small number of patients in the cohort, which was further stratified into three groups. Small cohorts may not accurately represent the diversity of the patient population, and there is a higher risk of sampling bias. Additionally, small sample sizes may not provide enough statistical power to detect true differences in survival accurately. Therefore, the observed significant differences could be due to chance rather than a genuine effect. Extremity cases were underrepresented in the DDLPS cohort, 5% in our cohort

compared to 25% incidence^{113,544}. The anatomical location of DDLPS shows association with survival, with extremity cases generally performing better⁶². As a result, it is challenging to extend the findings to the general DDLPS population, particularly individuals with DDLPS occurring in the extremities. Although promising results for DDLPS risk stratification were presented in this chapter, an independent validation in a larger cohort with patients from other research institutes is required.

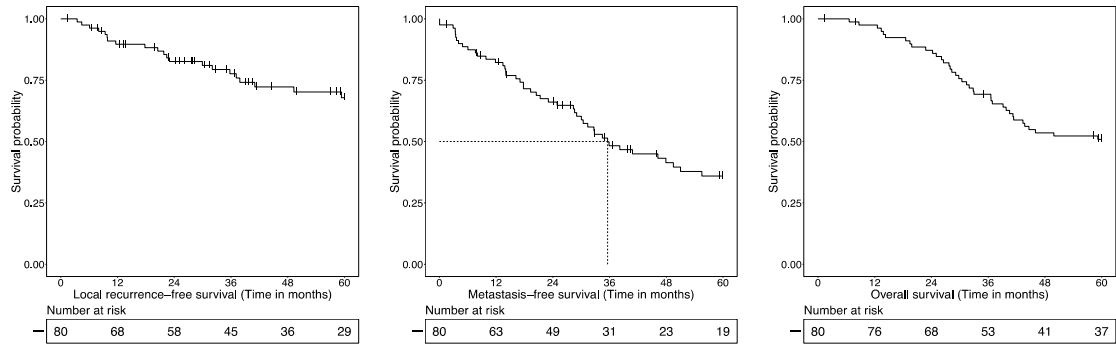
UPS-specific limitations

Similar to the DDLPS analysis, the limitation of UPS-specific analysis lies in the small number of patients in the subgroup analysis. In subgroup analyses, controlling for potential confounding variables (such as known prognostic factors) becomes challenging with a small number of patients. In addition, information on some clinicopathological features, such as performance status, was incomplete. Although performance status was included in the multivariable analysis in UPS, the missing data could mean that the compounding effect of the performance status on UPS survival was not accurately modelled.

Overall, the analysis in this chapter was preliminary and can be considered the first step towards characterising and assessing the clinical implications of matrix signalling heterogeneity in LMS, DDLPS and UPS. Before the findings in this chapter can be applied in clinical practice, they need to undergo external validation. Specifically, future cohorts should incorporate a representative number of anatomical location cases to provide a more comprehensive understanding of matrix signalling heterogeneity in the general population of LMS, DDLPS and UPS patients. Secondly, the DDLPS and UPS cohorts should be larger to increase statistical power in the subgroup analyses. Future efforts should focus on collaborations with external institutions to collate a larger multi-institutional cohort of LMS, DDLPS and UPS to avoid overfitting the results and account for centre-specific differences in STS management. After externally validating the findings in primary LMS, DDLPS and UPS, further work can include incorporating recurrence and metastatic samples to assess the applicability of the findings in advanced settings.

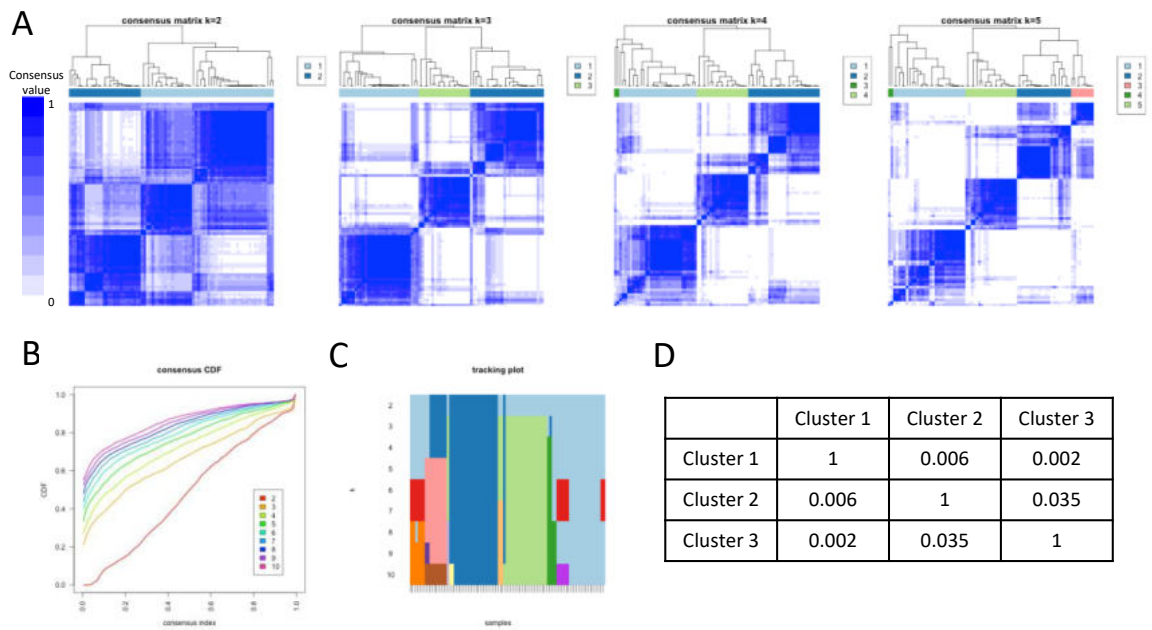
4.4. Supplementary Material

4.4.1. Supplementary Figures



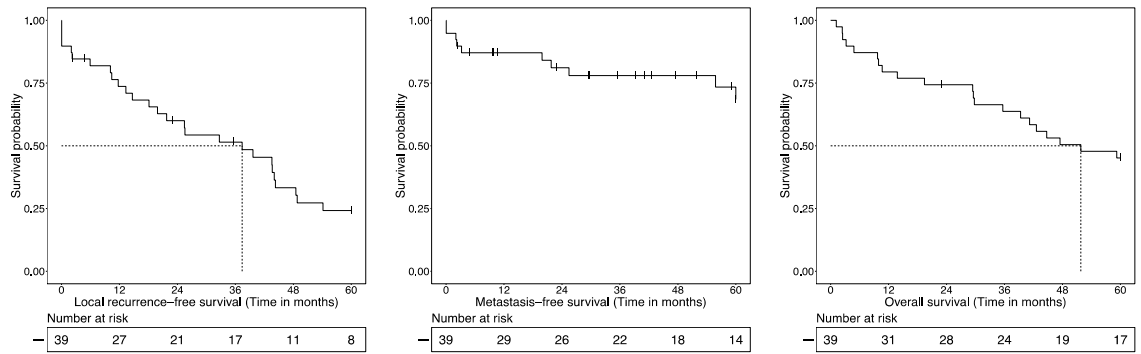
Supplementary Figure 4.1 Kaplan-Meier plots of local recurrence-free survival (LRFS), metastasis-free survival (MFS) and overall survival (OS) in a cohort of n = 80 leiomyosarcoma (LMS).

Horizontal and vertical lines indicate the median MFS.



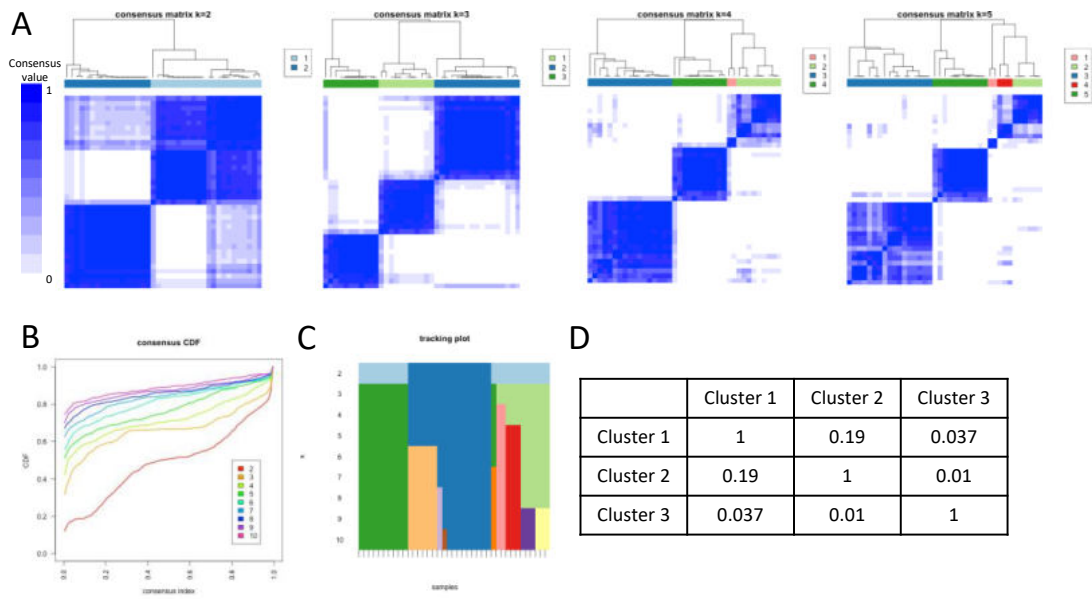
Supplementary Figure 4.2 Identification of leiomyosarcoma (LMS) subgroups based on the protein expression of matrisome and adhesome.

A) Consensus matrices for k up to 5, B) Consensus CDF plot for k = 2-10, C) Tracking plot for k = 2-10, D) Results of Sigclust are summarised in a matrix, showing p-value for pairwise comparisons between the three consensus clusters. A pair of clusters with p-values < 0.05 was considered as significantly different.



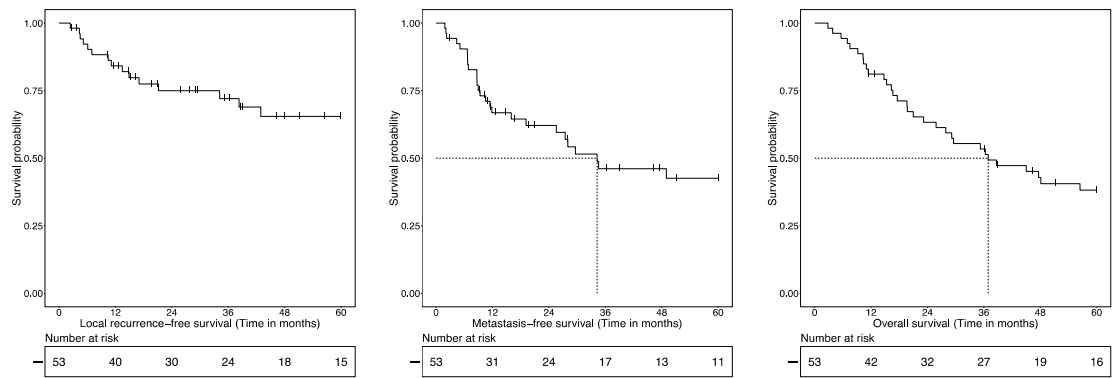
Supplementary Figure 4.3 Kaplan-Meier plots of local recurrence-free survival (LRFS), metastasis-free survival (MFS) and overall survival (OS) in a cohort of n = 39 dedifferentiated liposarcoma (DDLPS).

Horizontal and vertical lines indicate the median LRFS and OS.



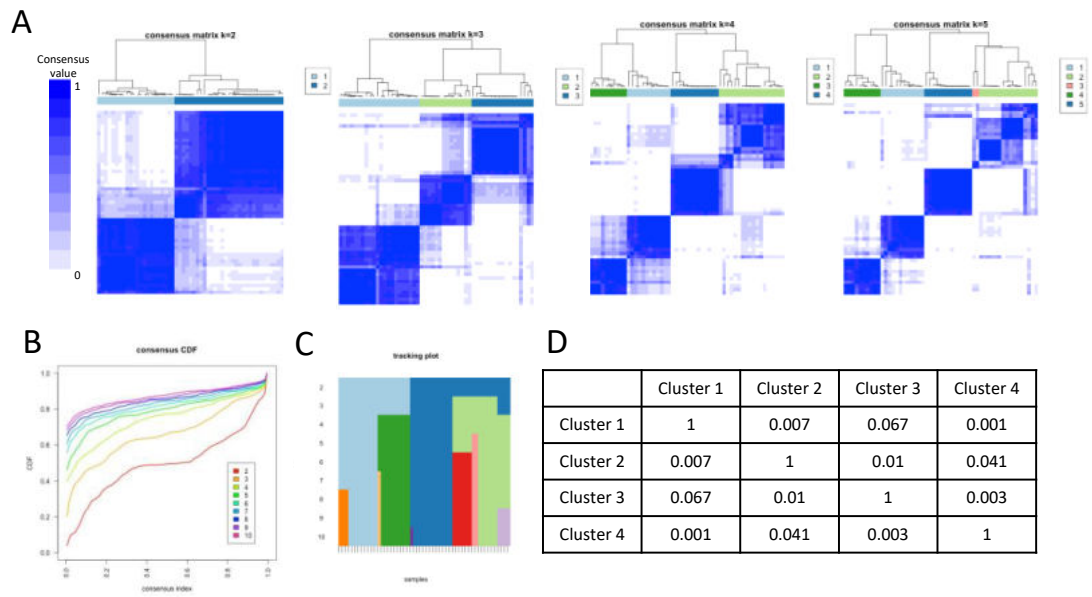
Supplementary Figure 4.4 Identification of dedifferentiated liposarcoma (DDLPS) subgroups based on the protein expression of matrisome and adhesome.

A) Consensus matrices for k up to 5, B) Consensus CDF plot for k = 2-10, C) Tracking plot for k = 2-10, D) Results of Sigclust are summarised in a matrix, showing p-value for pairwise comparisons between the three clusters. A pair of clusters with p-values < 0.05 was considered as significantly different.



Supplementary Figure 4.5 Kaplan-Meier plots of local recurrence-free survival (LRFS), metastasis-free survival (MFS) and overall survival (OS) in a cohort of n = 53 undifferentiated pleomorphic sarcoma (UPS).

Horizontal and vertical lines indicate the median MFS and OS.



Supplementary Figure 4.6 Identification of undifferentiated pleomorphic sarcoma (UPS) subgroups based on matrisome and adhesome protein expression.

- 5) Consensus matrices for k up to 5, B) Consensus CDF plot for k = 2-10, C) Tracking plot for k = 2-10, D) Results of Sigclust are 215ecellular in a matrix, showing p-value for pairwise comparisons between the four clusters. A pair of clusters with p-values < 0.05 was considered as significantly different.

4.4.2. Supplementary Tables

Supplementary Table 4.1 A summary of 112 differentially expressed matrisome and adhesome proteins (DEPs) uniquely upregulated in each of the three LMS subgroups.

LMS1		LMS2		LMS3	
Upregulated proteins	Description	Upregulated proteins	Description	Upregulated proteins	Description
COL4A2	Basement membrane	F13A1	ECM regulator	ACAN	Basement membrane
HSPG2	Basement membrane	PTPN6	Tyrosine phosphatase	AGRN	Basement membrane
ITGA1	Basement membrane/Adhesion receptor	CTSS	ECM Regulators	BGN	Basement membrane
ITGA5	Basement membrane/Adhesion receptor			COL12A1	Basement membrane
ITGA7	Basement membrane/Adhesion receptor			COL14A1	Basement membrane
ITGB1	Basement membrane/Adhesion receptor			COL5A1	Basement membrane
LAMA4	Basement membrane			DCN	Basement membrane
LAMB2	Basement membrane			ECM1	Basement membrane
NID1	Basement membrane			FBN2	Basement membrane
SDC4	Basement membrane/Adhesion receptor			FMOD	Basement membrane
TINAGL1	Basement membrane			LAMA5	Basement membrane
VWA1	Basement membrane			LAMB1	Basement membrane
SVIL	Actin regulation			LUM	Basement membrane
FLNA	Actin regulation			MATN2	Basement membrane
TNS1	Adaptor			POSTN	Basement membrane
SORBS1	Adaptor			PXDN	Basement membrane
SYNM	Adaptor			SPARC	Basement membrane
LPP	Adaptor			SPARCL1	Basement membrane
CSRP1	Adaptor			TGFBI	Basement membrane
PALLD	Adaptor			TNC	Basement membrane
TGFB11	Adaptor			VCAN	Basement membrane
FABP3	Adaptor			VTN	Basement membrane
PARVA	Adaptor			HSPA2	Chaperone
VCL	Adaptor			COL5A2	Collagens
LDB3	Adaptor			COL3A1	Collagens
HSPB1	Chaperone			COL1A2	Collagens
SBSPON	ECM Glycoproteins			COL1A1	Collagens
EMILIN1	ECM Glycoproteins			IGFBP5	ECM Glycoproteins
CTSG	ECM Regulators			IGFBP3	ECM Glycoproteins
ILK	Serine/threonine kinase			IGFALS	ECM Glycoproteins
				IGFBP7	ECM Glycoproteins
				LRG1	ECM Glycoproteins
				THSD4	ECM Glycoproteins
				MFAP5	ECM Glycoproteins
				DPT	ECM Glycoproteins
				AEBP1	ECM Glycoproteins
				ABI3BP	ECM Glycoproteins
				TNXB	ECM Glycoproteins
				FGL2	ECM Glycoproteins
				ITIH3	ECM Regulators
				ITIH1	ECM Regulators
				HRG	ECM Regulators
				AGT	ECM Regulators
				SERPINA1	ECM Regulators
				ITIH2	ECM Regulators
				SERPINC1	ECM Regulators
				SERPINA6	ECM Regulators
				SERPING1	ECM Regulators
				KNG1	ECM Regulators
				PLG	ECM Regulators
				A2M	ECM Regulators
				SERPINF2	ECM Regulators
				HTRA1	ECM Regulators
				SERPINF1	ECM Regulators
				SERPINA4	ECM Regulators
				PLOD1	ECM Regulators
				SERPINH1	ECM Regulators
				F2	ECM Regulators
				ITIH4	ECM Regulators
				F12	ECM Regulators
				CPN2	ECM Regulators
				SERPINA7	ECM Regulators
				LOX	ECM Regulators
				AMBP	ECM Regulators
				P4HA1	ECM Regulators
				P4HA2	ECM Regulators
				SERPINA3	ECM Regulators
				PLOD2	ECM Regulators
				SERPIND1	ECM Regulators
				CLEC3B	ECM-affiliated Proteins
				CLEC11A	ECM-affiliated Proteins
				GPC6	ECM-affiliated Proteins
				HAPLN3	Proteoglycans
				PRELP	Proteoglycans
				ASPN	Proteoglycans
				OGN	Proteoglycans
				FSTL1	Secreted Factors
				S100A4	Secreted Factors
				SFRP4	Secreted Factors

Supplementary Table 4.2 Summary of statistical tests to assess the association between clinicopathological features and leiomyosarcoma (LMS) subgroups.

Significant associations are highlighted in red. F = female; M = male.

Characteristic	Category	LMS subgroups			Test results			
		LMS1	LMS2	LMS3	Test performed	χ -squared	Degrees of freedom	p-value
Age at excision (years)	Median	61.3	66.3	66.8	Kruskal Wallis	0.92	2	0.631
	Min	30.5	29.3	36.8				
	Max	86.9	83.5	86.3				
Anatomical site	Extremity	10	11	10	Chi-square	10.84	10	0.3702
	Intra-abdominal	5	4	1				
	Pelvic	3	3	3				
	Retroperitoneal	12	4	3				
	Trunk	1	1	0				
Grade	Uterine	1	5	3	Chi-square	18.38	2	0.0001
	2	27	15	5				
Tumour depth	3	5	13	15	Chi-square	0.32	2	0.8515
	Deep	26	24	16				
Tumour size (mm)	Superficial	6	4	4	Kruskal Wallis	3.28	2	0.2937
	Median	98.5	110	76				
	Min	35	5	25				
Tumour margins	Max	400	250	260	Chi-square	3.98	6	0.6792
	R0	19	14	9				
	R1	11	13	11				
	R2	1	0	0				
Performance status	Rx	1	1	0	Chi-square	10.67	6	0.099
	0	17	16	7				
	1	5	3	8				
	2-3	4	1	3				
Sex	unknown	6	8	2	Chi-square	1.84	2	0.3992
	F	20	20	16				
	M	12	8	4				

Supplementary Table 4.3 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and leiomyosarcoma (LMS) subgroups with metastasis-free survival (MFS).

Clinicopathological factors which were significantly associated with MFS in UVA analysis (significant p-values are in red) were included in the MVA model. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

Variable	Groups	n	Univariable analysis (MFS)		Multivariable analysis (MFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1 (0.98-1.02)	0.707	-	-
Grade	2 (reference)	47	-	-	-	-
	3	33	2.46 (1.35-4.46)	0.003	2.12 (1.02-4.44)	0.045
Anatomical location	Intra-abdominal/Retroperitoneal/Pelvic (reference)	38	-	-	-	-
	Extremity/Trunk	33	0.48 (0.25-0.93)	0.03	0.37 (0.17-0.80)	0.011
	Uterine	9	1.02 (0.39-2.67)	0.965	0.90 (0.32-2.59)	0.849
Log [tumour size] (mm)	4-5 (reference)	50	-	-	-	-
	<4	12	0.23 (0.07-0.74)	0.014	0.43 (0.20-1.56)	0.199
	>5	17	0.64 (0.30-1.33)	0.228	0.46 (0.19-1.1)	0.08
Tumour depth	Deep (reference)	66	-	-	-	-
	Superficial	14	0.29 (0.09-0.94)	0.039	0.59 (0.16-2.22)	0.434
Tumour margin	R0 (reference)	42	-	-	-	-
	R1&2 (R1-55cases; R2-1case)	36	1.07 (0.59-1.93)	0.824	-	-
Sex	F (reference)	56	-	-	-	-
	M	24	0.60 (0.30-1.21)	0.15	-	-
Performance status	0 (reference)	40	-	-	-	-
	1	16	1.25 (0.58-2.68)	0.571	-	-
	2-3	8	2.12 (0.72-6.25)	0.175	-	-
	unknown	16	0.79 (0.34-1.84)	0.589	-	-
Consensus cluster	LMS1 (reference)	32	-	-	-	-
	LMS2	28	1.65 (0.81-3.34)	0.168	1.82 (0.82-4.04)	0.143
	LMS3	20	2.44 (1.15-5.15)	0.02	2.23 (0.88-5.64)	0.091

Supplementary Table 4.4 Summary of multivariable (MVA) Cox regression analysis assessing the association of clinicopathological factors and leiomyosarcoma (LMS) subgroups with local recurrence-free survival (LRFS) and overall survival (OS).

Significant p-values are in red. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by multivariate Cox regression with a two-sided Wald test.

Variable	Groups	n	Multivariable analysis (LRFS)		Multivariable analysis (OS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
Grade	2 (reference)	47	-	-	-	-
	3	33	0.98 (0.33-2.91)	0.975	1.39 (0.66-2.94)	0.391
Anatomical location	Intra-abdominal/Retroperitoneal/Pelvic (reference)	38	-	-	-	-
	Extremity/Trunk	33	0.25 (0.07-0.86)	0.028	0.42 (0.17-1)	0.051
	Uterine	9	0.23 (0.03-1.77)	0.157	1.1 (0.40-3.05)	0.859
Log [tumour size] (mm)	4-5 (reference)	50	-	-	-	-
	<4	12	0.90 (0.18-4.62)	0.898	0.46 (0.12-1.78)	0.258
	>5	17	1.10 (0.38-3.16)	0.857	0.99 (0.45-2.17)	0.98
Tumour depth	Deep (reference)	66	-	-	-	-
	Superficial	14	0.50 (0.05-4.82)	0.545	1.88 (0.56-6.33)	0.311
Consensus cluster	LMS1 (reference)	32	-	-	-	-
	LMS2	28	2.39 (0.81-7.06)	0.115	1.27 (0.54-2.98)	0.58
	LMS3	20	1.97 (0.51-7.61)	0.325	1.54 (0.60-3.99)	0.371

Supplementary Table 4.5 Summary of statistical tests to assess the association between clinicopathological features and matrisome and adhesome clusters in dedifferentiated liposarcoma (DDLPS).

Significant associations are highlighted in red. F = female; M = male.

Characteristic	Category	DDLPS subgroup			Test results			
		DDLPS1	DDLPS2	DDLPS3	Test performed	χ -squared	Degrees of freedom	p-value
Age at excision (years)	median	64.3	62.3	63.3	Kruskal Wallis	0.85	2	0.654
	min	40.7	35.1	51.4				
	max	80.3	80.9	81.3				
Tumour size (mm)	median	190	270	170	Kruskal Wallis	6.54	2	0.038
	min	35	100	70				
	max	410	1090	450				
Anatomical site	Extremity	1	1	0	Chi-square	5.91	6	0.434
	Intra-abdominal	1	0	2				
	Retroperitoneal	8	16	8				
	Trunk	1	0	1				
Grade	2	7	9	3	Chi-square	3.13	2	0.209
	3	4	8	8				
Performance status	0	9	6	2	Chi-square	17.51	6	0.008
	1	1	5	6				
	2-3	1	0	2				
	unknown	0	6	1				
Sex	F	1	7	7	Chi-square	7.01	2	0.030
	M	10	10	4				
Tumour margins	R0	4	3	2	Chi-square	4.66	4	0.324
	R1	7	12	6				
	Rx	0	2	3				

Supplementary Table 4.6 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and dedifferentiated liposarcoma (DDLPS) subgroups with local recurrence-free survival (LRFS).

Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariable and multivariable Cox regression with a two-sided Wald test. Significant p-values are in red.

Variable	Groups	n	Univariable analysis (LRFS)		Multivariable analysis (LRFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1.03 (0.99-1.06)	0.134	1.04 (0.99-1.08)	0.107
Grade	3 (reference)	20	-	-	-	-
	2	19	0.70 (0.33-1.49)	0.352	1.18 (0.49-2.84)	0.707
Log [tumour size] (mm)	>5 (reference)	29	-	-	-	-
	≤5	10	1.24 (0.52-2.95)	0.625	0.64 (0.22-1.87)	0.414
Tumour margin	R1 (reference)	25	-	-	-	-
	R0	9	1.20 (0.49-2.89)	0.693	-	-
	Rx	5	0.96 (0.28-3.28)	0.942	-	-
Sex	M (reference)	24	-	-	-	-
	F	15	1.56 (0.73-3.34)	0.253	0.73 (0.30-1.82)	0.504
Performance status	0 (reference)	17	-	-	-	-
	1	12	2.19 (0.92-5.20)	0.076	0.76 (0.22-2.68)	0.672
	2-3	3	0.81 (0.10-6.31)	0.841	1.01 (0.12-8.56)	0.99
	unknown	7	1.27 (0.44-3.67)	0.655	0.80 (0.23-2.75)	0.725
Consensus cluster	DDLPS3 (reference)	11	-	-	-	-
	DDLPS1	11	0.19 (0.06-0.60)	0.005	0.11 (0.02-0.56)	0.008
	DDLPS2	17	0.40 (0.16-0.99)	0.049	0.28 (0.08-0.98)	0.046

Supplementary Table 4.7 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and DDLPS subgroups with overall survival (OS).

Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariable and multivariable Cox regression with a two-sided Wald test. Significant p-values are in red.

Variable	Groups	n	Univariable analysis (OS)		Multivariable analysis (OS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1.05 (1.01-1.09)	0.025	1.05 (0.99-1.1)	0.08
Grade	3 (reference)	20	-	-	-	-
	2	19	0.38 (0.15-0.95)	0.038	0.71 (0.24-2.11)	0.534
Log [tumour size] (mm)	>5 (reference)	29	-	-	-	-
	≤5	10	1.59 (0.64-3.96)	0.318	0.48 (0.15-1.5)	0.206
Tumour margin	R1 (reference)	25	-	-	-	-
	R0	9	1.2 (0.46-3.13)	0.708	-	-
	Rx	5	0.34 (0.05-2.61)	0.301	-	-
Sex	M (reference)	24	-	-	-	-
	F	15	1.06 (0.44-2.56)	0.9	0.56 (0.20-1.6)	0.279
Performance status	0 (reference)	17	-	-	-	-
	1	12	2.71 (1.01-7.31)	0.049	1.53 (0.42-5.52)	0.516
	2-3	3	2.95 (0.61-14.3)	0.178	2.21 (0.37-13.2)	0.387
	unknown	7	1.16 (0.3-4.49)	0.83	1.34 (0.25-7.21)	0.735
Consensus cluster	DDLPS3 (reference)	11	-	-	-	-
	DDLPS1	11	0.22 (0.07-0.66)	0.007	0.15 (0.03-0.69)	0.015
	DDLPS2	17	0.21 (0.08-0.59)	0.003	0.13 (0.03-0.53)	0.005

Supplementary Table 4.8 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and DDLPS subgroups with metastasis-free survival (MFS).

Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariable and multivariable Cox regression with a two-sided Wald test. In univariable Cox regression analysis, the survival estimates of the performance status could not be estimated due to extreme hazard ratios and infinite confidence intervals from the models and, therefore, performance status was not included for multivariable analysis. Significant p-values are in red.

Variable	Groups	n	Univariable analysis (MFS)		Multivariable analysis (MFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1.02 (0.96-1.07)	0.569	1 (0.93-1.08)	0.932
Grade	3 (reference)	20	-	-	-	-
	2	19	0.34 (0.09-1.34)	0.125	0.35 (0.07-1.82)	0.213
Log [tumour size] (mm)	>5 (reference)	29	-	-	-	-
	≤5	10	0.78 (0.16-3.68)	0.749	0.29 (0.05-1.81)	0.186
Tumour margin	R1 (reference)	25	-	-	-	-
	R0	9	0.98 (0.20-4.87)	0.983	-	-
	Rx	5	1.73 (0.35-8.62)	0.504	-	-
Sex	M (reference)	24	-	-	-	-
	F	15	4.44 (1.15-17.20)	0.031	2.57 (0.55-12)	0.231
Performance status	0 (reference)	17	-	-	-	-
	1	12	2.62 (0.70-9.89)	0.155	-	-
	2-3	3	1.5e-08 (0-Inf)	0.999	-	-
	unknown	7	0.64 (0.07-5.76)	0.692	-	-
Consensus cluster	DDLPS3 (reference)	11	-	-	-	-
	DDLPS1	11	0.08 (0.01-0.74)	0.044	0.14 (0.01-1.86)	0.135
	DDLPS2	17	0.23 (0.05-0.96)	0.026	0.18 (0.03-1.01)	0.052

Supplementary Table 4.9 Summary of statistical tests to assess the association between clinicopathological features and undifferentiated pleomorphic sarcoma (UPS) subgroups.

Significant associations are highlighted in red. F = female; M = male.

Characteristic	Category	UPS subgroup				Test results			
		UPS1	UPS2	UPS3	UPS4	Test performed	χ -squared	Degrees of freedom	p-value
Age at excision (years)	Median	73.9	72.3	71.9	77.4	Kruskal Wallis	0.5	3	0.925
	Min	58.3	55.9	56.1	28.2				
	Max	81.8	86.3	84.3	90				
Anatomical site [n (%)]	Extremity	11	10	12	5	Chi-square	14.2	12	0.29
	Head/neck	0	2	0	2				
	Trunk	0	2	4	2				
	Pelvic	1	0	0	1				
	Intra-abdominal	0	0	1	0				
Grade [n (%)]	3	11	14	17	7	Chi-square	6.5	3	0.089
	2	1	0	0	2				
Tumour depth [n (%)]	Deep	12	9	14	8	Chi-square	5.4	3	0.144
	Superficial	0	5	3	2				
Tumour size (mm)	Median	115	50.5	95	69	Kruskal Wallis	11.4	3	0.01
	Min	53	15	20	20				
	Max	360	150	254	110				
Tumour margins [n (%)]	R0	6	7	9	4	Chi-square	0.4	3	0.932
	R1	6	7	8	6				
Performance status [n (%)]	0	2	6	7	7	Chi-square	9.3	9	0.414
	1	5	4	4	2				
	2-3	2	1	4	0				
	unknown	3	3	2	1				
Sex [n (%)]	F	4	11	10	3	Chi-square	7.9	3	0.048
	M	8	3	7	7				

Supplementary Table 4.10 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and undifferentiated pleomorphic sarcoma (UPS) subgroups with metastasis-free survival (MFS).

(significant p-values are in red). Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

Variable	Groups	n	Univariable analysis (MFS)		Multivariable analysis (MFS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1.01 (0.98-1.05)	0.459	1.02 (0.96-1.09)	0.474
Anatomical site	Extremity (reference)	38	-	-	-	-
	Other	15	0.52 (0.2-1.37)	0.186	0.93 (0.31-2.79)	0.901
Log [tumour size] (mm)	4-5 (reference)	30	-	-	-	-
	<4	15	0.58 (0.22-1.51)	0.262	0.65 (0.19-2.29)	0.506
	>5	8	3.62 (1.23-10.7)	0.02	3.33 (0.89-12.5)	0.075
Tumour margin	R1 (reference)	27	-	-	-	-
	R0	26	1.24 (0.57-2.72)	0.586	-	-
Tumour depth	Deep (reference)	43	-	-	-	-
	Superficial	10	0.37 (0.11-1.26)	0.111	-	-
Sex	F (reference)	28	-	-	-	-
	M	25	1.41 (0.65-3.04)	0.387	1.13 (0.41-3.1)	0.819
Performance status	0 (reference)	22	-	-	-	-
	1	15	1.15 (0.44-2.97)	0.776	0.76 (0.25-2.25)	0.617
	2-3	7	1.43 (0.38-5.33)	0.598	0.56 (0.11-3.02)	0.502
	unknown	9	2.97 (1-8.79)	0.05	2.05 (0.51-8.33)	0.315
Consensus cluster	UPS1 (reference)	12	-	-	-	-
	UPS2	14	0.47 (0.17-1.26)	0.133	0.51 (0.12-2.06)	0.342
	UPS3	17	0.37 (0.14-0.99)	0.048	0.36 (0.11-1.12)	0.078
	UPS4	10	0.19 (0.05-0.71)	0.014	0.18 (0.04-0.84)	0.029

Supplementary Table 4.11 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and undifferentiated pleomorphic (UPS) subgroups with local recurrence-free survival (LRFS).

(significant p-values are in red). Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

Variable	Groups	Univariable analysis (LRFS)			Multivariable analysis (LRFS)	
		n	HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1.03 (0.98-1.09)	0.224	1.05 (0.97-1.13)	0.227
Anatomical site	Extremity (reference)	38	-	-	-	-
	Other	15	1.41 (0.48-4.14)	0.528	3.31 (0.88-12.5)	0.077
Log [tumour size] (mm)	4-5 (reference)	30	-	-	-	-
	<4	15	1.15 (0.36-3.63)	0.814	0.94 (0.21-4.19)	0.934
	>5	8	3.79 (0.94-15.3)	0.062	2.44 (0.51-11.6)	0.262
Tumour margin	R1 (reference)	27	-	-	-	-
	R0	26	0.95 (0.34-2.62)	0.913	-	-
Tumour depth	Deep (reference)	43	-	-	-	-
	Superficial	10	0.53 (0.12-2.33)	0.397	-	-
Sex	F (reference)	28	-	-	-	-
	M	25	1.78 (0.63-5)	0.275	1.32 (0.30-5.71)	0.712
Performance status	0 (reference)	22	-	-	-	-
	1	15	1.82 (0.53-6.32)	0.344	1.13 (0.24-5.43)	0.876
	2-3	7	1.82 (0.35-9.62)	0.48	0.83 (0.1-7.12)	0.865
	unknown	9	3.23 (0.76-13.7)	0.111	3.43 (0.59-19.9)	0.169
Consensus cluster	UPS1 (reference)	12	-	-	-	-
	UPS2	14	0.54 (0.15-2.03)	0.364	0.39 (0.05-2.99)	0.365
	UPS3	17	0.35 (0.08-1.45)	0.146	0.28 (0.05-1.6)	0.152
	UPS4	10	0.51 (0.12-2.13)	0.354	0.3 (0.04-2.02)	0.215

Supplementary Table 4.12 Summary of univariable (UVA) and multivariable (MVA) Cox regression analyses assessing the association of clinicopathological factors and undifferentiated pleomorphic sarcoma (UPS) subgroups with overall survival (OS).

(significant p-values are in red). Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test.

Variable	Groups	n	Univariable analysis (OS)		Multivariable analysis (OS)	
			HR (95% CI)	p-value	HR (95% CI)	p-value
	Age	-	1.04 (1.01-1.08)	0.025	1.07 (1.01-1.14)	0.031
Anatomical site	Extremity (reference)	38	-	-	-	-
	Other	15	0.91 (0.41-2.05)	0.828	3.39 (1.22-9.42)	0.019
Log [tumour size] (mm)	4-5 (reference)	30	-	-	-	-
	<4	15	0.50 (0.2-1.28)	0.148	0.23 (0.07-0.75)	0.015
	>5	8	6.02 (2.4-15.1)	0.0001	3.14 (1.14-8.68)	0.027
Tumour margin	R1 (reference)	27	-	-	-	-
	R0	26	0.6 (0.29-1.22)	0.157	-	-
Tumour depth	Deep (reference)	43	-	-	-	-
	Superficial	10	0.48 (0.17-1.39)	0.176	-	-
Sex	F (reference)	28	-	-	-	-
	M	25	1.26 (0.62-2.55)	0.521	1.05 (0.37-2.97)	0.922
Performance status	0 (reference)	22	-	-	-	-
	1	15	2.51 (0.99-6.38)	0.053	1.87 (0.60-5.78)	0.279
	2-3	7	5.1 (1.72-15.1)	0.003	1.23 (0.29-5.23)	0.784
	unknown	9	6.1 (2.18-17)	0.001	6.38 (1.73-23.5)	0.005
Consensus cluster	UPS1 (reference)	12	-	-	-	-
	UPS2	14	0.66 (0.26-1.65)	0.37	1.14 (0.29-4.48)	0.847
	UPS3	17	0.64 (0.26-1.58)	0.337	0.82 (0.26-2.53)	0.724
	UPS4	10	0.27 (0.07-1)	0.05	0.21 (0.04-1.12)	0.068

Chapter 5 Generation and characterisation of ECM preclinical models

5.1. Background and objectives

Cell cultures where cancer cells are grown on plastic in 2D remain the most common *in vitro* model in STS research and preclinical drug screening due to their low cost, high reproducibility and ease of use^{545,546}. However, only a small percentage of drug candidates identified on plastic demonstrate efficacy in clinical trials⁵⁴⁷. This low success rate might be due, in part, to the failure of conventional plastic cultures to incorporate essential TME components such as the ECM⁵⁴⁸. Existing cell-based assays that model the ECM involvement traditionally rely on commercially available products such as Matrigel, Geltrex and Cultrex basement membrane extract (BME)^{351,549}. These products are soluble forms of basement membrane purified from murine tumours. They are extensively used across various cell types and cell culture applications as they better support cell functions than plain plasticware. Matrigel has been used as a thin gel coating to culture and expand human pluripotent stem cells, maintaining self-renewal and pluripotency⁵⁵⁰ or support the growth of organoids in 3D⁵⁵¹. Although commercial BMEs support current cell culture, they have several limitations. Firstly, commercial BMEs are derived from mouse tumours and thus can contain animal-derived factors that might affect cell behaviour or experimental results. Moreover, the BMEs lack comprehensive analysis of their composition, and coupled with batch-to-batch variability and their animal-derived origin, it introduces experimental uncertainty and impedes reproducibility^{337–339}. Consequently, there is a pressing need to develop novel ECM models to enhance the translational success of preclinical findings in clinical trials.

Chapter 3 showed considerable heterogeneity in the ECM and integrin adhesion networks across the STS subtypes, with several histologies having subtype-specific unique profiles. Moreover, Chapter 4 highlighted intra-subtype heterogeneity in ECM-integrin signalling within LMS and DDLPS.

Collectively, these results support and strengthen the argument that STS subtype-specific and even patient-specific preclinical ECM models are necessary to study the functional role of ECM rather than relying on the 'one-size-fits-all' approach with commercial BME.

Using LMS as an example, Chapter 5 will discuss establishing patient-derived preclinical ECM models (ECM scaffold and ECM hydrogel). The models were biochemically characterised to determine if they recapitulate LMS-specific components. To assess if the LMS ECM impacts the functional biology of LMS cells, LMS patient-derived ECM were used to assess LMS cell adhesion, migration and drug responses.

5.2. Results

5.2.1. Generation of preclinical ECM models

To investigate the functional biology of LMS ECM, I first established preclinical ECM models derived from patient tumours (as described in Section 2.3.1 and Section 2.3.2). Extraction of ECM for cancer research relies on the strategies borrowed from regenerative medicine and tissue engineering and, as a first step, involves tissue decellularisation^{340,552}. Decellularisation is a process used to remove cellular components (cells, DNA, cellular debris) from tissues or organs, leaving behind the insoluble ECM scaffold. Typical strategies involve physical techniques such as freeze-thawing cycles and chemical methods using ionic, non-ionic detergents or enzymes. The choice of decellularisation agent depends on tissue cellularity, density, lipid content and tissue thickness, and typically, a combination of physical and chemical techniques are employed. Ionic detergent such as SDS solubilises cell and nucleic membranes and is preferred for dense tissues such as tumours^{340,553}. Freeze-thawing cycles decellularise the tissue by disrupting the cell membrane by intracellular ice crystal formation and were shown to improve the decellularisation efficiency when combined with SDS⁵⁵⁴. As such, the decellularisation method in this thesis involved freeze-thawing cycles, followed by incubations with SDS. In brief, LMS flash frozen tumours from 7 patients

were decellularised with freeze-thaw cycles, cut into smaller pieces to expose more surface to the detergent and incubated with SDS detergent for 4 days at 4 °C to obtain LMS-specific decellularised ECM scaffolds (Figure 5.1).

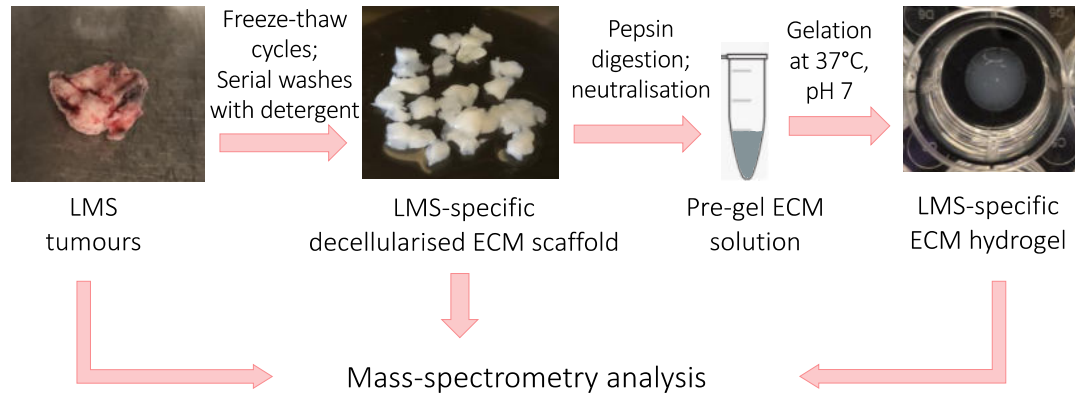


Figure 5.1 A pipeline for generation and characterisation of LMS patient-derived ECM scaffolds and LMS patient-derived ECM hydrogels.

An optimised workflow to generate LMS decellularised scaffolds from fresh frozen LMS tumours by extensive washes with detergent, homogenisation of decellularised scaffold in acidified pepsin solution using Precellys to obtain pre-gel ECM solution. Solidification of the pre-gel ECM solution is achieved by incubation at 37°C, pH 7. Paired tumour, scaffold and hydrogel samples were characterised by mass spectrometry.

Decellularised ECM scaffolds can be recellularised with cancer cells^{353,354,356} or further processed to obtain another type of ECM model, tissue-derived ECM hydrogel³⁸⁰. The most common strategy entails lyophilising the decellularised ECM scaffold, cryo-milling it into a fine powder, and enzymatic digestion with pepsin under acidic conditions^{347,555,556}. Pepsin is an endopeptidase with broad proteolytic activity and typically cleaves peptide bonds, following aromatic amino acids (tryptophan, tyrosine and phenylalanine)⁵⁵⁷. Most ECM proteins, particularly collagen type I, are highly insoluble⁴²⁷. As such, enzymatic protein proteolysis is essential for solubilising ECM scaffolds. At its core, collagen type I consists of collagen fibrils characterised by intertwining three α helices, linked by telopeptides and cross-linkages⁵⁵⁸. During pepsin digestion, collagen type I becomes susceptible to cleavage by pepsin at the telopeptide region, resulting in the unfolding of the collagen protein⁵⁵⁹. Ultimately, this unfolding process substantially increases the solubility of

collagen type I, forming a viscous pre-gel solution. The pepsin digestion is quenched by adding cold sodium hydroxide to adjust the pH of the pre-gel solution to 7. Incubation of the pre-gel solution at 37 °C induces self-assembly of the collagen network, regulated by the presence of glycoproteins, proteoglycans and other ECM molecules^{380,382}. This leads to the formation of ECM hydrogel, a material with high water content and gel-like consistency³²⁸.

In the absence of the cryo-milling machine at the Institute of Cancer Research and other universities in London, homogenisation of the ECM scaffold to a fine powder for pepsin digestion proved difficult. As a result, other homogenisation methods were trialled, and I used decellularised murine lungs to optimise the protocol before applying it to patient samples. These included evaluating manual homogenisation methods such as using mortar and pestle and hand-held homogeniser. Firstly, decellularised scaffolds were dried in a SpeedVac vacuum concentrator set at 35 °C. Unfortunately, the homogenisation of dried ECM scaffold with mortar and pestle resulted in > 50% loss of the sample, which was unacceptable due to the limited and precious starting material. Next, I tried to homogenise dried ECM scaffolds directly in the acidified pepsin solution to reduce material loss. The homogenisation using a hand-held homogeniser produced a homogeneous solution; however, the solution did not gelate. Possibly, the homogenisation was too disruptive for the ECM structure (even at the low-speed setting), damaging the ECM molecules and preventing the gelation. Following these unsuccessful attempts, an extensive literature search was performed for an alternative homogenisation method. A protocol described by Nehrenheim *et al.* utilised bead-based homogenisation using a commercial Precellys machine instead of the cryo-milling procedure to produce a fine ECM powder suitable for pepsin digestion³⁹⁹. Bead-based homogenisation involves placing small beads within a tube along with the samples and subjecting the tube to vigorous shaking. This mechanical agitation disrupts and breaks down the samples, reducing particle size. As a result, Nehrenheim *et al.* protocol was modified to include an in-solution homogenisation (dried ECM scaffold in acidic pepsin solution) using Precellys to reduce the loss of the powdered ECM scaffold on the tube surfaces. As such, to obtain ECM hydrogels, dried decellularised ECM scaffolds were

homogenised in pepsin solution to obtain a pre-gel solution using the Precellys system, which, when subjected to physiological pH and temperature, gelled into hydrogels. After successfully optimising the protocol on mouse tissue, the protocol was applied on the ECM scaffolds obtained from 7 LMS tumours to generate LMS-specific ECM hydrogels (Figure 5.1).

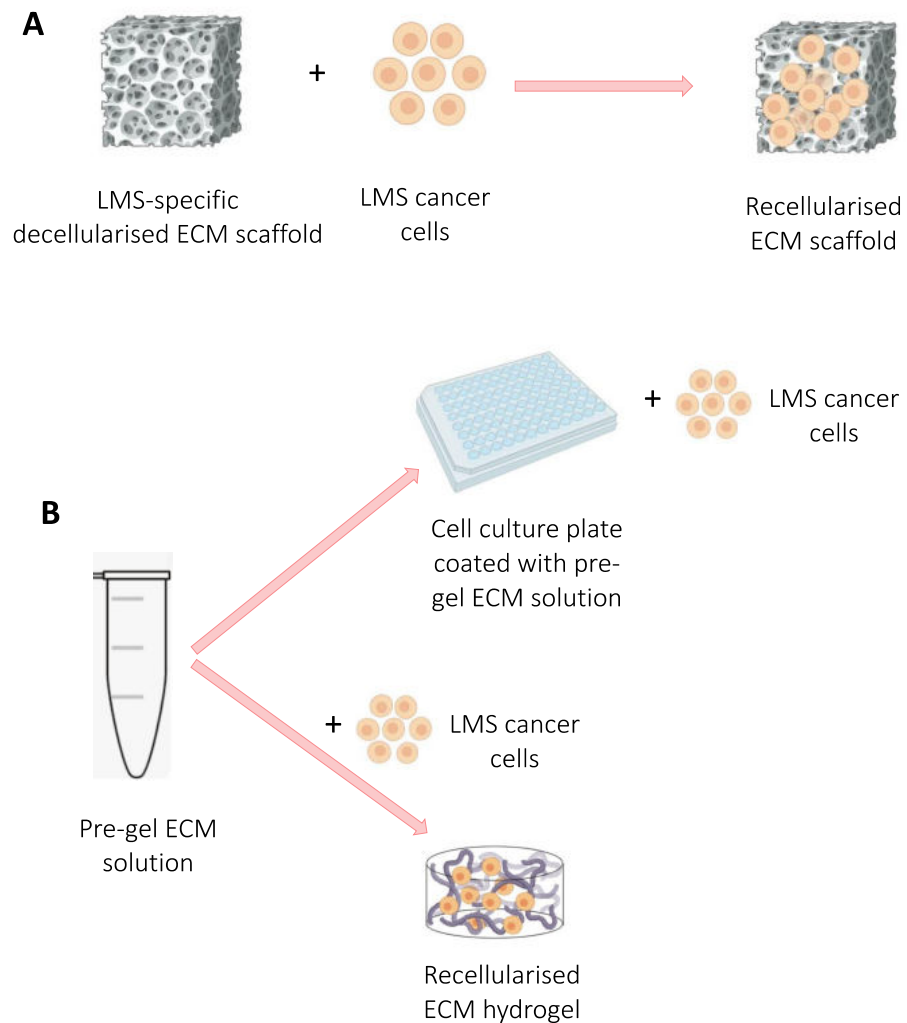


Figure 5.2 A workflow for research application of ECM scaffolds and pre-gel ECM solution.

A) Recellularisation of LMS decellularised ECM scaffold, B) Coating of cell culture plasticware to study the role of LMS cancer cells on a thin layer of ECM in 2D (top), encapsulation of LMS cancer cells and gelation of the pre-gel solution to produce recellularised ECM hydrogel for 3D cell culture. Figure created with BioRender.com.

As mentioned previously, one research application of decellularised ECM scaffolds involves recellularisation with cells of interest (Figure 5.2A). Regarding the ECM hydrogel, the pre-gel solution can be used to coat the plasticware to study the role of the ECM in 2D (Figure 5.2B). Alternatively, cells of interest can be resuspended in the pre-gel ECM solution, incubated at 37 °C to induce temperature-mediated gelation, and studied in 3D.

5.2.2. Characterisation of the composition of preclinical ECM models

To assess if the two derived preclinical ECM models retained components of the matrisome found in LMS patient tumours and to understand the limitations of the models, 7 matched sets of original tumours, ECM scaffolds and ECM hydrogels were analysed by mass spectrometry to evaluate their protein composition (Figure 5.1). The acquisition of proteomic data and initial data processing were performed by Dr Lukas Krasny. Across the 7 patient specimens, 3101, 2025 and 690 proteins were detected in the tumours, ECM scaffolds and ECM hydrogel, respectively (Table 5.1).

Table 5.1 Numbers of matrisome, non-matrisome and total number of protein identities (IDs) identified across 7 tumours, 7 scaffolds and 7 hydrogels.

An IDs was counted if it was detected in at least 1 of 7 samples. Each sample was run in two technical replicates, a protein ID was considered expressed in a sample if it was detected in at least 1 technical replicate.

Protein ID type	Tumour	Scaffold	Hydrogel
Matrisome	203	177	102
Non-matrisome	2898	1848	588
All	3101	2025	690

The proteomic data was mapped onto the human matrisome database¹⁷⁴. Using this approach, 203 matrisome proteins were identified across the 7 tumour specimens, 177 matrisome proteins in ECM scaffolds and 102 in ECM hydrogels (Table 5.1). All of the 102 matrisome proteins identified in ECM hydrogels were present in ECM scaffolds and originating LMS tumours (Figure

5.3A). These overlapping proteins included 35 glycoproteins, 22 regulators, 17 collagen chains, 15 ECM-affiliated proteins, 9 proteoglycans and 4 secreted factors. Forty-four of those proteins (43%) are known to localise to the basement membrane²⁵⁴ (Figure 5.3B).

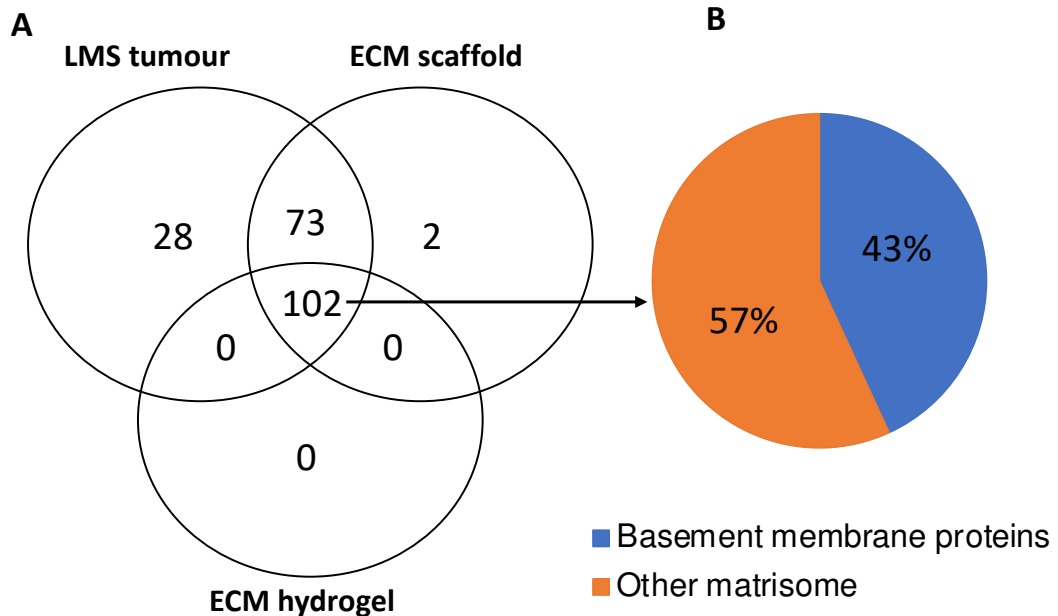


Figure 5.3 Leiomyosarcoma (LMS) ECM scaffolds and hydrogels retain matrisome components from their native counterparts.

(A) Venn diagram showing overlap of matrisome protein IDs detected across 7 tumours, 7 scaffolds and 7 hydrogels. (B) Pie chart showing the proportion of basement membrane proteins in the overlapped matrisome proteins.

This finding validated the results in Section 3.2.2, confirming that LMS tumours were rich in basement membrane proteins. Furthermore, it demonstrated that the key ECM biology of LMS was captured in ECM scaffold and ECM hydrogels and could be used for *in vitro* studies.

Decellularisation of tumours is known to preferentially reduce intracellular proteins over matrisomal proteins^{32,33}; therefore, the decellularised sample is predominantly enriched for matrisome proteins. In order to confirm that ECM scaffolds and ECM hydrogels comprised mainly of ECM proteins and not other cellular components, the percentage of relative matrisome abundance was calculated for each scaffold and hydrogel sample and compared to the

matched patient tumours (Figure 5.4A). Mean relative matrisome abundance in scaffolds was 79% (range: 76 – 88%) and 59% (range: 44 – 79%) in hydrogels, which were significantly higher (both $p < 0.0001$) than in the matched tumours with a median of 22% (range: 12 – 45%) (Figure 5.4B).

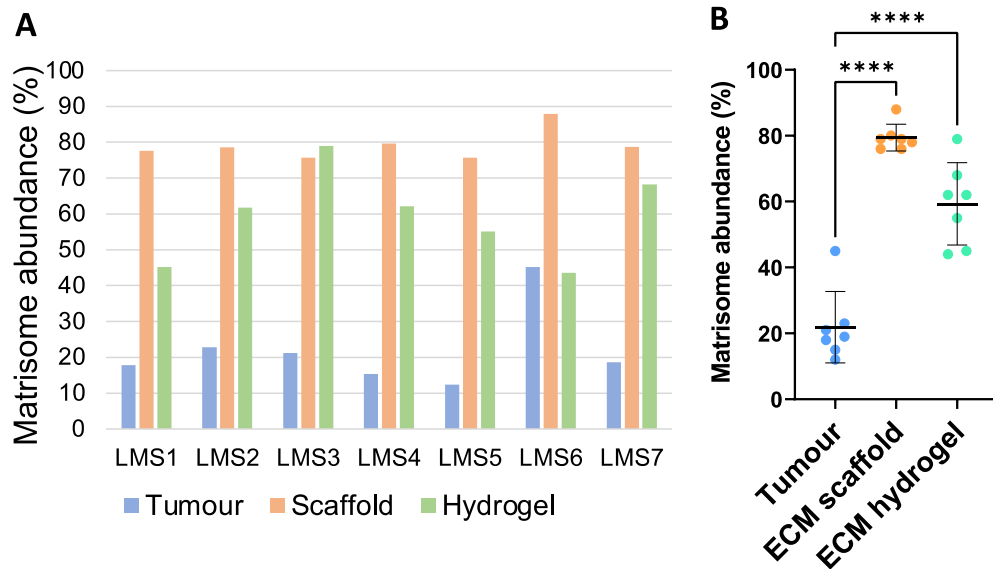


Figure 5.4 Decellularisation and hydrogel preparation enrich for matrisome content.

A) Bar charts showing % relative matrisome abundance in originating tumours, ECM scaffolds and ECM hydrogels for each leiomyosarcoma (LMS) patient LMS1-7. B) Percentage of relative matrisome abundance grouped by sample type (tumour, scaffold and hydrogel). Significance is shown following a one-way ANOVA with Tukey's Honest Significant Difference (HSD) tests; **** $p < 0.0001$.

Overall, this analysis showed that although non-matrisome proteins were present in ECM scaffolds and ECM hydrogels, their relative abundance was much lower compared to originating tumours, and the scaffold and hydrogel samples were predominantly comprised of matrisome proteins. This implies that cellular remnants or components within the ECM scaffold and ECM hydrogel samples could potentially affect the cellular phenotypes when used in functional experiments. Nevertheless, the data on the protein composition suggests that the dominant factor influencing cellular behaviour in the experiments is the ECM itself rather than any residual cellular proteins.

To evaluate the patterns of matrisome protein retention in LMS hydrogels, proteins were analysed by the matrisome classes: core matrisome (glycoproteins, collagens, proteoglycans) and matrisome-associated (affiliated proteins, ECM regulators and secreted factors). The protein ID counts identified in ECM scaffolds and hydrogels were compared to the originating tumours for each matrisome class (Figure 5.5).

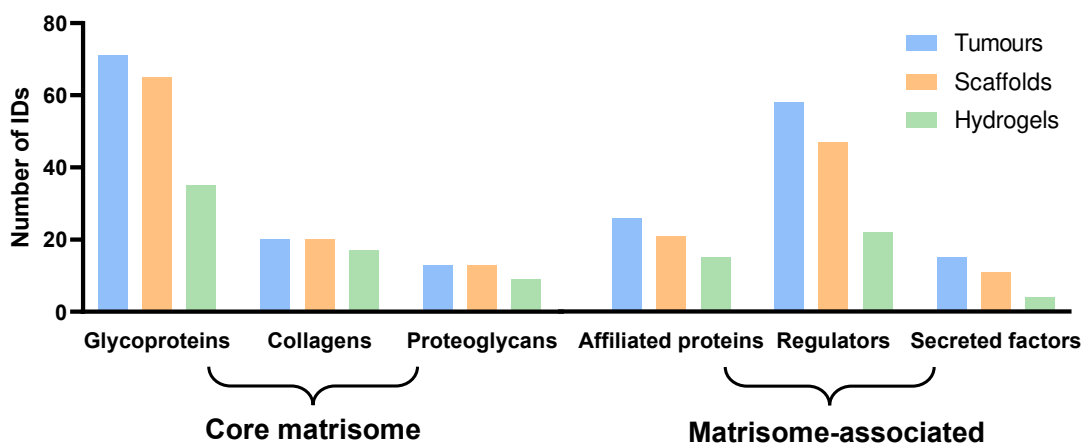


Figure 5.5 Decellularisation and ECM hydrogel preparation retain collagen chains.

Bar charts showing the breakdown of number of protein IDs in matrisome classes for leiomyosarcoma tumours, scaffolds and hydrogels.

After the decellularisation process, ECM scaffolds retained 92% of glycoproteins (65 out of 71 IDs), 100% of collagen chains (20 out of 20 IDs), and 100% of proteoglycans (13 out of 13 IDs) from the originating tumour. Thus, ECM scaffolds retained most of the core matrisome proteins (94%), consistent with previous reports^{368,560}. Compared to originating tumours, ECM hydrogels retained 49% of glycoproteins (35 out of 71 IDs), 85% of collagen chains (17 out of 20 IDs) and 69% of proteoglycans (9 out of 13 IDs). Overall, 59% of core tumour matrisome proteins were preserved in ECM hydrogels.

In terms of the matrisome-associated proteins, decellularisation retained 81% of ECM-affiliated proteins (21 out of 26 IDs), 81% of ECM regulators (47 out of 58 IDs) and 73% of secreted factors (11 out of 15 IDs) from the originating tumours. Therefore, 80% of tumour matrisome-associated proteins were detected in ECM scaffolds. The matrisome-associated proteins which were not

detected in ECM scaffolds after the decellularisation included ADAM9 and ADAM10, members of the Serpin protease inhibitors family (SERPINA6/7/10, SERPINB7), members of S100 family (S100A13 and S100B), cathepsins family members (CTSA and CTSL) and other secreted proteins. In contrast, ECM hydrogels preserved 58% of the ECM-affiliated proteins (15 out of 26 IDs), 38% of ECM regulators (22 out of 58 IDs) and 27% of secreted factors (4 out of 15 IDs) from the tumours, meaning that 41% of the original tumour matrisome-associated proteins were present in ECM hydrogels.

Next, I set out to understand which steps of the LMS ECM hydrogel preparation protocol might contribute to the observed decrease in matrisome proteins in ECM hydrogels compared to the originating tumour samples. A comparison of matrisome classes between ECM scaffolds and ECM hydrogels was performed to understand which matrisome proteins were detected in ECM scaffolds and not in ECM hydrogels. ECM hydrogels retained 62% of core matrisome proteins and 52% of matrisome-associated proteins detected in ECM scaffolds. Examining the core matrisome components, it was found that hydrogels retained 85% of collagen chains (17 out of 20), 69% of proteoglycans (9 out of 13) and 54% of glycoproteins (35 out of 65 IDs) identified in ECM scaffolds. Regarding the matrisome-associated proteins, ECM hydrogels preserved 71% of ECM-affiliated proteins (15 out of 21 IDs), 47% of ECM regulators (22 out of 47 IDs) and 36% of secreted factors (4 out of 11 IDs), which were detected in ECM scaffolds. As such, it is evident that additional losses of core matrisome and matrisome-associated proteins occurred during the preparation of ECM hydrogels from the dry decellularised ECM scaffolds. This suggests that the reduction of matrisome proteins in ECM hydrogel compared to LMS frozen tumour samples could not be solely attributed to the decellularisation process. Other factors, such as ECM scaffold homogenisation, pepsin digestion, and ECM pre-gel solution gelation, could play a role in the observed decline of matrisome protein identification.

Further analysis focused on identifying the consensus matrisome proteins shared across all 7 patients and consistently detected in all ECM scaffolds and all ECM hydrogels. Of all the matrisome proteins, 125 were found across

tumours, 77 were in ECM scaffolds, and 25 were in ECM hydrogels (Figure 5.6). Twenty-three matrixome proteins overlapped between the tumours, scaffolds and hydrogels. They included 9 collagen chains (COL1A1/2, COL3A1, COL4A1/2/3, COL6A1/2/3), 8 glycoproteins (FN1, LAMB2, LAMC1, DPT, EFEMP1, FGA, POSTN, TGFBI), two ECM-affiliated proteins (ANXA5 and LGALS1), two proteoglycans (HSPG2 and LUM) and two secreted factors (S100A11 and S100A6).

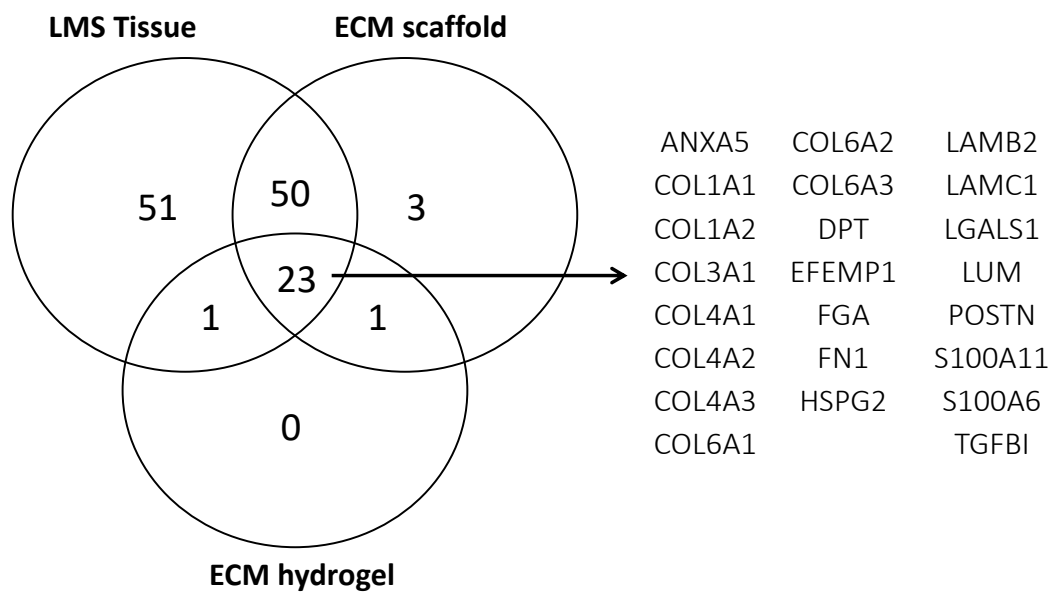


Figure 5.6 Consensus matrixome proteins in leiomyosarcoma (LMS).

Venn diagram showing overlap of matrixome protein IDs consistently detected in all tumours, scaffolds and hydrogels. Identities of 23 overlapping proteins are shown on the right.

To determine if the identified proteins are specifically expressed in LMS compared to other STS subtypes, the data from the STS FFPE tissue in Section 3.2.2 was analysed (Figure 5.7A-D). All 23 consensus matrixome proteins identified in the fresh frozen LMS samples and ECM models were expressed in the LMS FFPE samples. Thirteen out of 23 consensus matrixome proteins were significantly expressed at higher levels in LMS versus other STS subtypes. These proteins included collagen type VI chains (COL6A1/2/3) (Figure 5.7A), heparan sulfate proteoglycan 2 (HSPG2) (Figure 5.7B), collagen type IV alpha-1 chains (COL4A1/2) (Figure 5.7C), fibronectin (FN1) (Figure 5.7D), laminin subunit beta-2 (LAMB2) and laminin subunit gamma-1

(LAMC1) (Figure 5.7E). As expected and previously discussed in Chapter 3, Collagen I chains were not significantly enriched in LMS but rather were enriched in DES (Figure 5.7F).

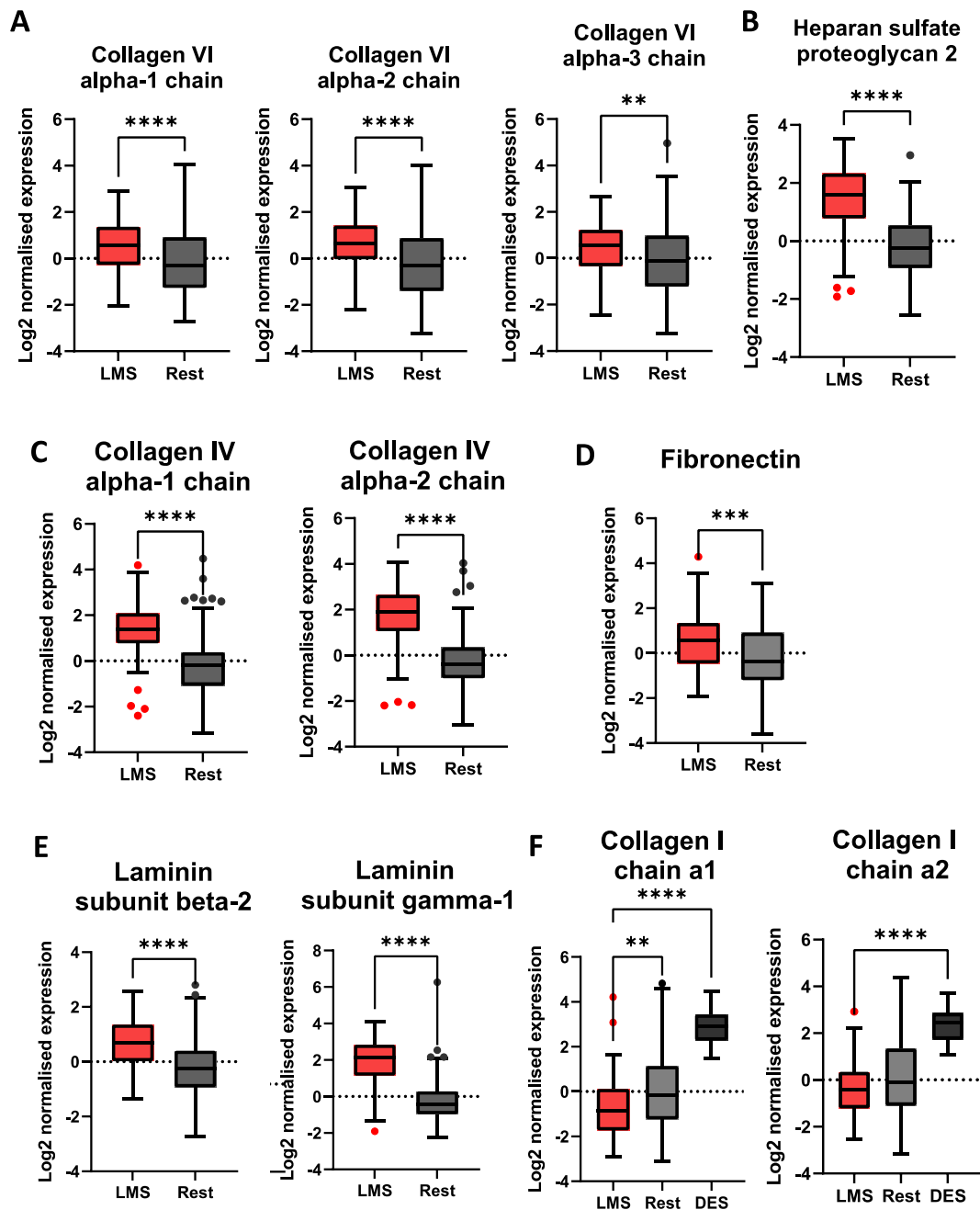


Figure 5.7 Consensus matrisome proteins are leiomyosarcoma (LMS) specific.

(A-F) Box plots showing Log₂ normalised protein expression of A) collagen VI chains, B) heparan sulfate proteoglycan 2, C) collagen IV chains, D) fibronectin, E) laminin chains in leiomyosarcoma (LMS) (n = 80) vs rest of STS subtypes (n = 241), F) collagen I chains in LMS (n = 80) vs DES (n = 37) and other STS subtypes (n = 204). Significance is shown following

Mann–Whitney U test for (A-E), or Kruskal Wallis tests with Dunn’s multiple testing correction for (F), ** $p < 0.01$; **** $p < 0.0001$. DES: desmoid tumours.

5.2.3. The biological role of LMS ECM

The biochemical and biophysical properties of the ECM play a pivotal role in virtually all biological processes, impacting cell migration, invasiveness, and drug resistance, thereby significantly contributing to cancer progression⁴¹⁰. It is widely established that the interaction between cells and ECM proteins, predominantly mediated by the integrin receptors, is a significant factor that modulates the response to various therapeutic approaches in oncology, including chemotherapy, targeted therapies, and other treatments like immunotherapy and radiation therapy^{561–564}. For instance, growing cancer cells on purified or cell-derived ECM coatings in 2D compared to plastic differentially modulated chemotherapy response and resistance in carcinoma models^{541,565,566}. To investigate the role of LMS ECM in drug response, cell adhesion and drug testing were performed using an established LMS cell line, SK-UT-1, to investigate the biological function of LMS ECM.

Furthermore, the ECM has been shown to promote the migration of carcinoma and sarcoma cells^{567–570}. Therefore, I also evaluated the cell migration phenotype in a panel of LMS cell lines, including established lines (SK-UT-1 and SK-UT-1b) and patient-derived (ICR-LMS-1, SHEF-LMS w1 and SHEF-LMS ws) lines (Table 5.2). SK-UT-1 and SK-UT-1b are morphological variants isolated from the same LMS patient^{401,402}. Similarly, SHEF-LMS w1 and SHEF-LMS ws are two distinct morphologic LMS cell populations established from the other LMS patient³⁸³. Given that LMS ECM is patient-derived and, therefore, is a limited resource, dried LMS ECM scaffolds from 7 LMS patients were pooled (in equal mass ratio) and digested with pepsin to produce a combined pre-gel LMS ECM solution (referred to as LMS ECM), which was used to coat plasticware for phenotypic assays.

Table 5.2 Summary of established and patient-derived uterine leiomyosarcoma cell lines.

Overview of characteristics of uterine leiomyosarcoma (LMS) cell lines, including LMS subtype, originating site, Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC) grade or International Federation of Gynecology and Obstetrics (FIGO) stage of originating tumour/biopsy. * and ** indicate that the cell lines are morphological variants derived from the same patient.

Cell line ID	Site of biopsy /resection	Primary site	Grade	FIGO stage	Source
SK-UT-1*	Uterus	Uterus	3	-	Dr Priya Chudasama, German Cancer Research Centre
SK-UT-1b*	Uterus	Uterus	3	-	Dr Priya Chudasama, German Cancer Research Centre
ICR-LMS-1	Pelvis	Uterus	3	-	Derived in Huang laboratory
SHEF-LMS w1**	Pelvis	Uterus	-	IIA	Dr Karen Sisley, University of Sheffiled
SHEF-LMS ws**	Pelvis	Uterus	-	IIA	Dr Karen Sisley, University of Sheffiled

5.2.3.1. Adhesion of SK-UT-1 cells to LMS ECM

Cells interact with the ECM via several types of receptors, mainly via integrin receptors, whose activation enables the formation of cell adhesions⁵⁷¹. To assess the adhesion properties of SK-UT-1 cells to LMS ECM, the cells were plated on plastic dishes or plastic pre-coated with a thin layer of LMS ECM and allowed to adhere for 1 h. Quantification of the adhesion assay showed that at 1h, significantly (p -value < 0.0001) more SK-UT-1 cells adhered to plastic (mean = 3317 cells) compared to LMS ECM (mean = 296 cells) (Figure 5.8).

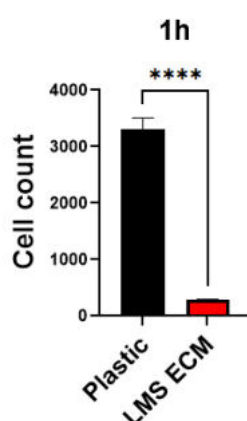


Figure 5.8 Adhesion of SK-UT-1 cells to plastic and leiomyosarcoma derived pre-gel extracellular matrix solution (LMS ECM).

Bars represent the mean number of cells adhering to the bottom of 96-well plastic plates or plastic plates pre-coated with LMS ECM after 1 h (n = 3). The mean and standard error of the mean (SEM) are shown. Significance is shown following an unpaired t -test; **** p < 0.0001.

5.2.3.2. Drug response of SK-UT-1 cells grown on LMS ECM

Cell density in a monolayer is known to affect drug resistance⁵⁷². Given that I observed reduced adhesion of SK-UT-1 to LMS ECM compared to plastic at 1h, the differences in the drug responses could be caused by different cell numbers at the start of the assay. To ensure that the number of SK-UT-1 cells at the beginning of the assay was comparable on plastic and LMS ECM, the cells were allowed to adhere for 24 h. Then, the plates were washed with 0.2% DMSO in media, and the plate was fixed and stained (Figure 5.9A). The number of adherent cells on LMS ECM was not statistically different to the number of cells on plastic. The growth curve analysis over 7 days showed that the growth of SK-UT-1 cells on LMS ECM was similar to growth on plastic (Figure 5.9B).

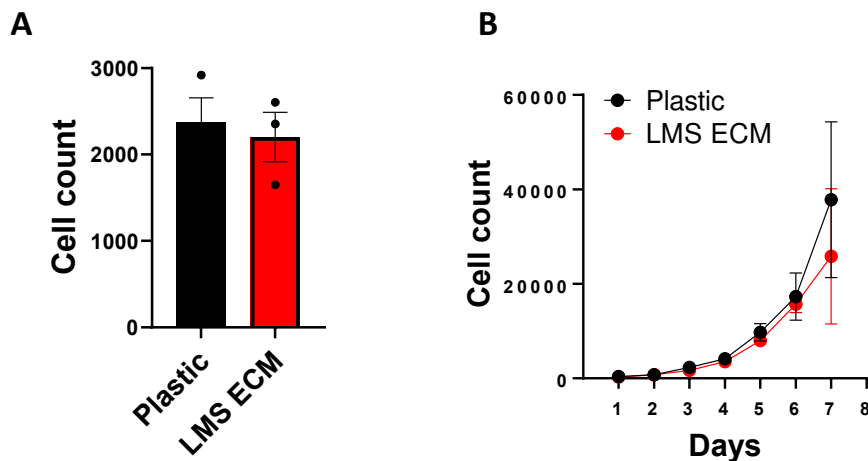


Figure 5.9 SK-UT-1 cell adhesion at 24 hour and cell growth on plastic compared to leiomyosarcoma pre-gel extracellular matrix (LMS ECM).

A) Bar charts showing the number of SK-UT-1 cells adhering to plastic and plastic pre-coated with LMS ECM before the drug addition. The mean and standard error of the mean are shown ($n = 3$). Significance was tested with an unpaired t -test. B) Growth curves of SK-UT-1 cells plated on plastic and LMS ECM over 7 days ($n = 2$). The mean and standard deviations are shown.

A small molecule inhibitor screen was used to evaluate whether the LMS ECM can modulate drug response in the SK-UT-1 model compared to plastic. SK-UT-1 cells were plated on plastic plates or plastic with pre-coated LMS ECM. The cells were treated (500 nM) with a panel of 48 small molecule inhibitors (Table 2.2). The screen comprised mostly kinase inhibitors targeting major cellular signalling pathways associated with cancer survival and progression. Additionally, the screen also contained non-kinase inhibitors targeting heat shock protein 90 (Hsp90) (luminespib (NVP-AUY922)), STAT (niclosamide and SH-4-54), integrins (cilengitide trifluoroacetate), B-cell lymphoma 2 protein (Bcl-2) (navitoclax), and poly (adenosine diphosphate(ADP)-ribose) polymerase (PARP) (rucaparib and talazoparib), which are undergoing preclinical and/or clinical assessment in cancer. The cell viability was assessed after 72 hours by direct cell counting.

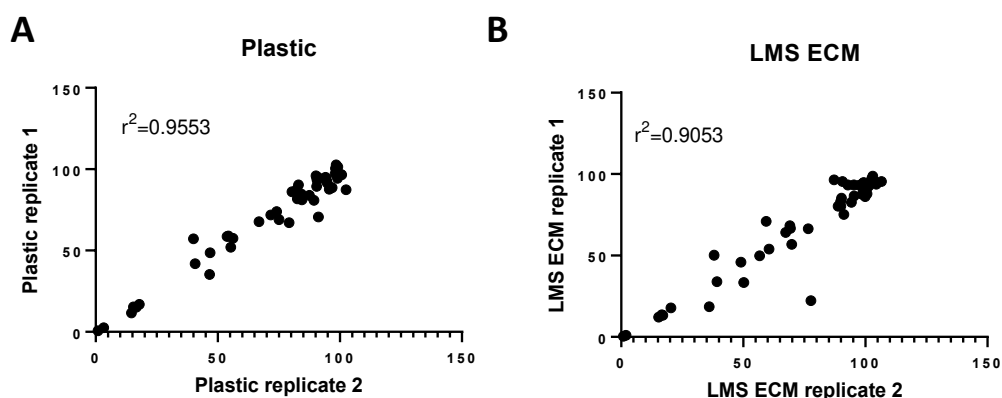


Figure 5.10 Pearson's correlation analysis of biological replicates.

A) Plastic and B) LMS ECM. Y-axes show the results of biological replicate 1 normalised to DMSO, and the x-axes show the results of biological replicate 2 normalised to DMSO. DMSO; Dimethyl sulfoxide.

Pearson's correlation analysis assessed the reproducibility of biological replicates ($n = 2$) within the small molecule inhibitor screens (Figure 5.10). Pearson's analysis showed a strong correlation ($r^2 > 0.90$) between biological replicates, indicating high reproducibility. The unsupervised clustering of normalised cell viability data showed that SK-UT-1 responses on plastic were similar to LMS ECM responses for most drugs (Figure 5.11).

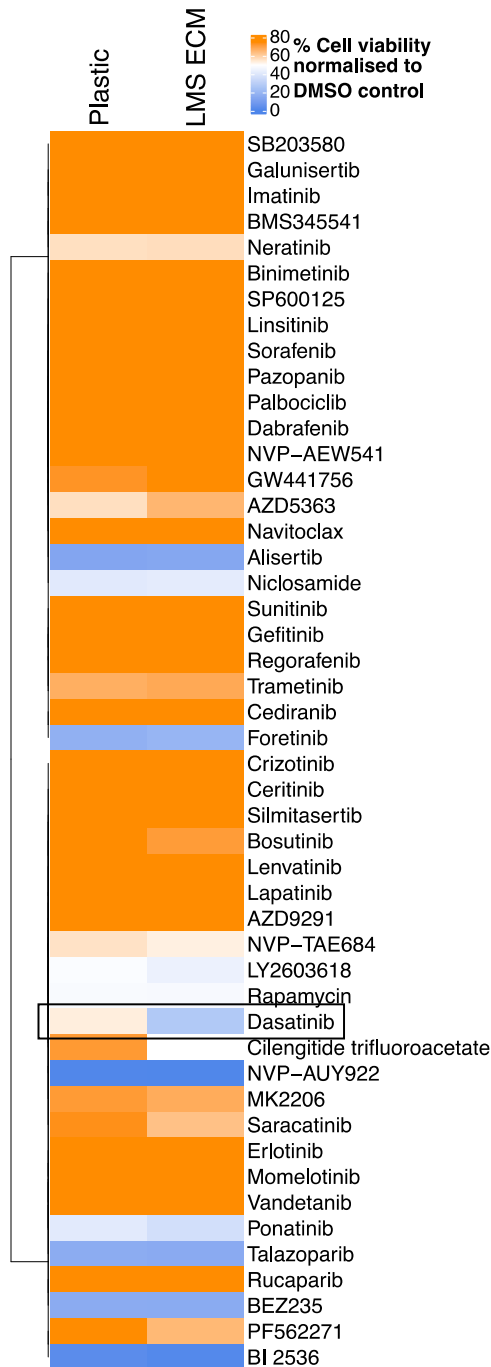


Figure 5.11 Small molecule inhibitor screen of SK-UT-1 on plastic and leiomyosarcoma pre-gel extracellular matrix (LMS ECM).

A hierarchically clustered heatmap showing average normalised percentage cell viability on plastic and LMS ECM after 72 h treatment with 500 nM drug (or 50 nM for NVP-AUY922). Viability values for plastic and LMS ECM were normalised to their own DMSO values (n=2). DMSO; Dimethyl sulfoxide.

A multi-tyrosine kinase inhibitor dasatinib showed the highest percentage change in viability compared to plastic (a 26% decrease).

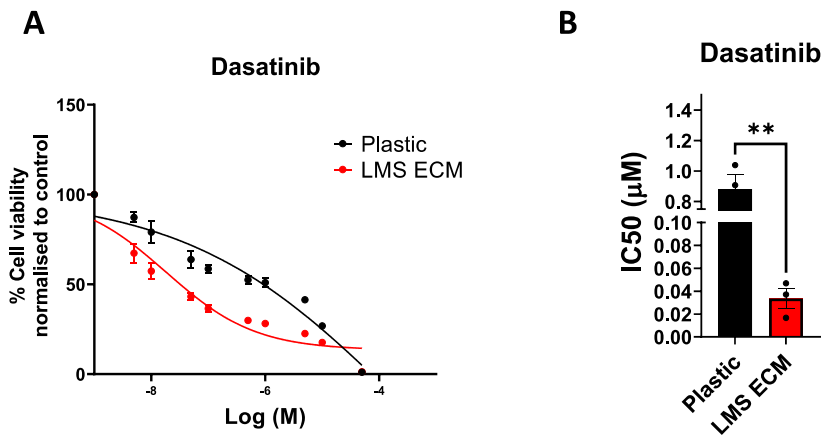


Figure 5.12 Confirmation of SK-UT-1 sensitivity to dasatinib on leiomyosarcoma pre-gel extracellular matrix (LMS ECM).

A) Cell viability assays of SK-UT-1 cells on plastic and LMS ECM treated with increasing doses of dasatinib, B) Bar plots displaying IC₅₀ values for dasatinib, Viability values for plastic and LMS ECM were normalised to their own DMSO values (n=3). DMSO; Dimethyl sulfoxide. Significance is shown following an unpaired *t*-test; ** *p* < 0.01.

To confirm the increased sensitivities of SK-UT-1 cells on LMS ECM to dasatinib, full dose-response cell viability assays (Figure 5.12) and colony formation assays were performed (Figure 5.13). Cells on LMS ECM (IC₅₀ = 0.03 µM) were more sensitive to dasatinib and had significantly (*p*-value = 0.001) lower IC₅₀ values than plastic (IC₅₀ = 0.88 µM) (Figure 5.12A,B). Consistent with the short-term assay, in the long-term colony formation assay, SK-UT-1 cells on LMS ECM were significantly more sensitive to dasatinib than on plastic across 50 nM (*p*-value = 0.022), 100 nM (*p*-value = 0.027), and 1000 nM (*p*-value = 0.006) dasatinib concentrations (Figure 5.13A, B).

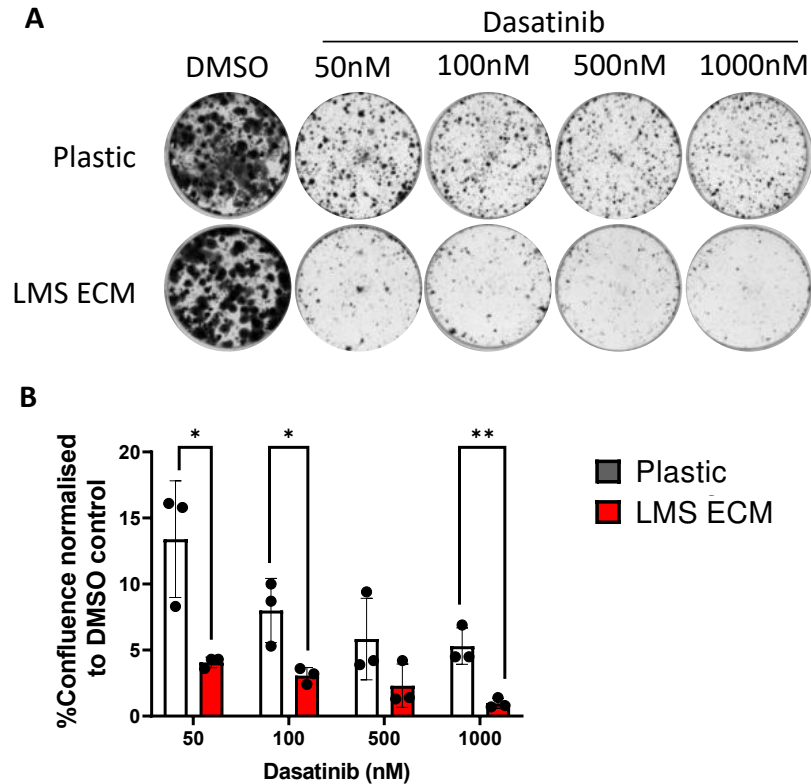


Figure 5.13 Confirmation of SK-UT-1 sensitivity to dasatinib on leiomyosarcoma pre-gel extracellular matrix (LMS ECM).

A) Representative images of crystal violet stained wells from the colony formation assay of SK-UT-1 cells on plastic and LMS ECM treated with increasing doses of dasatinib. B) Mean % cell confluence on plastic and LMS ECM was normalised to their own DMSO values (n=3) DMSO; Dimethyl sulfoxide. The mean and standard error of the mean (SEM) are shown. Significance is shown following an unpaired *t*-test for each concentration; * $p < 0.05$; ** $p < 0.01$.

5.2.3.3. Migration of LMS cells on LMS ECM

The metastatic potential of cancer cells is profoundly influenced by their ability to migrate⁵⁷³. The ECM was shown to promote the migration of carcinoma as well as sarcoma cells^{567–570}. Given that LMS is characterised by high metastasis rates, I hypothesised that LMS ECM could promote cell migration in LMS. As such, this section investigated the role of LMS ECM in the 2D migration of LMS cells.

Baseline migratory characteristics of LMS cells

To migrate effectively, cancer cells are required to have persistent movement in a specific direction⁵⁷⁴. Here, migration properties of LMS cells in 2D were quantified with two primary metrics: cell migration speed and directionality (persistence). The directionality index, also called persistence, is the Euclidean distance between the start and finish divided by the total distance travelled⁵⁷⁵ (Figure 5.14B). It informs whether a cell follows a relatively straight path (index closer to 1) or has a confined migration and follows a random trajectory (index closer to 0).

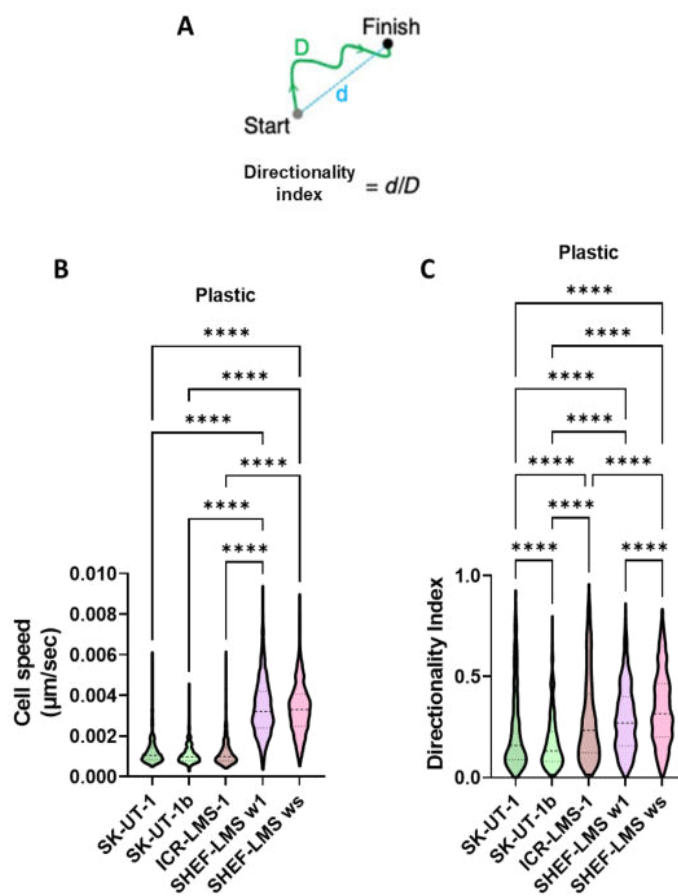


Figure 5.14 Baseline migration properties of leiomyosarcoma (LMS) cell lines.

A) a diagram, explaining the directionality index, adapted from Gorelik and Gautrea (2014), with permission from Springer Nature⁴⁰⁸. B) Violin plots showing the distribution of 2D migration cell speed on plastic in GFP+ SK-UT-1 (n = 910), SK-UT-1b (n = 495), ICR-LMS-1 (n = 528), SHEF-LMS w1 (n = 832) and SHEF-LMS ws (n = 887) over 18 h, C) Violin plots show the directionality index on plastic in the 5 GFP+ LMS cell lines. The cells were tracked for 18 h. Data were pooled from three independent experiments for all graphs. Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile-(1.5*

interquartile range (IQR)) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Kruskal Wallis tests with Dunn's multiple testing correction, **** p < 0.0001.

To understand the baseline migratory properties of established LMS cell lines (SK-UT-1 and SK-UT-1b) and patient-derived cell lines (ICR-LMS-1, SHEF-LMS w1 and SHEF-LMS ws), all 5 cell lines were transduced with H2B-GFP construct to enable nuclei tracking over time. The GFP+ LMS cell lines were monitored by time-lapse microscopy over 18 h and tracked to determine their migration speed and directionality when plated on plastic (Figure 5.14).

Cell speed

In terms of cell speed, SHEF-LMS w1 (median = 0.0032 $\mu\text{m}/\text{sec}$) cells showed significantly faster migration compared to ICR-LMS-1 (median = 0.0010 $\mu\text{m}/\text{sec}$; p-value < 0.0001), SK-UT-1 (median = 0.0011 $\mu\text{m}/\text{sec}$; p-value < 0.0001) and SK-UT-1b (median = 0.0010 $\mu\text{m}/\text{sec}$; p-value < 0.0001) (Figure 5.14A). Similarly, SHEF-LMS ws (median = 0.0033 $\mu\text{m}/\text{sec}$) migrated significantly faster than ICR-LMS-1 (p-value < 0.0001), SK-UT-1 (p-value < 0.0001) and SK-UT-1b (p-value < 0.0001) (Figure 5.14A). As such, I considered ICR-LMS-1, SK-UT-1 and SK-UT-1b intrinsically 'slow-migrating', while SHEF-LMS w1 and SHEF-LMS ws were 'fast-migrating'.

Directionality

Intrinsically, every LMS cell line had significantly different mean directionality indices than every other cell line, apart from comparing ICR-LMS-1 and SHEF-LMS w1. All the significant pair-wise comparisons were at p-value < 0.0001. SK-UT-1b cells had the lowest directionality index (median = 0.14), meaning these cells were displaying the most random migration relative to others. SK-UT-1b were followed by SK-UT-1 (median = 0.16), ICR-LMS-1 (median = 0.23), SHEF-LMS w1 (median = 0.27), whilst SHEF-LMS ws cells had the highest index (median = 0.32), suggesting that relative to others, SHEF-LMS ws displayed the most directionally persistent (oriented in a single direction) migration (Figure 5.14C).

Influence of LMS ECM matrix on cell speed

The speed of LMS cells migrating on LMS ECM coating was compared to the speed on plastic to assess if LMS ECM increased the speed of LMS cells. A more than 15% increase in the median cell speed (and p-value < 0.05) compared to plastic was considered a significant change. Compared to plastic, the median speed of SHEF-LMS w1 cells was 74% higher on LMS ECM (p-value < 0.0001), 22% higher in SK-UT-1 cells (p-value < 0.0001) and 16% higher in SHEF-LMS ws (p-value < 0.0001) (Figure 5.15A-C). LMS ECM coating did not induce a meaningful increase in the cell speed of ICR-LMS-1 (11% increase; p-value = 0.0080) and SK-UT-1b (5% increase; p-value = 0.0450) cells (Figure 5.15D-E).

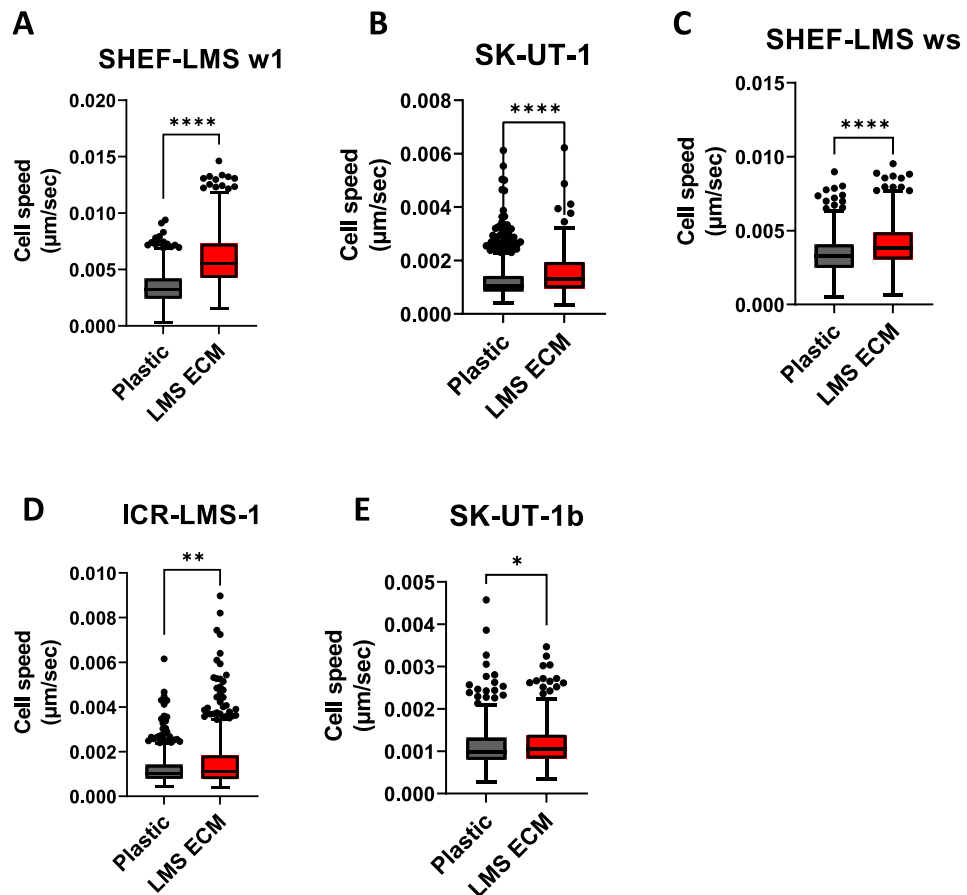


Figure 5.15 Average cell speed of leiomyosarcoma (LMS) cell lines on leiomyosarcoma pre-gel extracellular matrix (LMS ECM) and plastic.

Box plots showing the cell speed of GFP+ LMS cell lines when grown in 2D conditions A) SHEF-LMS w1, B) SK-UT-1, C) SHEF-LMS ws, D) ICR-LMS-1 and E) SK-UT-1b. The GFP+ cells were tracked for 18 h. Data were pooled from three independent experiments for all

graphs. Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile-(1.5* interquartile range (IQR)) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Mann–Whitney *U* tests, * $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$.

Influence of LMS ECM on cell directionality

To determine whether LMS ECM promoted the directionality of LMS cell migration, the directionality index of migration on the LMS ECM coating was compared to plastic (Figure 5.16). Similarly to the speed threshold, a > 15% change in the median directionality compared to plastic (and p -value < 0.05) was considered substantial. The LMS ECM induced a 36% increase in the median directionality of ICR-LMS-1 cells (p -value < 0.0001), a 67% increase in SK-UT-1 (p -value < 0.0001) and an 18% in SK-UT-1b (p -value = 0.0010) compared to plastic (Figure 5.16A-C). In contrast, the LMS ECM coating did not affect the directionality of the ‘fast-migrating’ SHEF-LMS w1 (6% increase; p -value = 0.0480) and SHEF-LMS ws (9% increase; p -value = 0.0040) cells (Figure 5.16D-E).

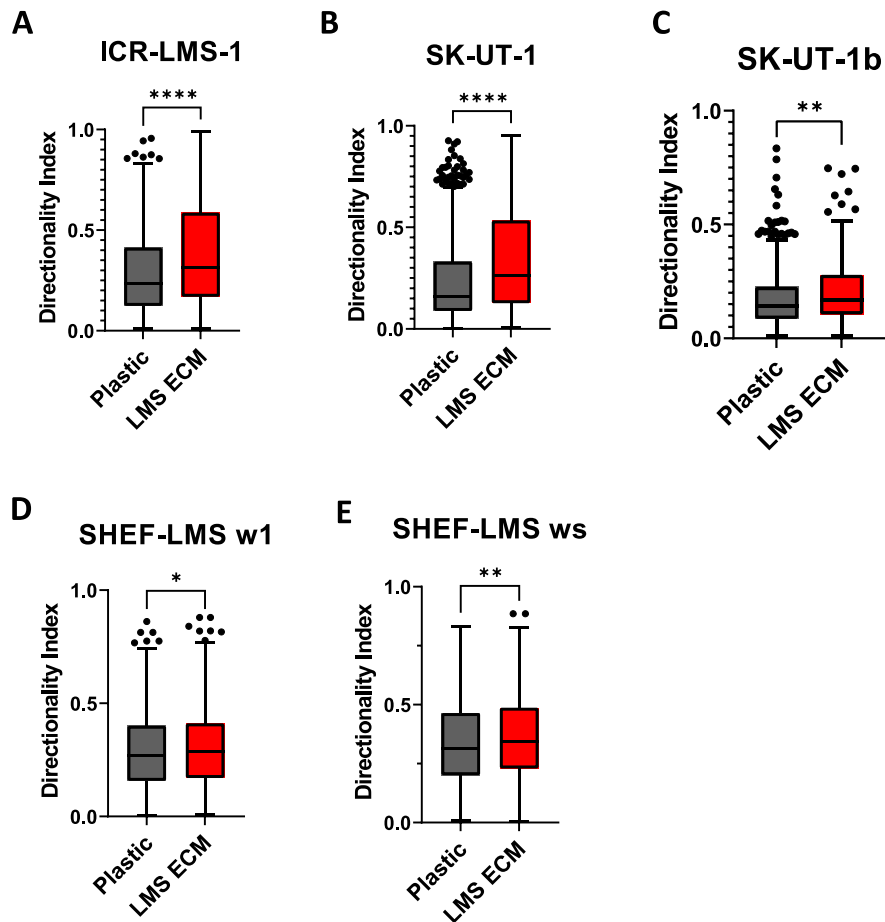


Figure 5.16 The directionality of leiomyosarcoma (LMS) cell lines on leiomyosarcoma pre-gel extracellular matrix (LMS ECM) and plastic.

Box plots showing the distribution of directionality indices in GFP+ A) ICR-LMS-1, B) SK-UT-1, C) SK-UT-1b, D) SHEF-LMS w1 and E) SHEF-LMS ws. The GFP+ cells were tracked for 18 h. Data were pooled from three independent experiments for all graphs. Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile-(1.5*IQR) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Mann–Whitney *U* tests, * $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$.

5.2.3.4. Comparison of LMS migration on LMS ECM and commercially purified matrices

Various ECM components can engage distinct cell surface receptors²⁹¹ and can differentially influence migration^{576–579}. For instance, a lung adenocarcinoma cell line preferentially uses $\alpha 3\beta 1$ integrin to bind to laminin and $\alpha 5\beta 1$ integrin to bind to fibronectin, leading to the activation of differential intracellular signalling pathways with opposite effects on the cell migration⁵⁷⁷.

This section investigated whether the impact of LMS ECM on cell migration differed from that of commonly employed ECM coatings like fibronectin, laminin, and collagen type IV.

Table 5.3 Comparison of the cell speed of leiomyosarcoma (LMS) cell lines plated on a coating of leiomyosarcoma pre-gel extracellular matrix (LMS ECM) and commercial purified matrices.

Summary showing % changes in the median cell speed in ICR-LMS-1, SK-UT-1, SK-UT-1b, SHEF-LMS w1 and SHEF-LMS ws plated on fibronectin, laminin and collagen type IV compared to LMS ECM. Data were pooled from three independent experiments. The key shows colour coding to indicate the significance and the direction of the change. Significant change is defined as > 15 % change in the median cell speed compared to LMS ECM at p-value < 0.05.

	Fibronectin	Laminin	Collagen IV	Key
ICR-LMS-1	2	11	4	
SK-UT-1	48	7	13	
SK-UT-1b	26	38	6	
SHEF-LMS w1	8	13	2	
SHEF-LMS ws	6	6	2	

	significantly increased compared to LMS ECM
	significantly decreased compared to LMS ECM
	not significant

Cell speed

To compare the LMS cell speed on different ECM coatings, ICR-LMS-1, SK-UT-1, SK-UT-1b, SHEF-LMS w1 and SHEF-LMS ws GFP+ cells were plated on plastic dishes pre-coated with LMS ECM pre-gel, fibronectin, laminin and collagen IV and the GFP+ cells were tracked for 18 h (Table 5.3; Supplementary Figure 5.1). Fibronectin induced a 48% increase in the median cell speed of SK-UT-1 cells (p-value < 0.0001) and a 26% increase in SK-UT-1b cells (p-value < 0.0001) compared to LMS ECM. There was a 38% increase in the SK-UT-1b (p-value < 0.0001) cell speed on laminin compared to LMS ECM. LMS cell speeds on collagen IV were similar to those on LMS ECM.

Table 5.4 Comparison of the cell directionality of leiomyosarcoma (LMS) cell lines plated on a coating of leiomyosarcoma pre-gel extracellular matrix (LMS ECM) and commercial purified matrices.

Summary showing % changes in the median directionality indices in ICR-LMS-1, SK-UT-1, SK-UT-1b, SHEF-LMS w1 and SHEF-LMS ws plated on fibronectin, laminin and collagen type IV compared to LMS ECM. Data were pooled from three independent experiments. The key shows colour coding to indicate the significance and the direction of the change. Significant change is defined as > 15 % change in the median directionality index compared to LMS ECM at p-value < 0.05.

	Fibronectin	Laminin	Collagen IV	Key
ICR-LMS-1	2	55	4	
SK-UT-1	24	42	1	
SK-UT-1b	34	8	13	
SHEF-LMS w1	2	17	0	
SHEF-LMS ws	1	23	7	
				significantly increased compared to LMS ECM
				significantly decreased compared to LMS ECM
				not significant

Directionality

To compare the directional persistence of LMS cells during migration on different ECM coatings, the change in the median directionality indices was calculated for each LMS cell line compared to LMS ECM (Table 5.4; Supplementary Figure 5.2). Fibronectin promoted a 34% increase in the directionality index of SK-UT-1b (p < 0.0001) compared to the LMS ECM. In contrast, laminin significantly reduced the directionality of 4/5 LMS cell lines compared to LMS ECM, implying that LMS cells on laminin show more random motion in terms of trajectories than the cells on LMS ECM. Apart from SK-UT-

1b, there was a 55% reduction in the directionality index of ICR-LMS-1 ($p < 0.0001$), a 42% decrease in SK-UT-1 ($p < 0.0001$), a 17% reduction in SHEF-LMS w1 ($p < 0.0001$) and a 23% decrease in SHEF-LMS ws ($p < 0.0001$) on laminin compared to LMS ECM. The directional persistence of all LMS cells was not affected by collagen IV compared to LMS ECM.

5.3. Discussion

Chapter 5 described the optimisation of a pipeline for a successful generation of preclinical ECM models (ECM scaffolds and ECM hydrogels) from LMS fresh frozen tumours. The composition of ECM scaffolds and ECM hydrogels were comprehensively analysed using mass spectrometry and compared to the native specimens. The effect of incorporating LMS ECM pre-gel coatings in migration and drug screening assays was assessed.

5.3.1. Generation and characterisation of preclinical ECM models

Both decellularised ECM scaffold and ECM hydrogel models retained the core structural matrix proteins of the originating LMS tumours. Almost half (43%) of the matrix proteins retained in both the scaffolds and hydrogels were constituents of the basement membrane. This is consistent with the results obtained from LMS FFPE samples in Section 3.2.2 and serves as an orthogonal validation, confirming that the LMS tumours are rich in basement membrane proteins.

5.3.1.1. Decellularised ECM scaffolds

Post decellularisation with SDS detergent, LMS ECM scaffolds retained all collagen chains and all proteoglycans, compared to the originating tumours. Overall, ECM scaffolds preserved 94% of the tumour core matrix proteins. This aligns with previous reports demonstrating retention of the core matrix proteins in decellularised samples^{560,580–582}. For example, a study by Calle *et al.* showed nearly complete preservation of collagen and laminin chains in SDS/Triton decellularised murine lungs compared to native specimens⁵⁸¹. In contrast, ECM scaffolds retained 80% of the matrix-

associated protein from the originating tumours. Several studies reported a loss of those components during decellularisation protocols^{368,580,583–585}. A study by Schiller *et al.* extracted matrisome proteins from murine lungs by employing sequential extractions with detergents of increasing stringency to obtain three soluble fractions and the last fraction with insoluble matrisome proteins⁵⁸⁵. The study revealed that almost half of murine lung matrisome-associated proteins were found in the first highly soluble fraction and not an insoluble fraction at the end of the protocol⁵⁸⁵. The study suggested that this differential retention of matrisome-associated proteins is because most matrisome-associated proteins are not tightly bound to the core structural ECM network and, thus, can be readily washed away during incubations with a detergent solution. In this study, proteins absent from the insoluble fraction included members of ADAM, Serpin, S100 and cathepsin families. This is consistent with the observed loss of some of these members after the decellularisation procedure in my ECM scaffolds. However, whilst Schiller *et al.* study reported a loss of annexin (ANXA) family proteins, 7 ANXA proteins were preserved in my ECM scaffolds. This discrepancy could be due to a more aggressive ECM extraction procedure that Schiller *et al.* study employed. Overall, the ECM scaffolds generated in this thesis represent an attractive preclinical ECM model, capturing LMS-specific core matrisome proteins and most matrisome-associated proteins. ECM scaffolds would allow studying LMS pathobiology in an environment closely recapitulating the native ECM TME of LMS tumours.

5.3.1.2. ECM hydrogels

Chapter 5 discussed the optimisation of ECM scaffold homogenisation to enable pepsin digestion and generation of ECM pre-gel, which solidified under 37 °C and pH 7 to ECM hydrogel. The lack of cryo-milling equipment, standardly used for ECM scaffold homogenisation, posed a significant challenge for generating pre-gel ECM capable of gelation. This challenge was addressed by introducing an alternative method involving an in-solution digestion (with acidified pepsin solution) using a bead-based homogenisation

system. In this alternative method, mechanical disruption homogenised the ECM scaffolds into suitable sizes for pepsin cleavage.

Mass spectrometry characterisation found that 59% of the core matrisome and 41% of matrisome-associated proteins were conserved in LMS ECM hydrogels compared to native tumour specimens. The losses of matrisome proteins might have occurred during various stages of the LMS ECM hydrogel preparation, including the LMS tumour decellularisation, ECM scaffold homogenisation, pepsin digestion, and ECM pre-gel solution gelation processes. Compared to ECM scaffolds, there was a decrease in core matrisome, with 46% of glycoproteins lost, and a decrease in matrisome-associated, with 53% ECM regulators and 64% secreted factors lost in ECM hydrogels. This implies that the SDS decellularisation could not fully account for the matrisome protein losses in ECM hydrogels compared to originating tumour samples. It is possible that the ECM scaffold homogenisation, pepsin digestion, and gelation could have additionally contributed to the observed decrease of matrisome protein identification. Furthermore, this finding highlights that the matrisome protein composition of decellularised ECM scaffolds does not directly translate into the composition of ECM hydrogels. A few studies characterised the protein composition of solidified ECM hydrogels by mass spectrometry^{399,586}; however, one study lacked the comparisons with decellularised tissue or native counterparts³⁹⁹. Simsa *et al.* study used mass spectrometry to characterise the matrisome protein composition of porcine brain, decellularised brain and brain hydrogel⁵⁸⁶. The authors reported a loss of nearly all glycoproteins, proteoglycans and a complete loss of ECM regulators in brain hydrogel compared to decellularised brain samples⁶², corroborating an observation of substantial loss of glycoproteins and ECM regulators in my LMS ECM hydrogels. Other studies on tissue decellularisation and ECM hydrogels characterised only ECM scaffolds and not ECM hydrogels, and the authors tended to assume that the ECM scaffolds' protein composition directly transfers to the protein composition in the ECM hydrogels^{355,587,588}. The result of this chapter suggests otherwise, highlighting the importance of characterising the products of every step during the ECM

hydrogel preparation protocol. Nevertheless, more studies will be necessary to confirm my finding.

5.3.1.3. Limitations of generated preclinical ECM models

LMS ECM scaffolds and ECM hydrogels generated in this thesis have several limitations. Firstly, although mainly composed of matrisome proteins, both scaffolds and hydrogels are not pure products. As shown in the literature, regardless of the decellularisation protocol used to process tissues, intracellular proteins are always present and detected during the characterisation of decellularised tissue^{368,581,585} or ECM hydrogels^{399,586}. Moreover, limited studies which assessed proteins present within the commercial BME extract, Matrigel, reported the identification of structural cellular components such as actin, tubulin and dynactin, nuclear, mitochondrial and other cellular proteins^{338,339}. As such, it's unsurprising that other non-matrisomal proteins were also detected in ECM scaffolds and ECM hydrogels in this thesis. In my analysis, cytoskeletal proteins such as beta-tubulin, intermediate filament proteins vimentin and desmin and nuclear histone H4 were present in all ECM scaffolds and all ECM hydrogels. Non-matrisome proteins might functionally affect cells cultured on ECM scaffolds and ECM hydrogels. However, the role of other cellular proteins in ECM scaffolds, ECM hydrogels or Matrigel has never been formally assessed in other studies or in this thesis. Future work could focus on dissecting the effect of other cellular proteins, inherently present in ECM scaffolds and ECM hydrogels, on the cellular phenotypes.

Secondly, the characterisation of preclinical ECM models showed that ECM scaffolds and, to a greater extent, the ECM hydrogels do not completely replicate the entire spectrum of matrisome-associated proteins, including ECM-affiliated proteins, ECM regulators, and secreted factors present in the originating tumours. While this might be viewed as a limitation, it can be harnessed as an opportunity. It could enable a systematic and controlled supplementation of ECM hydrogels or ECM scaffolds. As such, intentionally incorporating specific purified matrisome-associated components would allow

studying their role in LMS pathobiology. Additionally, ECM model supplementation can potentially improve the models to more accurately mirror the complexity of the TME. This refinement could lead to the creation of preclinical ECM models that better represent the biochemical composition of the ECM TME, offering a more comprehensive and representative platform for studying various aspects of LMS tumour behaviour, progression, and response to therapies.

Finally, a limitation in the analysis of ECM hydrogels is the absence of characterisation for the intermediate products between ECM scaffolds and the fully formed ECM hydrogels, referred to as pre-gel ECM solutions. Unfortunately, in this thesis, mass spectrometry was not employed to investigate the protein composition of LMS ECM pre-gel solutions. The gelation process could potentially influence the retention of ECM components. Introducing the characterisation of the pre-gel ECM solutions would enable a more comprehensive understanding of how converting ECM scaffolds to ECM hydrogels impacts the retention of matrix proteins. This step could shed light on factors contributing to the observed matrix protein loss, potentially leading to a more informed optimisation of the ECM hydrogel generation process.

5.3.2. The biological role of LMS ECM

The biochemical and biophysical characteristics of the ECM are central to nearly all biological processes. They affect various aspects, such as cell proliferation, migration, invasiveness, angiogenesis and drug resistance, substantially contributing to cancer progression⁴¹⁰. Chapter 5 investigated the pre-gel LMS ECM in 2D LMS cell migration and drug response.

Drug screening in 2D on plastic and LMS ECM pre-gel coating showed similar sensitivities to most tested drugs in SK-UT-1 cells. The drug screening identified increased sensitivities to dasatinib of SK-UT-1 cells grown on the LMS ECM, compared to plastic. The result was further validated in a short-term full dose-response cell viability assay and a long-term colony formation

assay. Future work could investigate the mechanism of LMS ECM-induced sensitivity to dasatinib and assess if LMS ECM increases sensitivity to dasatinib or other drugs in other LMS cell lines. This work would be important to inform future preclinical drug testing designs in LMS, advocating for a shift from screening potential compounds on cells cultivated on plastic surfaces to integrating ECM components into drug testing platforms.

Chapter 5 evaluated changes in the migration of LMS cells in response to LMS ECM and commercially available purified fibronectin, laminin and collagen IV. LMS ECM significantly increased the directional persistence of ICR-LMS-1, SK-UT-1 and SK-UT-1b cells compared to plastic but not SHEF-LMS w1 and SHEF-LMS ws cells. Those changes could be related to LMS cell lines' baseline directional migration (i.e. on plastic). Of the 5 tested cell lines, SHEF-LMS ws had the highest, and SHEF-LMS w1 had the second-highest median directionality indices at the baseline, which could explain why adding LMS ECM did not further increase the directionality of these cells. In contrast, LMS ECM induced a more persistent directional migration in LMS cell lines with low baseline displacement indices. Regarding changes in the cell speed, LMS ECM increased the median cell speed of SK-UT-1, SHEF-LMS ws and SHEF-LMS w1 compared to plastic. The increased cancer cell speed and directional persistence potentially relate to the increased metastatic potential of cancer cells^{589,590}. Lui *et al.* study generated a subline of breast cancer cell line MDA-MB-231, called lung-metastatic cells (LMCs), which was isolated from the metastatic lung colonies of MDA-MB-231 xenograft⁵⁸⁹. LMCs had a significantly higher speed and directionality index than the parental cell line in a 2D *in vitro* migration assay⁵⁸⁹. In a Matrigel spheroid invasion assay, it was observed that U87-MG glioblastoma cells that spread further away from the spheroid core displayed significantly increased directionality compared to cells remaining near the spheroid core⁵⁹⁰. This observation suggests that the increased directionality is associated with increased invasive potential in glioblastoma cells. Cell migration is a complex process whereby the cell integrates multiple signals to produce a coordinated movement⁵⁹¹. The underlying molecular mechanisms that control cell speed and migration directionality can be coordinated through actin dynamics, cellular contractility

(mediated by the actin and myosin cytoskeleton), cellular adhesion and adhesion-mediated signalling, and matrix remodeling^{592–595}. The molecular mechanisms responsible for the increase in the directional persistence and cell speed of LMS cells on LMS ECM coating compared to plastic are currently unclear and likely involve a combination of the above-mentioned processes. The precise mechanisms governing the LMS cell migration on LMS ECM could be further investigated using antibodies blocking specific integrin subunits, inhibitors targeting contractility (e.g. rho-associated protein kinase inhibitor Y-27632) or inhibitors targeting MMPs. While the relationship between cancer cell motility and metastasis is complex and can vary depending on cancer type and context, directional and fast-moving cancer cells are potentially more likely to produce metastasis. Understanding the underlying mechanisms driving these behaviours in LMS could aid in developing targeted therapies to inhibit metastasis in LMS cancer patients.

Additionally, Chapter 6 showed that the choice of ECM matrix impacted the LMS cell migration. This chapter compared the speed and directional persistence of LMS cell lines cultured on LMS ECM and commercially purified ECM fibronectin, laminin and collagen IV. The most striking example was a significant reduction in the directionality indices of 4/5 LMS cell lines on laminin compared to LMS ECM. The studies assessing the migration properties of LMS or other STS in response to different ECM coatings are lacking. In breast cancer, ECM coatings differentially affected cell speed and directional persistence^{357,576}. In MDA-MB-231 and MDA-MB-468 cell lines, collagen I, collagen IV⁵⁷⁶ and collagen VI³⁵⁷ have promoted faster migration than fibronectin, and the ECM-driven changes in directionality were cell line specific. In a lung cancer cell line model A549, laminin coating stimulated faster cell migration than fibronectin⁵⁷⁷. The underlying mechanism involved the engagement of different integrins and differential downstream activation of the Rho family of small GTPases⁵⁷⁷, which control the actin cytoskeleton organisation and are one of the key players in migration regulation⁵⁹⁶. Gu *et al.* results indicated that laminin and $\alpha 3\beta 1$ integrin-mediated signals preferentially activated Rac1, increasing cell migration⁵⁷⁷. In contrast, fibronectin and $\alpha 5\beta 1$ integrin-mediated signals led to RhoA activation, which

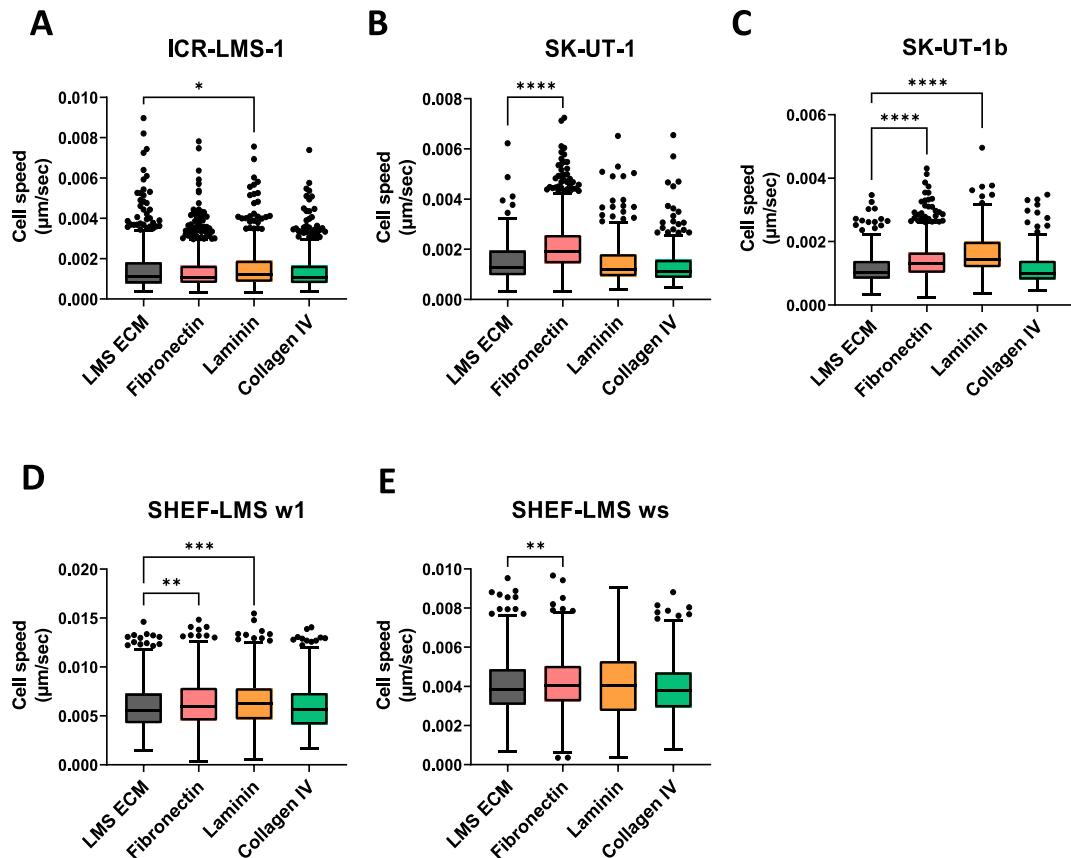
promoted focal adhesions and decreased cell migration⁵⁷⁷. Further work could investigate if the Rho family of small GTPases controls cell speed and directionality in LMS cells.

One of the limitations associated with this chapter is that my investigations focused on the ECM's role concerning its biochemical composition. It is well known that the ECM also affects biological processes via its biophysical properties, such as stiffness, 3D architecture, alignment of ECM fibres and pore size of a 3D substrate⁵⁹⁷. Future studies could investigate the role of biophysical properties of the ECM in LMS pathobiology using natural or synthetic hydrogels with tunable properties^{598–600}. For instance, polyacrylamide hydrogels can be engineered to have different stiffnesses by varying concentrations of acrylamide monomers, crosslinker N,N-methylene-bis-acrylamide and PBS^{601,602}.

Overall, Chapter 5 emphasises the significance of including LMS-specific ECM elements when evaluating cellular phenotypes in LMS. However, further follow-up studies are required to understand the molecular mechanisms underlying the increased sensitivity of SK-UT-1 cells to dasatinib in response to the LMS ECM and the mechanisms driving faster and more directional migration of LMS cells in response to the LMS ECM.

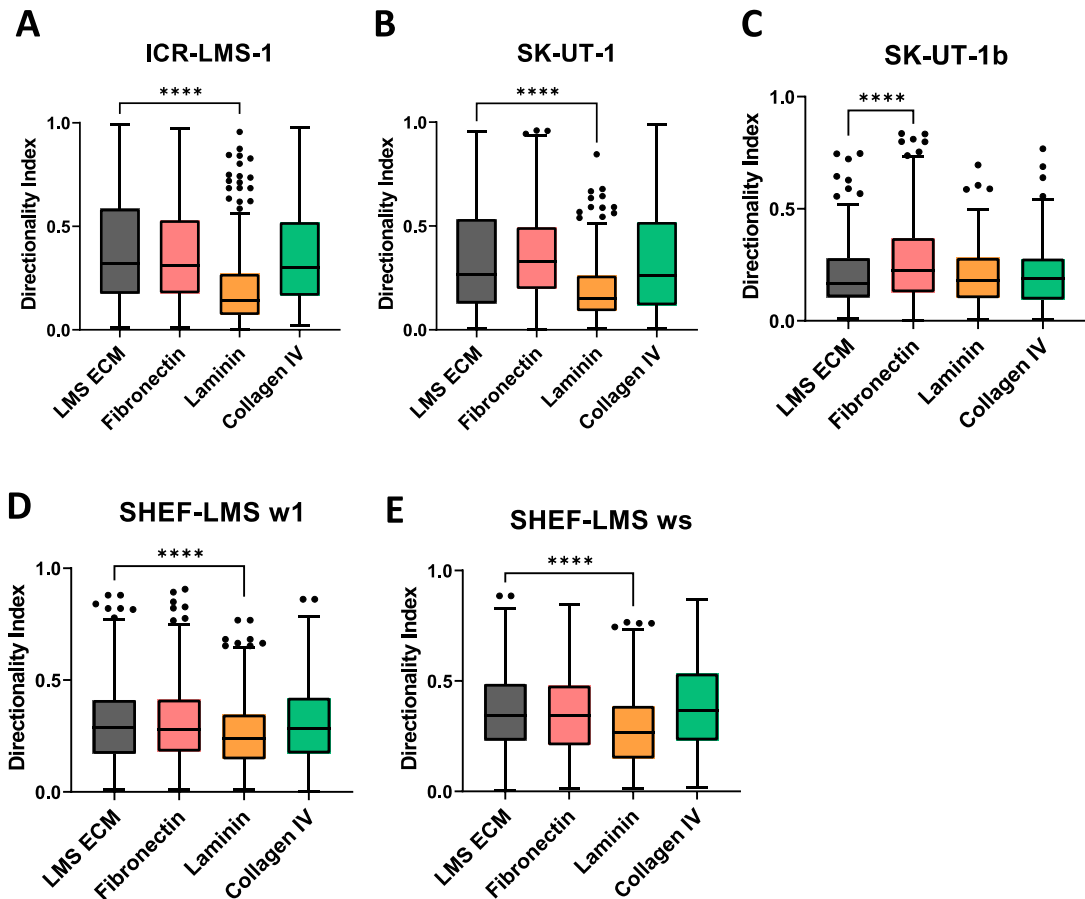
5.4. Supplementary Material

5.4.1. Supplementary Figures



Supplementary Figure 5.1 Comparison of cell speed in leiomyosarcoma (LMS) cell lines plated on a coating of leiomyosarcoma pre-gel extracellular matrix (LMS ECM) and commercial purified matrices.

Box plots showing the distribution of cell speed in GFP+ A) ICR-LMS-1, B) SK-UT-1, C) SK-UT-1b, D) SHEF-LMS w1 and E) SHEF-LMS ws. The GFP+ cells were tracked for 18 h. Data were pooled from three independent experiments for all graphs. Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile-(1.5* interquartile range (IQR)) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Kruskal Wallis tests with Dunn's multiple testing correction, * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001.



Supplementary Figure 5.2 Comparison of directionality indices in leiomyosarcoma (LMS) cell lines plated on a coating of leiomyosarcoma pre-gel extracellular matrix (LMS ECM) and commercial purified matrices.

Box plots showing the directionality indices in GFP+ A) ICR-LMS-1, B) SK-UT-1, C) SK-UT-1b, D) SHEF-LMS w1, and E) SHEF-LMS ws. The GFP+ cells were tracked for 18 h. Data were pooled from three independent experiments for all graphs. Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile-(1.5*IQR) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Kruskal Wallis tests with Dunn's multiple testing correction, **** p < 0.0001.

Chapter 6 Identification and functional validation of matrix signalling hits from proteomic data

6.1. Background and objectives

Previous studies in other cancer types examining the composition of the ECM and ECM and integrin-mediated signalling in tumour microenvironments unveiled previously unknown ECM proteins implicated in driving cancer progression and metastasis⁴¹¹. Chapter 3 provided the first-of-its-kind comprehensive characterisation of ECM and integrin adhesion networks across multiple STS subtypes. The development of the proteoglycan score for patient risk stratification in Chapter 3 showcased the value of the characterisation of ECM-integrin signalling proteins in STS, highlighting that it is a rich resource for data mining.

In addition to identifying candidate biomarkers, the analysis of matrisome and adhesome in STS patient proteomic data has the potential to reveal novel disease biology. Leveraging on the rich proteomic data coupled with matched clinical survival data, matrisome and adhesome analysis can pinpoint novel proteins within the ECM-integrin signalling pathway that are important in cancer progression. Given that LMS was our largest patient cohort (n = 80), Chapter 6 focuses on identifying the proteins in the ECM-integrin signalling pathway associated with LMS survival outcomes and thus potentially implicated in LMS pathobiology. One of those proteins was then functionally assessed in preclinical cell line models of LMS.

6.2. Results

6.2.1. Identification of the ECM and integrin adhesion signalling proteins for functional validation

A statistically significant association of the protein expression with cancer patient survival time implies that the expression levels of such proteins may have a meaningful impact on the progression of the disease. For instance, if a

high protein expression is significantly associated with poorer LRFS, MFS or OS, it suggests that the protein may promote tumour growth or metastasis. It is important to recognise that such statistical relationships represent associations, and establishing the causation requires experimental validation. Nevertheless, the significant association serves as a strong starting point for investigating the protein's potential involvement in cancer progression. The functional validation might involve manipulating the protein level in cancer cell lines and performing cellular assays *in vitro*. For instance, proliferation and cell growth assays serve as a proxy for evaluating the protein's role in tumour growth, while migration and invasion assays can assess its potential involvement in metastasis. A summary of a workflow to identify the LMS matrisome and adhesome proteins potentially involved in LMS tumour growth and metastasis and to select proteins for functional validation (in cell growth and migration assays) is shown in Figure 6.1.

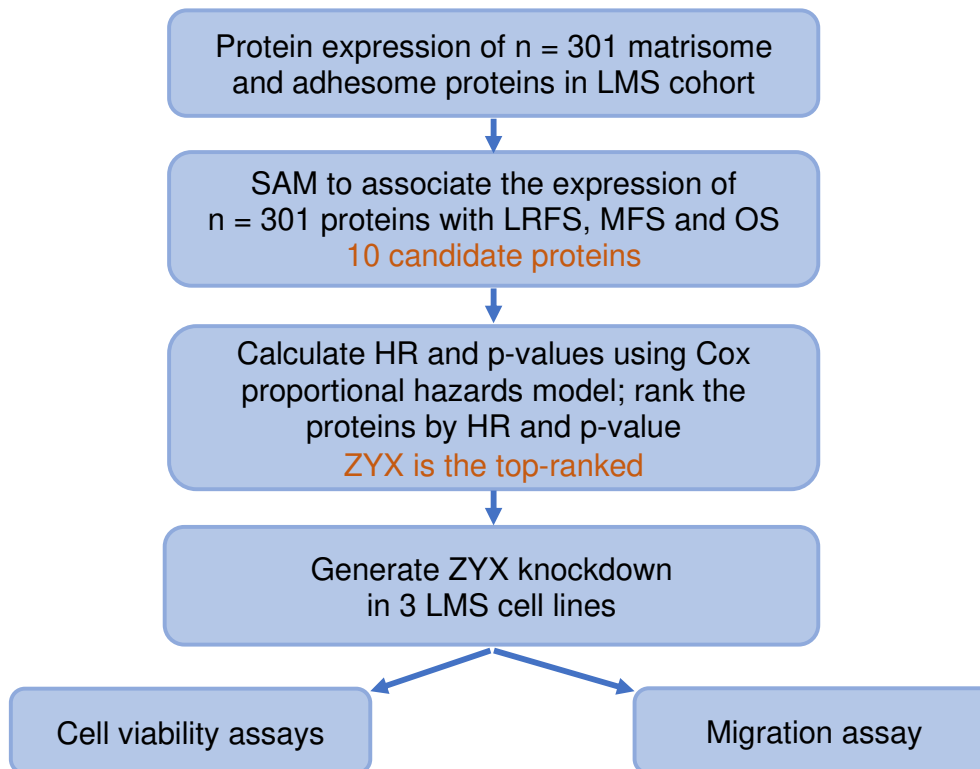


Figure 6.1 Summary of a workflow to identify ZYX as a protein for functional validation in cell viability and migration assays in LMS cell lines.

Significance of microarray analysis (SAM) was performed to associate the protein expression of n = 301 matrisome and adhesome proteins with local recurrence-free survival (LRFS), metastasis-free survival (MFS) and overall survival (OS) in a cohort of n = 80 leiomyosarcoma (LMS) patients. Ten proteins showed significant association with LRFS, MFS or OS. To nominate a protein for functional validation in cellular assays, 10 proteins were ranked based on the hazard ratio (HR) and p-values. Zyxin (ZYX) was the top-ranked protein with the lowest HR and second-lowest p-value, suggesting it has the highest influence on the survival outcome of the 10 identified proteins. ZYX was associated with MFS and could potentially influence LMS tumour growth and metastasis. To validate the functional role of ZYX in LMS tumorigenesis, ZYX shRNA-mediated knockdown was introduced in three LMS cell lines based on their high expression of ZYX. The role of ZYX *in vitro* was assessed in cell viability and migration assays using ZYX knockdown LMS cell lines.

The proteomic data derived from n = 80 FFPE LMS samples and associated patient survival outcomes (LRFS, MFS and OS) clinical data (from Section 3.2.1) were used. The SAM analysis was performed to statistically test the association of survival time (LRFS, MFS and OS) and the expression of matrisome and adhesome proteins (n = 301) in a cohort of n = 80 LMS cases.

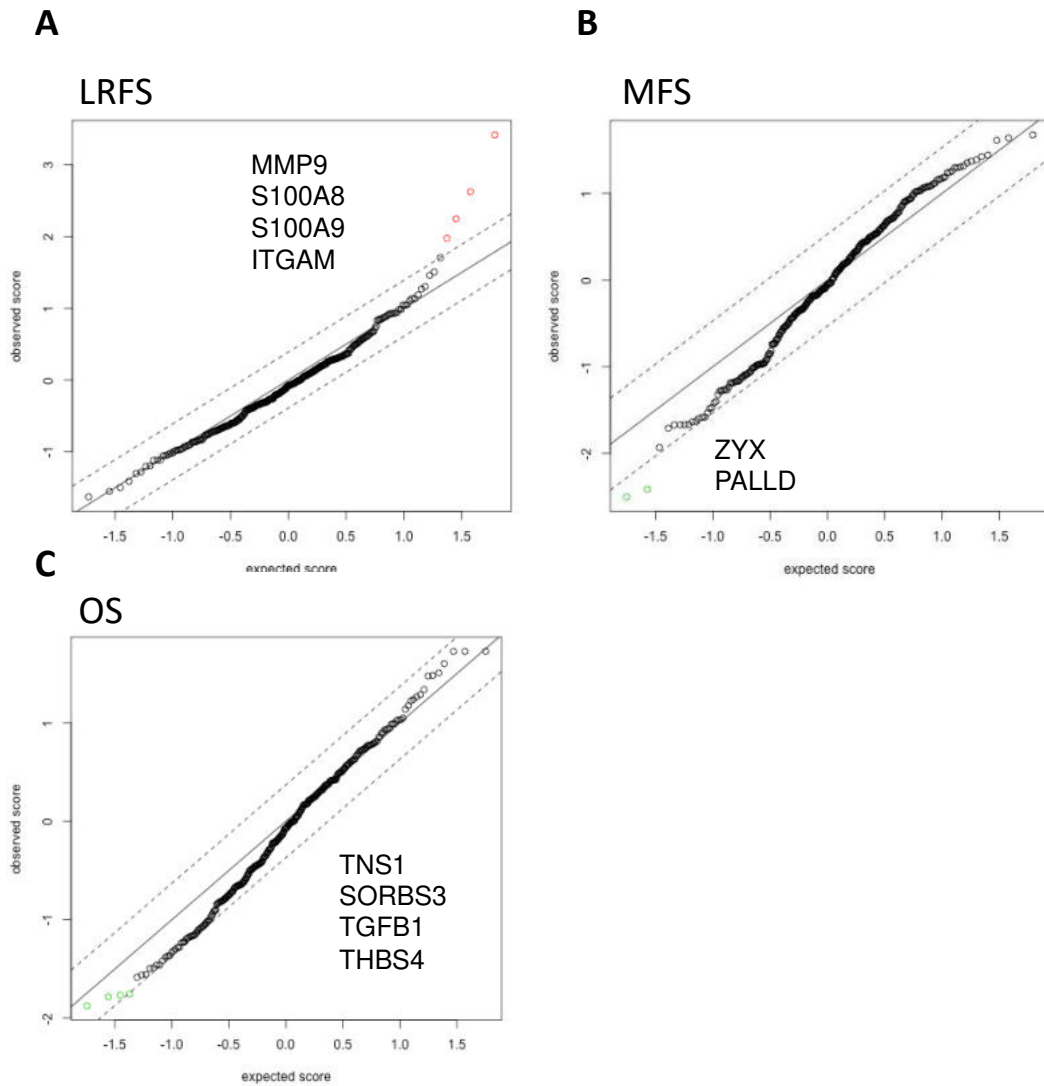


Figure 6.2 Significance analysis of microarray (SAM) analysis for (A) local recurrence-free survival (LRFS), (B) metastasis-free survival (MFS) and overall survival (OS) in leiomyosarcoma.

Proteins within the dashed lines had an FDR ≥ 0.01 and therefore were not significantly associated with LRFS, MFS or OS. Red genes in the SAM plot mean that higher expression correlates with higher risk (shorter survival). Green genes mean that higher expression correlates with lower risk (longer survival). ZYX = zyxin; MMP9 = matrix metalloproteinase 9; ITGAM = integrin alpha M chain; SORBS3 = sorbin and SH3 domain containing 3; PALLD = palladin; S100A8 = S100 calcium-binding protein A8; THBS4 = thrombospondin 4; TGFB1 = transforming growth factor-beta 1; S100A9 = S100 calcium-binding protein A9; TNS1 = tensin 1.

The SAM analysis identified MMP9, S100A8, S100A9 and ITGAM, the higher expression of which was significantly associated with shorter LRFS (Figure 6.2A). Higher expression of ZYX and PALLD proteins was significantly

associated with longer MFS (Figure 6.2B). Higher TNS1, SORBS3, TGF β -1 and THBS4 expression were associated with longer OS (Figure 6.2C). To calculate HR and p-values, univariable Cox proportional hazards regression analysis was performed for each protein separately. In practical terms, within the confines of this project, I could examine only one of the candidate proteins in functional experiments. As such, it was important to prioritise the proteins for in-depth functional follow-up studies, and all significant proteins across all survival outcomes were ranked by HR values and p-values (Table 6.1).

Table 6.1 Summary of univariable (UVA) Cox regression analysis assessing the association of Significance analysis of microarray (SAM) significant proteins with local recurrence-free survival, metastasis-free survival and overall survival in leiomyosarcoma patients.

Hazard ratio (HR), 95% confidence intervals (CI) and p-value were determined by univariable Cox regression with Wald tests. Proteins were ranked by absolute HR, meaning proteins were sorted based on the magnitude of their HR values, without considering the direction of the effect (whether it increases or decreases the risk of survival event). ZYX = zyxin; MMP9 = matrix metalloproteinase 9; ITGAM = integrin alpha M chain; SORBS3 = sorbin and SH3 domain containing 3; PALLD = palladin; S100A8 = S100 calcium-binding protein A8; THBS4 = thrombospondin 4; TGFB1 = transforming growth factor-beta 1; S100A9 = S100 calcium-binding protein A9; TNS1 = tensin 1.

Protein	HR (95% CI)	p-value	Absolute HR
ZYX	0.4 (0.29-0.64)	0.000021	2.3
MMP9	1.8 (1.4-2.3)	0.000008	1.8
ITGAM	1.8 (1.2-2.6)	0.0025	1.8
SORBS3	0.6 (0.37-0.97)	0.0039	1.7
PALLD	0.7 (0.51-0.83)	0.0005	1.5
S100A8	1.4 (1.2-1.7)	0.0005	1.4
THBS4	0.7 (0.53-0.92)	0.011	1.4
TGFB1	0.7 (0.56-0.92)	0.0093	1.4
S100A9	1.4 (1.1-1.7)	0.0029	1.4
TNS1	0.8 (0.61-0.93)	0.0097	1.3

Of the 10 evaluated proteins, ZYX had the highest estimated effect on MFS (HR = 0.4, 95% CI 0.29 – 0.64), which was statistically significant (p-value = 0.000021). The median protein expression of ZYX was used as a cut-off to generate a Kaplan-Meier plot to visualise the association between ZYX expression and MFS (log-rank p-value = 0.006) (Figure 6.3). Median MFS in

patients with high ZYX expression was 56 months (95% CI 46 – NA) compared to 28 months (95% CI 18 – 33) in patients with low ZYX expression.

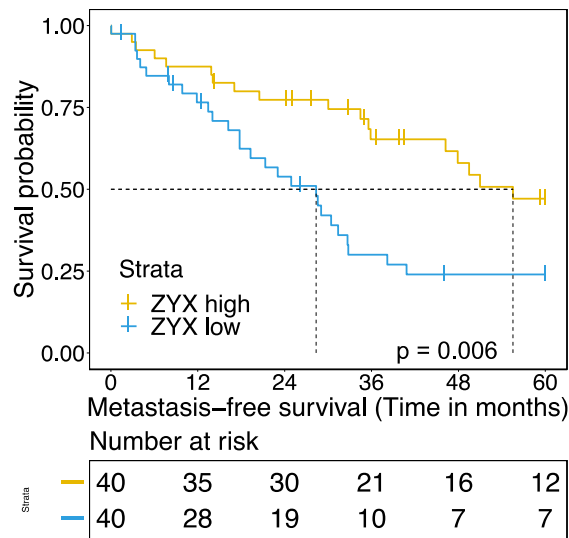


Figure 6.3 ZYX is associated with metastasis-free survival in leiomyosarcoma.

Kaplan-Meier plot of metastasis-free survival (MFS) with leiomyosarcoma (LMS) patients stratified by median zyxin (ZYX) protein expression. p-value (p) was estimated with the log-rank test. Horizontal and vertical lines indicate the median MFS for each stratum.

A multivariable analysis was performed to assess if the ZYX expression was independent of baseline prognostic clinicopathological variables. After adjusting for tumour grade, size, depth and anatomical location, the low ZYX expression remained an independent prognostic factor for inferior MFS (HR = 0.43, 95% CI 0.23 – 0.81, p-value = 0.009) (Supplementary Table 6.1). As expected, tumour grade and anatomical location were significantly associated with MFS in the multivariable analysis. The association of ZYX with MFS was LMS-specific, and ZYX expression was not associated with LRFS, MFS or OS in the whole n = 271 STS cohort (Table 6.2). Considering the above analysis, I nominated ZYX as a target for further functional validation.

Table 6.2 Univariable Cox regression analysis for ZYX in soft tissue sarcoma n = 271 cohort.

The association of zyxin (ZYX) expression and local recurrence-free survival (LRFS), metastasis-free survival (MFS), and overall survival (OS) was assessed. Hazard ratio (HR), 95% confidence intervals (CI) and p-value determined by univariable Cox regression with Wald tests.

	OS		LRFS		MFS	
Variable	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
ZYX	1 (0.85-1.2)	0.88	0.83 (0.68-1)	0.06	1 (0.84-1.2)	0.93

6.2.2. ZYX biology

Zyxin (ZYX) is an intracellular protein concentrated at the cell-ECM adhesion sites (focal adhesions) and cell-cell junctions, where it primarily acts as a scaffolding protein for other proteins involved in cytoskeletal dynamics and signalling⁶⁰³. ZYX colocalises with integrins at focal adhesion sites⁶⁰⁴ and has several domains which interact with actin-regulating proteins. For instance, ZYX has been reported to bind the actin filament crosslinker α -actinin (ACTN1), the actin assembly modulator Ena/vasodilator-stimulated phosphoprotein (VASP), the cytoskeletal proteins LIM and SH3 domain protein 1 (LASP-1) and the LIM-nebulette (LAPS-2)^{605,606}. Additionally, ZYX contains a nuclear export signal motif, which is sufficient to export nuclear ZYX to the cytoplasm⁶⁰⁷. ZYX is generally not present within the cell nuclei; however, it was shown that ZYX is able to shuttle between focal adhesions and the nucleus, potentially relaying signals from cell adhesions to regulate gene expression⁶⁰⁷⁻⁶¹⁰. Based on the literature search and STRINGdb, the protein-protein interaction network for ZYX was constructed (Figure 6.4). Most of ZYX's upstream and downstream interacting partners were present in our LMS proteomic dataset, indicating that the ZYX-interacting network is present in LMS tumours.

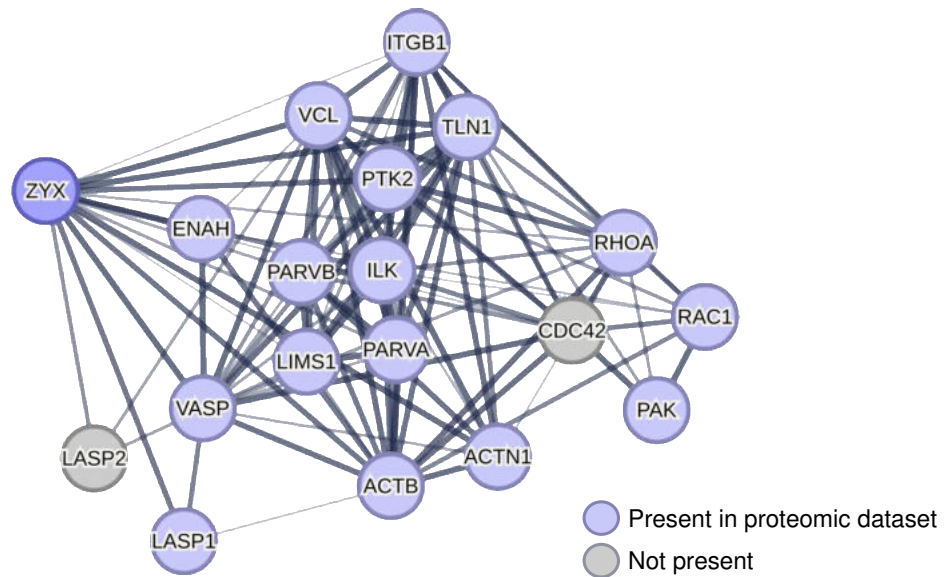


Figure 6.4 ZYX protein-protein interaction network.

Protein-protein (PPI) interaction network showing proteins interacting with zyxin (ZYX). Grey indicates proteins that not present in the leiomyosarcoma (LMS) proteomic dataset. The PPI was constructed with STRINGbd.

To investigate the biological pathways associated with ZYX expression in LMS patients, a two-class unpaired SAM test (with FDR < 0.01 and fold change \geq 2) was performed to identify DEPs in ZYX high compared to ZYX low LMS patients, classified based on whether their ZYX protein expression was above or below the median. There were no significantly downregulated proteins. The over-representation analysis against the Reactome pathway database was then performed for the upregulated DEPs with the 3,290 proteins found within the full proteomic dataset as a background (Figure 6.5). Apart from the muscle-specific signatures ('smooth muscle contraction' and 'muscle contraction'), the top-upregulated pathways included 'cell-extracellular matrix interactions', 'cell junction organisation', 'RHO GTPases activate PAKs' and 'Localization of the PINCH-ILK-PARVIN complex to focal adhesions'. This suggests a potential mechanistic link between ZYX and the proteins involved in those pathways and potential ZYX-mediated signalling in LMS.

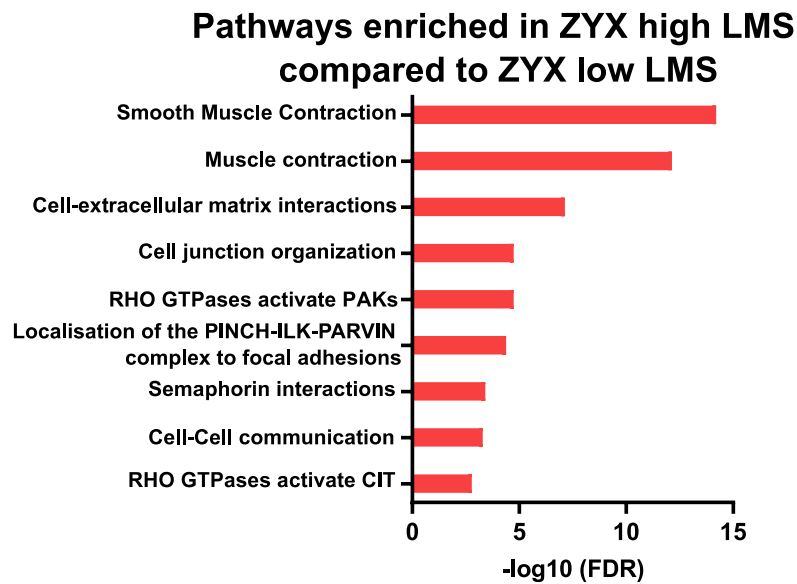


Figure 6.5 Biological pathways enriched in ZYX high compared to ZYX low leiomyosarcoma (LMS) samples.

Summary plot showing over-representation analysis of upregulated Reactome pathways in ZYX high versus ZYX low leiomyosarcoma (LMS) samples. The False Discovery Rate (FDR) was calculated using the Benjamini-Hochberg method.

6.2.3. Functional assessment of the role of ZYX in LMS cell lines

In my analysis, low ZYX protein expression was significantly associated with inferior MFS, suggesting that low levels of ZYX may promote LMS tumour growth or metastasis. The survival analysis provided a rationale for investigating the role of ZYX in LMS tumorigenesis. Considering that LMS samples profiled by mass-spectrometry had > 75% tumour cell content (Section 2.1.1), an assumption was made that ZYX is primarily expressed by LMS cancer cells as opposed to stromal cells. As such, to study the causative relationship between ZYX expression and LMS progression, ZYX-knockdown LMS cell lines were generated. Then, ZYX-knockdown cell lines were assessed in a cell viability assay as a proxy for tumour growth. Metastasis is a complex, multistep process involving tumour cell migration and invasion at the initial stages⁶¹¹. To begin to understand the role of ZYX in LMS metastasis, ZYX's influence on a metastasis-related process, migration, was investigated.

6.2.3.1. Selection of LMS cell lines for genetic knockdown of ZYX

First, to study the role of ZYX in LMS, ZYX expression levels were assessed in a panel of LMS cell lines. The panel consisted of high-grade, mostly of uterine origin, established (SK-UT-1 and SK-UT-1b) and patient-derived (SHEF-LMS w1, SHEF-LMS ws, ICR-LMS-1, SARC-393, ICR-LMS-4 and ICR-LMS-6) LMS cell lines (Table 6.3).

Table 6.3 Summary of established and patient-derived leiomyosarcoma cell lines.

Overview of characteristics of leiomyosarcoma (LMS) cell lines, including LMS subtype, originating site, Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC) grade or International Federation of Gynecology and Obstetrics (FIGO) stage of originating tumour/biopsy. * and ** indicates that the cell lines are morphological variants derived from the same patient; uLMS: uterine leiomyosarcoma.

Cell line ID	LMS subtype	Site of biopsy/resection	Primary site	Grade	FIGO stage	Source
SK-UT-1*	uLMS	Uterus	Uterus	3	-	Dr Priya Chudasama, German Cancer Research Centre
SK-UT-1b*	uLMS	Uterus	Uterus	3	-	Dr Priya Chudasama, German Cancer Research Centre
ICR-LMS-1	uLMS	Pelvis	Uterus	3	-	Derived in Huang laboratory
SHEF-LMS w1**	uLMS	Pelvis	Uterus	-	IIA	Dr Karen Sisley, University of Sheffiled
SHEF-LMS ws**	uLMS	Pelvis	Uterus	-	IIA	Dr Karen Sisley, University of Sheffiled
SARC-393	uLMS	Lung	Uterus	-	-	Derived in Huang laboratory
ICR-LMS-4	uLMS	Liver	Uterus	3	-	Derived in Huang laboratory
ICR-LMS-6	LMS	Peritoneum	Stomach	-	-	Derived in Huang laboratory

SHEF-LMS w1 and SHEF-LMS ws were two distinct morphologic LMS cell populations established from the same LMS patient⁴⁰³. Similarly, SK-UT-1 and SK-UT-1b were morphological variants isolated from the other LMS patients^{401,402}. Protein expression of ZYX was evaluated using western blotting (Figure 6.6). SK-UT-1b did not show ZYX protein expression. ICR-LMS-1 and SARC-393 had the lowest expression of ZYX relative to the rest of the cell lines. SK-UT-1, SHEF-LMS w1 and SHEF-LMS ws had a high ZYX expression and were chosen for shRNA-based ZYX knockdown.

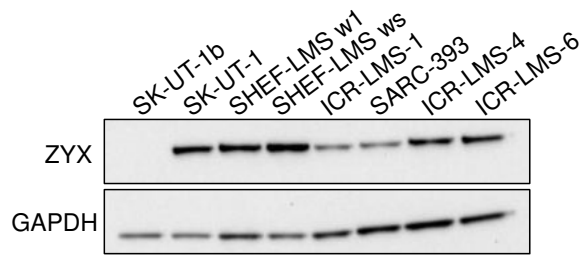


Figure 6.6 ZYX expression in established and patient-derived leiomyosarcoma cell lines.

Immunoblot of total levels of zyxin (ZYX) in SK-UT-1b, SK-UT-1, SHEF-LMS w1, SHEF-LMS ws, ICR-LMS-1, SARC-393, ICR-LMS-4 and ICR-LMS-6 cell lines (n = 1). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control.

6.2.3.2. Confirmation of ZYX knockdown in LMS cell lines

To silence ZYX expression, SK-UT-1, SHEF-LMS w1 and SHEF-LMS ws cell lines were transduced with three different doxycycline-inducible shRNA sequences targeting ZYX and a scramble shRNA control sequence. Western blotting confirmed ZYX knockdown after 72 h of doxycycline induction (Figure 6.7A-C).

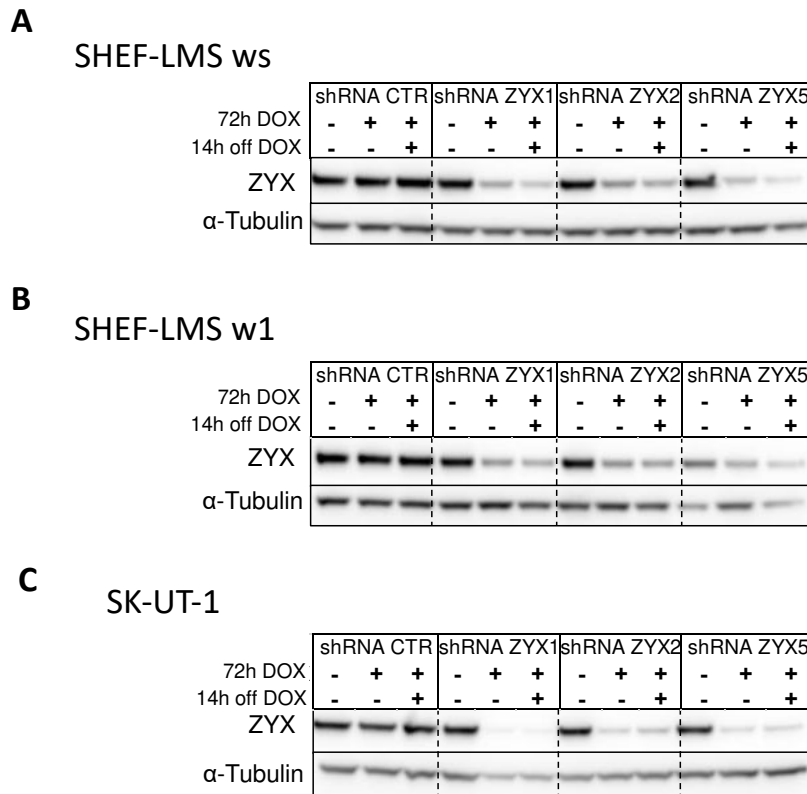


Figure 6.7 Confirmation of ZYX knockdown in leiomyosarcoma cell lines.

Immunoblot of total levels of zyxin (ZYX) with three doxycycline (DOX)-inducible short hairpin RNA (shRNA) constructs in (A) SHEF-LMS ws, (B) SHEF-LMS w1 and (C) SK-UT-1 cell lines 72h after dox induction or 14h after dox was removed from the media. α-Tubulin was used as a loading control. Images are representative of two separate experiments (n = 2).

In SHEF-LMS ws and SK-UT-1, ZYX1, ZYX2 and ZYX5 shRNA sequences resulted in a substantial reduction in the protein expression levels of ZYX in the doxycycline-induced (+DOX) samples compared to the non-induced control (-DOX) samples, confirming the successful suppression of ZYX (Figure 6.7 A,C). In SHEF-LMS w1, ZYX1 and ZYX2 sequences showed ZYX knockdown, but the ZYX5 sequence induced a less robust ZYX knockdown (Figure 6.7B). Moreover, the ZYX knockdown was sustained after the doxycycline withdrawal for 14 h in all tested shRNA sequences in the three LMS cell lines. There was no difference in ZYX expression between the shRNA scramble control-expressing cells (CTR-DOX) and scramble control cells treated with doxycycline (CTR+DOX), indicating that the doxycycline treatment on its own did not affect ZYX expression.

6.2.3.3. *Role of ZYX in cell growth and migration of LMS*

ZYX knockdown reduces LMS cell growth

In other cancer types, ZYX silencing was shown to reduce cell growth^{612,613}. To investigate whether ZYX affects LMS cell growth, I assessed the cell population growth of SHEF-LMS ws, SHEF-LMS w1 and SK-UT-1 control and ZYX-silenced cells (Figure 6.8-6.10). Three unique ZYX-targeting shRNA sequences (ZYX1, ZYX2 and ZYX5) were introduced to each cell line. Scramble shRNA (CTR) served as a control for doxycycline treatment, and cells with ZYX knockdown were compared to their respective non-doxycycline-treated cells.

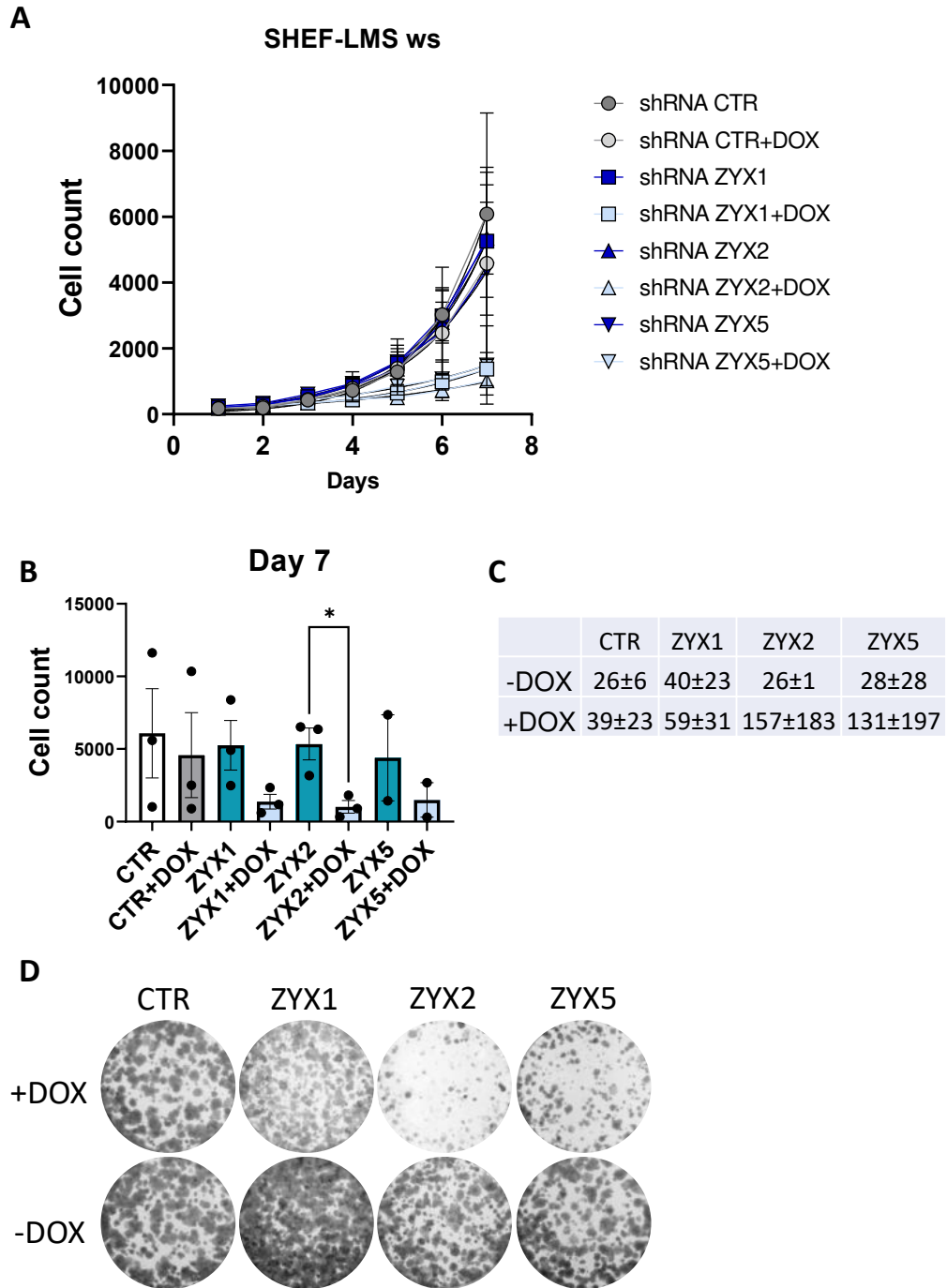


Figure 6.8 ZYX knockdown reduced SHEF-LMS ws cell population growth.

A) Growth curves of SHEF-LMS-ws cells stably expressing either control shRNA CTR or zyxin-targeting shRNA (ZYX1, ZYX2 and ZYX5) constructs. SHEF-LMS ws cells were grown with 2 $\mu\text{g/ml}$ doxycycline (+DOX) to induce ZYX knockdown or without DOX for 72h prior to the assay. Cells in +DOX conditions were grown in media-supplemented with DOX for the duration of the assay. Cell population growth was assessed by direct count of nuclei stained with Hoechst 33342. The graph shows the mean and standard error of the mean (SEM) from $n = 3$ independent experiments. B) Bar charts showing the number of cells on the 7th day of

the growth assay across all conditions. Significance is shown following an unpaired *t*-test, * $p < 0.05$. C) Comparison of doubling time (in hours) across the conditions. D) Representative images of crystal violet stained wells from the colony formation assay of SHEF-LMS ws cells expressing either control shRNA CTR or zyxin-targeting shRNA (ZYG1, ZYG2 and ZYG5) constructs. The cells were grown either in dox-supplemented media (+DOX) or without dox (-DOX) for 14 days ($n = 2$).

SHEF-LMS ws

Upon doxycycline-mediated knockdown of ZYG for 7 days, the ZYG1, ZYG2 and ZYG5 SHEF-LMS ws cells showed a reduction in cell growth compared to their corresponding non-doxycycline treated controls (Figure 6.8A, B). At day 7, the number of ZYG2+DOX cells was significantly lower than that of ZYG2 cells (p -value = 0.021). The mean doubling time of ZYG-silenced cells was 1.5 fold higher in ZYG1+DOX (59 h) compared to ZYG1 (40 h), 6 fold higher in ZYG2+DOX (157 h) compared to ZYG2 (26 h) and 4.7 fold higher in ZYG5+DOX (131 h) compared to ZYG5 (28 h) (Figure 6.8C). Consistent with the short-term cell growth assay, ZYG-silenced cells showed reduced clonogenic properties after 14 days compared to their corresponding non-doxycycline-treated controls (Figure 6.8D). The doubling time of CTR+DOX (39 h) was 1.5 fold higher than in CTR (26 h), and at day 7, there was no significant difference between CTR+DOX and CTR, indicating that treatment with doxycycline did not reduce SHEF-LMS w1 cell growth.

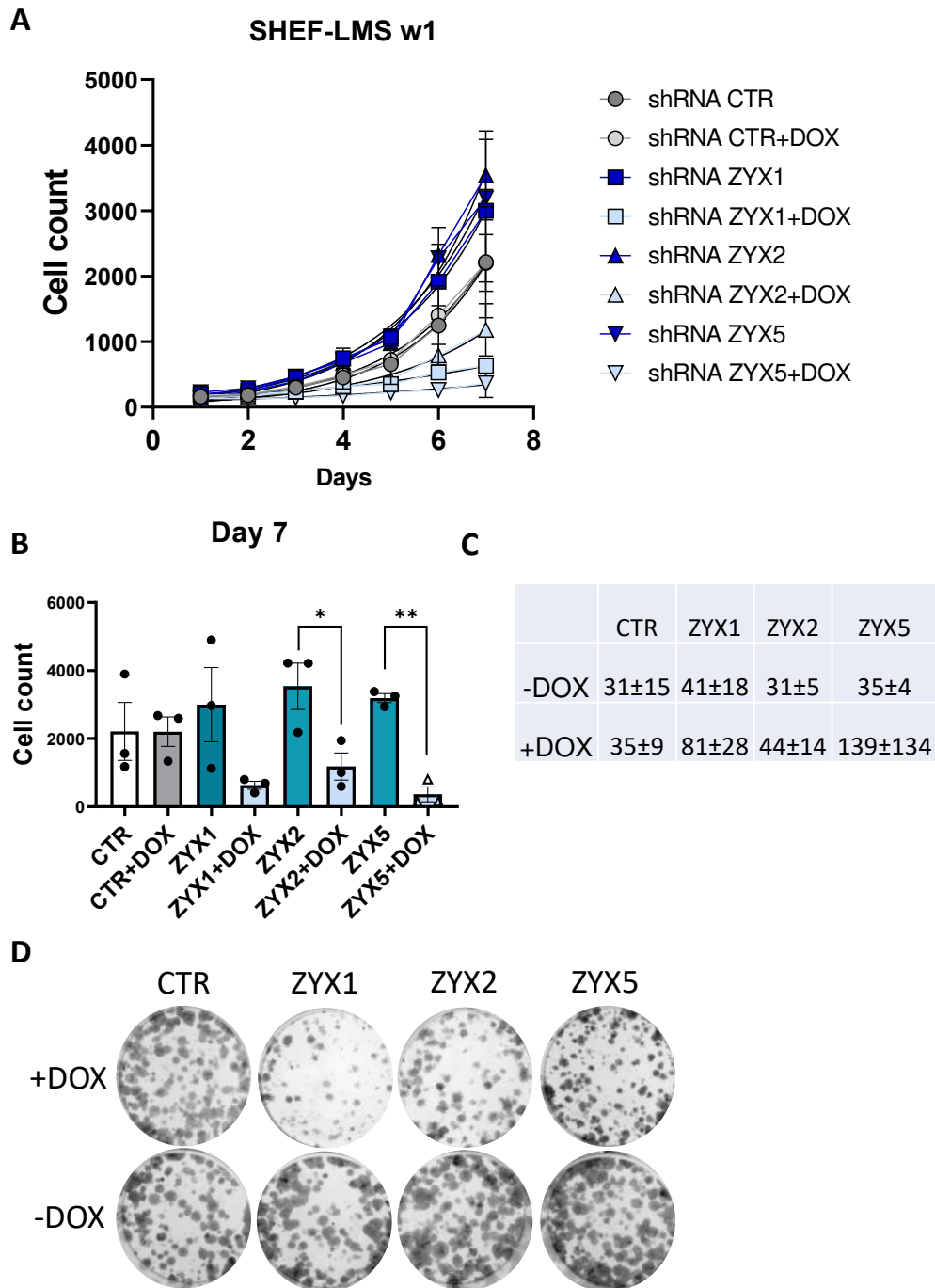


Figure 6.9 ZYX knockdown reduced SHEF-LMS w1 cell population growth.

A) Growth curves of SHEF-LMS-w1 cells stably expressing either control shRNA CTR or zyxin-targeting shRNA (ZYX1, ZYX2 and ZYX5) constructs. SHEF-LMS w1 cells were grown with 2 $\mu\text{g/ml}$ doxycycline (+DOX) to induce ZYX knockdown or without DOX for 72h prior to the assay. Cells in +DOX conditions were grown in media-supplemented with DOX for the duration of the assay. Cell population growth was assessed by direct count of nuclei stained with Hoechst 33342. The graph shows the mean and standard error of the mean (SEM) from $n = 3$ independent experiments. B) Bar charts showing the number of cells on the 7th day of the growth assay across all conditions. Significance is shown following an unpaired t -test, * p

< 0.05; ** $p < 0.01$. C) Comparison of doubling time (in hours) across the conditions. D) Representative images of crystal violet stained wells from the colony formation assay of SHEF-LMS w1 cells expressing either control shRNA CTR or zyxin-targeting shRNA (ZYG1, ZYG2 and ZYG5) constructs. The cells were grown either in dox-supplemented media (+DOX) or without dox (-DOX) for 14 days ($n = 2$).

SHEF-LMS w1

Silencing of ZYG in SHEF-LMS w1 led to the suppression of cell growth (Figure 6.9A). By day 7, the numbers of ZYG2-silenced (p -value = 0.0402) and ZYG5-silenced (p -value = 0.0004) cells were significantly lower than their non-doxycycline-treated controls (Figure 6.9B). There was a 2-fold increase in the average doubling time of ZYG1+DOX (81 h) compared to ZYG1 (41 h) and a 4-fold increase in ZYG5+DOX (139 h) compared to ZYG5 (35 h), whilst ZYG2+DOX (44 h) produced a modest 1.4 fold increase compared to ZYG2 (31 h) (Figure 6.9C). The doubling time of CTR+DOX (33 h) and CTR (31 h) were similar, and at day 7, there was no significant difference in the cell number between CTR+DOX and CTR, indicating that treatment with doxycycline did not affect SHEF-LMS w1 cell growth (Figure 6.9B, C). The inhibition of SHEF-LMS w1 cell growth after ZYG knockdown was additionally confirmed in a 14-day colony formation assay (Figure 6.9D).

SK-UT-1

Cell growth inhibition was observed only for ZYG5-silenced cells in a 7-day assay testing ZYG knockdown in the SK-UT-1 cell line (Figure 6.10A). At day 7, the number of ZYG5-silenced cells was significantly (p -value = 0.017) lower than its non-doxycycline-treated controls (Figure 6.10B). The doubling times were similar across all the conditions, and doxycycline treatment did not affect cell growth (Figure 6.10C). Nevertheless, suppression of cell growth was observed with the ZYG2 and ZYG5 shRNA sequences in a 14-day colony formation assay (Figure 6.10D).

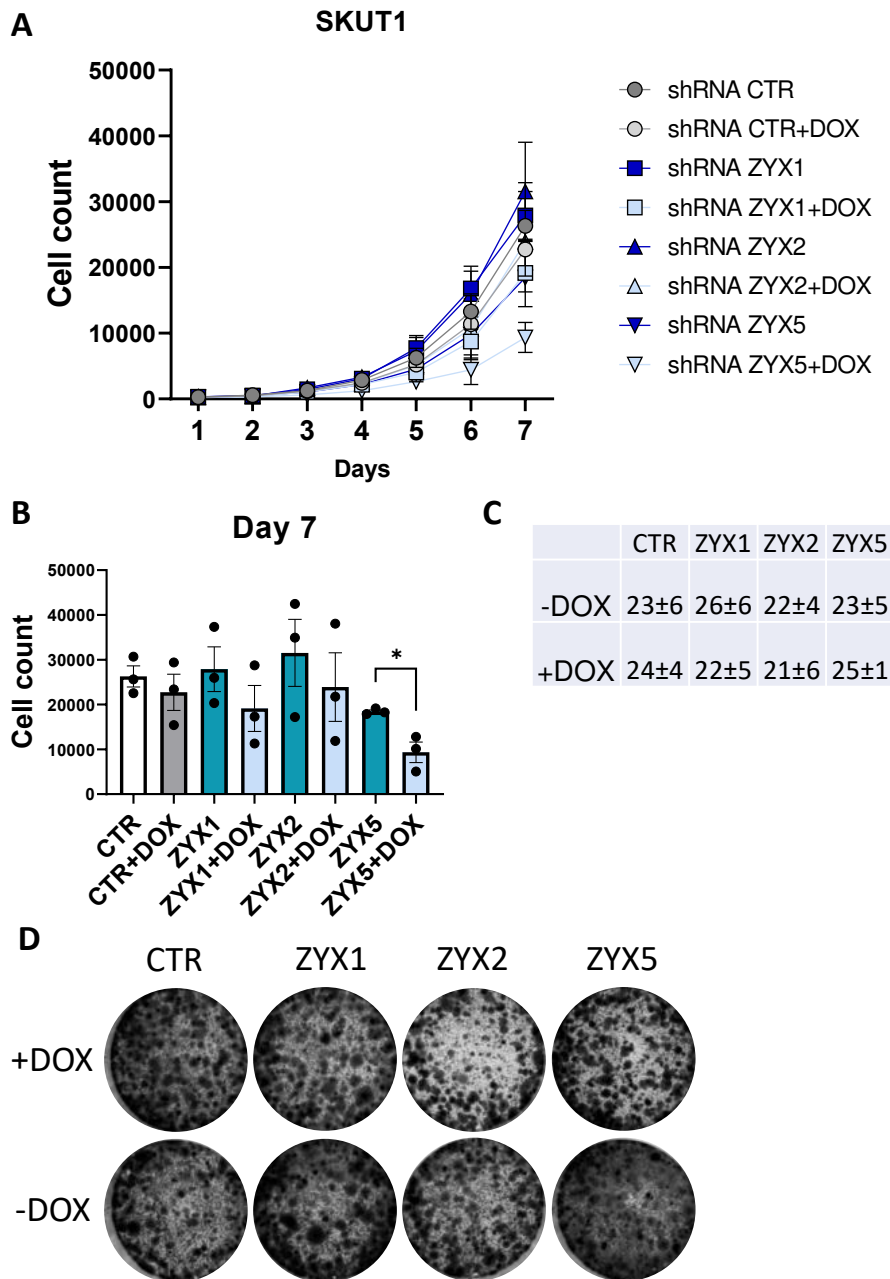


Figure 6.10 ZYX knockdown in SK-UT-1 cell population growth.

A) Growth curves of SK-UT-1 cells stably expressing either control shRNA CTR or zyxin-targeting shRNA (ZYX1, ZYX2 and ZYX5) constructs. SK-UT-1 cells were grown with 2 $\mu\text{g/ml}$ doxycycline (+DOX) to induce ZYX knockdown or without DOX for 72h prior to the assay. Cells in +DOX conditions were grown in media-supplemented with DOX for the duration of the assay. Cell population growth was assessed by direct count of nuclei stained with Hoechst 33342. The graph shows the mean and standard error of the mean (SEM) from $n = 3$ independent experiments. B) Bar charts showing the number of cells on the 7th day of the growth assay across all conditions. Significance is shown following an unpaired t -test, * $p < 0.05$. C) Comparison of doubling time (in hours) across the conditions. D) Representative images of crystal violet stained wells from the colony formation assay of SK-UT-1 cells

expressing either control shRNA CTR or zyxin-targeting shRNA (ZYG1, ZYG2 and ZYG5) constructs. The cells were grown either in dox-supplemented media (+DOX) or without dox (-DOX) for 14 days (n = 2).

ZYG knockdown increases SHEF-LMS ws cell migration

In addition to affecting cell growth, ZYG modulates cancer cell migration and correlates with metastasis in other cancer types^{614,615}. Given that low ZYG expression was significantly associated with shorter MFS in our LMS cohort, I hypothesised that lowered ZYG expression could promote metastasis-related processes such as migration. In line with that, I hypothesised that a ZYG knockdown in LMS cell lines would increase cell migration.

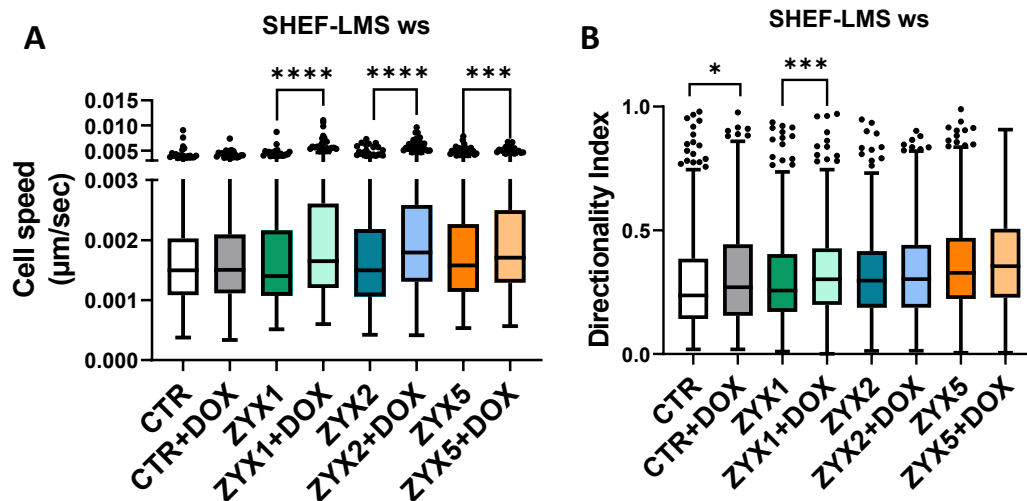


Figure 6.11 ZYG knockdown increases cell speed of SHEF-LMS ws cells

A) Box plots showing the cell speed of GFP+ cell lines in 2D on plastic over 12 h. SHEF-LMS-ws cells stably express either scramble shRNA (CTR) or zyxin-targeting shRNA (ZYG1, ZYG2 and ZYG5) constructs. SHEF-LMS ws cells were grown with 2 µg/ml doxycycline (+DOX) to induce ZYG knockdown or without DOX for 72h prior to the assay. DOX was removed for the duration of the assay. (B) Box plots show the directionality indexes on plastic. Data were pooled from n = 3 independent experiments, apart from ZYG1 and ZYG1+DOX (n = 2). Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile- (1.5*IQR) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Mann–Whitney *U* tests, * p < 0.05; *** p < 0.001; **** p < 0.0001.

To establish the role of ZYX in LMS cell motility, ZYX-silenced and the control SK-UT-1, SHEF-LMS w1 and SHEF-LMS ws cells were monitored by time-lapse microscopy over 12 h and tracked to determine their migration speed and directionality (Figure 6.11-6.13). The cells were pre-treated with doxycycline for 72 h prior to the assay, and doxycycline was removed from the media for the duration of the assay. As shown in Figure 6.7, SK-UT-1, SHEF-LMS w1, and SHEF-LMS ws cells pre-treated with doxycycline maintained ZYX knockdown for the duration of the migration experiment.

SHEF-LMS ws

Silencing ZYX expression with the ZYX1, ZYX2 and ZYX5 shRNA sequences increased the SHEF-LMS ws cell speed compared to their corresponding non-doxycycline-treated controls (Figure 6.11A). ZYX knockdown with ZYX1 shRNA induced an 18% increase (p-value < 0.0001) in the median cell speed, ZYX2 shRNA induced a 20% increase (p-value < 0.0001) in the median speed, whilst only an 8% increase was observed with ZYX5 shRNA. An 8% increase was not considered meaningful. Nevertheless, two distinct ZYX shRNA constructs increased SHEF-LMS ws cell speed. Regarding the directionality of migration, there was an 18% increase (p-value = 0.001) in the median directionality index between ZYX1-silenced cells and their corresponding non-doxycycline-treated control (Figure 6.11B). However, the increase is likely due to the doxycycline pre-treatment, given that there was a 14% increase (p-value = 0.021) in the directionality index between CTR and CTR+DOX (Figure 6.11B).

SK-UT-1

In SK-UT-1, only the ZYX2 shRNA construct significantly elevated the cell speed of doxycycline-induced ZYX knockdown cells by 15% (p-value < 0.0001) compared to the intrinsic control (Figure 6.12A). The ZYX1+DOX (p-value = 0.003) and ZYX2+DOX (p-value < 0.0001) increased the directionality indices compared to their intrinsic controls, likely due to the doxycycline pre-treatment as there was a statistical difference between CTR+DOX and CTR (p-value = 0.007) (Figure 6.12B).

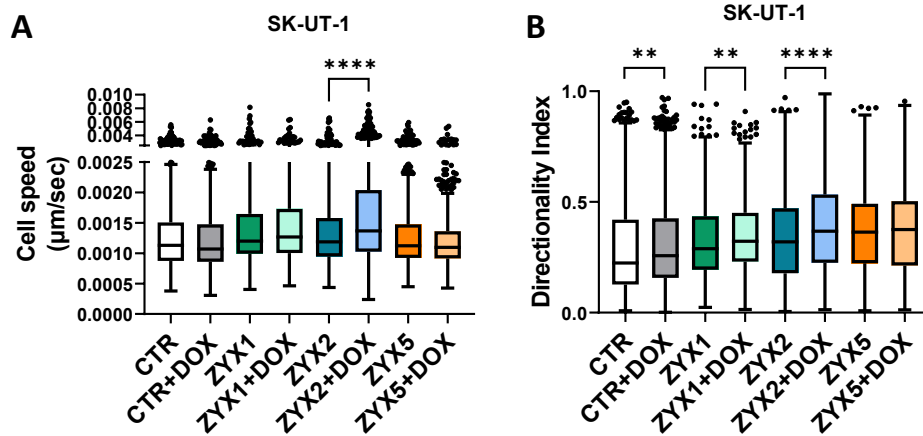


Figure 6.12 Cell speed and directionality of SK-UT-1 ZYX knockdown cells.

A) Box plots showing the cell speed of GFP⁺ cell lines in 2D on plastic over 12 h. SK-UT-1 cells stably express either scramble shRNA (CTR) or zyxin-targeting shRNA (ZYX1, ZYX2 and ZYX5) constructs. SK-UT-1 cells were grown with 2 µg/ml doxycycline (+DOX) to induce ZYX knockdown or without DOX for 72h prior to the assay. DOX was removed for the duration of the assay. (B) Box plots show the directionality indexes on plastic. Data were pooled from $n = 3$ independent experiments, apart from ZYX1 and ZYX1+DOX ($n = 2$). Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile-(1.5*IQR) to the 75th percentile+(1.5*IQR), and outliers plotted as points. Significance is shown following Mann–Whitney U tests, ** $p < 0.01$; **** $p < 0.0001$.

SHEF-LMS w1

ZYX knockdown with the ZYX2 and ZYX5 shRNA constructs did not affect the cell speed of SHEF-LMS w1 cells (Figure 6.13A). In contrast, ZYX1-mediated knockdown reduced the cell speed of knockdown cells by 18% (p -value < 0.0001) compared to their non-doxycycline-treated control (Figure 6.13A). None of the ZYX shRNA constructs affected the directionality of SHEF-LMS w1 cells (Figure 6.13B).

Overall, sufficient evidence (i.e. two separate shRNA constructs showing a significant increase) for an increase in the migration speed upon ZYX knockdown was only available for the SHEF-LMS ws cell line. As such, further experiments with other ZYX-targeting shRNA constructs or assessments in other migration assays are required to confirm that ZYX knockdown also increases cell migration in SK-UT-1 and SHEF-LMS w1.

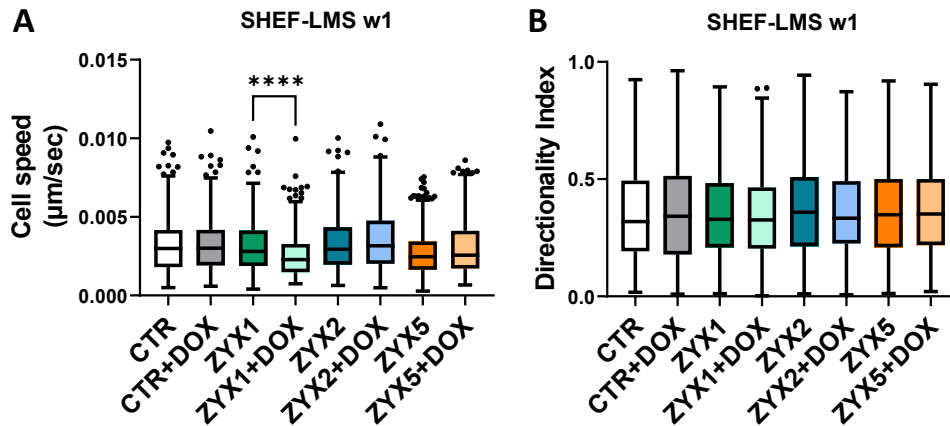


Figure 6.13 Cell speed and directionality of SHEF-LMS-w1 ZYX knockdown cells.

A) Box plots showing the cell speed of GFP+ cell lines in 2D on plastic over 12 h. SHEF-LMS-w1 cells stably express either scramble shRNA (CTR) or zyxin-targeting shRNA (ZYX1, ZYX2 and ZYX5) constructs. SHEF-LMS w1 cells were grown with 2 µg/ml doxycycline (+DOX) to induce ZYX knockdown or without DOX for 72 h prior to the assay. DOX was removed for the duration of the assay. (B) Box plots show the directionality indexes on plastic. Data were pooled from $n = 3$ independent experiments, apart from ZYX1 and ZYX1+DOX ($n = 2$). Boxes indicate the 25th, median, and 75th percentile, with whiskers extending from the 25th percentile- $(1.5 \times \text{IQR})$ to the 75th percentile+ $(1.5 \times \text{IQR})$, and outliers plotted as points. Significance is shown following Mann–Whitney U tests, **** $p < 0.0001$.

6.3. Discussion

Chapter 6 has shed light on the potential role of ECM and integrin adhesion signalling proteins, particularly ZYX, in the progression of LMS. Identifying ECM and integrin adhesion signalling proteins associated with LRFS, MFS, and OS in LMS suggests their significance as potential candidates in LMS progression. ZYX, in particular, emerged as a promising candidate for further investigation.

Notably, higher ZYX protein expression was associated with longer MFS in our LMS patient cohort, implicating ZYX as a metastasis suppressor in LMS. ZYX expression was not associated with survival outcome measures in our full STS cohort ($n=271$), including multiple histological subtypes, suggesting that the potential role of ZYX in tumour progression is a LMS-specific finding. There are no studies investigating the association of ZYX with survival in STS. The association of higher ZYX expression with better survival was in line with the

study on ZYX in osteosarcoma, a malignant bone cancer⁶¹⁶. The study reported that a higher mRNA expression level of ZYX was associated with a better prognosis in the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) osteosarcoma cohort⁶¹⁶. However, reports in other cancer types showed the opposite trend to the finding in this chapter. Higher ZYX gene expression was associated with shorter recurrence-free survival in a colorectal cancer cohort (n=145)⁶¹⁷, shorter OS in three public glioma databases and datasets of a subtype of glioma, glioblastoma⁶¹⁸. This discrepancy could be due to poor concordance of RNA expression and protein expression⁶¹⁹, or that the association of high ZYX expression with better survival could be a LMS-specific finding. Nevertheless, a significant association between ZYX protein expression and MFS in our LMS cohort justified studying the role of ZYX in functional follow-up experiments.

6.3.1. ZYX biology

ZYX, a cytoskeletal protein, plays a role in cell adhesion, migration, and signal transduction by interacting with actin filaments and multi-molecular structures known as focal adhesions, which link the ECM to the actin cytoskeleton^{615,620}. Analysis of differentially expressed proteins in ZYX high compared to ZYX low LMS patients and over-representation analysis revealed top-upregulated pathways such 'cell-extracellular matrix interactions', 'cell junction organisation', 'RHO GTPases activate PAKs' and 'Localization of the PINCH-ILK-PARVIN complex to focal adhesions'. Particularly interesting new cysteine-histidine-rich protein (PINCH), ILK and alpha-parvin (PARVIN) proteins assemble in a complex which binds to ECM-engaged integrins⁶²¹. This complex creates a molecular bridge facilitating bidirectional communication between the ECM and the cell's internal machinery. A previous report showed that the PINCH-ILK-PARVIN complex regulates ZYX recruitment to mature focal adhesions⁶²². Genetic knockout of *ILK* or *PINCH1* in murine fibroblasts prevented the formation of the PINCH-ILK-PARVIN complex and reduced the localisation of ZYX to focal adhesions⁶²². The enrichment of the 'Localization of the PINCH-ILK-PARVIN complex to focal adhesions' pathway in LMS patients with high ZYX expression compared to

patients with low ZYX suggests a potential mechanistic link between ZYX and the PINCH-ILK-PARVIN complex. Further investigations involving Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) or shRNA-mediated knockdown of ILK, PINCH or PARVIN would be necessary to assess if the PINCH-ILK-PARVIN complex regulates ZYX recruitment to focal adhesions in LMS cell lines.

My analysis also showed enrichment of the 'RHO GTPases activate PAKs' pathway in LMS patients with high ZYX expression. Rho GTPases function as molecular switches, cycling between an inactive GDP-bound state and an active GTP-bound state⁵⁹⁶. Rho GTPases are known to be activated upon cell adhesion to the ECM⁶²³. The Rho GTPases regulate key cellular motility and invasion aspects, encompassing cellular polarity, cytoskeletal reorganisation, and signal transduction pathways⁵⁹⁶. The p21-activated kinases (PAK) PAK1, PAK2, PAK3 and PAK4 are direct effectors of RHO GTPases such as RAC1 and Cell division control protein 42 homolog (CDC42)^{624,625}. PAKs are activated by RHO GTPase binding. The autoinhibitory conformation of PAK is disrupted upon binding to an active RHO GTPase, a process that triggers the activation of PAK's catalytic domain through phosphorylation. Once activated, PAK kinase assumes a crucial role in cellular regulation by phosphorylating myosin light chain kinase (MLCK), causing its inactivation^{626,627}. Consequently, this phosphorylation event inhibits myosin light chain (MLC) phosphorylation and decreases cellular contractility^{626,628}. Actin-myosin contraction regulates cell shape⁶²⁹, and generates forces to promote the cell's migration^{630,631}. As such, the loss of contractility can compromise the cell's ability to move. The upregulation of 'RHO GTPases activate PAKs' in LMS patients with high ZYX expression suggests that ZYX in LMS might inhibit LMS cellular migration through the downstream activation of RHO GTPases, the PAK signalling and loss of actin-myosin contraction.

In Chapter 6, my analysis alluded to the potential binding partners of ZYX and ZYX-associated signalling pathways in LMS. However, validation in patient-derived LMS cell lines is required to confirm that the identified proteins physically associate with ZYX *in vitro*. One approach to determine the actual

cell-specific adhesion partners can be achieved through isolation and proteomic characterisation of IACs^{430,632–634}. The process involves plating the cells on a 2D substrate, most commonly on fibronectin, to allow the formation of adhesion complexes. Next, membrane-permeable chemical reagents are applied that specifically modify and cross-links individual components of the assembled adhesion complex, capturing and preserving the complex during cell lysis and protein extraction. The cross-linked complexes are isolated, enriching for adhesome proteins and analysed by mass spectrometry. An extension to the protocol involves a phosphopeptide enrichment step prior to the mass spectrometry analysis⁶³². Adhesome complexes recruit a large number of kinases (e.g. ILK, FAK, Src) and phosphatases (PP2CA, SHP1, SSH1), and thus, phosphorylation plays a major role in adhesion signalling^{295,635,636}. Characterising phosphorylation events within the adhesion complexes can clarify how the adhesion-associated kinases regulate the adhesion complexes formation and understand their role in regulating cell behaviour post the cell binding to the ECM. Future work could utilise ZYX shRNA knockdown LMS cells generated in this thesis and plating them on the LMS patient-derived pre-gel ECM coating (discussed in Chapter 5) to study the composition of ZYX adhesome and phosphoproteomic analysis to understand how ZYX impacts the adhesion signalling. Understanding the molecular mechanisms underlying these interactions and signalling pathways could provide valuable insights into the role of ZYX in LMS progression and may open up avenues for targeted therapeutic interventions.

6.3.2. ZYX role in cell growth

The association of ZYX expression and MFS in our LMS cohort provided the rationale for investigating the functional role of ZYX in LMS cell lines. Chapter 6 assessed the role of ZYX in LMS cell growth in three LMS cell lines derived from two patients. shRNA-mediated ZYX knockdown reduced cell growth of SHEF-LMS ws using two distinct shRNA constructs and SHEF-LMS w1 cells using three constructs. One construct reduced the growth of SK-UT-1 cells in a short-term assay. Potentially, ZYX knockdown produced less growth inhibition of SK-UT-1 than in SHEF-LMS cells due to the faster-growing rate

of the SK-UT-1 cell line. Nevertheless, cell growth inhibition was observed in a longer-term assay with at least two shRNA constructs in each cell line, increasing our confidence that ZYX can promote LMS cell growth. Whilst there are no studies investigating the role of ZYX in sarcoma, ZYX was shown to promote cell growth in other cancer types. Silencing ZYX with siRNA decreased cell growth in oral squamous carcinoma cells, compared to siRNA control cells⁶¹². shRNA mediated knockdown of ZYX in colon cancer cells, impaired cell proliferation *in vitro* and reduced tumour formation *in vivo*⁶¹³. In the same study, ZYX was shown to promote colon cancer growth via activating Hippo–yes-associated protein (YAP) signalling activity by protein kinase cyclin-dependent kinase 8 (CDK8). Further experiments are required to determine the mechanism whereby ZYX promotes the LMS cell growth.

6.3.3. ZYX role in migration

In Chapter 6, I silenced ZYX expression using three distinct shRNA constructs in three LMS cell lines and evaluated the effect on migration speed and directionality. Knockdown of ZYX did not affect the directionality of SK-UT-1, SHEF-LMS w1 and SHEF-LMS ws. Regarding the cell speed, ZYX knockdown with two distinct shRNA constructs significantly increased the cell speed of the SHEF-LMS ws cell line. In SHEF-LMS w1, silencing ZYX with one shRNA construct (ZYX1) significantly decreased the cell speed. However, as the other two shRNA constructs did not affect the SHEF-LMS w1 cell speed, there is insufficient evidence to support the finding. The knockdown of ZYX with only one shRNA construct (ZYX2) promoted SK-UT-1 cell speed, and as such, there is a lack of evidence for the role of ZYX in this cell line. My results indicate that ZYX does not influence the directionality of LMS cells. However, as it currently stands, the results of the role of ZYX in influencing cell speed in LMS cell lines are inconclusive. Further investigations of the migration of LMS cell lines with other shRNA constructs or CRISPR-mediated ZYX knockout would provide more concrete evidence.

Studies investigating the role of ZYX in the migration of LMS and other STS are lacking, whilst the evidence in other cancer types is contradicting. The

association of high ZYX expression and longer MFS in our LMS cohort and the evidence from the SHEF-LMS ws cell line suggest that ZYX may suppress LMS metastasis. A study on osteosarcoma reported increased migration in the wound-healing assay and increased invasion in a transwell assay when ZYX was silenced with shRNA in 143B and U2OS cells⁶¹⁶. Moreover, ZYX overexpression in the same study produced the opposite effect on osteosarcoma migration and invasion. Whilst these results are consistent with increased cell speed upon ZYX knockdown in SHEF-LMS ws cell line, the authors only investigated one construct to silence and one to overexpress ZYX expression⁶¹⁶. As such, more robust evidence is required to confirm the role of ZYX as a suppressor of migration and invasion. Another study aligned with my results, where silencing ZYX with siRNA increased cell velocity in a lung cancer cell line⁶³⁷. Opposite to my results, two studies implicated ZYX as a promoter of cancer cell migration. ZYX-knockdown in breast cancer MDA-MB-231 cells showed impaired cell motility⁶³⁸. Similarly, ZYX silencing in the HCT116 colorectal cancer cell line caused decreased cell migration and invasion⁶³⁹.

A limitation of this study was that the association of ZYX expression and MFS was identified in a cohort of LMS, predominantly comprising eLMS (89% of the cohort), whereas the functional validation of the role of ZYX was performed using LMS cell lines of uterine origin. For a holistic understanding of ZYX in LMS pathobiology, follow-up studies employing LMS cell lines of extra-uterine origin are required. For example, further investigations could include introducing ZYX knockdown in the ICR-LMS-6 cell line (from LMS originating in the stomach) derived in our laboratory and subjecting the cells to similar assays described in this chapter.

Overall, this chapter contributes valuable insights into the potential role of ZYX in LMS growth and metastasis, offering a foundation for future research to unravel the molecular mechanisms underlying ZYX's impact on LMS pathobiology and its potential as a therapeutic target in this challenging STS subtype. Future work should include assessments of ZYX knockdown on LMS cell invasion and on tumour growth and metastasis formation *in vivo*.

Additionally, other proteins (e.g. MMP9, ITGAM and SORBS3) on the list of pre-ranked matrisome and adhesome proteins significantly associated with survival outcome measures in LMS should be investigated.

6.4. Supplementary Material

6.4.1. Supplementary Tables

Supplementary Table 6.1 Summary of multivariable (MVA) Cox regression analysis assessing the association of leiomyosarcoma clinicopathological factors and ZYX expression with metastasis-free survival (MFS).

Clinicopathological factors which were significantly associated with MFS in univariable analysis were included in the MVA model. Hazard ratio (HR), 95% confidence intervals (CI) and p-values were determined by univariate Cox regression with a two-sided Wald test. Significant associations are in red.

Variable	Groups	n	Multivariable analysis (MFS)	
			HR (95% CI)	p-value
Grade	2 (reference)	47	-	-
	3	33	2.83 (1.45-5.54)	0.002
Anatomical location	Intra-abdominal/Retroperitoneal/Pelvic	38	-	-
	Extremity/Trunk	33	0.40 (0.19-0.82)	0.013
	Uterine	9	1.31 (0.46-3.72)	0.617
Log [tumour size] (mm)	4-5 (reference)	50	-	-
	<4	12	0.87 (0.25-2.99)	0.828
	>5	17	0.51 (0.23-1.1)	0.087
Tumour depth	Deep (reference)	66	-	-
	Superficial	14	0.38 (0.09-1.53)	0.173
ZYX expression	ZYX low (reference)	40	-	-
	ZYX high	40	0.43 (0.23-0.81)	0.009

Chapter 7 Conclusions and Future Directions

The sarcoma community has undertaken considerable efforts to understand the STS pathobiology, as evidenced by large-scale genomic, transcriptomic and epigenetic studies¹⁻⁵. Whilst these studies have provided valuable insights into the molecular drivers of STS, they have centred on the tumour cells without a specific focus on the components of TME. In particular, the interaction between tumour cells and the ECM and the composition of the ECM remains insufficiently investigated in the context of STS.

To address this gap in the knowledge of STS pathobiology, this thesis leveraged a recent large-scale proteomic study of STS by our laboratory⁶ and provided a comprehensive analysis of the ECM composition and associated integrin adhesion signalling across multiple STS subtypes in Chapter 3. The analysis revealed significant diversity in the expression profiles of matrisome and adhesome proteins among various histological subtypes of STS, highlighting the unique composition of the ECM and integrin adhesion-related proteins in each subtype. These distinct protein profiles may be associated with specific biological processes governed by the ECM in STS. Overall, three main networks of co-regulated ECM and integrin adhesion-related proteins described biological processes related to the innate immune system in UPS and DDLPS, laminin interactions and integrin signalling in LMS and collagen biosynthesis and degradation in DES. In addition to unique matrisome and adhesome profiles, LMS, DDLPS and UPS shared common ECM-related pathobiology. Specifically, a proteoglycan score based on the protein expression of 11 proteoglycan molecules was negatively associated with OS and inversely correlated with TILs levels in a sub-cohort with LMS, DDLPS and UPS histologies. Adding the CD8+ TILs score into the proteoglycan score identified a high-risk group of STS patients with significantly inferior 5-year OS.

In addition to the inter-subtype heterogeneity, substantial intra-subtype heterogeneity of matrix signalling was observed in LMS, DDLPS and UPS in Chapter 4. This heterogeneity can potentially underpin the clinical heterogeneity observed in these STS subtypes. Specifically, intra-subtype

analyses identified subgroups of LMS, DDLPS and UPS patients based on similarities of matrisome and adhesome expression profiles. In LMS, the subgroups were associated with tumour grade. In DDLPS, the DDLPS3 subgroup was an independent prognostic factor for LRFS and OS. In UPS, subgroups were associated with TIL levels and MFS. To enable studying the role of the ECM *in vitro*, two preclinical ECM models were successfully derived and characterised in Chapter 5. Chapter 6 identified ZYX as a putative suppressor of metastasis in LMS, and phenotypic assays using ZYX knockdown cell lines implicated ZYX as a potential promoter of LMS cell growth. The future directions and implications associated with the key findings of this thesis are discussed below.

7.1. Proteoglycan and CD8+ TILs score is prognostic in LMS, DDLPS and UPS cohort

The heterogeneity of STS, as seen in the clinic is well established^{3,20,462,463}. Despite having the same histological diagnosis, STS patients may experience different symptoms, disease progression rates, and treatment responses and thus have strikingly different prognoses^{20,40,62,640}. Common clinical approaches to disease risk stratification include staging and grading, and while these have improved STS management, they have several limitations. For instance, although the FNCLCC grading stands as an important prognostic factor of MFS and OS in STS and is incorporated in numerous nomograms, its reliability remains constrained by variations in interpretation among different pathologists, particularly regarding the differentiation score^{46,47}. As such, there is a scope for improving the risk stratification in STS. Novel molecular biomarkers can potentially capture additional aspects of STS tumour behaviour, which are not addressed by the traditional prognostic clinicopathological factors. The work in this thesis investigated the ECM and integrin adhesion-associated proteins as a source of prognostic biomarkers, which could improve risk stratification in STS.

The results in Chapter 3 were hypothesis-generating and provided the basis for future prognostic biomarker development. As discussed in Chapter 3, low

proteoglycan score and low CD8+ TILs score represent a putative biomarker that identified the STS patient subgroup with significantly shorter OS than the other patient subgroups. The score was prognostic of OS even after adjusting for known prognostic risk factors such as age, grade, STS subtype, tumour size, sex and performance status. However, as with other retrospective studies, the derivation of proteoglycan/CD8+ TILs score from retrospectively collected samples and clinical data presents some limitations⁶⁴¹. Those include a lack of prospectively determined eligibility criteria and power calculations⁶⁴¹. In this thesis, there were no prospectively determined inclusion criteria for clinicopathological features, which resulted in missing data entries. For instance, grade and performance status information were missing for some patients, meaning those confounding factors could not be accurately accounted for in our multivariable analysis model. Moreover, the patient inclusion in this study was based on the availability of tumour tissue for mass spectrometry analysis, and it could have introduced selection bias. Finally, the current proteoglycan/CD8+ TILs score was developed on samples with LMS, DDLPS and UPS diagnoses. Consequently, due to clinical STS heterogeneity, the prognostic applicability of proteoglycan/CD8+ TILs score to other histological subtypes remains uncertain.

Future research efforts should prioritise validating the proteoglycan/CD8+ TILs score in an independent prospective patient cohort, including other histological subtypes, and assess its performance compared to existing risk stratification tools employed in clinical practice for STS. The prospective study design offers the advantage of implementing rigorous controls over specimen collection, processing, storage, strict eligibility criteria, and collection of all the necessary clinicopathological and endpoint survival data. The validation in an independent cohort will ensure that the proteoglycan/CD8+ TILs score biomarker is not a result of data noise or sample selection bias, demonstrating that it is robust, generalisable and has a true clinical value.

Mass spectrometry proteomics is predominantly utilised for research due to complex sample processing, specialised expertise and training required for data analysis, and high associated costs^{642,643}. These factors present

challenges in integrating mass spectrometry into routine clinical STS practice. To facilitate the translation of the prognostic biomarker based on the proteoglycan score and CD8+ levels, it will be necessary to develop a standardised, reliable, easy-to-implement and relatively cheap assay for measuring the biomarker in STS clinical samples. Firstly, the approaches might involve reducing the list of 11 proteoglycans to fewer proteoglycans, individually the most prognostic, and ensuring that it would not compromise the risk stratification value of the biomarker. Then, the orthogonal validation methods such as IHC could be used to validate that the expression of 1-2 selected proteoglycans and CD8+ score can stratify STS patients according to OS risk. The scoring system should then be established to determine the criteria for classifying STS patients as having a 'low' or 'high' proteoglycan/CD8+ TILs score. IHC is a routinely used tool in the clinical diagnosis of STS, and IHC-based biomarker assay development will accelerate the translation of proteoglycan/CD8-based prognostic biomarker in clinical STS management.

Regarding the long-term future implications, the proteoglycan/CD8+ TILs score can be integrated into nomograms, allowing for more precise risk stratification and improving the prognostic power of nomograms. This can help identify high-risk patients who may benefit from more aggressive treatments and spare low-risk patients from unnecessary interventions.

On top of that, proteoglycan/CD8+ TILs score biomarker discovery advanced our understanding of the underlying pathobiology of STS, and can potentially open up novel treatment options. Our data suggest that proteoglycans play a tumour-suppressive role, with higher proteoglycan expression associated with longer OS. At present, increasing the protein level of proteoglycans in tumours is a therapeutic challenge. A potential alternative strategy to increase the level of proteoglycans in a tumour can involve targeting proteoglycan-degrading enzymes such as MMPs, heparinase and a Disintegrin And Metalloproteinase with Thrombospondin motifs (ADAMTS) enzymes⁶⁴⁴⁻⁶⁴⁶. Additionally, our data shows that high proteoglycan expression might be associated with low TILs levels. Further work could experimentally investigate the molecular

mechanism of proteoglycans and CD8+ TILs interaction. Understanding the molecular mechanisms underlying the interaction between proteoglycans and CD8+ TILs could inform the development of immunotherapy strategies. If these interactions are inhibitory, finding ways to modulate the proteoglycans and CD8+ TILs interactions could enhance the effectiveness of immunotherapies in STS.

7.2. Implications of matrix signalling heterogeneity in DDLPS and UPS

The analysis in Chapter 4 showed that LMS, DDLPS and UPS are biologically heterogeneous diseases displaying intra-subtype differences in the expression of ECM and integrin adhesion signalling proteins. All three subtypes were classified based on the matrix signalling features into subgroups associated with clinicopathological features and survival outcomes. Additionally, DDLPS and UPS subgroups were associated with TIL levels.

7.2.1. DDLPS

Intra-subtype analysis of matrix signalling in DDLPS identified three subgroups (DDLPS1, DDLPS2 and DDLPS3), which were associated with tumour size, patient sex and performance status. Moreover, the DDLPS3 subgroup was prognostic for LRFS and OS, even after adjusting for other known prognostic factors. In the near term, it would be crucial to validate the molecular DDLPS classification based on matrisome and adhesome profiles, ensuring its reproducibility and assessing the prognostic value of the DDLPS3 group for LRFS and OS in an independent cohort. Following that, future efforts would benefit from developing a prognostic classifier in DDLPS based on matrisome and adhesome features. In DDLPS, local recurrence is a common event contributing to poor OS. By identifying DDLPS patients at higher risk of local recurrence, clinicians can implement more vigilant monitoring and potentially more aggressive treatments. Conversely, for patients at lower risk, the classifier can spare them from unnecessary aggressive therapies, minimising treatment-related burdens. Given that DDLPS matrix subtypes were identified

from surgical excision samples, the prognostic classification proposed in this thesis is most applicable in a postoperative context. The findings would need to be assessed in DDLPS biopsy samples to extend its utility to risk stratification at the diagnosis. Likewise, the molecular subtypes would need to be assessed in recurrent or metastatic disease samples to evaluate the prognostic significance in advanced settings.

The DDLPS3 group was characterised by a higher expression of 9 matrisome and adhesome proteins (CTSA/B/Z, PLAUR, ITGB2, ACTN1, TNC, SLC3A2 and PLOD2) compared to DDLPS1 and DDLPS2, and low expression of DDLPS1- and DDLPS2-specific matrisome and adhesome proteins. The development of a prognostic classifier could involve the identification of matrisome and adhesome proteins which best discriminate the DDLPS3 group from the rest of the DDLPS patients in the cohort presented in this thesis, which could be considered a training cohort. Next, a reliable, standardised and feasible for routine clinical use assay would need to be developed to classify DDLPS patients based on the pre-specified cut-offs of expression of selected proteins in FFPE samples. As discussed in Section 7.1.1, mass-spectrometry-based proteomics is not currently used in the clinical setting, and the development of the classifier should involve using IHC-based biomarker assays. The developed classifier assay should be validated on an independent cohort of DDLPS, which is prospectively collected and includes samples representative of the general DDLPS patient population.

DDLPS is considered relatively chemo-resistant with a low response rate to systemic chemotherapy⁷⁸. Recently developed CDK4 and MDM2 targeting inhibitors showed limited efficacy in the clinical trials^{117,122}. Consequently, there is a growing need to identify other molecular targets for the development of novel therapies. The molecular stratification in DDLPS presented in this thesis supports the use of targeted therapies, such as strategies targeting a component of the DDLPS adhesome PLAUR encoding uPAR. uPAR is a membrane-bound protein, and it serves as the receptor for urokinase-type plasminogen activator (uPA)⁶⁴⁷. Once bound to uPAR, uPA catalyses the conversion of plasminogen, generating the protease plasmin, which breaks

down the ECM during processes like fibrinolysis, wound healing, and tumorigenesis. Additionally, uPAR operates as a signalling receptor, facilitating the mobility, infiltration, and survival of cancer cells⁶⁴⁷. The high-risk patient group DDLPS3 was characterised by high expression of uPAR, a target for novel CAR T-cell and ARMs immunotherapies showing promising results in preclinical settings in other cancer types^{523,524}. Future directions could include preclinical testing of uPAR-targeted immunotherapy in co-cultures of peripheral blood mononuclear cells (PBMCs) from DDLPS patients and matched patient-derived DDLPS cell lines. Additional preclinical work could investigate whether uPAR-targeted immunotherapy can enhance the efficacy of existing CDK4 and MDM2 inhibitors. Overall, next-generation targeted therapies, such as uPAR-targeted immunotherapies, could provide a much-needed shift away from the 'one-size-fits-all' treatment approach in DDLPS towards a personalised medicine with stratification to therapies based on molecular characteristics and underlying DDLPS biology.

7.2.2. UPS

Intra-subtype analysis of matrix signalling in UPS identified four subgroups (UPS1, UPS2, UPS3 and UPS4). The UPS1 subgroup was characterised by high expression of fibrillar collagens and collagen synthesis, maturation and degradation enzymes. In addition, the UPS1 was associated with lower CD3+, CD4+ and CD8+ TILs levels than the rest of the subgroups. In a multivariable analysis after adjusting for age, anatomical site, tumour size, sex and performance status, the UPS1 subgroup showed significantly shorter MFS than the reference UPS4 subgroup.

Importantly, the matrix signalling analysis identified putative therapeutic targets for the high-risk UPS1 subgroup of patients. The analysis nominated MMP14 based on a higher expression of MMP14 in UPS1 compared to the rest of the UPS subtypes. MMP14 is a membrane-bound zinc-dependent protease with a wide range of ECM substrates that mainly targets extracellular collagens and activates other MMPs, such as MMP2 and MMP13^{238,648}. MMP14 is highly expressed in many cancer types⁶⁴⁹, including sarcoma⁶⁵⁰,

and MMP14 activity promotes cancer cell migration and invasion, metastasis and angiogenesis^{651–655}. A peptide drug conjugate BT1718 targeting MMP14 is undergoing assessment in a phase I/II clinical trial targeting solid cancers, including sarcoma; however, the information about sarcoma histological subtypes is not provided (NCT03486730). At the progress report update, 13 of 24 evaluable patients (54%) had stable disease at eight weeks, including a patient who experienced a 45% reduction in a target lesion⁵³⁰. However, the progress report did not report whether any responders had a sarcoma diagnosis. Nevertheless, the drug is well tolerated, and the dose escalation study is ongoing. The results in this thesis suggest that the high-risk UPS1 subgroup might benefit from treatment with BT1718. Provided that the NCT03486730 trial shows the efficacy of BT1718 in sarcoma patients, MMP14 expression could potentially be used as a predictive biomarker to stratify UPS patients to BT1718 treatment.

Another potential target in the UPS1 subgroup is related to reducing collagen density and cross-linking by targeting collagen-modifying enzymes. The density of fibrillar collagens negatively correlates with TILs levels in other cancers^{413,534–536}. The UPS1 subgroup in my analysis, which showed high expression of fibrillar collagens and collagen biosynthesis enzymes, was associated with the lowest TILs infiltrate compared to other UPS subtypes. As it stands, this finding represents an observation without establishing a causative relationship. Further work is required to experimentally confirm whether increased fibrillar density or collagen-crosslinking regulates TILs infiltration in UPS tumours. Increased collagen deposition and crosslinking were therapeutically targeted to enhance CD8+ TILs infiltration and improve the response rates to PD-1 and PD-L1 blockage in preclinical models of lung and pancreatic cancers^{537,539}. A study by Peng *et al.* showed that mouse lung cancer models with inherent or acquired resistance to PD-1/PD-L1 blockage had increased collagen density and decreased levels of CD8+ TILs⁵³⁷. Inhibiting the LOXL2 crosslinking enzyme, which stabilises and enhances collagen deposition, with shRNA or LOXL2 inhibitor, ellagic acid, significantly decreased collagen deposition and cross-linking whilst increasing CD8+ TILs levels. Moreover, the combination of ellagic acid and anti-PD-L1 or anti-PD1

antibodies resensitised the resistant lung tumours, diminishing the primary tumour growth compared to the control and individual treatment⁵³⁷. Another study showed that inhibition of LOX, another collagen-crosslinking enzyme, with LOX inhibitor beta-aminopropionitrile (BAPN) reduced collagen density in 3/5 mouse pancreatic cancer models⁵³⁹. Moreover, a combined treatment with BAPN and anti-PD-1 antibody increased CD8 T cell accumulation in the tumour and reduced tumour growth compared to the control and individual treatments. Overall, these studies demonstrate the potential to enhance the efficiency of PD-1/PD-L1 blockage by inhibiting collagen deposition and crosslinking. Further work could investigate if inhibition of collagen-crosslinking enzymes in UPS would decrease collagen density and increase TILs levels. An immunocompetent genetically engineered mouse model of UPS has been previously reported^{656,657}. Primary UPS tumours are formed using the Cre-loxP system to activate conditional mutations in oncogenic *Kras* and the tumour suppressor *p53* in gastrocnemius muscle of LSL-*Kras*^{G12D/+}; *p53*^{flox/flox} (KP) mice⁶⁵⁶. Given that this UPS model has an intact immune system, it can be used to investigate the effect of inhibitors of collagen-crosslinking enzymes on TILs levels and the combination treatment of the inhibitors and PD-1/PD-L1 inhibitors in UPS.

In phase II clinical studies, a subset of UPS patients show responses to an ICI, pembrolizumab, which targets PD1 on T-cells^{79,151}. Pembrolizumab is a monoclonal antibody that blocks PD1, preventing PD1 to interact with the inhibitory PD-L1 on tumour cells. In the absence of a negative regulatory signal from PD-L1, T-cells can mount an anti-tumour immune response. Generally, STS tumours with increased T cell infiltration (i.e., 'immune-hot') have higher response rates to ICI. As such, the UPS1 subgroup, with the lowest TILs levels (i.e., 'immune-cold'), might show poor response to ICI. For 'immune-cold' UPS1 patients, combination options with ICI and other compounds, such as inhibitors of collagen-modifying enzymes, could be assessed. Further work could investigate the potential to increase TILs levels and improve the response rates to ICI in UPS mouse models treated with collagen cross-linking inhibitors. While further extensive experimental research is warranted, inhibiting collagen deposition and cross-linking might be a possible

combination strategy with ICIs, including pembrolizumab, for UPS ‘immune-cold’ patients.

In summary, the analysis in Chapter 4 showed that proteomic heterogeneity of matrix signalling in LMS, DDLPS and UPS might underpin some of the clinical heterogeneity observed in those subtypes. Furthermore, proteomic stratification based on matrix signalling features in DDLPS is prognostic of LRFS and OS and nominated several ECM and adhesion-related therapeutic options for high-risk DDLPS and UPS patients.

7.3. Future work related to preclinical ECM models

Components of the TME, such as the ECM, play pivotal functional roles in promoting numerous cancer hallmarks⁴¹⁰. Achieving a comprehensive understanding of cancer-ECM interactions and effectively utilising them as targets for drug development demands preclinical model systems that closely recapitulate the ECM and allow control over matrix properties.

The results in Chapter 5 highlight the importance of incorporating patient-relevant ECM when assessing LMS cellular phenotypes. This is evidenced by the finding that compared to standard plastic culture surface, LMS-patient-derived ECM pre-gel coating significantly increased the directionality of LMS cell lines with low baseline cell migration speed and directionality index. The results also showed that the choice of ECM coating affects the speed and directionality of LMS cells. The most prominent example was a significant decrease of directional persistence in 4/5 LMS cell lines on laminin compared to LMS ECM coating. While this serves as an initial step towards understanding the functional roles of specific ECM components, further work is required to dissect the exact mechanistic roles of individual ECM components and integrin signalling components in shaping the LMS migration, invasion and drug response. Further work should include rational selection of candidate ECM and integrin adhesion components for follow-up experiments, for instance, based on their association with LMS survival outcome measures.

Understanding the mechanism of ECM-integrin interactions will be crucial for developing targeted interventions for LMS.

The drawback of my work is the lack of assessment of LMS cellular phenotypes in 3D, with LMS cells grown in ECM gel or scaffold. Considering that LMS patient-derived preclinical ECM hydrogels were obtained from a limited resource of patient specimens, the experiments in Chapter 5 were restricted to the use of thin coatings, and all phenotypic assays were performed in 2D. It is known that the external cues experienced by a cell vary significantly when comparing an ECM-coated surface (2D) to a 3D ECM gel⁶⁵⁸. For instance, cellular adhesions in 2D are restricted to the x-y plane, whilst in 3D, adhesions are distributed in all three dimensions. Adhesions in 3D were also shown to resemble adhesions in vivo more closely than adhesions on 2D substrates^{659,660}. Previous studies showed differences in the composition of adhesion structures and adhesion signalling on 2D ECM coatings versus 3D ECM gels^{661–663}. Moreover, dimensionality (2D vs. 3D) and the composition of ECM substrates (cell-derived matrix vs. collagen I, fibrin and Matrigel) significantly influenced fibroblast speed⁶⁶¹. For instance, whilst in 3D, fibroblasts moved faster on cell-derived matrix (which is a matrix secreted by cells in 2D culture) than in other matrices, whereas in 2D, Matrigel induced the fastest migration⁶⁶¹. As such, biochemical composition and dimensionality both play important roles in cell-matrix interactions and influence cellular phenotypes such as migration. Future work should include an assessment of LMS cell phenotypes in a 3D context. This can be accomplished by encapsulating cancer cells in LMS ECM pre-gel solution and polymerising to form a 3D hydrogel or repopulating the decellularised ECM scaffolds with LMS cells.

Chapter 5 established the feasibility of generating LMS-specific preclinical ECM models (decellularised ECM scaffolds and ECM hydrogels). Characterisation of the protein composition of the models confirmed that the models recapitulate key LMS matrisome components, such as basement membrane proteins. The analysis also enabled me to understand the limitation of the models, namely the presence of other cellular components and loss of

soluble matrisome-associated components in decellularised scaffolds and hydrogels. Therefore, although the major structural ECM proteins of LMS can be replicated in the ECM models, achieving a comprehensive representation of all LMS matrisome components is challenging, and the functional implications of other cellular proteins on cell phenotypes are unknown. To increase the complexity of the preclinical ECM models in the future, decellularised ECM scaffolds and ECM pre-gel solutions can be supplemented with additional soluble ECM components, which are lost during the generation of the models.

Chapter 4 described intra-subtype heterogeneity in matrix signalling in LMS tissue samples. The analysis identified two distinct LMS subgroups with different compositions of basement membrane proteins and a third LMS subgroup that lacked the basement membrane proteins. Due to the limited patient tumour material, the pre-gel ECM solutions from LMS patients were combined, and the implications of intra-subtype matrix heterogeneity on LMS cellular phenotypes could not be assessed. In future assessments, a prospective approach could involve the classification of LMS samples into one of the LMS subgroups, as identified in Chapter 4, before the generation of ECM scaffolds and ECM hydrogels. Then, it could be investigated whether ECM models, classified according to the LMS subgroups definition, would elicit differential LMS cell migration, invasion and drug response. This would establish the causal relationship between the molecular composition of LMS ECM and LMS pathobiology.

7.3.1. Synthetic extracellular matrices

Another limitation of ECM preclinical models generated in Chapter 5 is that they are a limited material. Given the recent advances in bioengineering, it is feasible to design synthetic hydrogels incorporating functional ECM peptides⁶⁶⁴, which are easily scalable and can be generated on demand. Other advantages of synthetic versus natural hydrogels are high reproducibility due to the chemical synthesis of base matrix and synthetic peptide and the absence of other cellular proteins inherent to the tissue-derived hydrogel

preparations. For instance, polyethylene glycol (PEG) can be used as a base matrix, which can be further customised by incorporating ECM peptide motifs with various functions^{665,666}. These ECM motifs can either promote integrin binding^{667,668}, facilitate cell-mediated matrix degradation⁶⁶⁹, enable the binding of growth factors⁶⁷⁰ or mimic the growth factors and other molecules.

Recently, a tissue-inspired PEG hydrogel containing 20 unique ECM-related peptides was engineered⁶⁶⁶. Peptides for 13 cell adhesion integrin-binding and 7 MMP-degradable peptides have been designed to mimic the diversity of ECM proteins in human bone marrow. The peptides were selected from native tissue through a proteomic-based and biomechanics method. The specific combination of bone-marrow-derived adhesion cues and cleavable sites provided a biomimetic niche for MSC growth and differentiation⁶⁶⁶. Similarly, a PEG-based hydrogel was engineered for the brain, containing 9 integrin-binding peptides and 5 MMP-degradable peptides found in native brain tissue⁶⁶⁵. This hydrogel enabled the precise *in vitro* control and maintenance of astrocyte quiescence with minimal activation.

Synthetic hydrogels have important advantages over natural hydrogels. However, due to the intricate nature of ECM composition and structure, fabricating a synthetic STS ECM material that completely replicates the biochemistry and architecture of native STS tissue ECM is likely an unattainable goal. The future of preclinical ECM models in STS likely involves a combination of synthetic and STS-derived ECM material to overcome the limitations associated with each technology.

7.4. Zyxin in LMS metastases

LMS has a high tendency for distant metastasis, which is associated with increased mortality risk²⁰. However, the underlying mechanisms and the drivers of metastatic spread in LMS are poorly understood. Consequently, there is a lack of in-depth research into potential therapies that could specifically target and inhibit the metastatic spread of LMS.

Chapter 6 highlighted putative matrisome and adhesome components which could be involved in LMS progression and metastatic spread. The analysis nominated ZYX for validation in preclinical LMS models, and the role of ZYX was investigated in LMS cell growth and migration. In our LMS cohort, high ZYX expression was significantly associated with longer MFS, implicating ZYX as a metastasis suppressor. Metastasis is a complex multistep process in cancer involving local invasion, intravasation into blood or lymphatic vessels, survival in circulation, arrest at distant organs, extravasation, micrometastasis formation, and eventual macroscopic metastasis growth⁶⁷¹. This thesis investigated the role of ZYX in the growth and 2D migration of LMS cell lines and was a first step in dissecting the role of ZYX in LMS tumour growth and metastasis. It is important to note that the two phenotypic assays used in Chapter 6 do not sufficiently reflect metastasis-related processes, and further investigations are warranted. Other functional validations are currently underway in our laboratory, including invasion assays and mouse experiments. Spheroid invasion assays in collagen I and LMS-derived ECM hydrogels are currently assessing the role of ZYX in LMS cell invasion. Additionally, subcutaneous LMS xenografts have been established using SK-UT-1 cells, and mice were randomised on a 20 g/kg doxycycline-containing or control diet to induce ZYX knockdown *in vivo*. The high-concentration doxycycline-diet was well tolerated, and ZYX knockdown will be investigated in the FFPE samples by IHC and in fresh frozen tumour samples by western blotting. Provided *in vivo* ZYX knockdown is achieved, further mouse experiments will be planned to assess the implication of ZYX knockdown on tumour growth *in vivo* of subcutaneously injected cells. ZYX shRNA-carrying SK-UT-1 cells will be transduced with luciferase construct to enable the *in vivo* monitoring of metastasis formation and the metastasis model will be established using a tail-vein injection.

7.4.1. Therapeutic targeting of ZYX

Chapter 6 implicated ZYX as a regulator of LMS cell growth *in vitro*, and as such, ZYX represents a potential target for anti-tumour drug development for LMS. ZYX is an intracellular adaptor protein which regulates the actin

cytoskeleton. ZYX is a scaffolding protein and lacks enzymatic domains (e.g. tyrosine or serine/threonine kinase domains), and poses a challenge to designing traditional small-molecule inhibitor drugs that directly target and inhibit ZYX. For alternative ways to inhibit ZYX, the ZYX-interacting protein network could be investigated to identify targetable kinases upstream or downstream of ZYX. It would be important to determine which protein-protein interactions are crucial for ZYX-mediated LMS pathogenesis to prioritise putative targets for drug design. Analysis of pathways enriched in LMS patients with high ZYX protein expression compared to patients with low ZYX in Chapter 6 can be considered an initial step towards decoding ZYX binding partners. The analysis identified ILK kinase, which could act upstream of ZYX⁶²², RHOA GTPases and PAK kinase, although it is less clear how the latter interact with ZYX due to a lack of experimental evidence of their interactions. Several PAK inhibitors were developed and are investigated in preclinical studies, and two compounds targeting PAK4 were evaluated in clinical trials for solid cancers and blood malignancies^{672,673}. One of the PAK inhibitors did not show an objective response rate and had unexpected toxicities, and the trial was terminated⁶⁷⁴. The other compound is being investigated in phase I clinical trials⁶⁷² and can potentially be repurposed for LMS.

7.4.2. Proteolysis-targeting chimaera technology

Proteolysis-targeting chimaera (PROTAC) technology is an innovative approach to drug development and allows targeting previously 'undruggable' proteins. PROTACs form a ternary complex within cells, consisting of a ligand that binds to the target protein and another ligand that recruits an E3 ubiquitin ligase. This complex leads to the ubiquitination of the target protein, subsequently marking it for degradation by the cellular proteasome machinery. Multiple PROTAC entered clinical trials for solid cancers, including sarcoma, and many more are in preclinical development⁶⁷⁵. FHD-609 PROTAC targeting chromatin remodelling protein, bromodomain-containing protein 9 (BRD9), is under investigation in a phase I clinical trial, which recruits advanced synovial sarcoma patients (NCT04965753), indicating that the

PROTAC technology is already being used in sarcoma patients. In Chapter 6, shRNA-mediated knockdown of ZYX reduced LMS cell growth and using PROTAC to degrade ZYX in LMS patient tumours could be a viable option in the future.

7.5. Concluding remarks

In conclusion, my work has comprehensively characterised the ECM composition and associated integrin adhesion signalling across multiple STS subtypes, identifying subtype-specific ECM biology. Moreover, my project uncovered matrix signalling heterogeneity in LMS, DDLPS and UPS, which was associated with the clinical heterogeneity of these subtypes. The analysis in this thesis nominated several ECM-related prognostic biomarkers for future development and identified putative therapies for specific subtypes. Using LMS patient-derived ECM in phenotypic assays reiterated the importance of studying STS pathobiology in preclinical models capturing physiologically relevant TME components. Finally, the characterisation of matrix signalling components in STS subtypes provided a valuable resource for future analysis to identify prognostic biomarkers and drug targets and could be used to inform the development of next-generation preclinical ECM models in STS.

Chapter 8 References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin.* 2015;65(1):5-29.
2. Kannan S, Lock I, Ozenberger BB, Jones KB. Genetic drivers and cells of origin in sarcomagenesis. *J Pathol.* 2021;254(4):474-493.
3. Brennan MF, Antonescu CR, Moraco N, Singer S. Lessons Learned From the Study of 10,000 Patients With Soft Tissue Sarcoma. *Ann Surg.* 2014;260(3):416-422.
4. WHO Classification of Tumours Editorial Board eds. *World Health Organization Classification of Soft Tissue and Bone Tumours.* Vol 3rd. 5th ed. IARC Press; 2020.
5. Smith GM, Johnson GD, Grimer RJ, Wilson S. Trends in presentation of bone and soft tissue sarcomas over 25 years: little evidence of earlier diagnosis. *The Annals of The Royal College of Surgeons of England.* 2011;93(7):542-547.
6. Gamboa AC, Gronchi A, Cardona K. Soft-tissue sarcoma in adults: An update on the current state of histiotype-specific management in an era of personalized medicine. *CA Cancer J Clin.* 2020;70(3):200-229.
7. Howlader N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA NAM. SEER Cancer Statistics Review, 1975-2018. Published online 2020.
8. Stiller CA, Botta L, Brewster DH, et al. Survival of adults with cancers of bone or soft tissue in Europe—Report from the EURO CARE-5 study. *Cancer Epidemiol.* 2018;56:146-153.

9. Harris SJ, Maruzzo M, Thway K, et al. Metastatic soft tissue sarcoma, an analysis of systemic therapy and impact on survival. *Journal of Clinical Oncology*. 2015;33(15_suppl):10545.
10. Tap WD, Wagner AJ, Schöffski P, et al. Effect of Doxorubicin Plus Olaratumab vs Doxorubicin Plus Placebo on Survival in Patients With Advanced Soft Tissue Sarcomas: The ANNOUNCE Randomized Clinical Trial. *JAMA*. 2020;323(13):1266.
11. Rodriguez R, Rubio R, Menendez P. Modeling sarcomagenesis using multipotent mesenchymal stem cells. *Cell Res*. 2012;22(1):62-77.
12. Bi P, Yue F, Karki A, et al. Notch activation drives adipocyte dedifferentiation and tumorigenic transformation in mice. *Journal of Experimental Medicine*. 2016;213(10):2019-2037.
13. Salter DM, Griffin M, Muir M, et al. Development of mouse models of angiosarcoma driven by p53. *Dis Model Mech*. Published online January 1, 2019.
14. Guijarro M V., Dahiya S, Danielson LS, et al. Dual Pten/Tp53 Suppression Promotes Sarcoma Progression by Activating Notch Signaling. *Am J Pathol*. 2013;182(6):2015-2027.
15. Burningham Z, Hashibe M, Spector L, Schiffman JD. The Epidemiology of Sarcoma. *Clin Sarcoma Res*. 2012;2(1):14.
16. Davidson T, Westbury G, Harmer CL. Radiation-induced soft-tissue sarcoma. *British Journal of Surgery*. 1986;73(4):308-309.
17. Preston DL, Ron E, Tokuoka S, et al. Solid Cancer Incidence in Atomic Bomb Survivors: 1958–1998. *Radiat Res*. 2007;168(1):1-64.

18. Katz SC, Baldwin K, Karakousis G, et al. Soft tissue sarcoma diagnosed subsequent to lymphoma is associated with prior radiotherapy and decreased survival. *Cancer*. 2011;117(20):4756-4763.
19. Dineen SP, Roland CL, Feig R, et al. Radiation-Associated Undifferentiated Pleomorphic Sarcoma is Associated with Worse Clinical Outcomes than Sporadic Lesions. *Ann Surg Oncol*. 2015;22(12):3913-3920.
20. Gladdy RA, Qin LX, Moraco N, Agaram NP, Brennan MF, Singer S. Predictors of survival and recurrence in primary leiomyosarcoma. *Ann Surg Oncol*. 2013;20(6):1851-1857.
21. Correa H. Li-Fraumeni Syndrome. *J Pediatr Genet*. 2016;05(02):084-088.
22. Kleihues P, Schäuble B, zur Hausen A, Estève J, Ohgaki H. Tumors associated with p53 germline mutations: a synopsis of 91 families. *Am J Pathol*. 1997;150(1):1-13.
23. Bougeard G, Renaux-Petel M, Flaman JM, et al. Revisiting Li-Fraumeni Syndrome From *TP53* Mutation Carriers. *Journal of Clinical Oncology*. 2015;33(21):2345-2352.
24. Eulo V, Lesmana H, Doyle LA, Nichols KE, Hirbe AC. Secondary Sarcomas: Biology, Presentation, and Clinical Care. *American Society of Clinical Oncology Educational Book*. 2020;(40):463-474.
25. Evans DGR. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J Med Genet*. 2002;39(5):311-314.

26. Ballinger ML, Goode DL, Ray-Coquard I, et al. Monogenic and polygenic determinants of sarcoma risk: an international genetic study. *Lancet Oncol.* 2016;17(9):1261-1271.
27. Ballinger ML, Pinese M, Thomas DM. Translating genomic risk into an early detection strategy for sarcoma. *Genes Chromosomes Cancer.* 2019;58(2):130-136.
28. Chang Y, Cesarman E, Pessin MS, et al. Identification of Herpesvirus-Like DNA Sequences in AIDS-Associated Kaposi's Sarcoma. *Science (1979).* 1994;266(5192):1865-1869.
29. Boshoff C, Schulz TF, Kennedy MM, et al. Kaposi's sarcoma-associated herpesvirus infects endothelial and spindle cells. *Nat Med.* 1995;1(12):1274-1278.
30. Mesri EA, Cesarman E, Boshoff C. Kaposi's sarcoma and its associated herpesvirus. *Nat Rev Cancer.* 2010;10(10):707-719.
31. Taylor BS, Barretina J, Maki RG, Antonescu CR, Singer S, Ladanyi M. Advances in sarcoma genomics and new therapeutic targets. *Nat Rev Cancer.* 2011;11(8):541-557.
32. Skytting B, Nilsson G, Brodin B, et al. A Novel Fusion Gene, SYT-SSX4, in Synovial Sarcoma. *JNCI Journal of the National Cancer Institute.* 1999;91(11):974-975.
33. Ladanyi M, Antonescu CR, Leung DH, et al. Impact of SYT-SSX fusion type on the clinical behavior of synovial sarcoma: a multi-institutional retrospective study of 243 patients. *Cancer Res.* 2002;62(1):135-140.
34. Dickson BC, Chung CT -S., Hurlbut DJ, et al. Genetic diversity in alveolar soft part sarcoma: A subset contain variant fusion genes,

- highlighting broader molecular kinship with other MiT family tumors. *Genes Chromosomes Cancer*. 2020;59(1):23-29.
35. Abeshouse A, Adebamowo C, Adebamowo SN, et al. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell*. 2017;171(4):950-965.e28.
 36. Gounder MM, Agaram NP, Trabucco SE, et al. Clinical genomic profiling in the management of patients with soft tissue and bone sarcoma. *Nat Commun*. 2022;13(1):3406.
 37. Binh MBN, Sastre-Garau X, Guillou L, et al. MDM2 and CDK4 Immunostainings Are Useful Adjuncts in Diagnosing Well-Differentiated and Dedifferentiated Liposarcoma Subtypes. *American Journal of Surgical Pathology*. 2005;29(10):1340-1347.
 38. Gronchi A, Miah AB, Dei Tos AP, et al. Soft tissue and visceral sarcomas: ESMO–EURACAN–GENTURIS Clinical Practice Guidelines for diagnosis, treatment and follow-up☆ *Annals of Oncology*. 2021;32(11):1348-1365.
 39. Coindre JM, Terrier P, Bui NB, et al. Prognostic factors in adult patients with locally controlled soft tissue sarcoma. A study of 546 patients from the French Federation of Cancer Centers Sarcoma Group. *Journal of Clinical Oncology*. 1996;14(3):869-877.
 40. Stefanovski PD, Bidoli E, De Paoli A, et al. Prognostic factors in soft tissue sarcomas: a study of 395 patients. *European Journal of Surgical Oncology (EJSO)*. 2002;28(2):153-164.
 41. Coindre JM, Terrier P, Guillou L, et al. Predictive value of grade for metastasis development in the main histologic types of adult soft tissue sarcomas: a study of 1240 patients from the French

- Federation of Cancer Centers Sarcoma Group. *Cancer*. 2001;91(10):1914-1926.
42. Costa J, Wesley RA, Glatstein E, Rosenberg SA. The grading of soft tissue sarcomas results of a clinicohistopathologic correlation in a series of 163 cases. *Cancer*. 1984;53(3):530-541.
 43. Trojani M, Contesso G, Coindre JM, et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int J Cancer*. 1984;33(1):37-42.
 44. Guillou L, Coindre JM, Bonichon F, et al. Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. *Journal of Clinical Oncology*. 1997;15(1):350-362.
 45. Hashimoto H, Daimaru Y, Takeshita S, Tsuneyoshi M, Enjoji M. Prognostic significance of histologic parameters of soft tissue sarcomas. *Cancer*. 1992;70(12):2816-2822.
 46. Coindre JM. Grading of Soft Tissue Sarcomas: Review and Update. *Arch Pathol Lab Med*. 2006;130(10):1448-1453.
 47. Danieli M, Gronchi A. Staging Systems and Nomograms for Soft Tissue Sarcoma. *Current Oncology*. 2023;30(4):3648-3671.
 48. Hoerber I, Spillane AJ, Fisher C, Thomas JM. Accuracy of Biopsy Techniques for Limb and Limb Girdle Soft Tissue Tumors. *Ann Surg Oncol*. 2001;8(1):80-87.
 49. Ray-Coquard I, Ranchère-Vince D, Thiesse P, et al. Evaluation of core needle biopsy as a substitute to open biopsy in the diagnosis of soft-tissue masses. *Eur J Cancer*. 2003;39(14):2021-2025.

50. Almond LM, Tirotta F, Tattersall H, et al. Diagnostic accuracy of percutaneous biopsy in retroperitoneal sarcoma. *British Journal of Surgery*. 2019;106(4):395-403.
51. Schneider N, Strauss DC, Smith MJ, et al. The Adequacy of Core Biopsy in the Assessment of Smooth Muscle Neoplasms of Soft Tissues. *American Journal of Surgical Pathology*. 2017;41(7):923-931.
52. Cates JMM. The AJCC 8th Edition Staging System for Soft Tissue Sarcoma of the Extremities or Trunk: A Cohort Study of the SEER Database. *Journal of the National Comprehensive Cancer Network J Natl Compr Canc Netw*. 2018;16(2):144-152.
53. Cates JMM. Performance Analysis of the American Joint Committee on Cancer 8th Edition Staging System for Retroperitoneal Sarcoma and Development of a New Staging Algorithm for Sarcoma-Specific Survival. *Ann Surg Oncol*. 2017;24(13):3880-3887.
54. Gronchi A, Miceli R, Shurell E, et al. Outcome Prediction in Primary Resected Retroperitoneal Soft Tissue Sarcoma: Histology-Specific Overall Survival and Disease-Free Survival Nomograms Built on Major Sarcoma Center Data Sets. *Journal of Clinical Oncology*. 2013;31(13):1649-1655.
55. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. *CA Cancer J Clin*. 2017;67(2):93-99.
56. HARRELL Jr. FE, LEE KL, MARK DB. MULTIVARIABLE PROGNOSTIC MODELS: ISSUES IN DEVELOPING MODELS, EVALUATING ASSUMPTIONS AND ADEQUACY, AND

MEASURING AND REDUCING ERRORS. *Stat Med.* 1996;15(4):361-387.

57. Danieli M, Barretta F, Fiore M, et al. Refining the Approach to Patients with Primary Soft Tissue Sarcoma of the Extremities and Trunk Wall: Outcome Improvement Over Time at a Single Institution. *Ann Surg Oncol.* 2022;29(5):3274-3286.
58. Gronchi A, Strauss DC, Miceli R, et al. Variability in Patterns of Recurrence After Resection of Primary Retroperitoneal Sarcoma (RPS): A Report on 1007 Patients From the Multi-institutional Collaborative RPS Working Group. *Ann Surg.* 2016;263(5).
59. DIGITAL FOREST. Sarculator.
60. Squires MH, Ethun CG, Donahue EE, et al. Extremity Soft Tissue Sarcoma: A Multi-Institutional Validation of Prognostic Nomograms. *Ann Surg Oncol.* 2022;29(5):3291-3301.
61. Voss RK, Callegaro D, Chiang YJ, et al. Sarculator is a Good Model to Predict Survival in Resected Extremity and Trunk Sarcomas in US Patients. *Ann Surg Oncol.* 2022;29(7):4376-4385.
62. Gladdy RA, Gupta A, Catton CN. Retroperitoneal Sarcoma. *Surg Oncol Clin N Am.* 2016;25(4):697-711.
63. Raut CP, Miceli R, Strauss DC, et al. External validation of a multi-institutional retroperitoneal sarcoma nomogram. *Cancer.* 2016;122(9):1417-1424.
64. Wong RX, Koh YS, Ong F, Farid M, Tay TKY, Teo M. Applicability of the Sarculator and MSKCC nomograms to retroperitoneal sarcoma prognostication in an Asian tertiary center. *Asian J Surg.* 2020;43(11):1078-1085.

65. Callegaro D, Barretta F, Swallow CJ, et al. Longitudinal prognostication in retroperitoneal sarcoma survivors: Development and external validation of two dynamic nomograms. *Eur J Cancer*. 2021;157:291-300.
66. Callegaro D, Miceli R, Bonvalot S, et al. Development and external validation of a dynamic prognostic nomogram for primary extremity soft tissue sarcoma survivors. *EClinicalMedicine*. 2019;17:100215.
67. Rueten-Budde AJ, van Praag VM, Jeys LM, et al. Dynamic prediction of overall survival for patients with high-grade extremity soft tissue sarcoma. *Surg Oncol*. 2018;27(4):695-701.
68. Demicco EG, Boland GM, Brewer Savannah KJ, et al. Progressive loss of myogenic differentiation in leiomyosarcoma has prognostic value. *Histopathology*. 2015;66(5):627-638.
69. Carneiro A, Francis P, Bendahl PO, et al. Indistinguishable genomic profiles and shared prognostic markers in undifferentiated pleomorphic sarcoma and leiomyosarcoma: different sides of a single coin? *Laboratory Investigation*. 2009;89(6):668-675.
70. George S, Serrano C, Hensley ML, Ray-Coquard I. Soft Tissue and Uterine Leiomyosarcoma. *Journal of Clinical Oncology*. 2018;36(2):144-150.
71. Deneve JL, Messina JL, Bui MM, et al. Cutaneous Leiomyosarcoma: Treatment and Outcomes with a Standardized Margin of Resection. *Cancer Control*. 2013;20(4):307-312.
72. Farid M, Ong WS, Tan MH, et al. The Influence of Primary Site on Outcomes in Leiomyosarcoma: A Review of Clinicopathologic

Differences Between Uterine and Extrauterine Disease. *Am J Clin Oncol.* 2013;36(4).

73. Major FJ, Blessing JA, Silverberg SG, et al. Prognostic factors in early-stage uterine sarcoma: A gynecologic oncology group study. *Cancer.* 1993;71(S4):1702-1709.
74. Kasper B, Achee A, Schuster K, et al. Unmet Medical Needs and Future Perspectives for Leiomyosarcoma Patients—A Position Paper from the National Leiomyosarcoma Foundation (NLMSF) and Sarcoma Patients EuroNet (SPAEN). *Cancers (Basel).* 2021;13(4).
75. Mbatani N, Olawaiye AB, Prat J. Uterine sarcomas. *International Journal of Gynecology & Obstetrics.* 2018;143(S2):51-58.
76. Dangoor A, Seddon B, Gerrand C, Grimer R, Whelan J, Judson I. UK guidelines for the management of soft tissue sarcomas. *Clin Sarcoma Res.* 2016;6(1):20.
77. George S, Barysaukas C, Serrano C, et al. Retrospective cohort study evaluating the impact of intraperitoneal morcellation on outcomes of localized uterine leiomyosarcoma. *Cancer.* 2014;120(20):3154-3158.
78. Linch M, Miah AB, Thway K, Judson IR, Benson C. Systemic treatment of soft-tissue sarcoma—gold standard and novel therapies. *Nat Rev Clin Oncol.* 2014;11(4):187-202.
79. Tawbi HA, Burgess M, Bolejack V, et al. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *Lancet Oncol.* 2017;18(11):1493-1501.

80. Ben-Ami E, Barysaukas CM, Solomon S, et al. Immunotherapy with single agent nivolumab for advanced leiomyosarcoma of the uterus: Results of a phase 2 study. *Cancer*. 2017;123(17):3285-3290.
81. D'Angelo SP, Mahoney MR, Van Tine BA, et al. Nivolumab with or without ipilimumab treatment for metastatic sarcoma (Alliance A091401): two open-label, non-comparative, randomised, phase 2 trials. *Lancet Oncol*. 2018;19(3):416-426.
82. Chudasama P, Mughal SS, Sanders MA, et al. Integrative genomic and transcriptomic analysis of leiomyosarcoma. *Nat Commun*. 2018;9(1):144.
83. Nacev BA, Sanchez-Vega F, Smith SA, et al. Clinical sequencing of soft tissue and bone sarcomas delineates diverse genomic landscapes and potential therapeutic targets. *Nat Commun*. 2022;13(1):3405.
84. Schaefer I, Lundberg MZ, Demicco EG, et al. Relationships between highly recurrent tumor suppressor alterations in 489 leiomyosarcomas. *Cancer*. 2021;127(15):2666-2673.
85. Hu J, Rao UNM, Jasani S, Khanna V, Yaw K, Surti U. Loss of DNA copy number of 10q is associated with aggressive behavior of leiomyosarcomas: a comparative genomic hybridization study. *Cancer Genet Cytogenet*. 2005;161(1):20-27.
86. Huang J, Manning BD. A complex interplay between Akt, TSC2 and the two mTOR complexes. *Biochem Soc Trans*. 2009;37(1):217-222.

87. Vanhaesebroeck B, Stephens L, Hawkins P. PI3K signalling: the path to discovery and understanding. *Nat Rev Mol Cell Biol.* 2012;13(3):195-203.
88. Jonsson P, Bandlamudi C, Cheng ML, et al. Tumour lineage shapes BRCA-mediated phenotypes. *Nature.* 2019;571(7766):576-579.
89. Hensley ML, Chavan SS, Solit DB, et al. Genomic Landscape of Uterine Sarcomas Defined Through Prospective Clinical Sequencing. *Clinical Cancer Research.* 2020;26(14):3881-3888.
90. Rosenbaum E, Jonsson P, Seier K, et al. DNA damage response pathway alterations and clinical outcome in leiomyosarcoma. *Journal of Clinical Oncology.* 2019;37(15_suppl):11048.
91. Ford D, Easton DF, Stratton M, et al. Genetic Heterogeneity and Penetrance Analysis of the BRCA1 and BRCA2 Genes in Breast Cancer Families. *The American Journal of Human Genetics.* 1998;62(3):676-689.
92. Xia F, Taghian DG, DeFrank JS, et al. Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. *Proceedings of the National Academy of Sciences.* 2001;98(15):8644-8649.
93. Moynahan ME, Chiu JW, Koller BH, Jasin M. Brca1 Controls Homology-Directed DNA Repair. *Mol Cell.* 1999;4(4):511-518.
94. Rosenbaum E, Jonsson P, Seier K, et al. Clinical Outcome of Leiomyosarcomas With Somatic Alteration in Homologous Recombination Pathway Genes. *JCO Precis Oncol.* 2020;(4):1350-1360.

95. Seligson ND, Kautto EA, Passen EN, et al. BRCA1/2 Functional Loss Defines a Targetable Subset in Leiomyosarcoma. *Oncologist*. 2019;24(7):973-979.
96. Morales J, Li L, Fattah FJ, et al. Review of Poly (ADP-ribose) Polymerase (PARP) Mechanisms of Action and Rationale for Targeting in Cancer and Other Diseases. *Crit Rev Eukaryot Gene Expr*. 2014;24(1):15-28.
97. Ingham M, Allred JB, Gano K, et al. NCI protocol 10250: A phase II study of temozolomide and olaparib for the treatment of advanced uterine leiomyosarcoma. *Journal of Clinical Oncology*. 2021;39(15_suppl):11506.
98. Lee ATJ, Thway K, Huang PH, Jones RL. Clinical and Molecular Spectrum of Liposarcoma. *Journal of Clinical Oncology*. 2018;36(2):151-159.
99. Henricks WH, Chu YC, Goldblum JR, Weiss SW. Dedifferentiated Liposarcoma. *Am J Surg Pathol*. 1997;21(3):271-281.
100. Thway K. Well-differentiated liposarcoma and dedifferentiated liposarcoma: An updated review. *Semin Diagn Pathol*. 2019;36(2):112-121.
101. Dantey K, Schoedel K, Yergiyev O, Bartlett D, Rao UNM. Correlation of histological grade of dedifferentiation with clinical outcome in 55 patients with dedifferentiated liposarcomas. *Hum Pathol*. 2017;66:86-92.
102. ZHANG WD, LIU DR, QUE RS, et al. Management of retroperitoneal liposarcoma: A case report and review of the literature. *Oncol Lett*. 2015;10(1):405-409.

103. Crago AM, Singer S. Clinical and molecular approaches to well differentiated and dedifferentiated liposarcoma. *Curr Opin Oncol.* 2011;23(4):373-378.
104. Fairweather M, Gonzalez RJ, Strauss D, Raut CP. Current principles of surgery for retroperitoneal sarcomas. *J Surg Oncol.* 2018;117(1):33-41.
105. Livingston JA, Bugano D, Barbo A, et al. Role of chemotherapy in dedifferentiated liposarcoma of the retroperitoneum: defining the benefit and challenges of the standard. *Sci Rep.* 2017;7(1):11836.
106. Hull MA, Molina G, Niemierko A, et al. Improved local control with an aggressive strategy of preoperative (with or without intraoperative) radiation therapy combined with radical surgical resection for retroperitoneal sarcoma. *J Surg Oncol.* 2017;115(6):746-751.
107. Chen J, Hang Y, Gao Q, Huang X. Surgical Diagnosis and Treatment of Primary Retroperitoneal Liposarcoma. *Front Surg.* 2021;8.
108. Keung EZ, Hornick JL, Bertagnolli MM, Baldini EH, Raut CP. Predictors of Outcomes in Patients with Primary Retroperitoneal Dedifferentiated Liposarcoma Undergoing Surgery. *J Am Coll Surg.* 2014;218(2):206-217.
109. Gronchi A, Miceli R, Colombo C, et al. Frontline extended surgery is associated with improved survival in retroperitoneal low- to intermediate-grade soft tissue sarcomas. *Annals of Oncology.* 2012;23(4):1067-1073.

110. Lee SY, Goh BKP, Teo MCC, et al. Retroperitoneal liposarcomas: the experience of a tertiary Asian center. *World J Surg Oncol.* 2011;9(1):12.
111. Koczkowska M, Lipska-Ziętkiewicz BS, Iliszko M, et al. Application of high-resolution genomic profiling in the differential diagnosis of liposarcoma. *Mol Cytogenet.* 2017;10(1):7.
112. Somaiah N, Beird HC, Barbo A, et al. Targeted next generation sequencing of well-differentiated/dedifferentiated liposarcoma reveals novel gene amplifications and mutations. *Oncotarget.* 2018;9(28):19891-19899.
113. Tyler R, Wanigasooriya K, Taniere P, et al. A review of retroperitoneal liposarcoma genomics. *Cancer Treat Rev.* 2020;86:102013.
114. Qin JJ, Nag S, Voruganti S, Wang W, Zhang R. Natural Product MDM2 Inhibitors: Anticancer Activity and Mechanisms of Action. *Curr Med Chem.* 2012;19(33):5705-5725.
115. Konopleva M, Martinelli G, Daver N, et al. MDM2 inhibition: an important step forward in cancer therapy. *Leukemia.* 2020;34(11):2858-2874.
116. Traweek RS, Cope BM, Roland CL, Keung EZ, Nassif EF, Erstad DJ. Targeting the MDM2-p53 pathway in dedifferentiated liposarcoma. *Front Oncol.* 2022;12.
117. Gounder MM, Bauer TM, Schwartz GK, et al. A First-in-Human Phase I Study of Milademetan, an MDM2 Inhibitor, in Patients With Advanced Liposarcoma, Solid Tumors, or Lymphomas. *Journal of Clinical Oncology.* 2023;41(9):1714-1724.

118. Gounder M, Patel MR, Yamamoto N, et al. 1548P A phase Ia/Ib, dose-escalation/expansion study of the MDM2–p53 antagonist BI 907828 in patients (pts) with advanced/metastatic sarcoma. *Annals of Oncology*. 2021;32:S1124-S1125.
119. LoRusso P, Yamamoto N, Patel MR, et al. The MDM2–p53 Antagonist Brigimadlin (BI 907828) in Patients with Advanced or Metastatic Solid Tumors: Results of a Phase Ia, First-in-Human, Dose-Escalation Study. *Cancer Discov*. 2023;13(8):1802-1813.
120. Schöffski P, Lahmar M, Lucarelli A, Maki RG. Brightline-1: phase II/III trial of the MDM2–p53 antagonist BI 907828 versus doxorubicin in patients with advanced DDLPS. *Future Oncology*. 2023;19(9):621-629.
121. Adon T, Shanmugarajan D, Kumar HY. CDK4/6 inhibitors: a brief overview and prospective research directions. *RSC Adv*. 2021;11(47):29227-29246.
122. Dickson MA, Koff A, D'Angelo SP, et al. Phase 2 study of the CDK4 inhibitor abemaciclib in dedifferentiated liposarcoma. *Journal of Clinical Oncology*. 2019;37(15_suppl):11004-11004.
123. Chen S, Huang W, Luo P, et al. Undifferentiated Pleomorphic Sarcoma: Long-Term Follow-Up from a Large Institution. *Cancer Manag Res*. 2019;Volume 11:10001-10009.
124. Fletcher CDM. The evolving classification of soft tissue tumours: an update based on the new WHO classification. *Histopathology*. 2006;48(1):3-12.
125. Fletcher CDM, Gustafson P, Rydholm A, Willén H, Åkerman M. Clinicopathologic Re-Evaluation of 100 Malignant Fibrous

- Histiocytomas: Prognostic Relevance of Subclassification. *Journal of Clinical Oncology*. 2001;19(12):3045-3050.
126. Hornick JL. Subclassification of pleomorphic sarcomas: How and why should we care? *Ann Diagn Pathol*. 2018;37:118-124.
127. Vodanovich DA, Spelman T, May D, Slavin J, Choong PFM. Predicting the prognosis of undifferentiated pleomorphic soft tissue sarcoma: a 20-year experience of 266 cases. *ANZ J Surg*. 2019;89(9):1045-1050.
128. Ozcelik M, Seker M, Eraslan E, et al. Evaluation of prognostic factors in localized high-grade undifferentiated pleomorphic sarcoma: report of a multi-institutional experience of Anatolian Society of Medical Oncology. *Tumor Biology*. 2016;37(4):5231-5237.
129. Savina M, Le Cesne A, Blay JY, et al. Patterns of care and outcomes of patients with METAstatic soft tissue SARCOMa in a real-life setting: the METASARC observational study. *BMC Med*. 2017;15(1):78.
130. Hames-Fathi S, Nottley SWG, Pillay N. Unravelling undifferentiated soft tissue sarcomas: insights from genomics. *Histopathology*. 2022;80(1):109-121.
131. Steele CD, Tarabichi M, Oukrif D, et al. Undifferentiated Sarcomas Develop through Distinct Evolutionary Pathways. *Cancer Cell*. 2019;35(3):441-456.e8.
132. Hélias-Rodzewicz Z, Pérot G, Chibon F, et al. YAP1 and VGLL3, encoding two cofactors of TEAD transcription factors, are amplified and overexpressed in a subset of soft tissue sarcomas. *Genes Chromosomes Cancer*. 2010;49(12):1161-1171.

133. Pollack SM, He Q, Yearley JH, et al. T-cell infiltration and clonality correlate with programmed cell death protein 1 and programmed death-ligand 1 expression in patients with soft tissue sarcomas. *Cancer*. 2017;123(17):3291-3304.
134. Dancsok AR, Setsu N, Gao D, et al. Expression of lymphocyte immunoregulatory biomarkers in bone and soft-tissue sarcomas. *Modern Pathology*. 2019;32(12):1772-1785.
135. Petitprez F, de Reyniès A, Keung EZ, et al. B cells are associated with survival and immunotherapy response in sarcoma. *Nature*. 2020;577(7791):556-560.
136. Burns J, Wilding CP, Krasny L, et al. The proteomic landscape of soft tissue sarcomas. *Nat Commun*. 2023;14(1):3834.
137. Lazcano R, Barreto CM, Salazar R, et al. The immune landscape of undifferentiated pleomorphic sarcoma. *Front Oncol*. 2022;12.
138. Toulmonde M, Lucchesi C, Verbeke S, et al. High throughput profiling of undifferentiated pleomorphic sarcomas identifies two main subgroups with distinct immune profile, clinical outcome and sensitivity to targeted therapies. *EBioMedicine*. 2020;62.
139. Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. *J Cell Sci*. 2012;125(23):5591-5596.
140. He X, Yang Y, Han Y, et al. Extracellular matrix physical properties govern the diffusion of nanoparticles in tumor microenvironment. *Proceedings of the National Academy of Sciences*. 2023;120(1).
141. Jin MZ, Jin WL. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct Target Ther*. 2020;5(1):166.

142. Bergers G, Benjamin LE. Tumorigenesis and the angiogenic switch. *Nat Rev Cancer*. 2003;3(6):401-410.
143. Anderson NM, Simon MC. The tumor microenvironment. *Current Biology*. 2020;30(16):R921-R925.
144. Labani-Motlagh A, Ashja-Mahdavi M, Loskog A. The Tumor Microenvironment: A Milieu Hindering and Obstructing Antitumor Immune Responses. *Front Immunol*. 2020;11.
145. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun*. 2020;11(1):5120.
146. Kumar B V., Connors TJ, Farber DL. Human T Cell Development, Localization, and Function throughout Life. *Immunity*. 2018;48(2):202-213.
147. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat Rev Immunol*. 2020;20(11):651-668.
148. Vignali DAA, Collison LW, Workman CJ. How regulatory T cells work. *Nat Rev Immunol*. 2008;8(7):523-532.
149. Sousa LM, Almeida JS, Fortes-Andrade T, et al. Tumor and Peripheral Immune Status in Soft Tissue Sarcoma: Implications for Immunotherapy. *Cancers (Basel)*. 2021;13(15):3885.
150. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med*. 2017;9(1):34.
151. Burgess MA, Bolejack V, Schuetze S, et al. Clinical activity of pembrolizumab (P) in undifferentiated pleomorphic sarcoma

(UPS) and dedifferentiated/pleomorphic liposarcoma (LPS): Final results of SARC028 expansion cohorts. *Journal of Clinical Oncology*. 2019;37(15_suppl):11015.

152. Dancsok AR, Setsu N, Gao D, et al. Expression of lymphocyte immunoregulatory biomarkers in bone and soft-tissue sarcomas. *Modern Pathology*. 2019;32(12):1772-1785.
153. Smolle MA, Herbsthofer L, Granegger B, et al. T-regulatory cells predict clinical outcome in soft tissue sarcoma patients: a clinico-pathological study. *Br J Cancer*. 2021;125(5):717-724.
154. Boxberg M, Steiger K, Lenze U, et al. PD-L1 and PD-1 and characterization of tumor-infiltrating lymphocytes in high grade sarcomas of soft tissue – prognostic implications and rationale for immunotherapy. *Oncoimmunology*. 2018;7(3):e1389366.
155. D'Angelo SP, Shoushtari AN, Agaram NP, et al. Prevalence of tumor-infiltrating lymphocytes and PD-L1 expression in the soft tissue sarcoma microenvironment. *Hum Pathol*. 2015;46(3):357-365.
156. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity*. 2013;39(4):782-795.
157. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat Rev Cancer*. 2019;19(6):307-325.
158. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. *J Cell Sci*. 2010;123(24):4195-4200.

159. Lu P, Weaver VM, Werb Z. The extracellular matrix: A dynamic niche in cancer progression. *Journal of Cell Biology*. 2012;196(4):395-406.
160. Hastings JF, Skhinas JN, Fey D, Croucher DR, Cox TR. The extracellular matrix as a key regulator of intracellular signalling networks. *Br J Pharmacol*. 2019;176(1):82-92.
161. Schultz GS, Wysocki A. Interactions between extracellular matrix and growth factors in wound healing. *Wound Repair and Regeneration*. 2009;17(2):153-162.
162. Tracy LE, Minasian RA, Caterson EJ. Extracellular Matrix and Dermal Fibroblast Function in the Healing Wound. *Adv Wound Care (New Rochelle)*. 2016;5(3):119-136.
163. Watt FM, Huck WTS. Role of the extracellular matrix in regulating stem cell fate. *Nat Rev Mol Cell Biol*. 2013;14(8):467-473.
164. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol*. 2014;15(12):786-801.
165. Provenzano PP, Inman DR, Eliceiri KW, et al. Collagen density promotes mammary tumor initiation and progression. *BMC Med*. 2008;6(1):11.
166. Levental KR, Yu H, Kass L, et al. Matrix Crosslinking Forces Tumor Progression by Enhancing Integrin Signaling. *Cell*. 2009;139(5):891-906.
167. Tian C, Öhlund D, Rickelt S, et al. Cancer Cell–Derived Matrisome Proteins Promote Metastasis in Pancreatic Ductal Adenocarcinoma. *Cancer Res*. 2020;80(7):1461-1474.

168. Izzi V, Lakkala J, Devarajan R, et al. Pan-Cancer analysis of the expression and regulation of matrisome genes across 32 tumor types. *Matrix Biol Plus*. 2019;1:100004.
169. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics*. 2012;11(4):M111.014647-M111.014647.
170. Krasny L, Bland P, Burns J, et al. A mouse SWATH-MS reference spectral library enables deconvolution of species-specific proteomic alterations in human tumour xenografts. *Dis Model Mech*. Published online July 5, 2020:dmm.044586.
171. Naba A, Clauser KR, Whittaker CA, Carr SA, Tanabe KK, Hynes RO. Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver. *BMC Cancer*. 2014;14(1):518.
172. Naba A, Clauser KR, Lamar JM, Carr SA, Hynes RO. Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *Elife*. 2014;3.
173. Papanicolaou M, Parker AL, Yam M, et al. Temporal profiling of the breast tumour microenvironment reveals collagen XII as a driver of metastasis. *Nat Commun*. 2022;13(1):4587.
174. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The Matrisome: In Silico Definition and In Vivo Characterization by Proteomics of Normal and Tumor Extracellular Matrices. *Molecular & Cellular Proteomics*. 2012;11(4):M111.014647.
175. Hohenester E, Engel J. Domain structure and organisation in extracellular matrix proteins. *Matrix Biology*. 2002;21(2):115-128.

176. Hynes RO, Naba A. Overview of the Matrisome--An Inventory of Extracellular Matrix Constituents and Functions. *Cold Spring Harb Perspect Biol.* 2012;4(1):a004903-a004903.
177. Krasny L, Huang PH. Advances in the proteomic profiling of the matrisome and adhesome. *Expert Rev Proteomics.* 2021;18(9):781-794.
178. Shao X, Taha IN, Clauser KR, Gao Y (Tom), Naba A. MatrisomeDB: the ECM-protein knowledge database. *Nucleic Acids Res.* 2020;48(D1):D1136-D1144.
179. Shao X, Taha IN, Clauser KR, Gao Y (Tom), Naba A. MatrisomeDB: the ECM-protein knowledge database. *Nucleic Acids Res.* 2020;48(D1):D1136-D1144.
180. Shao X, Gomez CD, Kapoor N, et al. MatrisomeDB 2.0: 2023 updates to the ECM-protein knowledge database. *Nucleic Acids Res.* 2023;51(D1):D1519-D1530.
181. Gordon MK, Hahn RA. Collagens. *Cell Tissue Res.* 2010;339(1):247-257.
182. Brodsky B, Persikov A V. Molecular Structure of the Collagen Triple Helix. In: ; 2005:301-339.
183. Ricard-Blum S. The Collagen Family. *Cold Spring Harb Perspect Biol.* 2011;3(1):a004978-a004978.
184. Ortega N, Werb Z. New functional roles for non-collagenous domains of basement membrane collagens. *J Cell Sci.* 2002;115(22):4201-4214.
185. Xu ER, Blythe EE, Fischer G, Hyvönen M. Structural analyses of von Willebrand factor C domains of collagen 2A and CCN3 reveal

an alternative mode of binding to bone morphogenetic protein-2. *Journal of Biological Chemistry*. 2017;292(30):12516-12527.

186. Aigner T, Hambach L, Söder S, Schlötzer-Schrehardt U, Pöschl E. The C5 Domain of Col6A3 Is Cleaved Off from the Col6 Fibrils Immediately after Secretion. *Biochem Biophys Res Commun*. 2002;290(2):743-748.
187. Colombo M, Brittingham RJ, Klement JF, et al. Procollagen VII Self-Assembly Depends on Site-Specific Interactions and Is Promoted by Cleavage of the NC2 Domain with Procollagen C-Proteinase. *Biochemistry*. 2003;42(39):11434-11442.
188. Bella J, Hulmes DJS. Fibrillar Collagens. In: ; 2017:457-490.
189. KADLER KE, HOLMES DF, TROTTER JA, CHAPMAN JA. Collagen fibril formation. *Biochemical Journal*. 1996;316(1):1-11.
190. Shaw LM, Olsen BR. FACIT collagens: diverse molecular bridges in extracellular matrices. *Trends Biochem Sci*. 1991;16:191-194.
191. Olsen BR. Collagen IX. *Int J Biochem Cell Biol*. 1997;29(4):555-558.
192. Knupp C, Squire JM. Molecular Packing in Network-Forming Collagens. *The Scientific World JOURNAL*. 2003;3:558-577.
193. Jayadev R, Sherwood DR. Basement membranes. *Current Biology*. 2017;27(6):R207-R211.
194. Franzke CW, Bruckner P, Bruckner-Tuderman L. Collagenous Transmembrane Proteins: Recent Insights into Biology and Pathology*. *Journal of Biological Chemistry*. 2005;280(6):4005-4008.

195. Franzke CW, Tasanen K, Schäcke H, et al. Transmembrane collagen XVII, an epithelial adhesion protein, is shed from the cell surface by ADAMs. *EMBO J.* 2002;21(19):5026-5035.
196. Walia A, Yang JF, Huang Y hui, Rosenblatt MI, Chang JH, Azar DT. Endostatin's emerging roles in angiogenesis, lymphangiogenesis, disease, and clinical applications. *Biochimica et Biophysica Acta (BBA) - General Subjects.* 2015;1850(12):2422-2438.
197. Ramchandran R, Dhanabal M, Volk R, et al. Antiangiogenic Activity of Restin, NC10 Domain of Human Collagen XV: Comparison to Endostatin. *Biochem Biophys Res Commun.* 1999;255(3):735-739.
198. O'Reilly MS, Boehm T, Shing Y, et al. Endostatin: An Endogenous Inhibitor of Angiogenesis and Tumor Growth. *Cell.* 1997;88(2):277-285.
199. Myllyharju J. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends in Genetics.* 2004;20(1):33-43.
200. Bella J, Humphries MJ. Cα-H...O = C hydrogen bonds contribute to the specificity of RGD cell-adhesion interactions. *BMC Struct Biol.* 2005;5(1):4.
201. Leitinger B, Hohenester E. Mammalian collagen receptors. *Matrix Biology.* 2007;26(3):146-155.
202. Shrivastava A, Radziejewski C, Campbell E, et al. An Orphan Receptor Tyrosine Kinase Family Whose Members Serve as Nonintegrin Collagen Receptors. *Mol Cell.* 1997;1(1):25-34.

203. Vogel W, Gish GD, Alves F, Pawson T. The Discoidin Domain Receptor Tyrosine Kinases Are Activated by Collagen. *Mol Cell*. 1997;1(1):13-23.
204. Leitinger B, Kwan APL. The discoidin domain receptor DDR2 is a receptor for type X collagen. *Matrix Biology*. 2006;25(6):355-364.
205. Leitinger B. Molecular Analysis of Collagen Binding by the Human Discoidin Domain Receptors, DDR1 and DDR2. *Journal of Biological Chemistry*. 2003;278(19):16761-16769.
206. Brockhausen I, Schutzbach J, Kuhns W. Glycoproteins and Their Relationship to Human Disease. *Cells Tissues Organs*. 1998;161(1-4):36-78.
207. Hynes RO. The Extracellular Matrix: Not Just Pretty Fibrils. *Science (1979)*. 2009;326(5957):1216-1219.
208. M MP. Basement Membrane Proteins: Structure, Assembly, and Cellular Interactions. *Crit Rev Biochem Mol Biol*. 1992;27(1-2):93-127.
209. Chi-Rosso G, Gotwals PJ, Yang J, et al. Fibronectin Type III Repeats Mediate RGD-independent Adhesion and Signaling through Activated β 1 Integrins. *Journal of Biological Chemistry*. 1997;272(50):31447-31452.
210. Rahman S, Patel Y, Murray J, et al. Novel hepatocyte growth factor (HGF) binding domains on fibronectin and vitronectin coordinate a distinct and amplified Met-integrin induced signalling pathway in endothelial cells. *BMC Cell Biol*. 2005;6(1):8.
211. Couchman JR, Pataki CA. An Introduction to Proteoglycans and Their Localization. *Journal of Histochemistry & Cytochemistry*. 2012;60(12):885-897.

212. Iozzo R V, Murdoch AD. Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J.* 1996;10(5):598-614.
213. Prydz K. Determinants of Glycosaminoglycan (GAG) Structure. *Biomolecules.* 2015;5(3):2003-2022.
214. Soares da Costa D, Reis RL, Pashkuleva I. Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. *Annu Rev Biomed Eng.* 2017;19(1):1-26.
215. Iozzo R V., Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biology.* 2015;42:11-55.
216. Freemont AJ, Denton J. Synovial fluid. In: *Diagnostic Cytopathology.* Elsevier; 2010:809-817.
217. Kirkbride KC, Ray BN, Blobel GC. Cell-surface co-receptors: emerging roles in signaling and human disease. *Trends Biochem Sci.* 2005;30(11):611-621.
218. Myhre K, Blobel GC. Proteoglycan signaling co-receptors: Roles in cell adhesion, migration and invasion. *Cell Signal.* 2009;21(11):1548-1558.
219. Whitelock JM, Melrose J, Iozzo R V. Diverse Cell Signaling Events Modulated by Perlecan. *Biochemistry.* 2008;47(43):11174-11183.
220. Aviezer D, Hecht D, Safran M, Eisinger M, David G, Yayon A. Perlecan, basal lamina proteoglycan, promotes basic fibroblast growth factor-receptor binding, mitogenesis, and angiogenesis. *Cell.* 1994;79(6):1005-1013.

221. Iozzo R V., Sanderson RD. Proteoglycans in cancer biology, tumour microenvironment and angiogenesis. *J Cell Mol Med.* 2011;15(5):1013-1031.
222. Neill T, Schaefer L, Iozzo R V. Oncosuppressive functions of decorin. *Mol Cell Oncol.* 2015;2(3):e975645.
223. Santra M, Reed CC, Iozzo R V. Decorin Binds to a Narrow Region of the Epidermal Growth Factor (EGF) Receptor, Partially Overlapping but Distinct from the EGF-binding Epitope. *Journal of Biological Chemistry.* 2002;277(38):35671-35681.
224. Csordás G, Santra M, Reed CC, et al. Sustained Down-regulation of the Epidermal Growth Factor Receptor by Decorin. *Journal of Biological Chemistry.* 2000;275(42):32879-32887.
225. Moscatello DK, Santra M, Mann DM, McQuillan DJ, Wong AJ, Iozzo R V. Decorin suppresses tumor cell growth by activating the epidermal growth factor receptor. *Journal of Clinical Investigation.* 1998;101(2):406-412.
226. Zhu JX, Goldoni S, Bix G, et al. Decorin Evokes Protracted Internalization and Degradation of the Epidermal Growth Factor Receptor via Caveolar Endocytosis. *Journal of Biological Chemistry.* 2005;280(37):32468-32479.
227. Colin Hughes R. Galectins as modulators of cell adhesion. *Biochimie.* 2001;83(7):667-676.
228. Aspberg A, Miura R, Bourdoulous S, et al. The C-type lectin domains of lecticans, a family of aggregating chondroitin sulfate proteoglycans, bind tenascin-R by protein–protein interactions independent of carbohydrate moiety. *Proceedings of the National Academy of Sciences.* 1997;94(19):10116-10121.

229. Strongin AY. Mislocalization and unconventional functions of cellular MMPs in cancer. *Cancer and Metastasis Reviews*. 2006;25(1):87-98.
230. Jodele S, Blavier L, Yoon JM, DeClerck YA. Modifying the soil to affect the seed: role of stromal-derived matrix metalloproteinases in cancer progression. *Cancer and Metastasis Reviews*. 2006;25(1):35-43.
231. Payne SL, Hendrix MJC, Kirschmann DA. Paradoxical roles for lysyl oxidases in cancer—A prospect. *J Cell Biochem*. 2007;101(6):1338-1354.
232. Olson OC, Joyce JA. Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer*. 2015;15(12):712-729.
233. Cox TR, Bird D, Baker AM, et al. LOX-Mediated Collagen Crosslinking Is Responsible for Fibrosis-Enhanced Metastasis. *Cancer Res*. 2013;73(6):1721-1732.
234. Kirschmann DA, Seftor EA, Fong SFT, et al. A molecular role for lysyl oxidase in breast cancer invasion. *Cancer Res*. 2002;62(15):4478-4483.
235. Payne SL, Fogelgren B, Hess AR, et al. Lysyl Oxidase Regulates Breast Cancer Cell Migration and Adhesion through a Hydrogen Peroxide-Mediated Mechanism. *Cancer Res*. 2005;65(24):11429-11436.
236. Werb Z. ECM and Cell Surface Proteolysis: Regulating Cellular Ecology. *Cell*. 1997;91(4):439-442.

237. Cabral-Pacheco GA, Garza-Veloz I, Castruita-De la Rosa C, et al. The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. *Int J Mol Sci.* 2020;21(24):9739.
238. Loffek S, Schilling O, Franzke CW. Biological role of matrix metalloproteinases: a critical balance. *European Respiratory Journal.* 2011;38(1):191-208.
239. Mott JD, Werb Z. Regulation of matrix biology by matrix metalloproteinases. *Curr Opin Cell Biol.* 2004;16(5):558-564.
240. Xu J, Rodriguez D, Petitclerc E, et al. Proteolytic exposure of a cryptic site within collagen type IV is required for angiogenesis and tumor growth in vivo. *J Cell Biol.* 2001;154(5):1069-1080.
241. Hangai M, Kitaya N, Xu J, et al. Matrix Metalloproteinase-9-Dependent Exposure of a Cryptic Migratory Control Site in Collagen is Required before Retinal Angiogenesis. *Am J Pathol.* 2002;161(4):1429-1437.
242. Ferreras M, Felbor U, Lenhard T, Olsen BR, Delaissé JM. Generation and degradation of human endostatin proteins by various proteinases. *FEBS Lett.* 2000;486(3):247-251.
243. Hinz B. The extracellular matrix and transforming growth factor- β 1: Tale of a strained relationship. *Matrix Biology.* 2015;47:54-65.
244. Crawford SE, Stellmach V, Murphy-Ullrich JE, et al. Thrombospondin-1 Is a Major Activator of TGF- β 1 In Vivo. *Cell.* 1998;93(7):1159-1170.
245. Munger JS, Harpel JG, Gleizes PE, Mazzieri R, Nunes I, Rifkin DB. Latent transforming growth factor- β : Structural features and mechanisms of activation. *Kidney Int.* 1997;51(5):1376-1382.

246. Vempati P, Popel AS, Mac Gabhann F. Extracellular regulation of VEGF: Isoforms, proteolysis, and vascular patterning. *Cytokine Growth Factor Rev.* 2014;25(1):1-19.
247. Bergers G, Brekken R, McMahon G, et al. Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nat Cell Biol.* 2000;2(10):737-744.
248. Pozzi A, Yurchenco PD, Iozzo R V. The nature and biology of basement membranes. *Matrix Biology.* 2017;57-58:1-11.
249. Jayadev R, Sherwood DR. Basement membranes. *Current Biology.* 2017;27(6):R207-R211.
250. Xu L, Nirwane A, Yao Y. Basement membrane and blood–brain barrier. *Stroke Vasc Neurol.* 2019;4(2):78-82.
251. McCarthy KJ, Wassenhove-McCarthy DJ. The Glomerular Basement Membrane as a Model System to Study the Bioactivity of Heparan Sulfate Glycosaminoglycans. *Microscopy and Microanalysis.* 2012;18(1):3-21.
252. Behrens DT, Villone D, Koch M, et al. The Epidermal Basement Membrane Is a Composite of Separate Laminin- or Collagen IV-containing Networks Connected by Aggregated Perlecan, but Not by Nidogens. *Journal of Biological Chemistry.* 2012;287(22):18700-18709.
253. Yurchenco PD, O'Rear JJ. Basal lamina assembly. *Curr Opin Cell Biol.* 1994;6(5):674-681.
254. Jayadev R, Morais MRPT, Ellingford JM, et al. A basement membrane discovery pipeline uncovers network complexity, regulators, and human disease associations. *Sci Adv.* 2023;8(20):eabn2265.

255. Jayadev R, Sherwood DR. Basement membranes. *Current Biology*. 2017;27(6):R207-R211.
256. Hotary K, Li XY, Allen E, Stevens SL, Weiss SJ. A cancer cell metalloprotease triad regulates the basement membrane transmigration program. *Genes Dev*. 2006;20(19):2673-2686.
257. Chang J, Chaudhuri O. Beyond proteases: Basement membrane mechanics and cancer invasion. *Journal of Cell Biology*. 2019;218(8):2456-2469.
258. Nam S, Chaudhuri O. Mitotic cells generate protrusive extracellular forces to divide in three-dimensional microenvironments. *Nat Phys*. 2018;14(6):621-628.
259. Naba A, Clauser KR, Ding H, Whittaker CA, Carr SA, Hynes RO. The extracellular matrix: Tools and insights for the “omics” era. *Matrix Biology*. 2016;49:10-24.
260. Yuzhalin AE, Urbonas T, Silva MA, Muschel RJ, Gordon-Weeks AN. A core matrisome gene signature predicts cancer outcome. *Br J Cancer*. 2018;118(3):435-440.
261. Lim S Bin, Tan SJ, Lim WT, Lim CT. An extracellular matrix-related prognostic and predictive indicator for early-stage non-small cell lung cancer. *Nat Commun*. 2017;8(1):1734.
262. Wu G, Yang Y, Ye R, et al. Development and validation of an ECM-related prognostic signature to predict the immune landscape of human hepatocellular carcinoma. *BMC Cancer*. 2022;22(1):1036.
263. Yang Z, Xue F, Li M, et al. Extracellular Matrix Characterization in Gastric Cancer Helps to Predict Prognosis and Chemotherapy Response. *Front Oncol*. 2021;11.

264. Parker AL, Bowman E, Zingone A, et al. Extracellular matrix profiles determine risk and prognosis of the squamous cell carcinoma subtype of non-small cell lung carcinoma. *Genome Med.* 2022;14(1):126.
265. Pearce OMT, Delaine-Smith RM, Maniati E, et al. Deconstruction of a Metastatic Tumor Microenvironment Reveals a Common Matrix Response in Human Cancers. *Cancer Discov.* 2018;8(3):304-319.
266. Langlois B, Saupe F, Rupp T, et al. AngioMatrix, a signature of the tumor angiogenic switch-specific matrisome, correlates with poor prognosis for glioma and colorectal cancer patients. *Oncotarget.* 2014;5(21):10529-10545.
267. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* 2009;583(24):3966-3973.
268. de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. *Mol Biosyst.* Published online 2009.
269. Kosti I, Jain N, Aran D, Butte AJ, Sirota M. Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. *Sci Rep.* 2016;6(1):24799.
270. Schiller HB, Friedel CC, Boulegue C, Fässler R. Quantitative proteomics of the integrin adhesome show a myosin II-dependent recruitment of LIM domain proteins. *EMBO Rep.* 2011;12(3):259-266.
271. Küttner V, Mack C, Rigbolt KT, et al. Global remodelling of cellular microenvironment due to loss of collagen VII. *Mol Syst Biol.* 2013;9(1).

272. Randles MJ, Woolf AS, Huang JL, et al. Genetic Background is a Key Determinant of Glomerular Extracellular Matrix Composition and Organization. *Journal of the American Society of Nephrology*. 2015;26(12):3021-3034.
273. Leeming DJ, Bay-Jensen AC, Vassiliadis E, Larsen MR, Henriksen K, Karsdal MA. Post-translational modifications of the extracellular matrix are key events in cancer progression: Opportunities for biochemical marker development. *Biomarkers*. 2011;16(3):193-205.
274. Choi H, Simpson D, Wang D, et al. Heterogeneity of proteome dynamics between connective tissue phases of adult tendon. *Elife*. 2020;9.
275. Watt FM. NEW EMBO MEMBER'S REVIEW: Role of integrins in regulating epidermal adhesion, growth and differentiation. *EMBO J*. 2002;21(15):3919-3926.
276. Guo W, Giancotti FG. Integrin signalling during tumour progression. *Nat Rev Mol Cell Biol*. 2004;5(10):816-826.
277. Brown, Marshall. Integrin-Mediated TGF β Activation Modulates the Tumour Microenvironment. *Cancers (Basel)*. 2019;11(9):1221.
278. Mould AP. Getting integrins into shape: recent insights into how integrin activity is regulated by conformational changes. *J Cell Sci*. 1996;109(11):2613-2618.
279. Askari JA, Buckley PA, Mould AP, Humphries MJ. Linking integrin conformation to function. *J Cell Sci*. 2009;122(2):165-170.

280. Qiong G, Zhaolin S, Deyu F, Peifang W. Integrins in human hepatocellular carcinoma tumorigenesis and therapy. *Chin Med J (Engl)*. 2023;136(03):253-268.
281. Liu S, Calderwood DA, Ginsberg MH. Integrin cytoplasmic domain-binding proteins. *J Cell Sci*. 2000;113(20):3563-3571.
282. Vinogradova O, Velyvis A, Velyviene A, et al. A Structural Mechanism of Integrin α IIb β 3 "Inside-Out" Activation as Regulated by Its Cytoplasmic Face. *Cell*. 2002;110(5):587-597.
283. Takagi J, Springer TA. Integrin activation and structural rearrangement. *Immunol Rev*. 2002;186(1):141-163.
284. Luo BH, Springer TA. Integrin structures and conformational signaling. *Curr Opin Cell Biol*. 2006;18(5):579-586.
285. Galbraith CG, Yamada KM, Sheetz MP. The relationship between force and focal complex development. *J Cell Biol*. 2002;159(4):695-705.
286. Ginsberg MH, Du X, Plow EF. Inside-out integrin signalling. *Curr Opin Cell Biol*. 1992;4(5):766-771.
287. Leiss M, Beckmann K, Girós A, Costell M, Fässler R. The role of integrin binding sites in fibronectin matrix assembly in vivo. *Curr Opin Cell Biol*. 2008;20(5):502-507.
288. Puklin-Faucher E, Sheetz MP. The mechanical integrin cycle. *J Cell Sci*. 2009;122(2):179-186.
289. Zhang Z, Vuori K, Wang HG, Reed JC, Ruoslahti E. Integrin Activation by R-ras. *Cell*. 1996;85(1):61-69.

290. Henning Stumpf B, Ambriović-Ristov A, Radenovic A, Smith AS. Recent Advances and Prospects in the Research of Nascent Adhesions. *Front Physiol.* 2020;11.
291. Hynes RO. Integrins: Bidirectional, Allosteric Signaling Machines. *Cell.* 2002;110(6):673-687.
292. Klapholz B, Brown NH. Talin – the master of integrin adhesions. *J Cell Sci.* Published online January 1, 2017.
293. Geiger B, Bershadsky A, Pankov R, Yamada KM. Transmembrane crosstalk between the extracellular matrix and the cytoskeleton. *Nat Rev Mol Cell Biol.* 2001;2(11):793-805.
294. Zamir E, Geiger B. Molecular complexity and dynamics of cell-matrix adhesions. *J Cell Sci.* 2001;114(20):3583-3590.
295. Zaidel-Bar R, Itzkovitz S, Ma'ayan A, Iyengar R, Geiger B. Functional atlas of the integrin adhesome. *Nat Cell Biol.* 2007;9(8):858-867.
296. Kuo JC, Han X, Hsiao CT, Yates III JR, Waterman CM. Analysis of the myosin-II-responsive focal adhesion proteome reveals a role for β -Pix in negative regulation of focal adhesion maturation. *Nat Cell Biol.* 2011;13(4):383-393.
297. Byron A, Humphries JD, Bass MD, Knight D, Humphries MJ. Proteomic Analysis of Integrin Adhesion ComplexesA presentation from the 6th British Society for Proteome Research (BSPR)–European Bioinformatics Institute (EBI) Meeting “Multiscale Proteomics: From Cells to Organisms” at the Wellcome Trust Conference Centre, Cambridge, UK, 14 to 16 July 2009. The Presentation also complements the *Science Signaling*

Research Article by Humphries *et al.* published 8 September 2009. *Sci Signal.* 2011;4(167).

298. Winograd-Katz SE, Fässler R, Geiger B, Legate KR. The integrin adhesome: from genes and proteins to human disease. *Nat Rev Mol Cell Biol.* 2014;15(4):273-288.
299. Hemler ME, Lobb RR. The leukocyte β 1 integrins. *Curr Opin Hematol.* 1995;2(1):61-67.
300. Pankova V, Thway K, Jones RL, Huang PH. The Extracellular Matrix in Soft Tissue Sarcomas: Pathobiology and Cellular Signalling. *Front Cell Dev Biol.* 2021;9.
301. Willems SM, Mohseny AB, Balog C, et al. Cellular/intramuscular myxoma and grade I myxofibrosarcoma are characterized by distinct genetic alterations and specific composition of their extracellular matrix. *J Cell Mol Med.* 2009;13(7):1291-1301.
302. West RB, Nuyten DSA, Subramanian S, et al. Determination of Stromal Signatures in Breast Carcinoma. Staudt LM, ed. *PLoS Biol.* 2005;3(6):e187.
303. Lian X, Bond JS, Bharathy N, et al. Defining the Extracellular Matrix of Rhabdomyosarcoma. *Front Oncol.* 2021;11.
304. Persson S, Willems JS, Kindblom LG, Angervall L. Alveolar soft part sarcoma: An immunohistochemical, cytologic and electron-microscopic study and a quantitative DNA analysis. *Virchows Arch A Pathol Anat Histopathol.* 1988;412(6):499-513.
305. Guarino M, Christensen L. Immunohistochemical analysis of extracellular matrix components in synovial sarcoma. *J Pathol.* 1994;172(3):279-286.

306. Guarino M. Immunohistochemical distribution of basement membrane type IV collagen and laminin in synovial sarcoma. *Tumori*. 1993;79(6):427-432.
307. Fukuda T, Tsuneyoshi M. Adhesion proteins, cellular morphology and fibrous components around the cell/extracellular-matrix interface in myxoid liposarcomas. *J Cancer Res Clin Oncol*. 2000;126(6):320-324.
308. Stenman S, Vaheri A. Fibronectin in human solid tumors. *Int J Cancer*. 1981;27(4):427-435.
309. Haraida S, Nerlich AG, Wiest I, Schleicher E, Löhrs U. Distribution of basement membrane components in normal adipose tissue and in benign and malignant tumors of lipomatous origin. *Mod Pathol*. 1996;9(2):137-144.
310. Ordóñez NG, Mahfouz SM, Mackay B. Synovial sarcoma: An immunohistochemical and ultrastructural study. *Hum Pathol*. 1990;21(7):733-749.
311. Ogawa K, Oguchi M, Yamabe H, Nakashima Y, Hamashima Y. Distribution of collagen type IV in soft tissue tumors. An immunohistochemical study. *Cancer*. 1986;58(2):269-277.
312. d'Ardenne AJ, Kirkpatrick P, Sykes BC. Distribution of laminin, fibronectin, and interstitial collagen type III in soft tissue tumours. *J Clin Pathol*. 1984;37(8):895-904.
313. Boudin L, de Nonneville A, Finetti P, et al. CSPG4 expression in soft tissue sarcomas is associated with poor prognosis and low cytotoxic immune response. *J Transl Med*. 2022;20(1):464.
314. Singh P, Carraher C, Schwarzbauer JE. Assembly of Fibronectin Extracellular Matrix. *Annu Rev Cell Dev Biol*. 2010;26(1):397-419.

315. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol.* 2008;20(5):495-501.
316. Sabatier L, Chen D, Fagotto-Kaufmann C, et al. Fibrillin Assembly Requires Fibronectin. *Mol Biol Cell.* 2009;20(3):846-858.
317. Benassi MS, Ragazzini P, Gamberi G, et al. Adhesion molecules in high-grade soft tissue sarcomas: correlation to clinical outcome. *Eur J Cancer.* 1998;34(4):496-502.
318. Miner JH, Yurchenco PD. LAMININ FUNCTIONS IN TISSUE MORPHOGENESIS. *Annu Rev Cell Dev Biol.* 2004;20(1):255-284.
319. Kleinman HK. Preparation of Basement Membrane Components from <scp>EHS</scp> Tumors. *Curr Protoc Cell Biol.* 1998;00(1).
320. Kim S, Min S, Choi YS, et al. Tissue extracellular matrix hydrogels as alternatives to Matrigel for culturing gastrointestinal organoids. *Nat Commun.* 2022;13(1):1692.
321. Barth T, Möller P, Mechttersheimer G. Differential expression of beta 1, beta 3 and beta 4 integrins in sarcomas of the small, round, blue cell category. *Virchows Arch.* 1995;426(1):19-25.
322. Okada T, Lee AY, Qin LX, et al. Integrin- α 10 Dependency Identifies RAC and RICTOR as Therapeutic Targets in High-Grade Myxofibrosarcoma. *Cancer Discov.* 2016;6(10):1148-1165.
323. Roma J, Masià A, Reventós J, Sánchez de Toledo J, Gallego S. Notch Pathway Inhibition Significantly Reduces Rhabdomyosarcoma Invasiveness and Mobility *In Vitro.* *Clinical Cancer Research.* 2011;17(3):505-513.

324. Masià A, Almazán-Moga A, Velasco P, et al. Notch-mediated induction of N-cadherin and $\alpha 9$ -integrin confers higher invasive phenotype on rhabdomyosarcoma cells. *Br J Cancer*. 2012;107(8):1374-1383.
325. Desai SS, Jambhekar NA. Pathology of Ewing's sarcoma/PNET: Current opinion and emerging concepts. *Indian J Orthop*. 2010;44(4):363-368.
326. Witherel CE, Sao K, Brisson BK, et al. Regulation of extracellular matrix assembly and structure by hybrid M1/M2 macrophages. *Biomaterials*. 2021;269:120667.
327. Liu T, Zhou L, Li D, Andl T, Zhang Y. Cancer-Associated Fibroblasts Build and Secure the Tumor Microenvironment. *Front Cell Dev Biol*. 2019;7.
328. Drury JL, Mooney DJ. Hydrogels for tissue engineering: scaffold design variables and applications. *Biomaterials*. 2003;24(24):4337-4351.
329. Benton G, Arnaoutova I, George J, Kleinman HK, Koblinski J. Matrigel: From discovery and ECM mimicry to assays and models for cancer research. *Adv Drug Deliv Rev*. 2014;79-80:3-18.
330. Passaniti A, Taylor RM, Pili R, et al. A simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin, and fibroblast growth factor. *Lab Invest*. 1992;67(4):519-528.
331. Lee SW, Lee HJ, Hwang HS, Ko K, Han DW, Ko K. Optimization of Matrigel-based culture for expansion of neural stem cells. *Anim Cells Syst (Seoul)*. 2015;19(3):175-180.

332. Ivascu A, Kubbies M. Rapid Generation of Single-Tumor Spheroids for High-Throughput Cell Function and Toxicity Analysis. *SLAS Discovery*. 2006;11(8):922-932.
333. Albini A, Noonan DM. The 'chemoinvasion' assay, 25 years and still going strong: the use of reconstituted basement membranes to study cell invasion and angiogenesis. *Curr Opin Cell Biol*. 2010;22(5):677-689.
334. Fridman R, Benton G, Aranoutova I, Kleinman HK, Bonfil RD. Increased initiation and growth of tumor cell lines, cancer stem cells and biopsy material in mice using basement membrane matrix protein (Cultrex or Matrigel) co-injection. *Nat Protoc*. 2012;7(6):1138-1144.
335. Rajan N, Habermehl J, Coté MF, Doillon CJ, Mantovani D. Preparation of ready-to-use, storable and reconstituted type I collagen from rat tail tendon for tissue engineering applications. *Nat Protoc*. 2006;1(6):2753-2758.
336. Bracke ME, Boterberg T, Bruyneel EA, Mareel MM. Collagen Invasion Assay. In: *Metastasis Research Protocols*. Humana Press; :081-089.
337. Aisenbrey EA, Murphy WL. Synthetic alternatives to Matrigel. *Nat Rev Mater*. 2020;5(7):539-551.
338. Hughes CS, Postovit LM, Lajoie GA. Matrigel: A complex protein mixture required for optimal growth of cell culture. *Proteomics*. 2010;10(9):1886-1890.
339. Hansen KC, Kiemele L, Maller O, et al. An In-solution Ultrasonication-assisted Digestion Method for Improved

Extracellular Matrix Proteome Coverage. *Molecular & Cellular Proteomics*. 2009;8(7):1648-1657.

340. Crapo PM, Gilbert TW, Badylak SF. An overview of tissue and whole organ decellularization processes. *Biomaterials*. 2011;32(12):3233-3243.
341. GILBERT T, SELLARO T, BADYLAK S. Decellularization of tissues and organs. *Biomaterials*. Published online March 7, 2006.
342. Moffat D, Ye K, Jin S. Decellularization for the retention of tissue niches. *J Tissue Eng*. 2022;13:204173142211011.
343. Fu RH, Wang YC, Liu SP, et al. Decellularization and Recellularization Technologies in Tissue Engineering. *Cell Transplant*. 2014;23(4-5):621-630.
344. Keane TJ, Swinehart IT, Badylak SF. Methods of tissue decellularization used for preparation of biologic scaffolds and in vivo relevance. *Methods*. 2015;84:25-34.
345. Badylak SF, Taylor D, Uygun K. Whole-Organ Tissue Engineering: Decellularization and Recellularization of Three-Dimensional Matrix Scaffolds. *Annu Rev Biomed Eng*. 2011;13(1):27-53.
346. Uriel S, Labay E, Francis-Sedlak M, et al. Extraction and Assembly of Tissue-Derived Gels for Cell Culture and Tissue Engineering. *Tissue Eng Part C Methods*. 2009;15(3):309-321.
347. Freytes DO, Martin J, Velankar SS, Lee AS, Badylak SF. Preparation and rheological characterization of a gel form of the porcine urinary bladder matrix. *Biomaterials*. 2008;29(11):1630-1637.

348. Lee JS, Shin J, Park HM, et al. Liver Extracellular Matrix Providing Dual Functions of Two-Dimensional Substrate Coating and Three-Dimensional Injectable Hydrogel Platform for Liver Tissue Engineering. *Biomacromolecules*. 2014;15(1):206-218.
349. Faulk DM, Londono R, Wolf MT, et al. ECM hydrogel coating mitigates the chronic inflammatory response to polypropylene mesh. *Biomaterials*. 2014;35(30):8585-8595.
350. DeQuach JA, Mezzano V, Miglani A, et al. Simple and High Yielding Method for Preparing Tissue Specific Extracellular Matrix Coatings for Cell Culture. *PLoS One*. 2010;5(9):e13039.
351. Kleinman HK, Martin GR. Matrigel: Basement membrane matrix with biological activity. *Semin Cancer Biol*. 2005;15(5):378-386.
352. Cheng MH, Uriel S, Moya ML, et al. Dermis-derived hydrogels support adipogenesis *in vivo*. *J Biomed Mater Res A*. 2009;9999A:NA-NA.
353. Hoshiba T. Decellularized Extracellular Matrix for Cancer Research. *Materials*. 2019;12(8):1311.
354. Sensi F, D'Angelo E, Piccoli M, et al. Recellularized Colorectal Cancer Patient-derived Scaffolds as *in vitro* Pre-clinical 3D Model for Drug Screening. *Cancers (Basel)*. 2020;12(3):681.
355. Rijal G, Li W. A versatile 3D tissue matrix scaffold system for tumor modeling and drug screening. *Sci Adv*. 2017;3(9).
356. Piccoli M, D'Angelo E, Crotti S, et al. Decellularized colorectal cancer matrix as bioactive microenvironment for *in vitro* 3D cancer research. *J Cell Physiol*. 2018;233(8):5937-5948.

357. Wishart AL, Conner SJ, Guarin JR, et al. Decellularized extracellular matrix scaffolds identify full-length collagen VI as a driver of breast cancer cell invasion in obesity and metastasis. *Sci Adv.* 2020;6(43).
358. Varinelli L, Guaglio M, Brich S, et al. Decellularized extracellular matrix as scaffold for cancer organoid cultures of colorectal peritoneal metastases. *J Mol Cell Biol.* 2023;14(11).
359. Gustafsson A, Garre E, Leiva MC, Salerno S, Ståhlberg A, Landberg G. Patient-derived scaffolds as a drug-testing platform for endocrine therapies in breast cancer. *Sci Rep.* 2021;11(1):13334.
360. Zhu X, Wu Q, He Y, et al. Fabrication of Size-Controllable and Arrangement-Orderly HepG2 Spheroids for Drug Screening via Decellularized Liver Matrix-Derived Micropattern Array Chips. *ACS Omega.* 2022;7(2):2364-2376.
361. Ferreira LP, Gaspar VM, Mendes L, Duarte IF, Mano JF. Organotypic 3D decellularized matrix tumor spheroids for high-throughput drug screening. *Biomaterials.* 2021;275:120983.
362. Dong H, Li Z, Bian S, et al. Culture of patient-derived multicellular clusters in suspended hydrogel capsules for pre-clinical personalized drug screening. *Bioact Mater.* 2022;18:164-177.
363. Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes.* 2004;1666(1-2):105-117.
364. Woods T, Gratzner PF. Effectiveness of three extraction techniques in the development of a decellularized bone–anterior cruciate ligament–bone graft. *Biomaterials.* 2005;26(35):7339-7349.

365. Cartmell JS, Dunn MG. Effect of chemical treatments on tendon cellularity and mechanical properties. *J Biomed Mater Res.* 2000;49(1):134-140.
366. Dahl SLM, Koh J, Prabhakar V, Niklason LE. Decellularized native and engineered arterial scaffolds for transplantation. *Cell Transplant.* 2003;12(6):659-666.
367. GRAUSS R, HAZEKAMP M, OPPENHUIZEN F, VANMUNSTEREN C, GITTENBERGERDEGROOT A, DERUITER M. Histological evaluation of decellularised porcine aortic valves: matrix changes due to different decellularisation methods. *European Journal of Cardio-Thoracic Surgery.* 2005;27(4):566-571.
368. Krasny L, Paul A, Wai P, Howard BA, Natrajan RC, Huang PH. Comparative proteomic assessment of matrixome enrichment methodologies. *Biochem J.* 2016;473(21):3979-3995.
369. Woods T, Gratzner PF. Effectiveness of three extraction techniques in the development of a decellularized bone–anterior cruciate ligament–bone graft. *Biomaterials.* 2005;26(35):7339-7349.
370. Faulk DM, Carruthers CA, Warner HJ, et al. The effect of detergents on the basement membrane complex of a biologic scaffold material. *Acta Biomater.* 2014;10(1):183-193.
371. Ren H, Shi X, Tao L, et al. Evaluation of two decellularization methods in the development of a whole-organ decellularized rat liver scaffold. *Liver International.* 2013;33(3):448-458.
372. Brown BN, Freund JM, Han L, et al. Comparison of Three Methods for the Derivation of a Biologic Scaffold Composed of Adipose

- Tissue Extracellular Matrix. *Tissue Eng Part C Methods*. 2011;17(4):411-421.
373. Gratzner PF, Harrison RD, Woods T. Matrix Alteration and Not Residual Sodium Dodecyl Sulfate Cytotoxicity Affects the Cellular Repopulation of a Decellularized Matrix. *Tissue Eng*. 2006;12(10):2975-2983.
374. Cebotari S, Tudorache I, Jaekel T, et al. Detergent Decellularization of Heart Valves for Tissue Engineering: Toxicological Effects of Residual Detergents on Human Endothelial Cells. *Artif Organs*. 2010;34(3):206-210.
375. Young DA, Ibrahim DO, Hu D, Christman KL. Injectable hydrogel scaffold from decellularized human lipoaspirate. *Acta Biomater*. 2011;7(3):1040-1049.
376. Yang M, Chen CZ, Wang XN, Zhu YB, Gu YJ. Favorable effects of the detergent and enzyme extraction method for preparing decellularized bovine pericardium scaffold for tissue engineered heart valves. *J Biomed Mater Res B Appl Biomater*. 2009;91B(1):354-361.
377. Spang MT, Christman KL. Extracellular matrix hydrogel therapies: In vivo applications and development. *Acta Biomater*. 2018;68:1-14.
378. Romero-López M, Trinh AL, Sobrino A, et al. Recapitulating the human tumor microenvironment: Colon tumor-derived extracellular matrix promotes angiogenesis and tumor cell growth. *Biomaterials*. 2017;116:118-129.

379. Mollica PA, Booth-Creech EN, Reid JA, et al. 3D bioprinted mammary organoids and tumoroids in human mammary derived ECM hydrogels. *Acta Biomater.* 2019;95:201-213.
380. Saldin LT, Cramer MC, Velankar SS, White LJ, Badylak SF. Extracellular matrix hydrogels from decellularized tissues: Structure and function. *Acta Biomater.* 2017;49:1-15.
381. Parkinson J. Simple physical model of collagen fibrillogenesis based on diffusion limited aggregation. *J Mol Biol.* 1995;247(4):823-831.
382. Brightman AO, Rajwa BP, Sturgis JE, McCallister ME, Robinson JP, Voytik-Harbin SL. Time-lapse confocal reflection microscopy of collagen fibrillogenesis and extracellular matrix assembly in vitro. *Biopolymers.* 2000;54(3):222-234.
383. Boso D, Maghin E, Carraro E, Giagante M, Pavan P, Piccoli M. Extracellular Matrix-Derived Hydrogels as Biomaterial for Different Skeletal Muscle Tissue Replacements. *Materials.* 2020;13(11):2483.
384. Hulmes DJS. Collagen Diversity, Synthesis and Assembly. In: *Collagen.* Springer US; :15-47.
385. Kim HW, Yeo IJ, Hwang KE, et al. Isolation and Characterization of Pepsin-soluble Collagens from Bones, Skins, and Tendons in Duck Feet. *Korean J Food Sci Anim Resour.* 2016;36(5):665-670.
386. Voytik-Harbin SL, Brightman AO, Waisner BZ, Robinson JP, Lamar CH. Small Intestinal Submucosa: A Tissue-Derived Extracellular Matrix That Promotes Tissue-Specific Growth and Differentiation of Cells *in Vitro.* *Tissue Eng.* 1998;4(2):157-174.

387. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Published online February 9, 2018.
388. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26(12):1572-1573.
389. Liu Y, Hayes DN, Nobel A, Marron JS. Statistical Significance of Clustering for High-Dimension, Low-Sample Size Data. *J Am Stat Assoc*. 2008;103(483):1281-1293.
390. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*. 2001;98(9):5116-5121.
391. Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res*. 2023;51(W1):W207-W212.
392. Barbie DA, Tamayo P, Boehm JS, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;462(7269):108-112.
393. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 2005;102(43):15545-15550.
394. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst*. 2015;1(6):417-425.

395. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023;51(D1):D638-D646.
396. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607-D613.
397. Shannon P, Markiel A, Ozier O, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 2003;13(11):2498-2504.
398. Xu H, Xu B, Yang Q, et al. Comparison of Decellularization Protocols for Preparing a Decellularized Porcine Annulus Fibrosus Scaffold. Pandit A, ed. *PLoS One.* 2014;9(1):e86723.
399. Nehrenheim L, Raschke S, Stefanski A, et al. Native aortic valve derived extracellular matrix hydrogel for three dimensional culture analyses with improved biomimetic properties. *Biomedical Materials.* 2019;14(3):035014.
400. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods.* 2020;17(1):41-44.
401. Fogh J, Wright WC, Loveless JD. Absence of HeLa Cell Contamination in 169 Cell Lines Derived From Human Tumors². *JNCI: Journal of the National Cancer Institute.* 1977;58(2):209-214.

402. Fogh J, Fogh JM, Orfeo T. One Hundred and Twenty-Seven Cultured Human Tumor Cell Lines Producing Tumors in Nude Mice²³. *JNCI: Journal of the National Cancer Institute*. 1977;59(1):221-226.
403. Salawu A, Fernando M, Hughes D, et al. Establishment and molecular characterisation of seven novel soft-tissue sarcoma cell lines. *Br J Cancer*. 2016;115(9):1058-1068.
404. Wiederschain D, Susan W, Chen L, et al. Single-vector inducible lentiviral RNAi system for oncology target validation. *Cell Cycle*. 2009;8(3):498-504.
405. Boettcher M, Tian R, Blau JA, et al. Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nat Biotechnol*. 2018;36(2):170-178.
406. Ershov D, Phan MS, Pylvänäinen JW, et al. TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat Methods*. 2022;19(7):829-832.
407. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9(7):676-682.
408. Gorelik R, Gautreau A. Quantitative and unbiased analysis of directional persistence in cell migration. *Nat Protoc*. 2014;9(8):1931-1943.
409. Kim M. Bidirectional Transmembrane Signaling by Cytoplasmic Domain Separation in Integrins. *Science (1979)*. 2003;301(5640):1720-1725.

410. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep.* 2014;15(12):1243-1253.
411. Socovich AM, Naba A. The cancer matrixome: From comprehensive characterization to biomarker discovery. *Semin Cell Dev Biol.* 2019;89:157-166.
412. Mushtaq MU, Papadas A, Pagenkopf A, et al. Tumor matrix remodeling and novel immunotherapies: the promise of matrix-derived immune biomarkers. *J Immunother Cancer.* 2018;6(1):65.
413. Salmon H, Franciszkievicz K, Damotte D, et al. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *Journal of Clinical Investigation.* 2012;122(3):899-910.
414. Kuczek DE, Larsen AMH, Thorseth ML, et al. Collagen density regulates the activity of tumor-infiltrating T cells. *J Immunother Cancer.* 2019;7(1):68.
415. Winograd-Katz SE, Fässler R, Geiger B, Legate KR. The integrin adhesome: from genes and proteins to human disease. *Nat Rev Mol Cell Biol.* 2014;15(4):273-288.
416. Oliver AJ, Lau PKH, Unsworth AS, et al. Tissue-Dependent Tumor Microenvironments and Their Impact on Immunotherapy Responses. *Front Immunol.* 2018;9.
417. Rosa F, Martinetti C, Piscopo F, et al. Multimodality imaging features of desmoid tumors: a head-to-toe spectrum. *Insights Imaging.* 2020;11(1):103.

418. Gounder MM, Thomas DM, Tap WD. Locally Aggressive Connective Tissue Tumors. *Journal of Clinical Oncology*. 2017;36(2):202-209.
419. Mitroulis I, Alexaki VI, Kourtzelis I, Ziogas A, Hajishengallis G, Chavakis T. Leukocyte integrins: Role in leukocyte recruitment and as therapeutic targets in inflammatory disease. *Pharmacol Ther*. 2015;147:123-135.
420. Gonzalez LL, Garrie K, Turner MD. Role of S100 proteins in health and disease. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. 2020;1867(6):118677.
421. Rossi E, Bernabeu C, Smadja DM. Endoglin as an Adhesion Molecule in Mature and Progenitor Endothelial Cells: A Function Beyond TGF- β . *Front Med (Lausanne)*. 2019;6.
422. Huang H, Liu Y, Yuan M, Marron JS. Statistical Significance of Clustering using Soft Thresholding. Published online May 24, 2013.
423. Fabregat A, Sidiropoulos K, Garapati P, et al. The Reactome pathway Knowledgebase. *Nucleic Acids Res*. 2016;44(D1):D481-D487.
424. Raudvere U, Kolberg L, Kuzmin I, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019;47(W1):W191-W198.
425. Orth MF, Buecklein VL, Kampmann E, et al. A comparative view on the expression patterns of PD-L1 and PD-1 in soft tissue sarcomas. *Cancer Immunology, Immunotherapy*. 2020;69(7):1353-1362.

426. Schroeder BA, LaFranzo NA, LaFleur BJ, et al. CD4+ T cell and M2 macrophage infiltration predict dedifferentiated liposarcoma patient outcomes. *J Immunother Cancer*. 2021;9(8):e002812.
427. Naba A. Ten Years of Extracellular Matrix Proteomics: Accomplishments, Challenges, and Future Perspectives. *Molecular & Cellular Proteomics*. 2023;22(4):100528.
428. McKee TJ, Perlman G, Morris M, Komarova S V. Extracellular matrix composition of connective tissues: a systematic review and meta-analysis. *Sci Rep*. 2019;9(1):10542.
429. McCabe MC, Saviola AJ, Hansen KC. Mass Spectrometry-Based Atlas of Extracellular Matrix Proteins across 25 Mouse Organs. *J Proteome Res*. 2023;22(3):790-801.
430. Byron A, Humphries JD, Craig SE, Knight D, Humphries MJ. Proteomic analysis of $\alpha 4\beta 1$ integrin adhesion complexes reveals α -subunit-dependent protein recruitment. *Proteomics*. 2012;12(13):2107-2114.
431. Horton ER, Byron A, Askari JA, et al. Definition of a consensus integrin adhesome and its dynamics during adhesion complex assembly and disassembly. *Nat Cell Biol*. 2015;17(12):1577-1587.
432. Geiger T, Zaidel-Bar R. Opening the floodgates: proteomics and the integrin adhesome. *Curr Opin Cell Biol*. 2012;24(5):562-568.
433. Jones MC, Humphries JD, Byron A, et al. Isolation of Integrin-Based Adhesion Complexes. *Curr Protoc Cell Biol*. 2015;66(1).
434. Nyström H, Jönsson M, Nilbert M, Carneiro A. Immune-cell infiltration in high-grade soft tissue sarcomas; prognostic

- implications of tumor-associated macrophages and B-cells. *Acta Oncol (Madr)*. 2023;62(1):33-39.
435. Beck AH, Lee CH, Witten DM, et al. Discovery of molecular subtypes in leiomyosarcoma through integrative molecular profiling. *Oncogene*. 2010;29(6):845-854.
436. Guo X, Jo VY, Mills AM, et al. Clinically Relevant Molecular Subtypes in Leiomyosarcoma. *Clinical Cancer Research*. 2015;21(15):3501-3511.
437. Abeshouse A, Adebamowo C, Adebamowo SN, et al. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell*. 2017;171(4):950-965.e28.
438. Hemming ML, Fan C, Raut CP, et al. Oncogenic Gene-Expression Programs in Leiomyosarcoma and Characterization of Conventional, Inflammatory, and Uterogenic Subtypes. *Molecular Cancer Research*. 2020;18(9):1302-1314.
439. Anderson ND, Babichev Y, Fuligni F, et al. Lineage-defined leiomyosarcoma subtypes emerge years before diagnosis and determine patient survival. *Nat Commun*. 2021;12(1):4496.
440. Angel PM, Schwamborn K, Comte-Walters S, et al. Extracellular Matrix Imaging of Breast Tissue Pathologies by MALDI–Imaging Mass Spectrometry. *Proteomics Clin Appl*. 2019;13(1).
441. Clift CL, Drake RR, Mehta A, Angel PM. Multiplexed imaging mass spectrometry of the extracellular matrix using serial enzyme digests from formalin-fixed paraffin-embedded tissue sections. *Anal Bioanal Chem*. 2021;413(10):2709-2719.
442. Li X, Truty MA, Kang Y, et al. Extracellular Lumican Inhibits Pancreatic Cancer Cell Growth and Is Associated with Prolonged

- Survival after Surgery. *Clinical Cancer Research*. 2014;20(24):6529-6540.
443. Jeanne A, Untereiner V, Perreau C, et al. Lumican delays melanoma growth in mice and drives tumor molecular assembly as well as response to matrix-targeted TAX2 therapeutic peptide. *Sci Rep*. 2017;7(1):7700.
444. Mongiat M, Sweeney SM, San Antonio JD, Fu J, Iozzo R V. Endorepellin, a Novel Inhibitor of Angiogenesis Derived from the C Terminus of Perlecan. *Journal of Biological Chemistry*. 2003;278(6):4238-4249.
445. Reed CC, Waterhouse A, Kirby S, et al. Decorin prevents metastatic spreading of breast cancer. *Oncogene*. 2005;24(6):1104-1110.
446. Grant DS, Yenisey C, Rose RW, Tootell M, Santra M, Iozzo R V. Decorin suppresses tumor cell-mediated angiogenesis. *Oncogene*. 2002;21(31):4765-4777.
447. Hu X, Villodre ES, Larson R, et al. Decorin-mediated suppression of tumorigenesis, invasion, and metastasis in inflammatory breast cancer. *Commun Biol*. 2021;4(1):72.
448. Niedworok C, Röck K, Kretschmer I, et al. Inhibitory Role of the Small Leucine-Rich Proteoglycan Biglycan in Bladder Cancer. *PLoS One*. 2013;8(11):e80084.
449. Weber CK, Sommer G, Michl P, et al. Biglycan is overexpressed in pancreatic cancer and induces G1-arrest in pancreatic cancer cell lines. *Gastroenterology*. 2001;121(3):657-667.
450. Vlodaysky I, Miao HQ, Medalion B, Danagher P, Ron D. Involvement of heparan sulfate and related molecules in

- sequestration and growth promoting activity of fibroblast growth factor. *Cancer and Metastasis Reviews*. 1996;15(2):177-186.
451. Chua CC, Rahimi N, Forsten-Williams K, Nugent MA. Heparan Sulfate Proteoglycans Function as Receptors for Fibroblast Growth Factor-2 Activation of Extracellular Signal-Regulated Kinases 1 and 2. *Circ Res*. 2004;94(3):316-323.
452. Hsu SHC, Nadesan P, Puvindran V, Stallcup WB, Kirsch DG, Alman BA. Effects of chondroitin sulfate proteoglycan 4 (NG2/CSPG4) on soft-tissue sarcoma growth depend on tumor developmental stage. *Journal of Biological Chemistry*. 2018;293(7):2466-2475.
453. Benassi MS, Pazzaglia L, Chiechi A, et al. NG2 expression predicts the metastasis formation in soft-tissue sarcoma patients. *Journal of Orthopaedic Research*. 2009;27(1):135-140.
454. Wight TN, Kang I, Evanko SP, et al. Versican—A Critical Extracellular Matrix Regulator of Immunity and Inflammation. *Front Immunol*. 2020;11.
455. Wei J, Hu M, Huang K, Lin S, Du H. Roles of Proteoglycans and Glycosaminoglycans in Cancer Development and Progression. *Int J Mol Sci*. 2020;21(17):5983.
456. Collins LE, Troeberg L. Heparan sulfate as a regulator of inflammation and immunity. *J Leukoc Biol*. 2019;105(1):81-92.
457. Hughes CE, Nibbs RJB. A guide to chemokines and their receptors. *FEBS J*. 2018;285(16):2944-2971.
458. Gray AL, Pun N, Ridley AJL, Dyer DP. Role of extracellular matrix proteoglycans in immune cell recruitment. *Int J Exp Pathol*. 2022;103(2):34-43.

459. Lipowsky HH. Role of the Glycocalyx as a Barrier to Leukocyte-Endothelium Adhesion. In: ; 2018:51-68.
460. Schmidt EP, Yang Y, Janssen WJ, et al. The pulmonary endothelial glycocalyx regulates neutrophil adhesion and lung injury during experimental sepsis. *Nat Med.* 2012;18(8):1217-1223.
461. Constantinescu AA, Vink H, Spaan JAE. Endothelial Cell Glycocalyx Modulates Immobilization of Leukocytes at the Endothelial Surface. *Arterioscler Thromb Vasc Biol.* 2003;23(9):1541-1547.
462. Wang Z, Liu J, Han J, Yang Z, Wang Q. Analysis of prognostic factors of undifferentiated pleomorphic sarcoma and construction and validation of a prediction nomogram based on SEER database. *Eur J Med Res.* 2022;27(1):179.
463. Mussi C, Collini P, Miceli R, et al. The prognostic impact of dedifferentiation in retroperitoneal liposarcoma. *Cancer.* 2008;113(7):1657-1665.
464. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun.* 2020;11(1):5120.
465. Henke E, Nandigama R, Ergün S. Extracellular Matrix in the Tumor Microenvironment and Its Impact on Cancer Therapy. *Front Mol Biosci.* 2020;6.
466. Smolle MA, Herbsthofer L, Goda M, et al. Influence of tumor-infiltrating immune cells on local control rate, distant metastasis, and survival in patients with soft tissue sarcoma. *Oncoimmunology.* 2021;10(1).

467. Takada Y, Ye X, Simon S. The integrins. *Genome Biol.* 2007;8(5):215.
468. Effenbein A, Simons M. Syndecan-4 signaling at a glance. *J Cell Sci.* Published online January 1, 2013.
469. Sepulveda JL, Wu C. The parvins. *Cellular and Molecular Life Sciences.* 2006;63(1):25.
470. Russell MA. Synemin Redefined: Multiple Binding Partners Results in Multifunctionality. *Front Cell Dev Biol.* 2020;8.
471. Hedberg-Oldfors C, Meyer R, Nolte K, et al. Loss of supervillin causes myopathy with myofibrillar disorganization and autophagic vacuoles. *Brain.* 2020;143(8):2406-2420.
472. Li D, Mukai K, Suzuki T, et al. Adrenocortical zonation factor 1 is a novel matricellular protein promoting integrin-mediated adhesion of adrenocortical and vascular smooth muscle cells. *FEBS Journal.* 2007;274(10):2506-2522.
473. Litteri G, Carnevale D, D'Urso A, et al. Vascular Smooth Muscle Emilin-1 Is a Regulator of Arteriolar Myogenic Response and Blood Pressure. *Arterioscler Thromb Vasc Biol.* 2012;32(9):2178-2184.
474. Yang Z, Xue F, Li M, et al. Extracellular Matrix Characterization in Gastric Cancer Helps to Predict Prognosis and Chemotherapy Response. *Front Oncol.* 2021;11.
475. Simon M, Mughal SS, Horak P, et al. Deconvolution of sarcoma methylomes reveals varying degrees of immune cell infiltrates with association to genomic aberrations. *J Transl Med.* 2021;19(1):204.

476. Chen S, Huang W, Luo P, et al. Undifferentiated Pleomorphic Sarcoma: Long-Term Follow-Up from a Large Institution. *Cancer Manag Res.* 2019;Volume 11:10001-10009.
477. Tumei PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature.* 2014;515(7528):568-571.
478. Keire PA, Bressler SL, Lemire JM, et al. A Role for Versican in the Development of Leiomyosarcoma. *Journal of Biological Chemistry.* 2014;289(49):34089-34103.
479. Qi Y, Xu R. Roles of PLODs in Collagen Synthesis and Cancer Progression. *Front Cell Dev Biol.* 2018;6.
480. Liburkin-Dan T, Toledano S, Neufeld G. Lysyl Oxidase Family Enzymes and Their Role in Tumor Progression. *Int J Mol Sci.* 2022;23(11):6249.
481. Egeblad M, Rasch MG, Weaver VM. Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol.* 2010;22(5):697-706.
482. Eisinger-Mathason TSK, Zhang M, Qiu Q, et al. Hypoxia-Dependent Modification of Collagen Networks Promotes Sarcoma Metastasis. *Cancer Discov.* 2013;3(10):1190-1205.
483. Gilkes DM, Bajpai S, Wong CC, et al. Procollagen Lysyl Hydroxylase 2 Is Essential for Hypoxia-Induced Breast Cancer Metastasis. *Molecular Cancer Research.* 2013;11(5):456-466.
484. Xu F, Zhang J, Hu G, Liu L, Liang W. Hypoxia and TGF- β 1 induced PLOD2 expression improve the migration and invasion of cervical cancer cells by promoting epithelial-to-mesenchymal

- transition (EMT) and focal adhesion formation. *Cancer Cell Int.* 2017;17(1):54.
485. Ferreira S, Saraiva N, Rijo P, Fernandes AS. LOXL2 Inhibitors and Breast Cancer Progression. *Antioxidants.* 2021;10(2):312.
486. Cox TR, Gartland A, Erler JT. Lysyl Oxidase, a Targetable Secreted Molecule Involved in Cancer Metastasis. *Cancer Res.* 2016;76(2):188-192.
487. Baldari S, Di Modugno F, Nisticò P, Toietta G. Strategies for Efficient Targeting of Tumor Collagen for Cancer Therapy. *Cancers (Basel).* 2022;14(19):4706.
488. Chan N, Willis A, Kornhauser N, et al. Influencing the Tumor Microenvironment: A Phase II Study of Copper Depletion Using Tetrathiomolybdate in Patients with Breast Cancer at High Risk for Recurrence and in Preclinical Models of Lung Metastases. *Clinical Cancer Research.* 2017;23(3):666-676.
489. Pennacchioli E, Tosti G, Barberis M, et al. Sarcoma spreads primarily through the vascular system: are there biomarkers associated with vascular spread? *Clin Exp Metastasis.* 2012;29(7):757-773.
490. van der Laan P, Tirota F, Pankova V, Ford S, Huang P, van Houdt WJ. Atypical Patterns of Metastases: How Do Sarcomas Metastasize? In: Leong SP, Nathanson SD, Zager JS, eds. *Cancer Metastasis Through the Lymphovascular System.* Springer International Publishing; 2022:629-637.
491. Ribatti D, Tamma R, Annese T. Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. *Transl Oncol.* 2020;13(6):100773.

492. Sannino G, Marchetto A, Kirchner T, Grünewald TGP. Epithelial-to-Mesenchymal and Mesenchymal-to-Epithelial Transition in Mesenchymal Tumors: A Paradox in Sarcomas? *Cancer Res.* 2017;77(17):4556-4561.
493. Stemmler MP, Eccles RL, Brabletz S, Brabletz T. Non-redundant functions of EMT transcription factors. *Nat Cell Biol.* 2019;21(1):102-112.
494. Yang J, Antin P, Bex G, et al. Guidelines and definitions for research on epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol.* 2020;21(6):341-352.
495. Saito T, Nagai M, Ladanyi M. SYT-SSX1 and SYT-SSX2 Interfere with Repression of E-Cadherin by Snail and Slug: A Potential Mechanism for Aberrant Mesenchymal to Epithelial Transition in Human Synovial Sarcoma. *Cancer Res.* 2006;66(14):6919-6927.
496. Yang J, Eddy JA, Pan Y, et al. Integrated Proteomics and Genomics Analysis Reveals a Novel Mesenchymal to Epithelial Reverting Transition in Leiomyosarcoma through Regulation of Slug. *Molecular & Cellular Proteomics.* 2010;9(11):2405-2413.
497. Armah HB, Parwani A V. Epithelioid Sarcoma. *Arch Pathol Lab Med.* 2009;133(5):814-819.
498. Nieto MA, Huang RYJ, Jackson RA, Thiery JP. EMT: 2016. *Cell.* 2016;166(1):21-45.
499. Tian W, Wang G, Yang J, Pan Y, Ma Y. Prognostic role of E-cadherin and Vimentin expression in various subtypes of soft tissue leiomyosarcomas. *Medical Oncology.* 2013;30(1):401.
500. Heinz A. Elastic fibers during aging and disease. *Ageing Res Rev.* 2021;66:101255.

501. Alkhouli N, Mansfield J, Green E, et al. The mechanical properties of human adipose tissues and their relationships to the structure and composition of the extracellular matrix. *American Journal of Physiology-Endocrinology and Metabolism*. 2013;305(12):E1427-E1435.
502. Spencer M, Unal R, Zhu B, et al. Adipose Tissue Extracellular Matrix and Vascular Abnormalities in Obesity and Insulin Resistance. *J Clin Endocrinol Metab*. 2011;96(12):E1990-E1998.
503. Sakai LY, Keene DR, Renard M, De Backer J. FBN1: The disease-causing gene for Marfan syndrome and other genetic disorders. *Gene*. 2016;591(1):279-291.
504. Muthu ML, Tiedemann K, Fradette J, Komarova S, Reinhardt DP. Fibrillin-1 regulates white adipose tissue development, homeostasis, and function. *Matrix Biology*. 2022;110:106-128.
505. Oike N, Kawashima H, Ogose A, et al. Human leukocyte antigen I is significantly downregulated in patients with myxoid liposarcomas. *Cancer Immunology, Immunotherapy*. 2021;70(12):3489-3499.
506. Dunkelberger JR, Song WC. Complement and its role in innate and adaptive immune responses. *Cell Res*. 2010;20(1):34-50.
507. Dzik S. Complement and Coagulation: Cross Talk Through Time. *Transfus Med Rev*. 2019;33(4):199-206.
508. Kolev M, Das M, Gerber M, Baver S, Deschatelets P, Markiewski MM. Inside-Out of Complement in Cancer. *Front Immunol*. 2022;13.

509. Kwak JW, Laskowski J, Li HY, et al. Complement Activation via a C3a Receptor Pathway Alters CD4+ T Lymphocytes and Mediates Lung Cancer Progression. *Cancer Res.* 2018;78(1):143-156.
510. Nabizadeh JA, Manthey HD, Steyn FJ, et al. The Complement C3a Receptor Contributes to Melanoma Tumorigenesis by Inhibiting Neutrophil and CD4+ T Cell Responses. *The Journal of Immunology.* 2016;196(11):4783-4792.
511. Markiewski MM, DeAngelis RA, Benencia F, et al. Modulation of the antitumor immune response by complement. *Nat Immunol.* 2008;9(11):1225-1235.
512. Neuhaus SJ, Barry P, Clark MA, Hayes AJ, Fisher C, Thomas JM. Surgical management of primary and recurrent retroperitoneal liposarcoma. *British Journal of Surgery.* 2005;92(2):246-252.
513. Singer S, Antonescu CR, Riedel E, Brennan MF. Histologic Subtype and Margin of Resection Predict Pattern of Recurrence and Survival for Retroperitoneal Liposarcoma. *Ann Surg.* 2003;238(3):358-371.
514. Fujiwara T, Zhang L, Chandler A, et al. Cathepsin protease expression in infiltrative soft tissue sarcomas: cathepsin-K correlates with infiltrative tumor growth and clinical outcomes. *Hum Pathol.* 2023;134:30-44.
515. Wurl P, Taubert H, Meye A, et al. Immunohistochemical and clinical evaluation of cathepsin expression in soft tissue sarcomas. *Virchows Archiv.* 1997;430(3):221-225.
516. Gocheva V, Zeng W, Ke D, et al. Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. *Genes Dev.* 2006;20(5):543-556.

517. Gopinathan A, DeNicola GM, Frese KK, et al. Cathepsin B promotes the progression of pancreatic ductal adenocarcinoma in mice. *Gut*. 2012;61(6):877-884.
518. Withana NP, Blum G, Sameni M, et al. Cathepsin B Inhibition Limits Bone Metastasis in Breast Cancer. *Cancer Res*. 2012;72(5):1199-1209.
519. Endo-Munoz L, Cai N, Cumming A, et al. Progression of Osteosarcoma from a Non-Metastatic to a Metastatic Phenotype Is Causally Associated with Activation of an Autocrine and Paracrine uPA Axis. *PLoS One*. 2015;10(8):e0133592.
520. Oh F, Todhunter D, Taras E, Vallera D, Borgatti A. Targeting EGFR and uPAR on human rhabdomyosarcoma, osteosarcoma, and ovarian adenocarcinoma with a bispecific ligand-directed toxin. *Clin Pharmacol*. 2018;Volume 10:113-121.
521. Pilbeam K, Wang H, Taras E, et al. Targeting pediatric sarcoma with a bispecific ligand immunotoxin targeting urokinase and epidermal growth factor receptors. *Oncotarget*. 2018;9(15):11938-11947.
522. Borgatti A, Koopmeiners JS, Sarver AL, et al. Safe and Effective Sarcoma Therapy through Bispecific Targeting of EGFR and uPAR. *Mol Cancer Ther*. 2017;16(5):956-965.
523. Amor C, Feucht J, Leibold J, et al. Senolytic CAR T cells reverse senescence-associated pathologies. *Nature*. 2020;583(7814):127-132.
524. Rullo AF, Fitzgerald KJ, Muthusamy V, et al. Re-engineering the Immune Response to Metastatic Cancer: Antibody-Recruiting

Small Molecules Targeting the Urokinase Receptor. *Angewandte Chemie International Edition*. 2016;55(11):3642-3646.

525. Ischenko I, Zhi J, Hayman MJ, Petrenko O. KRAS-dependent suppression of MYC enhances the sensitivity of cancer cells to cytotoxic agents. *Oncotarget*. 2017;8(11):17995-18009.
526. Maya-Mendoza A, Ostrakova J, Kosar M, et al. Myc and Ras oncogenes engage different energy metabolism programs and evoke distinct patterns of oxidative and DNA replication stress. *Mol Oncol*. 2015;9(3):601-616.
527. Byun J, Park M, Yun JW, et al. Oncogenic KRAS signaling activates mTORC1 through COUP-TFII-mediated lactate production. *EMBO Rep*. 2019;20(6).
528. Dong Y, Tu R, Liu H, Qing G. Regulation of cancer cell metabolism: oncogenic MYC in the driver's seat. *Signal Transduct Target Ther*. 2020;5(1):124.
529. Fukushiro-Lopes D, Hegel AD, Russo A, et al. Repurposing Kir6/SUR2 Channel Activator Minoxidil to Arrests Growth of Gynecologic Cancers. *Front Pharmacol*. 2020;11.
530. Cook N, Banerji U, Evans J, et al. Pharmacokinetic (PK) assessment of BT1718: A phase I/II a study of BT1718, a first in class bicycle toxin conjugate (BTC), in patients (pts) with advanced solid tumours. *Annals of Oncology*. 2019;30:v174.
531. Bennett G, Lutz R, Park P, Harrison H, Lee K. Abstract 1167: Development of BT1718, a novel Bicycle Drug Conjugate for the treatment of lung cancer. *Cancer Res*. 2017;77(13_Supplement):1167-1167.

532. Lopus M, Oroudjev E, Wilson L, et al. Maytansine and Cellular Metabolites of Antibody-Maytansinoid Conjugates Strongly Suppress Microtubule Dynamics by Binding to Microtubules. *Mol Cancer Ther.* 2010;9(10):2689-2699.
533. Gowland C, Berry P, Errington J, et al. Development of a LC–MS/MS method for the quantification of toxic payload DM1 cleaved from BT1718 in a Phase I study. *Bioanalysis.* 2021;13(2):101-113.
534. Kaur A, Ecker BL, Douglass SM, et al. Remodeling of the Collagen Matrix in Aging Skin Promotes Melanoma Metastasis and Affects Immune Cell Motility. *Cancer Discov.* 2019;9(1):64-81.
535. Hartmann N, Giese NA, Giese T, et al. Prevailing Role of Contact Guidance in Intrastromal T-cell Trapping in Human Pancreatic Cancer. *Clinical Cancer Research.* 2014;20(13):3422-3433.
536. Pruitt HC, Guan Y, Liu H, et al. Collagen VI deposition mediates stromal T cell trapping through inhibition of T cell motility in the prostate tumor microenvironment. *Matrix Biology.* 2023;121:90-104.
537. Peng DH, Rodriguez BL, Diao L, et al. Collagen promotes anti-PD-1/PD-L1 resistance in cancer through LAIR1-dependent CD8+ T cell exhaustion. *Nat Commun.* 2020;11(1):4520.
538. Monga V, Skubitz KM, Maliske S, et al. A Retrospective Analysis of the Efficacy of Immunotherapy in Metastatic Soft-Tissue Sarcomas. *Cancers (Basel).* 2020;12(7):1873.
539. Nicolas-Boluda A, Vaquero J, Vimeux L, et al. Tumor stiffening reversion through collagen crosslinking inhibition improves T cell migration and anti-PD-1 treatment. *Elife.* 2021;10.

540. Tirotta F, Bacon A, Collins S, et al. Primary retroperitoneal sarcoma: A comparison of survival outcomes in specialist and non-specialist sarcoma centres. *Eur J Cancer*. 2023;188:20-28.
541. Pietilä EA, Gonzalez-Molina J, Moyano-Galceran L, et al. Co-evolution of matrisome and adaptive adhesion dynamics drives ovarian cancer chemoresistance. *Nat Commun*. 2021;12(1):3904.
542. Krasny L, Bland P, Burns J, et al. A mouse SWATH-MS reference spectral library enables deconvolution of species-specific proteomic alterations in human tumour xenografts. *Dis Model Mech*. Published online January 1, 2020.
543. Toro JR, Travis LB, Wu HJ, Zhu K, Fletcher CDM, Devesa SS. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978-2001: An analysis of 26,758 cases. *Int J Cancer*. 2006;119(12):2922-2930.
544. Nakata E, Kunisada T, Hasei J, et al. What Are the Results of Resection of Localized Dedifferentiated Liposarcomas in the Extremities? *Clin Orthop Relat Res*. 2020;478(11):2550-2561.
545. Teicher BA, Polley E, Kunkel M, et al. Sarcoma Cell Line Screen of Oncology Drugs and Investigational Agents Identifies Patterns Associated with Gene and microRNA Expression. *Mol Cancer Ther*. 2015;14(11):2452-2462.
546. Zhao W, Li J, Chen MJM, et al. Large-Scale Characterization of Drug Responses of Clinically Relevant Proteins in Cancer Cell Lines. *Cancer Cell*. 2020;38(6):829-843.e4.

547. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? *Acta Pharm Sin B*. 2022;12(7):3049-3062.
548. Gaebler M, Silvestri A, Haybaeck J, et al. Three-Dimensional Patient-Derived In Vitro Sarcoma Models: Promising Tools for Improving Clinical Tumor Management. *Front Oncol*. 2017;7.
549. Benton G, Kleinman HK, George J, Arnaoutova I. Multiple uses of basement membrane-like matrix (BME/Matrigel) in vitro and in vivo with cancer cells. *Int J Cancer*. 2011;128(8):1751-1757.
550. Xu C, Inokuma MS, Denham J, et al. Feeder-free growth of undifferentiated human embryonic stem cells. *Nat Biotechnol*. 2001;19(10):971-974.
551. Lancaster MA, Renner M, Martin CA, et al. Cerebral organoids model human brain development and microcephaly. *Nature*. 2013;501(7467):373-379.
552. Gilpin A, Yang Y. Decellularization Strategies for Regenerative Medicine: From Processing Techniques to Applications. *Biomed Res Int*. 2017;2017:1-13.
553. López O, Cócera M, Pons R, Azemar N, de la Maza A. Kinetic Studies of Liposome Solubilization by Sodium Dodecyl Sulfate Based on a Dynamic Light Scattering Technique. *Langmuir*. 1998;14(16):4671-4674.
554. Burk J, Erbe I, Berner D, et al. Freeze-Thaw Cycles Enhance Decellularization of Large Tendons. *Tissue Eng Part C Methods*. 2014;20(4):276-284.

555. Giobbe GG, Crowley C, Luni C, et al. Extracellular matrix hydrogel derived from decellularized tissues enables endodermal organoid culture. *Nat Commun.* 2019;10(1):5658.
556. Pouliot RA, Link PA, Mikhael NS, et al. Development and characterization of a naturally derived lung extracellular matrix hydrogel. *J Biomed Mater Res A.* 2016;104(8):1922-1935.
557. Hamuro Y, Coales SJ, Molnar KS, Tuske SJ, Morrow JA. Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. *Rapid Communications in Mass Spectrometry.* 2008;22(7):1041-1046.
558. Amirrah IN, Lokanathan Y, Zulkiflee I, Wee MFMR, Motta A, Fauzi MB. A Comprehensive Review on Collagen Type I Development of Biomaterials for Tissue Engineering: From Biosynthesis to Bioscaffold. *Biomedicines.* 2022;10(9):2307.
559. Lynn AK, Yannas IV, Bonfield W. Antigenicity and immunogenicity of collagen. *J Biomed Mater Res.* 2004;71B(2):343-354.
560. Krasny L, Bland P, Kogata N, et al. SWATH mass spectrometry as a tool for quantitative profiling of the matrisome. *J Proteomics.* 2018;189:11-22.
561. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer.* 2013;13(10):714-726.
562. Holle AW, Young JL, Spatz JP. In vitro cancer cell–ECM interactions inform in vivo cancer treatment. *Adv Drug Deliv Rev.* 2016;97:270-279.

563. Dickreuter E, Cordes N. The cancer cell adhesion resistome: mechanisms, targeting and translational approaches. *Biol Chem.* 2017;398(7):721-735.
564. Jiang H, Hegde S, DeNardo DG. Tumor-associated fibrosis as a regulator of tumor immunity and response to immunotherapy. *Cancer Immunology, Immunotherapy.* 2017;66(8):1037-1048.
565. Jensen ARD, Horton ER, Blicher LH, et al. Organ-Specific, Fibroblast-Derived Matrix as a Tool for Studying Breast Cancer Metastasis. *Cancers (Basel).* 2021;13(13):3331.
566. Sherman-Baust CA, Weeraratna AT, Rangel LBA, et al. Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells. *Cancer Cell.* 2003;3(4):377-386.
567. Popova N V., Jücker M. The Functional Role of Extracellular Matrix Proteins in Cancer. *Cancers (Basel).* 2022;14(1):238.
568. Öhlund D, Franklin O, Lundberg E, Lundin C, Sund M. Type IV collagen stimulates pancreatic cancer cell proliferation, migration, and inhibits apoptosis through an autocrine loop. *BMC Cancer.* 2013;13(1):154.
569. Cattaruzza S, Nicolosi PA, Braghetta P, et al. NG2/CSPG4-collagen type VI interplays putatively involved in the microenvironmental control of tumour engraftment and local expansion. *J Mol Cell Biol.* 2013;5(3):176-193.
570. Gopal S, Veracini L, Grall D, et al. Fibronectin-guided migration of carcinoma collectives. *Nat Commun.* 2017;8(1):14105.

571. Parsons JT, Horwitz AR, Schwartz MA. Cell adhesion: integrating cytoskeletal dynamics and cellular tension. *Nat Rev Mol Cell Biol.* 2010;11(9):633-643.
572. Ohnuma T, Arkin H, Holland JF. Effects of cell density on drug-induced cell kill kinetics in vitro (inoculum effect). *Br J Cancer.* 1986;54(3):415-421.
573. Fares J, Fares MY, Khachfe HH, Salhab HA, Fares Y. Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduct Target Ther.* 2020;5(1):28.
574. Petrie RJ, Doyle AD, Yamada KM. Random versus directionally persistent cell migration. *Nat Rev Mol Cell Biol.* 2009;10(8):538-549.
575. Gorelik R, Gautreau A. Quantitative and unbiased analysis of directional persistence in cell migration. *Nat Protoc.* 2014;9(8):1931-1943.
576. Baskaran JP, Weldy A, Guarin J, et al. Cell shape, and not 2D migration, predicts extracellular matrix-driven 3D cell invasion in breast cancer. *APL Bioeng.* 2020;4(2).
577. Gu J, Sumida Y, Sanzen N, Sekiguchi K. Laminin-10/11 and Fibronectin Differentially Regulate Integrin- dependent Rho and Rac Activation via p130Cas-CrkII-DOCK180 Pathway. *Journal of Biological Chemistry.* 2001;276(29):27090-27097.
578. Graf F, Horn P, Ho AD, Boutros M, Maercker C. The extracellular matrix proteins type I collagen, type III collagen, fibronectin, and laminin 421 stimulate migration of cancer cells. *The FASEB Journal.* 2021;35(7).

579. Ramos G de O, Bernardi L, Lauxen I, Sant'Ana Filho M, Horwitz AR, Lamers ML. Fibronectin Modulates Cell Adhesion and Signaling to Promote Single Cell Migration of Highly Invasive Oral Squamous Cell Carcinoma. *PLoS One*. 2016;11(3):e0151338.
580. Hill RC, Calle EA, Dzieciatkowska M, Niklason LE, Hansen KC. Quantification of Extracellular Matrix Proteins from a Rat Lung Scaffold to Provide a Molecular Readout for Tissue Engineering. *Molecular & Cellular Proteomics*. 2015;14(4):961-973.
581. Calle EA, Hill RC, Leiby KL, et al. Targeted proteomics effectively quantifies differences between native lung and detergent-decellularized lung extracellular matrices. *Acta Biomater*. 2016;46:91-100.
582. Krasny L, Paul A, Wai P, Howard BA, Natrajan RC, Huang PH. Comparative proteomic assessment of matrixome enrichment methodologies. *Biochem J*. 2016;473(21):3979-3995.
583. Leng L, Ma J, Sun X, et al. Comprehensive proteomic atlas of skin biomatrix scaffolds reveals a supportive microenvironment for epidermal development. *J Tissue Eng*. 2020;11:204173142097231.
584. Barallobre-Barreiro J, Oklu R, Lynch M, et al. Extracellular matrix remodelling in response to venous hypertension: proteomics of human varicose veins. *Cardiovasc Res*. 2016;110(3):419-430.
585. Schiller HB, Fernandez IE, Burgstaller G, et al. Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. *Mol Syst Biol*. 2015;11(7).

586. Simsa R, Rothenbücher T, Gürbüz H, et al. Brain organoid formation on decellularized porcine brain ECM hydrogels. *PLoS One*. 2021;16(1):e0245685.
587. Sackett SD, Tremmel DM, Ma F, et al. Extracellular matrix scaffold and hydrogel derived from decellularized and delipidized human pancreas. *Sci Rep*. 2018;8(1):10452.
588. Giobbe GG, Crowley C, Luni C, et al. Extracellular matrix hydrogel derived from decellularized tissues enables endodermal organoid culture. *Nat Commun*. 2019;10(1):5658.
589. Liu Z, Lee SJ, Park S, et al. Cancer cells display increased migration and deformability in pace with metastatic progression. *The FASEB Journal*. 2020;34(7):9307-9315.
590. Nousi A, Søggaard MT, Audoin M, Jauffred L. Single-cell tracking reveals super-spreading brain cancer cells with high persistence. *Biochem Biophys Rep*. 2021;28:101120.
591. Ridley AJ, Schwartz MA, Burridge K, et al. Cell Migration: Integrating Signals from Front to Back. *Science (1979)*. 2003;302(5651):1704-1709.
592. Leineweber WD, Fraley SI. Adhesion tunes speed and persistence by coordinating protrusions and extracellular matrix remodeling. *Dev Cell*. 2023;58(15):1414-1428.e4.
593. Doyle AD, Carvajal N, Jin A, Matsumoto K, Yamada KM. Local 3D matrix microenvironment regulates cell migration through spatiotemporal dynamics of contractility-dependent adhesions. *Nat Commun*. 2015;6(1):8720.

594. Maiuri P, Rupprecht JF, Wieser S, et al. Actin Flows Mediate a Universal Coupling between Cell Speed and Cell Persistence. *Cell*. 2015;161(2):374-386.
595. Wortel IMN, Niculescu I, Koliijn PM, Gov NS, de Boer RJ, Textor J. Local actin dynamics couple speed and persistence in a cellular Potts model of cell migration. *Biophys J*. 2021;120(13):2609-2622.
596. Crosas-Molist E, Samain R, Kohlhammer L, et al. Rho GTPase signaling in cancer progression and dissemination. *Physiol Rev*. 2022;102(1):455-510.
597. Muncie JM, Weaver VM. The Physical and Biochemical Properties of the Extracellular Matrix Regulate Cell Fate. In: ; 2018:1-37.
598. Trappmann B, Chen CS. How cells sense extracellular matrix stiffness: a material's perspective. *Curr Opin Biotechnol*. 2013;24(5):948-953.
599. Li X, Sun Q, Li Q, Kawazoe N, Chen G. Functional Hydrogels With Tunable Structures and Properties for Tissue Engineering Applications. *Front Chem*. 2018;6.
600. Caliarì SR, Burdick JA. A practical guide to hydrogels for cell culture. *Nat Methods*. 2016;13(5):405-414.
601. Trappmann B, Gautrot JE, Connelly JT, et al. Extracellular-matrix tethering regulates stem-cell fate. *Nat Mater*. 2012;11(7):642-649.
602. Kadow CE, Georges PC, Janmey PA, Beningo KA. Polyacrylamide Hydrogels for Cell Mechanics: Steps Toward Optimization and Alternative Uses. In: ; 2007:29-46.
603. Beckerle MC. Zyxin: Zinc fingers at sites of cell adhesion. *BioEssays*. 1997;19(11):949-957.

604. Crawford AW, Beckerle MC. Purification and characterization of zyxin, an 82,000-dalton component of adherens junctions. *J Biol Chem.* 1991;266(9):5847-5853.
605. Drees B, Friederich E, Fradelizi J, Louvard D, Beckerle MC, Golsteyn RM. Characterization of the Interaction between Zyxin and Members of the Ena/Vasodilator-stimulated Phosphoprotein Family of Proteins. *Journal of Biological Chemistry.* 2000;275(29):22503-22511.
606. Hirata H, Tatsumi H, Sokabe M. Zyxin emerges as a key player in the mechanotransduction at cell adhesive structures. *Commun Integr Biol.* 2008;1(2):192-195.
607. Nix DA, Beckerle MC. Nuclear–Cytoplasmic Shuttling of the Focal Contact Protein, Zyxin: A Potential Mechanism for Communication between Sites of Cell Adhesion and the Nucleus. *J Cell Biol.* 1997;138(5):1139-1147.
608. Nix DA, Fradelizi J, Bockholt S, et al. Targeting of Zyxin to Sites of Actin Membrane Interaction and to the Nucleus. *Journal of Biological Chemistry.* 2001;276(37):34759-34767.
609. Wójtowicz A, Babu SS, Li L, Gretz N, Hecker M, Cattaruzza M. Zyxin Mediation of Stretch-Induced Gene Expression in Human Endothelial Cells. *Circ Res.* 2010;107(7):898-902.
610. Sabino F, Madzharova E, auf dem Keller U. Cell density-dependent proteolysis by HtrA1 induces translocation of zyxin to the nucleus and increased cell survival. *Cell Death Dis.* 2020;11(8):674.
611. Valastyan S, Weinberg RA. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell.* 2011;147(2):275-292.

612. YAMAMURA M, NOGUCHI K, NAKANO Y, et al. Functional analysis of Zyxin in cell migration and invasive potential of oral squamous cell carcinoma cells. *Int J Oncol*. 2013;42(3):873-880.
613. Zhou J, Zeng Y, Cui L, et al. Zyxin promotes colon cancer tumorigenesis in a mitotic phosphorylation-dependent manner and through CDK8-mediated YAP activation. *Proceedings of the National Academy of Sciences*. 2018;115(29).
614. Wen XM, Luo T, Jiang Y, et al. Zyxin (ZYX) promotes invasion and acts as a biomarker for aggressive phenotypes of human glioblastoma multiforme. *Laboratory Investigation*. 2020;100(6):812-823.
615. Hoffman LM, Jensen CC, Kloeker S, Wang CLA, Yoshigi M, Beckerle MC. Genetic ablation of zyxin causes Mena/VASP mislocalization, increased motility, and deficits in actin remodeling. *J Cell Biol*. 2006;172(5):771-782.
616. Wei Z, Xia K, Zhou B, Zheng D, Guo W. Zyxin Inhibits the Proliferation, Migration, and Invasion of Osteosarcoma via Rap1-Mediated Inhibition of the MEK/ERK Signaling Pathway. *Biomedicines*. 2023;11(8):2314.
617. Zhong C, Yu J, Li D, et al. Zyxin as a potential cancer prognostic marker promotes the proliferation and metastasis of colorectal cancer cells. *J Cell Physiol*. 2019;234(9):15775-15789.
618. Wen XM, Luo T, Jiang Y, et al. Zyxin (ZYX) promotes invasion and acts as a biomarker for aggressive phenotypes of human glioblastoma multiforme. *Laboratory Investigation*. 2020;100(6):812-823.

619. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* 2009;583(24):3966-3973.
620. Hirata H, Tatsumi H, Sokabe M. Zyxin emerges as a key player in the mechanotransduction at cell adhesive structures. *Commun Integr Biol.* 2008;1(2):192-195.
621. Legate KR, Montañez E, Kudlacek O, Füssler R. ILK, PINCH and parvin: the tIPP of integrin signalling. *Nat Rev Mol Cell Biol.* 2006;7(1):20-31.
622. Stanchi F, Grashoff C, Nguemeni Yonga CF, Grall D, Fässler R, Van Obberghen-Schilling E. Molecular dissection of the ILK-PINCH-parvin triad reveals a fundamental role for the ILK kinase domain in the late stages of focal-adhesion maturation. *J Cell Sci.* 2009;122(11):1800-1811.
623. DeMali KA, Burridge K. Coupling membrane protrusion and cell adhesion. *J Cell Sci.* 2003;116(12):2389-2397.
624. Rane CK, Minden A. P21 activated kinases: structure, regulation, and functions. *Small GTPases.* 2014;5.
625. Chan PM, Manser E. PAKs in Human Disease. In: ; 2012:171-187.
626. Sanders LC, Matsumura F, Bokoch GM, de Lanerolle P. Inhibition of Myosin Light Chain Kinase by p21-Activated Kinase. *Science (1979).* 1999;283(5410):2083-2085.
627. Goeckeler ZM, Masaracchia RA, Zeng Q, Chew TL, Gallagher P, Wysolmerski RB. Phosphorylation of Myosin Light Chain Kinase by p21-activated Kinase PAK2. *Journal of Biological Chemistry.* 2000;275(24):18366-18374.

628. González-Forero D, Montero F, García-Morales V, et al. Endogenous Rho-Kinase Signaling Maintains Synaptic Strength by Stabilizing the Size of the Readily Releasable Pool of Synaptic Vesicles. *The Journal of Neuroscience*. 2012;32(1):68-84.
629. Murrell M, Oakes PW, Lenz M, Gardel ML. Forcing cells into shape: the mechanics of actomyosin contractility. *Nat Rev Mol Cell Biol*. 2015;16(8):486-498.
630. Pandya P, Orgaz JL, Sanz-Moreno V. Modes of invasion during tumour dissemination. *Mol Oncol*. 2017;11(1):5-27.
631. Zaidel-Bar R, Zhenhuan G, Luxenburg C. The contractome – a systems view of actomyosin contractility in non-muscle cells. *J Cell Sci*. 2015;128(12):2209-2217.
632. Robertson J, Jacquemet G, Byron A, et al. Defining the phospho-adhesome through the phosphoproteomic analysis of integrin signalling. *Nat Commun*. 2015;6(1):6265.
633. Humphries JD, Byron A, Bass MD, et al. Proteomic Analysis of Integrin-Associated Complexes Identifies RCC2 as a Dual Regulator of Rac1 and Arf6. *Sci Signal*. 2009;2(87).
634. Kuo JC, Han X, Hsiao CT, Yates III JR, Waterman CM. Analysis of the myosin-II-responsive focal adhesion proteome reveals a role for β -Pix in negative regulation of focal adhesion maturation. *Nat Cell Biol*. 2011;13(4):383-393.
635. Panetti TS. Tyrosine phosphorylation of paxillin FAK AND p130CAS effects on cell spreading and migration. *Frontiers in Bioscience*. 2002;7(4):A771.

636. Iyer V V., Ballestrem C, Kirchner J, Geiger B, Schaller MD. Measurement of Protein Tyrosine Phosphorylation in Cell Adhesion. In: *Cell Migration*. Humana Press; :289-302.
637. Mise N, Savai R, Yu H, Schwarz J, Kaminski N, Eickelberg O. Zyxin Is a Transforming Growth Factor- β (TGF- β)/Smad3 Target Gene That Regulates Lung Cancer Cell Motility via Integrin $\alpha 5\beta 1$. *Journal of Biological Chemistry*. 2012;287(37):31393-31405.
638. Ma B, Cheng H, Gao R, et al. Zyxin-Siah2–Lats2 axis mediates cooperation between Hippo and TGF- β signalling pathways. *Nat Commun*. 2016;7(1):11123.
639. Zhong C, Yu J, Li D, et al. Zyxin as a potential cancer prognostic marker promotes the proliferation and metastasis of colorectal cancer cells. *J Cell Physiol*. 2019;234(9):15775-15789.
640. Pisters PW, Leung DH, Woodruff J, Shi W, Brennan MF. Analysis of prognostic factors in 1,041 patients with localized soft tissue sarcomas of the extremities. *Journal of Clinical Oncology*. 1996;14(5):1679-1689.
641. Simon RM, Paik S, Hayes DF. Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers. *JNCI Journal of the National Cancer Institute*. 2009;101(21):1446-1452.
642. Burns J, Wilding CP, L Jones R, H Huang P. Proteomic research in sarcomas - current status and future opportunities. *Semin Cancer Biol*. 2020;61:56-70.
643. Heaney LM, Jones DJ, Suzuki T. Mass spectrometry in medicine: a technology for the future? *Future Sci OA*. 2017;3(3):FSO213.
644. Stanton H, Melrose J, Little CB, Fosang AJ. Proteoglycan degradation by the ADAMTS family of proteinases. *Biochimica et*

Biophysica Acta (BBA) - Molecular Basis of Disease.
2011;1812(12):1616-1629.

645. Israel Vlodavsky, Neta Ilan, Annamaria Naggi, Benito Casu. Heparanase: Structure, Biological Functions, and Inhibition by Heparin-Derived Mimetics of Heparan Sulfate. *Curr Pharm Des.* 2007;13(20):2057-2073.
646. Cabral-Pacheco GA, Garza-Veloz I, Castruita-De la Rosa C, et al. The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. *Int J Mol Sci.* 2020;21(24):9739.
647. Smith HW, Marshall CJ. Regulation of cell signalling by uPAR. *Nat Rev Mol Cell Biol.* 2010;11(1):23-36.
648. Kessenbrock K, Plaks V, Werb Z. Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. *Cell.* 2010;141(1):52-67.
649. Gobin E, Bagwell K, Wagner J, et al. A pan-cancer perspective of matrix metalloproteases (MMP) gene expression profile and their diagnostic/prognostic potential. *BMC Cancer.* 2019;19(1):581.
650. Gonzalez-Molina J, Gramolelli S, Liao Z, Carlson JW, Ojala PM, Lehti K. MMP14 in Sarcoma: A Regulator of Tumor Microenvironment Communication in Connective Tissues. *Cells.* 2019;8(9):991.
651. Remacle AG, Cieplak P, Nam DH, Shiryaev SA, Ge X, Strongin AY. Selective function-blocking monoclonal human antibody highlights the important role of membrane type-1 matrix metalloproteinase (MT1-MMP) in metastasis. *Oncotarget.* 2017;8(2):2781-2799.

652. Cepeda MA, Evered CL, Pelling JJH., Damjanovski S. Inhibition of MT1-MMP proteolytic function and ERK1/2 signalling influences cell migration and invasion through changes in MMP-2 and MMP-9 levels. *J Cell Commun Signal*. 2017;11(2):167-179.
653. Roebuck MM, Helliwell TR, Chaudhry IH, et al. Matrix Metalloproteinase Expression Is Related to Angiogenesis and Histologic Grade in Spindle Cell Soft Tissue Neoplasms of the Extremities. *Am J Clin Pathol*. 2005;123(3):405-414.
654. Yan T, Lin Z, Jiang J, et al. MMP14 regulates cell migration and invasion through epithelial-mesenchymal transition in nasopharyngeal carcinoma. *Am J Transl Res*. 2015;7(5):950-958.
655. Zarrabi K, Dufour A, Li J, et al. Inhibition of Matrix Metalloproteinase 14 (MMP-14)-mediated Cancer Cell Migration. *Journal of Biological Chemistry*. 2011;286(38):33167-33177.
656. Dodd RD, Añó L, Blum JM, Li Z, Van Mater D, Kirsch DG. Methods to Generate Genetically Engineered Mouse Models of Soft Tissue Sarcoma. In: ; 2015:283-295.
657. Gutierrez WR, Scherer A, McGivney GR, et al. Divergent immune landscapes of primary and syngeneic Kras-driven mouse tumor models. *Sci Rep*. 2021;11(1):1098.
658. Baker BM, Chen CS. Deconstructing the third dimension – how 3D culture microenvironments alter cellular cues. *J Cell Sci*. Published online January 1, 2012.
659. Cukierman E, Pankov R, Stevens DR, Yamada KM. Taking Cell-Matrix Adhesions to the Third Dimension. *Science* (1979). 2001;294(5547):1708-1712.

660. Geraldo S, Simon A, Elkhatib N, Louvard D, Fetler L, Vignjevic DM. Do cancer cells have distinct adhesions in 3D collagen matrices and in vivo? *Eur J Cell Biol.* 2012;91(11-12):930-937.
661. Hakkinen KM, Harunaga JS, Doyle AD, Yamada KM. Direct Comparisons of the Morphology, Migration, Cell Adhesions, and Actin Cytoskeleton of Fibroblasts in Four Different Three-Dimensional Extracellular Matrices. *Tissue Eng Part A.* 2011;17(5-6):713-724.
662. Li S, Lao J, Chen BPC, et al. Genomic analysis of smooth muscle cells in three-dimensional collagen matrix. *The FASEB Journal.* 2003;17(1):97-99.
663. Harunaga JS, Yamada KM. Cell-matrix adhesions in 3D. *Matrix Biology.* 2011;30(7-8):363-368.
664. Ligorio C, Mata A. Synthetic extracellular matrices with function-encoding peptides. *Nature Reviews Bioengineering.* 2023;1(7):518-536.
665. Galarza S, Crosby AJ, Pak C, Peyton SR. Control of Astrocyte Quiescence and Activation in a Synthetic Brain Hydrogel. *Adv Healthc Mater.* 2020;9(4).
666. Jansen LE, Kim H, Hall CL, McCarthy TP, Lee MJ, Peyton SR. A poly(ethylene glycol) three-dimensional bone marrow hydrogel. *Biomaterials.* 2022;280:121270.
667. Salinas CN, Anseth KS. The influence of the RGD peptide motif and its contextual presentation in PEG gels on human mesenchymal stem cell viability. *J Tissue Eng Regen Med.* 2008;2(5):296-304.

668. Hersel U, Dahmen C, Kessler H. RGD modified polymers: biomaterials for stimulated cell adhesion and beyond. *Biomaterials*. 2003;24(24):4385-4415.
669. Gjorevski N, Sachs N, Manfrin A, et al. Designer matrices for intestinal stem cell and organoid culture. *Nature*. 2016;539(7630):560-564.
670. Xie AW, Murphy WL. Engineered biomaterials to mitigate growth factor cost in cell biomanufacturing. *Curr Opin Biomed Eng*. 2019;10:1-10.
671. Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. *Nat Rev Cancer*. 2004;4(6):448-456.
672. Li Y, Lu Q, Xie C, Yu Y, Zhang A. Recent advances on development of p21-activated kinase 4 inhibitors as anti-tumor agents. *Front Pharmacol*. 2022;13.
673. Semenova G, Chernoff J. Targeting PAK1. *Biochem Soc Trans*. 2017;45(1):79-88.
674. Mileschkin LR, Rosen LS, Blumenkopf T, et al. Phase I, dose-escalation, safety, pharmacokinetic, and pharmacodynamic study of single-agent PF-03758309, an oral PAK inhibitor, in patients with advanced solid tumors. *Journal of Clinical Oncology*. 2011;29(15_suppl):e13607-e13607.
675. Békés M, Langley DR, Crews CM. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov*. 2022;21(3):181-200.

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Dec 02, 2023

This Agreement between Valeriya Pankova ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5678291070895
License date	Nov 29, 2023
Licensed Content Publisher	Elsevier
Licensed Content Publication	Current Biology
Licensed Content Title	Basement membranes
Licensed Content Author	Ranjay Jayadev,David R. Sherwood
Licensed Content Date	Mar 20, 2017
Licensed Content Volume	27
Licensed Content Issue	6
Licensed Content Pages	5
Start Page	R207
End Page	R211
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations

Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of new work	Deconstructing the soft tissue sarcoma matrisome and adhesome for drug target and biomarker discovery
Institution name	Institute of Cancer Research
Expected presentation date	Dec 2023
Portions	Figure 1
	Valeriya Pankova 15 cotswold road
Requestor Location	Sutton, SM2 5NG United Kingdom Attn: Valeriya Pankova
Publisher Tax ID	GB 494 6272 12
Total	0.00 GBP
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your RightsLink account and that are available at any time at <https://myaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given. The material may not be reproduced or used in any other way, including use in combination with an artificial intelligence tool (including to train an algorithm, test, process, analyse, generate output and/or develop any form of artificial intelligence tool), or to create any derivative work and/or service (including resulting from the use of artificial intelligence tools).

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk [here](#)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. **No Transfer of License:** This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. **No Amendment Except in Writing:** This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. **Objection to Contrary Terms:** Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. **Revocation:** Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
 - via their non-commercial person homepage or blog
 - by updating a preprint in arXiv or RePEc with the accepted manuscript
 - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
 - directly by providing copies to their students or to research collaborators for their personal use
 - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
 - via non-commercial hosting platforms such as their institutional repository
 - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. For book authors the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not

represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.10

Questions? customercare@copyright.com.

SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Dec 02, 2023

This Agreement between Valeriya Pankova ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5680711364324
License date	Dec 02, 2023
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Protocols
Licensed Content Title	Quantitative and unbiased analysis of directional persistence in cell migration
Licensed Content Author	Roman Gorelik et al
Licensed Content Date	Jul 17, 2014
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Would you like a high resolution image with your order?	no

Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title of new work	Deconstructing the soft tissue sarcoma matrisome and adhesome for drug target and biomarker discovery
Institution name	Institute of Cancer Research
Expected presentation date	Dec 2023
Portions	Figure 1a
Requestor Location	Valeriya Pankova 15 cotswold road Sutton, SM2 5NG United Kingdom Attn: Valeriya Pankova
Total	0.00 USD

Terms and Conditions

Springer Nature Customer Service Centre GmbH Terms and Conditions

The following terms and conditions ("Terms and Conditions") together with the terms specified in your [RightsLink] constitute the License ("License") between you as Licensee and Springer Nature Customer Service Centre GmbH as Licensor. By clicking 'accept' and completing the transaction for your use of the material ("Licensed Material"), you confirm your acceptance of and obligation to be bound by these Terms and Conditions.

1. Grant and Scope of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, non-sublicensable, revocable, world-wide License to reproduce, distribute, communicate to the public, make available, broadcast, electronically transmit or create derivative works using the Licensed Material for the purpose(s) specified in your RightsLink Licence Details only. Licenses are granted for the specific use requested in the order and for no other use, subject to these Terms and Conditions. You acknowledge and agree that the rights granted to you under this License do not include the right to modify, edit, translate, include in collective works, or create derivative works of the

Licensed Material in whole or in part unless expressly stated in your RightsLink Licence Details. You may use the Licensed Material only as permitted under this Agreement and will not reproduce, distribute, display, perform, or otherwise use or exploit any Licensed Material in any way, in whole or in part, except as expressly permitted by this License.

1. 2. You may only use the Licensed Content in the manner and to the extent permitted by these Terms and Conditions, by your RightsLink Licence Details and by any applicable laws.

1. 3. A separate license may be required for any additional use of the Licensed Material, e.g. where a license has been purchased for print use only, separate permission must be obtained for electronic re-use. Similarly, a License is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the License.

1. 4. Any content within the Licensed Material that is owned by third parties is expressly excluded from the License.

1. 5. Rights for additional reuses such as custom editions, computer/mobile applications, film or TV reuses and/or any other derivative rights requests require additional permission and may be subject to an additional fee. Please apply to journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

2. Reservation of Rights

Licensor reserves all rights not expressly granted to you under this License. You acknowledge and agree that nothing in this License limits or restricts Licensor's rights in or use of the Licensed Material in any way. Neither this License, nor any act, omission, or statement by Licensor or you, conveys any ownership right to you in any Licensed Material, or to any element or portion thereof. As between Licensor and you, Licensor owns and retains all right, title, and interest in and to the Licensed Material subject to the license granted in Section 1.1. Your permission to use the Licensed Material is expressly conditioned on you not impairing Licensor's or the applicable copyright owner's rights in the Licensed Material in any way.

3. Restrictions on use

3. 1. Minor editing privileges are allowed for adaptations for stylistic purposes or formatting purposes provided such alterations do not alter the original meaning or intention of the Licensed Material and the new figure(s) are still accurate and representative of the Licensed Material. Any other changes including but not limited to, cropping, adapting, and/or omitting material that affect the meaning, intention or moral rights of the author(s) are strictly prohibited.

3. 2. You must not use any Licensed Material as part of any design or trademark.

3. 3. Licensed Material may be used in Open Access Publications (OAP), but any such reuse must include a clear acknowledgment of this permission visible at the same time as the figures/tables/illustration or abstract and which must indicate that the Licensed Material is not part of the governing OA license but has been reproduced with permission. This may be indicated according to any standard referencing system but must include at a minimum 'Book/Journal title, Author, Journal Name (if applicable), Volume (if applicable), Publisher, Year, reproduced with permission from SNCSC'.

4. STM Permission Guidelines

4. 1. An alternative scope of license may apply to signatories of the STM Permissions Guidelines ("STM PG") as amended from time to time and made available at <https://www.stm-assoc.org/intellectual-property/permissions/permissions-guidelines/>.

4. 2. For content reuse requests that qualify for permission under the STM PG, and which may be updated from time to time, the STM PG supersede the terms and conditions contained in this License.

4. 3. If a License has been granted under the STM PG, but the STM PG no longer apply at the time of publication, further permission must be sought from the Rightsholder. Contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

5. Duration of License

5. 1. Unless otherwise indicated on your License, a License is valid from the date of purchase ("License Date") until the end of the relevant period in the below table:

Reuse in a medical communications project	Reuse up to distribution or time period indicated in License
Reuse in a dissertation/thesis	Lifetime of thesis
Reuse in a journal/magazine	Lifetime of journal/magazine
Reuse in a book/textbook	Lifetime of edition
Reuse on a website	1 year unless otherwise specified in the License
Reuse in a presentation/slide kit/poster	Lifetime of presentation/slide kit/poster. Note: publication whether electronic or in print of presentation/slide kit/poster may require further permission.
Reuse in conference proceedings	Lifetime of conference proceedings
Reuse in an annual report	Lifetime of annual report
Reuse in training/CME materials	Reuse up to distribution or time period indicated in License
Reuse in newsmedia	Lifetime of newsmedia
Reuse in coursepack/classroom materials	Reuse up to distribution and/or time period indicated in license

6. Acknowledgement

6. 1. The Licensor's permission must be acknowledged next to the Licensed Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract and must be hyperlinked to the journal/book's homepage.

6. 2. Acknowledgement may be provided according to any standard referencing system and at a minimum should include "Author, Article/Book Title, Journal name/Book imprint, volume, page number, year, Springer Nature".

7. Reuse in a dissertation or thesis

7. 1. Where 'reuse in a dissertation/thesis' has been selected, the following terms apply: Print rights of the Version of Record are provided for; electronic rights for use only on institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/) and only up to what is required by the awarding institution.

7. 2. For theses published under an ISBN or ISSN, separate permission is required. Please contact journalpermissions@springernature.com or bookpermissions@springernature.com for these rights.

7. 3. Authors must properly cite the published manuscript in their thesis according to current citation standards and include the following acknowledgement: *'Reproduced with permission from Springer Nature'*.

8. License Fee

You must pay the fee set forth in the License Agreement (the "License Fees"). All amounts payable by you under this License are exclusive of any sales, use, withholding, value added or similar taxes, government fees or levies or other assessments. Collection and/or remittance of such taxes to the relevant tax authority shall be the responsibility of the party who has the legal obligation to do so.

9. Warranty

9. 1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. **You are solely responsible for ensuring that the material you wish to license is original to the Licensor and does not carry the copyright of another entity or third party (as credited in the published version).** If the credit line on any part of the Licensed Material indicates that it was reprinted or adapted with permission from another source, then you should seek additional permission from that source to reuse the material.

9. 2. EXCEPT FOR THE EXPRESS WARRANTY STATED HEREIN AND TO THE EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR PROVIDES THE LICENSED MATERIAL "AS IS" AND MAKES NO OTHER REPRESENTATION OR WARRANTY. LICENSOR EXPRESSLY DISCLAIMS ANY LIABILITY FOR ANY CLAIM ARISING FROM OR OUT OF THE CONTENT, INCLUDING BUT NOT LIMITED TO ANY ERRORS, INACCURACIES, OMISSIONS, OR DEFECTS CONTAINED THEREIN, AND ANY IMPLIED OR EXPRESS WARRANTY AS TO MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, PUNITIVE, OR EXEMPLARY DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE LICENSED MATERIAL REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION APPLIES NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

10. Termination and Cancellation

10. 1. The License and all rights granted hereunder will continue until the end of the applicable period shown in Clause 5.1 above. Thereafter, this license will be terminated and all rights granted hereunder will cease.

10. 2. Licensor reserves the right to terminate the License in the event that payment is not received in full or if you breach the terms of this License.

11. General

11. 1. The License and the rights and obligations of the parties hereto shall be construed, interpreted and determined in accordance with the laws of the Federal Republic of Germany without reference to the stipulations of the CISG (United Nations Convention on Contracts for the International Sale of Goods) or to Germany's choice-of-law principle.

11. 2. The parties acknowledge and agree that any controversies and disputes arising out of this License shall be decided exclusively by the courts of or having jurisdiction for Heidelberg, Germany, as far as legally permissible.

11. 3. This License is solely for Licensor's and Licensee's benefit. It is not for the benefit of any other person or entity.

Questions? For questions on Copyright Clearance Center accounts or website issues please contact springernaturesupport@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777. For questions on Springer Nature licensing please visit <https://www.springernature.com/gp/partners/rights-permissions-third-party-distribution>

Other Conditions:

Version 1.4 - Dec 2022

Questions? customer care@copyright.com.
