The Role of TWIST1 in Multiple Myeloma

Chee Man Cheong, BHSc (Hons)

Myeloma Research Group

Discipline of Physiology

Adelaide Medical School

Faculty of Health and Medical Sciences

The University of Adelaide

&

Cancer Theme

South Australian Health & Medical Research Institute (SAHMRI)

A thesis submitted for the degree of Doctor of Philosophy

December 2016

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 5

LIST OF TABLES

Chapter 4

ABSTRACT

Multiple myeloma (MM) is an incurable haematological malignancy characterised by the uncontrolled proliferation of clonal plasma cells (PCs) in the bone marrow (BM). MM disease progression relies on the continuous trafficking of MM PCs to distant BM sites leading to multiple focal tumours throughout the skeleton at diagnosis. While metastasis is not a concept generally applied to haematological malignancies, increasing evidence suggests that like solid tumours, an epithelial-to-mesenchymal (EMT)-like process is activated in MM PCs in order to disseminate. However, an association between high-risk MM subtypes, patient prognosis and the expression of an EMT-like gene expression signature in MM PCs had not been reported.

The chromosomal translocation t(4;14), characterised by overexpression of *MMSET* and *FGFR3*, is associated with poor prognosis and aggressive tumour dissemination. This project sought to determine whether the highly aggressive phenotype observed in t(4;14) is mediated by activation of EMT-like process. Using previously published microarray data from large cohorts of newly diagnosed MM patients and RNA-sequencing data from human MM cell lines, an EMT-like expression signature in t(4;14) MM was comprehensively evaluated. Among the mesenchymal genes identified, *TWIST1* was consistently upregulated in approximately 50% of the t(4;14) MM cases and was positively correlated with expression of *MMSET*.

Using RNA-sequencing technology, the transcriptome-wide effects of TWIST1 overexpression in MM PCs were determined. Furthermore, ectopic expression of TWIST1 was found to enhance MM PC migration *in vitro*, consistent with the enrichment of genes involved in cell motility from Gene Ontology (GO) analysis. To evaluate the role of TWIST1 overexpression *in vivo*, the 5TGM1/KaLwRij murine model of myeloma was utilised. TWIST1 overexpression in 5TGM1 cells increased total tumour burden and extramedullary growth of tumour in the spleens of recipient mice. Transcriptome analysis demonstrated that TWIST1 overexpression in 5TGM1 cells led to overexpression of genes involved in cytokine production, regulation of cell death and cell motility.

These studies highlight that TWIST1, downstream of MMSET promotes MM PC motility, which, in part contributes to the aggressive phenotype of t(4;14) MM patients. Taken together, this research reveals the role of TWIST1 in MM pathogenesis and adds to the current knowledge of mechanism underlying aggressive disease progression in t(4;14) patients.

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed:

……

Chee Man CHEONG Date:

ACKNOWLEDGEMENTS

Although there were times where I think finishing PhD seems like an impossible task, I would not have accomplished this thesis without the help and support I received throughout my PhD journey.

First, I would like to thank my supervisors Prof Andrew Zannettino and Dr Kate Vandyke for all the help, encouragement and support ever since my Honours year study. Thank you for your faith in me even at times when I doubted myself. Thank you Andrew for proof reading my thesis promptly especially near the end of submission. Thank you Kate for conceptualising this project and initiated my path on bioinformatics. I would also like to acknowledge Adelaide Graduate Research Scholarship and School of Medicine Short-Term Scholarship for financial assistant during my PhD study.

Thanks to all the members, both past and present, of the Myeloma Research Group and Mesenchymal Stem Cell Group who have helped and encouraged me over the years. You have made working within the group a wonderful experience. Special thanks Dr Duncan Hewett for his critical and attentive evaluation on my work, Dr Jacqueline Noll for her help in tail vein injection, Dr Stephen Fitter for his invaluable molecular expertise and advises (both professional and personal) and Dr Sally Martin for her assistance in histology. A special mention to Krzysztof Mrozik, Ankit Dutta and Jia Qi Ng for their companionship during out-of-hour laboratory work, coffee/lunch chats and friendship. They have broadened my view in scientific research and brighten up my PhD experience.

Many other people in SAHMRI have also assisted in this project. Thank you Dr Chung Hoow Kok for his expertise and advice on bioinformatic analysis, Manuel Bernal-Llinares for sharing his knowledge in computing and assistance in assessing the SAHMRI High-Performance Computing Cluster, Dr Randall Grose for FACS sorting the cell lines for my project, Mark van der Hoek for preparing cDNA libraries and running my RNA-seq, Carly Brune and the staff at SAHMRI Bioresources for providing training and taking care of the mice.

Finally, on a personal note, thanks to my parents for their love and endless support given me throughout pursuits in Adelaide. Most importantly, I would like to thank my partner, Soo Siang Ooi for his never-ending patience, companion, care and love throughout this journey.

Thank you!

LIST OF ABBREVIATIONS

5TGM1-EV	5TGM1-pLEGO-IRES-tdTomato2
5TGM1-TWIST1	5TGM1-pLEGO-IRES-tdTomato2-TWIST1
aCGH	array comparative genomic hybridization
bHLH	basic helix-loop-helix
BM	bone marrow
BMEC	bone marrow endothelial cell
BMSC	bone marrow stromal cell
bp	base pair
BSA	bovine serum albumin
cDNA	complimentary DNA
CM	conditioned media
CSR	class switch recombination
CXCL	CXC chemokine ligand
CXCR	CXC chemokine receptor
DEG	differentially expressed genes
DMEM	Dulbecco's modified eadle medium
ECF	enhanced chemifluorescence
ECM	extracellular matrix
EDTA	ethylenediaminetetra-acetic acid
EMD	extramedullary myeloma disease
EMT	epithelial-to-mesenchymal transition
FACS	fluorescence-activated cell sorting
FCS	fetal calves serum
FGFR3	fibroblast growth factor receptor 3
g/ mL/ mm/ mM	mmilligram/ millilitre/ millimetre/ millimolar
GAPDH	glutaraldehyde 3-phosphate dehydrogenase
GC	germinal centre
GEP	gene expression profile
GFP	green fluorescence protein

GSEA	gene set enrichment analysis
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMCL	human myeloma cell line
IGH	immunoglobulin heavy locus
IMDM	Iscove's Modified Dulbecco's Medium
IMiD	immunomodulatory drugs
IRES	internal ribosome entry site
ISS	International Staging System
kDa	kiloDalton
LDH	lactose dehydrogenase
M	molar
MET	mesenchymal-to-epithelial transition
MFI	mean fluroscence intensity
MGUS	monoclonal gammopathy of undetermined significance
MM	multiple myeloma
MMP	matrix metalloproteinase
MMSET	MM SET domain
mRNA	messenger RNA
mRNA	messenger ribonucleic acid
nm	nanometres
PBS	phosphate bufferred saline
PBS	phosphate buffered saline
PC	plasma cell
PCL	plasma cell leukemia
R-ISS	revised-International Staging System
RNA	ribonucleic acid
RPMI	Roswell Park Memorial Institute
RPMI8226-EV	RPMI8226-pRetroX-DsRed
RPMI8226-MMSET	RPMI8226-pRetroX-DsRed-MMSET
RPMI8226-MMSET-Y1118A	RPMI8226-pRetroX-DsRed-MMSET-Y1118A

RT-qPCR	reverse transcription-quantitative polymerase chain reaction
SD	standrad deviation
SDS-PAGE	sodium dodecyl sulfate polycrylamide gel electrophoresis
SEM	standard error of the mean
SHM	somatic hypermutation
siRNA	small interfering ribonucleic acid
SMM	smouldering multiple myeloma
SNP	single nucleotide polymorphism
Tween 20	polyethylene glycol sorbitan monolaurate
TWIST1	Twist family basic helix-loop-helix transcription factor 1
w/v	weight per volume
WES	whole exome sequencing
WHSC1	Wolf-Hirschhorn syndrome candidate 1
WL2-EV	WL2-pRUF-IRES-GFP
WL2-TWIST1	WL2-pRUF-IRES-GFP-TWIST1
WST-1	4-[3-(4-Iodophenyl)-2-(4-nitrophenyl)-2H-5-tetrazolio]-1,3-benzene disulphonate
x g	times gravity

LIST OF PUBLICATIONS

Scientific manuscripts

1. **Cheong CM**, Vandyke K, Mrozik KM, Kok CH, To LB, Licht JD, Zannettinno ACW. MMSET promotes the acquisition of an epithelial-to-mesenchymal-like gene expression signature in t(4;14) multiple myeloma (2017). *Manuscript in preparation.*
2. **Cheong CM**, Mrozik KM, Kok CH, Vandyke K, Hewett DR, Noll JE, Fitter S, Zannettino ACW. TWIST1 promotes tumour migration and dissemination in multiple myeloma (2017). *Manuscript in preparation.*
3. Vandyke K, Zeissig MN, Hewett DR, Martin SK, Mrozik KM, **Cheong CM**, Diamond P, To LB, Gronthos S, Peet DJ, Croucher PI, Zannettino ACW Hypoxia Inducible Factor 2 alpha (HIF-2α) drives plasma 1 cell dissemination in multiple myeloma by regulating CXCL12/CXCR4 and CCR1 (2016). *Submitted.*
4. Mrozik KM, **Cheong CM**, Hewett D, Chow AW, Blaschuk OW, Zannettino ACW, Vandyke K. Therapeutic targeting of N-cadherin is an effective treatment for multiple myeloma. *Br J Haematol* 2015 Nov; **171**(3)**:** 387-399
5. **Cheong CM**, Chow AW, Fitter S, Hewett DR, Martin SK, Williams SA, To LB, Zannettino ACW, Vandyke K. Tetraspanin 7 (TSPAN7) expression is upregulated in multiple myeloma patients and inhibits myeloma tumour development in vivo. *Exp Cell Res* 2015 Mar 1; **332**(1)**:** 24-38.

Conference proceedings

1. 10th Florey International Postgraduate Research Conference, Adelaide, Australia, September 2016. *TWIST1 promotes tumour progression and metastasis in multiple myeloma via cytoskeleton remodeling*. Poster. Awarded the John Barker Prize for Cancer Research
2. 2016 Australian Society for Medical Research (ASMR) SA Annual Scientific Meeting, Adelaide, Australia, June 2016. *RNA-sequencing reveals pathways regulated by TWIST1 to promote cell migration in multiple myeloma plasma cells.* Oral presentation.

3. TEMTIA-VII The EMT International Association Conference, Melbourne, Australia, October 2015. Meta-analysis of microarray datasets identifies EMT-like gene expression signature in t(4;14) multiple myeloma patients. Poster.
4. 2015 SAHMRI Research Showcase, Adelaide, Australia, September 2015. Meta-analysis of microarray datasets identifies EMT-like gene expression signature in t(4;14) multiple myeloma patients. Poster.
5. 2015 Australian Society for Medical Research (ASMR) SA Annual Scientific Meeting, Adelaide, Australia, June 2015. *Meta-analysis of microarray datasets identifies EMT-like gene expression signature in t(4;14) multiple myeloma patients.* Poster.
6. 2014 Annual Scientific Meetings of the HAA (Haematology Society of Australia and New Zealand, Australian & New Zealand Society of Blood Transfusion and the Australasian Society of Thrombosis and Haemostasis), Perth, Australia 2014. *Meta-analysis of microarray datasets identifies EMT-like gene expression signature in t(4;14) multiple myeloma patients.* Poster.
7. 2014 Australian Society for Medical Research (ASMR) SA Annual Scientific Meeting, Adelaide, Australia, June 2014. Tetraspanin 7 (TSPAN7) inhibits tumour development in vivo and regulates multiple myeloma cell transendothelial migration and adhesion in vitro. Poster.

Chapter 1:

Introduction

1.1 Overview

Multiple myeloma (MM) is an incurable haematological malignancy characterised by the uncontrolled proliferation of clonal plasma cells in the bone marrow (BM). In most cases, multiple focal tumours are evident throughout the skeleton at diagnosis, which result from the continuous trafficking of myeloma cells to distant BM sites via the peripheral blood.

While metastasis is not a concept generally applied to haematological malignancies, increasing evidence suggests that like solid tumours, MM plasma cells employ a metastasis-like process of invasion and dissemination in order to spread. Emerging evidence suggests that an epithelial-to-mesenchymal (EMT)-like process, in which polarized epithelial cells are converted into motile mesenchymal cells, maybe activated in MM PCs and associated with a more aggressive and migratory phenotype.

Recent study by Ezponda et al.[1] has identify a novel role for MM SET domain (MMSET), the key pathogenic histone methyltransferase overexpressed in t(4;14) MM, in driving EMT in prostate cancer via activation of the basic-helix-loop-helix (bHLH) transcription factor, TWIST1. The same study also found elevated TWIST1 expression in t(4;14) MM, which has never been reported in any other previously published gene expression profiling studies.[2-4] Concordant with this, previous findings from our laboratory and others[5, 6] have showed that the mesenchymal marker, N-cadherin, was upregulated in poor prognosis MM patients harbouring the t(4;14) translocation, suggesting involvement of EMT-like activity in t(4;14) MM. Although target genes downstream of MMSET-regulated global chromatin changes have been previously identified[7-9], the interplay between MMSET and EMT-activating transcription factor TWIST1, and the role of TWIST1 in MM progression remained to be determined.

The studies described in this thesis were designed to examine whether upregulation of TWIST1 and N-cadherin in t(4;14) MM is part of an EMT-like expression profile driven by MMSET using publicly available microarray data from large cohorts of newly diagnosed MM patients. Moreover, previously unrecognised roles of TWIST1 in MM disease progression are determined. Altogether, this research adds to the current knowledge of mechanism underlying aggressive disease progression in t(4;14) MM patients.

The introduction section that follows provides an overview of literature relating to (i) the biology of MM, (ii) tumour invasion and metastasis, (iii) dissemination and metastasis in MM, and (iv) role of EMT in MM.

1.2 Multiple Myeloma

Multiple myeloma (MM) is a B-cell malignancy characterised by clonal proliferation of the terminally differentiated, antibody-secreting, long-lived plasma cells (PCs) in the bone marrow (BM). The uncontrolled proliferation of malignant PCs results in aberrant production of non-functional monoclonal paraprotein (M proteins) or light chains detectable in both blood and urine.[10] Symptomatic MM is associated with presence of end-organ damage including hypercalcemia, renal insufficiency, anaemia and lytic bone lesions (referred to as CRAB features) at the time of diagnosis. While the aetiology of MM remains to be fully elucidated, environmental factors, age, race and gender are some of the risk factors associated with an increased propensity to develop MM.[11, 12]

1.2.1 Clinical manifestations of multiple myeloma

Pathogenesis of MM PCs are known to depend heavily on BM microenvironment (reviewed in [13-15]). The clinical features associated with MM result from the progressive accumulation of MM PCs in the BM and the interaction between the MM PC and the cells and synthetic products (cytokines, extracellular matrix proteins) of the bone marrow stroma.[14] The uncontrolled proliferation of clonal PCs crowds out the BM space, disrupting normal marrow function, leading to anaemia, immunosuppression and thrombocytopenia.[10, 16-18] Excessive M proteins secreted by MM PCs is released into the bloodstream, leading to increased plasma volume serum hyperviscosity.[19] M proteins and/or light chains can also be found in the urine and renal tubule reabsorption of the immunoglobulin proteins can contribute to renal failure.[16-18] Bone pain is also present at diagnosis in more than 75% of MM patients and skeletal complications, including lytic bone lesions, pathological fractures, compression fractures and osteoporosis, are common features in MM.[17] Adhesion of MM PCs to the stromal cells induces secretion of various osteoclast-activating factors, promoting osteoclast bone-resorbing activity at the sites of MM PCs infiltration.[14, 15] Elevated bone resorption leads to the release of excessive calcium into the blood, resulting in hypercalcemia, and increased renal burden.[20]

1.2.2 Epidemiology of multiple myeloma

MM is the third most prevalent haematological malignancy after non-Hodgkin's lymphoma and leukaemia, and accounts for 15% of all haematological cancers and 0.8% of all new cancer cases worldwide.[21, 22] Currently, there are nearly 230,000 people around the world living with MM. In Australia, over 1500 new MM cases are diagnosed every year, with the incidence rate of 6.2 new cases per 100,000 populations in 2011.[22] More than 80% of newly diagnosed MM patients are 60 years of age and older and only less than 1% of MM patients are younger than 40 years old.[22, 23] As the worldwide population age due to increasing life expectancy, the incidence of MM is predicted to rise to 390,000 cases in 2025.[23]

MM incidence varies by geographical region and race. Higher incidence is observed in more developed regions, with highest rates in Australia/New Zealand, Northern America and Western Europe (Figure 1.1).[23] This higher incidence of MM in developed countries most likely reflects the heightened awareness, superior surveillance measures and state-of-the art diagnostic techniques. Despite recent advances in cancer therapeutics, MM remains incurable, with a 10-year survival rate of 17%.[24]

1.2.3 Different stages of MM disease progression

Clonal PCs dyscrasias encompass a broad disease spectrum (Figure 1.2).[25] A discrete, solitary mass of malignant MM PCs can be present as a solitary plasmacytoma in either bone or at an extramedullary site. MM is almost always preceded by asymptomatic benign disease state known as monoclonal gammopathy of undetermined significance (MGUS).[26, 27] Diagnosis of MGUS requires detectable, but less than 10% clonal BM PCs , serum monoclonal paraprotein of less than 30 g/L and absence of CRAB features. About 3% of population, aged over 50 years, are estimated to have MGUS.[26, 28-30] Despite being relatively common in the elderly, the rate of MGUS progression to overt MM is approximately 1% per year [31], with a majority of the MGUS patients remaining stable for 20 years or more.[32] Previous studies have identified biological markers such as M protein level, M protein isotype and serum free light chain ratio as predictive risk factors for MGUS progression.[33] However, the precise mechanisms and genetic events leading to MGUS progression are yet to be fully elucidated.

Smouldering MM (SMM) is the intermediate clinical stage between MGUS and overt MM. In SMM patients, the diagnostic criteria for BM PCs activity and serum M

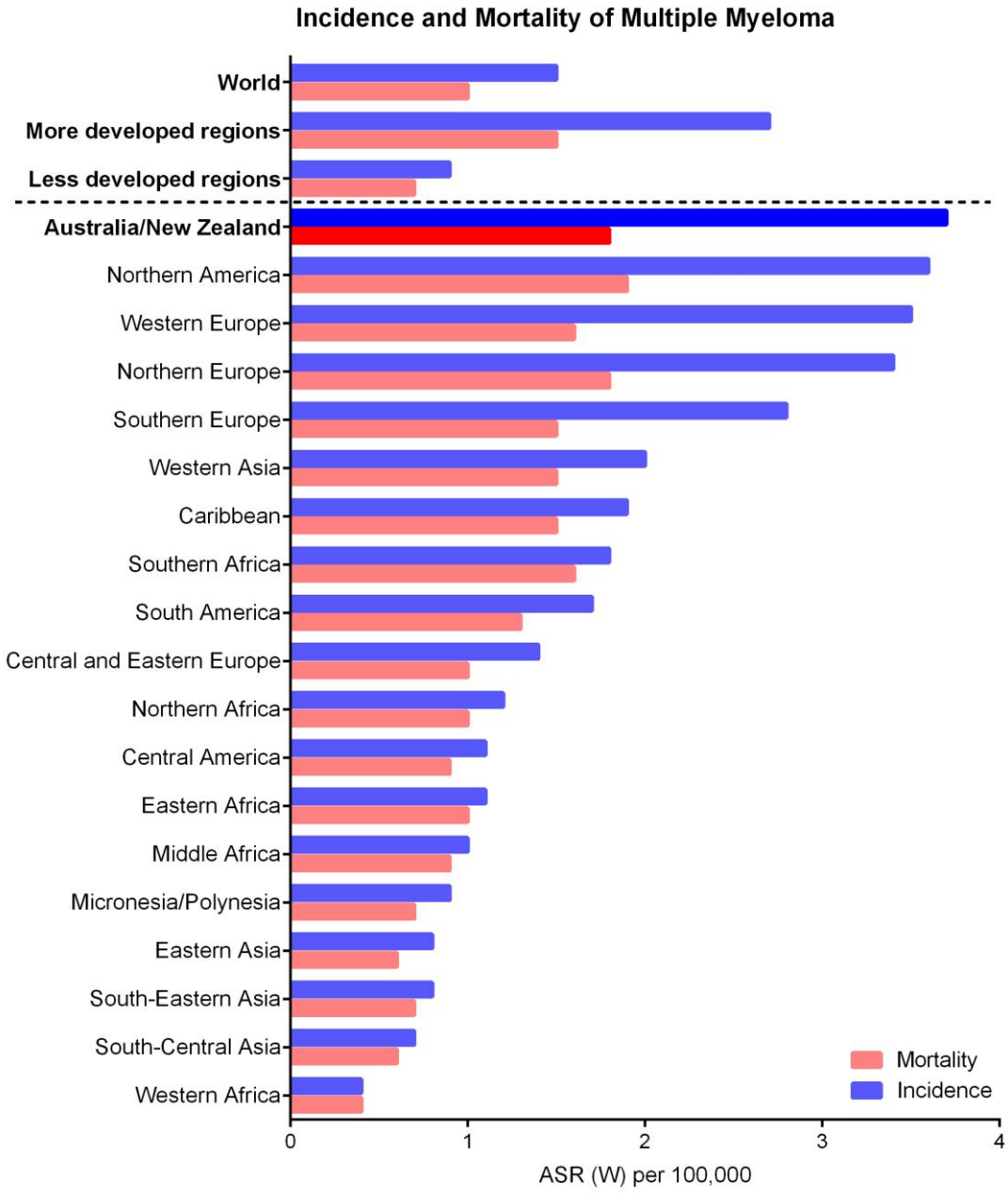


Figure 1.1 Incidence and mortality of multiple myeloma worldwide, includes all age and both gender. Plotted using data from Globocon 2012.[21]

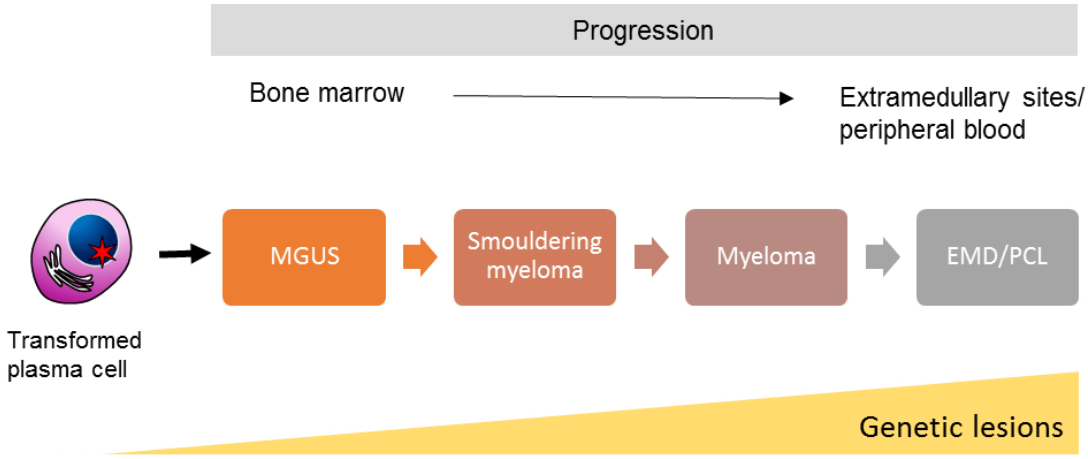


Figure 1.2 Pathogenic progression of multiple myeloma. Multiple myeloma arises from terminally differentiated plasma cells which reside in the BM. Monoclonal gammopathy of undetermined significance (MGUS) is the benign precursor stage of MM, that may progress into smouldering MM, symptomatic MM and the end-stage plasma cell leukemia (PCL) or extramedullary MM disease (EMD) that can grow outside of bone marrow. Accumulation of genetic lesions as MM cells progress through these stages leads to the development of clinical features associate with MM. *Adapted from Morgan et al.*[25]

Table 1.1 Disease spectrum of plasma-cell proliferative disorder. [34, 35]

Plasma-cell disorders	**Definition**
Monoclonal gammopathy of undetermined significance (MGUS)	Serum monoclonal protein: < 30 g/L Clonal BM plasma cells: < 10% Absence of CRAB feature
Smouldering myeloma (SMM)	Serum monoclonal protein: > 30 g/L Clonal BM plasma cells: 10 - 59% Absence of CRAB features
Active MM	Serum monoclonal protein: ≥ 30 g/L (except in non-secretory MM) Clonal BM plasma cells: > 10% (including solitary clonal plasma cells) CRAB features: - Hypercalcemia: serum calcium > 2.75 mmol/L (> 11mg/dL) - Renal failure: creatinine clearance < 40 ml or serum creatinince > 177 μmol/L (> 2 mg/dL) - Anaemia: haemaglobulin level < 100 g/L - Bone lesions: one or more osteolytic lesions on skeletal radiography, CT or PET-CT Any one of following new defining biomarkers (even with the absence of CRAB features): - Clonal BM plasma cells percentage ≥ 60% - Serum FLC- κ/λ ratio ≥ 100 - More than one focal lesion ≥ 5 mm on MRI
Solitary plasmacytoma	Absence of monoclonal protein Solitary lesion of bone or soft tissue with clonal plasma cells Normal bone marrow without clonal plasma cells Normal MRI or CT (except for the solitary lesion) Absence of CRAB features
Plasma cell leukaemia (PCL)	As per definition of active MM with peripheral blood involvement. Peripheral blood plasma cell count 2×10^9/L or > 20% of the differential white blood cell count. Primary PCL: without prior MM diagnosis Secondary PCL: leukaemic transformation of previously diagnosed MM

protein level for MM are met, but the CRAB features are absent. SMM is distinguished from MGUS due to the significant higher risk of progression into the malignant stage in the first 5 years post-diagnosis, at about 10% per year [36], as opposed to 1% per year of MGUS. Prognosis of SMM is heterogeneous, as it encompasses patients with benign disease who have a stable progression rate, as well as patients who develop symptoms and progress within the first 2 year of diagnosis. [37, 38, 39] Currently, treatment is only given to MM patients with evidence of CRAB features attributed to BM PCs proliferation, while MGUS and SMM patients are only monitored for further progression.

In MM cases displaying extramedullary disease (EMD), the monoclonal PCs colonise organs such as liver, lungs, skin and brain, forming macroscopic soft-tissues plasmacytomas and are capable to grow independent of the BM microenvironment. EMD is classified into skeletal EMD and extraosseous EMD, depending on origins and anatomical location of the extramedullary plasmacytomas. Extramedullary plasmacytomas of skeletal EMD originate from direct extension from BM lesions due to disruption of cortical bone and are therefore found in soft tissue surrounding proximal to the axial skeleton. On the other hand, extramedullary plasmacytomas of extraossseous EMD are derived from haematogenous metastasis and are commonly found in lymph nodes, central nervous system, skin and other soft tissue organs.[40, 41] EMD can also present with a leukaemic phase known as plasma cell leukemia (PCL) when more than 20% of the peripheral blood white cells are comprised of malignant PCs. Most human myeloma cell lines (HMCL), which provide convenient cell line systems to study MM biology *in* vitro, are derived from patients with extramedullary myeloma.[42, 43].

1.2.4 Updated diagnostic criteria

Historically, diagnosis of active MM required patients to display one or more of the CRAB features, accompanied with an expansion of clonal MM PCs in the BM. In 2014, the International Myeloma Working Group (IMWG) updated diagnostic criteria for MM due to improve treatment options and benefits of early intervention in high-risk asymptomatic patients.[35] In addition to the well-defined CRAB features, three new defining MM events were introduced and are based on biomarkers associated with 80% risk of progression within 2 years.[44] The new criteria include: 1) BM PCs of 60% or greater, 2) A serum free light chain ratio (FLC- κ/λ) of 100 or higher and 3) MRI

with more than one focal lesion.[35] Patients presenting with either one of these features are defined as having MM, even in the absence of CRAB features. These new criteria have identified a subset of MM patients who would otherwise be classified as ultra-high risk SMM and would otherwise remain untreated.

M protein requirements for MM diagnosis were also updated in the 2014 IMWG MM diagnostic criteria.[35] Monoclonal protein secretion is a hallmark of MM and can be easily detected by serum/urine electrophoresis (SEP/UEP) followed by immunofixation, except in the 3% cases of non-secretory MM. Non-secretory MM have similar prognosis to secretory MM of matched clinical features[45] and benefits from treatments for overt MM.[17] Therefore, according to 2014 IMWG, M protein is no longer a mandatory MM defining event, but serves to sub-classify MM into secretory or non-secretory types.

1.2.5 Staging MM

Developed in 1975, the Durie and Salmon staging system classifies MM patients into three stages and uses clinical parameters as surrogate measures of tumour burden. These clinical parameters include severity of lytic bone lesions, haemoglobin level, serum calcium level, monoclonal protein level and creatinine level.[46] In the effort to provide a more objective and widely available system, the International Staging System (ISS) was developed in 2005 based on clinical data of 10,750 MM patients worldwide.[47] The ISS system stratifies patients into three stages based on B_2-microglobulin level that reflects tumour burden and serum albumin level that reflects host fitness.[48, 49]

Beyond the ISS, other studies have suggested the inclusion of other prognostic factors such as serum lactase dehydrogenase (LDH) level [50], serum FLC ratio [51] and cytogenetic data by fluorescent *in situ* hybridisation (FISH) [52] to improve risk assessment of MM patients. As more follow-up data became available, the IMWG updated the ISS in 2015.[53] In addition to B_2-microglobulin and albumin level, the revised-ISS (R-ISS) incorporated cytogenetic abnormalities and LDH level.[53] The R-ISS overcomes the limitations of ISS by taking into account intrinsic PCs genetics, which play a major role in treatment resistance and disease evolution. The R-ISS has identified 26% patients who would otherwise been allocated to the good prognosis group, providing a better prognostic power compared to ISS.

Table 1.2 New International Staging System.[47]

Stage	Criteria	Median Survival (months)
I	Serum β_2-microglobulin < 3.5 mg/L Serum albumin ≥ 3.5 g/Dl	62
II	Not stage I or III (serum β_2-microglobulin < 3.5 mg/L but serum albumin < 3.5 g/dL; or serum β_2-microglobulin 3.5 to < 5.5 mg/L irrespective of the serum albumin level)	44
III	Serum β_2-microgoblulin ≥ 5.5 mg/L	29

Table 1.3 Revised International Staging System (R-ISS).[53]

Stage	Criteria	Median Overall Survival (months)
I	All the following: ISS stage I No high risk CA Normal LDH	Not reached
II	All other ISS/CA/LDH combinations not of R-ISS stage I or stage III	83
III	ISS stage III and High risk CA or high LDH	43
Cytogenetic abnormalities by iFISH		
High risk	Presence of del(17p) and/or Translocation t(4;14) and/or Translocation t(14;16)	
Standard risk	No high-risk CA	
LDH		
Normal	Serum LDH < the upper limit of normal	
High	Serum LDH > the upper limit of normal	

1.2.6 Development of malignant MM PCs

MM PCs originate from post-germinal centre B cells that have undergone genetic changes in the immunoglobulin (Ig) genes.[54, 55] Upon encounter with cognate antigen in secondary lymphoid organ, B cells migrate to primary follicles, proliferate rapidly and form germinal centre (GC).[56, 57] Two major DNA-modifying events, somatic hypermutation (SHM) and class switch recombination (CSR), which occur in the GC are thought to be the key processes that leads to oncogenic transformation of MM PCs (Figure 1.3).[58-60] While SHM introduces point mutations to variable region of Ig gene to produce antibodies with increased antigen affinity, CSR involves deletional rearrangement of immunoglobulin heavy chain (IGH) gene at chromosome 14q32 to alter Ig isotype to produce antibodies with diverse effector functions. Both SHM and CSR requires the activity of activation-induced deaminase to induce double-strand DNA breaks.[61, 62] Notably, errors can occur during the introduction of DNA breaks and during the DNA reassembling steps, both of are thought to drive the oncogenic processes in MM PCs oncogenesis (reviewed in [63]). To this end, recurrent chromosomal translocations of IGH locus often juxtapose oncogenic partners downstream of strong IGH enhancer, thereby dysregulating gene expression in MM PCs[59, 64] and accounting for the early immortalising events of MM PCs that occurs within GC. Like other post-GC B cells, the precursor MM PCs typically home to BM[65] where they may persist undiagnosed for many years. Acquisition of additional oncogenic events during the following passage through GC could further dysregulate MM PC behaviour, leading to clinically recognisable features of MM.

1.2.7 Cytogenetic and molecular markers of multiple myeloma

Historically, conventional metaphase karyotyping was used to detect cytogenetic abnormalities in MM. Due to the low proliferation index of MM PCs during early disease stage, only 30-40% of newly diagnose MM patients demonstrated detectable abnormal karyotypes using this method.[66-71] Over the years, incorporation of interphase FISH (iFISH) which analyse interphase nuclei has overcome this limitation, as it is independent of PC proliferation.

Prognostication impact of cytogenetic in predicting survival outcome has seen them been used to risk stratify MM patients and has been incorporated into staging

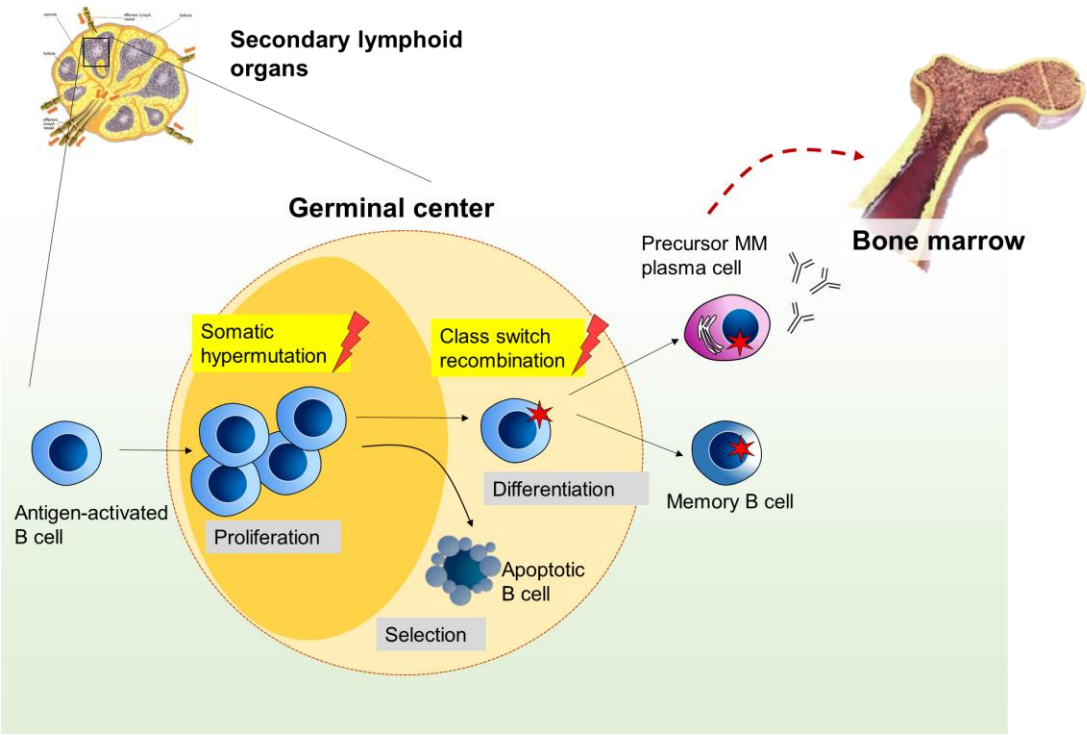


Figure 1.3 Initial malignant transforming events of MM PCs in germinal center. Upon interactions with antigen, mature B cell enters germinal center and undergone somatic hypermutation (SHM) to select for B cell clones that produce highly specific and avid antibodies. B cells bearing high-affinity antigen receptors are selected for survival and undergo class switch recombination (CSR) to produce antibodies of different isotypes. Aberrant DNA breakages and repairs that occur during SHM and CSR are the key genetic events that underlie malignant transformation of MM PCs. Terminally differentiated, antibody-secreting MM PCs typically home and reside in the bone marrow. *Adapted from de Silva et al.*[57]

system.[47, 53] In general, chromosomal abnormalities in MM are classified into two major groups: hyperdiploidy and non-hyperdiploidy. Hyperdiploid MM occurs in 50% of newly diagnosed MM[72, 73] and involves trisomies of chromosomes 3, 5, 7, 9, 11, 15, 19 and 21, gain of 1q and chromosome 13 deletion. Although its pathogenesis is unclear, in general, patients harbouring trisomies have a better prognosis compared with non-hyperdiploid MM.[74] In contrast, non-hyperdiploid MM, which accounts for approximately 40% of MM cases, is associated with poorer overall survival (OS) [75-77] and is associated with a high prevalence of chromosomal translocation involving the IGH locus 14q32.[78, 79] The 5 most common recurrent IgH translocations involve 11q13 (CCND1), 4p16 (FGFR3/MMSET), 16q23 (MAF), 20q12 (MAFB) and 6p21 (CCND3) and are characterised by overexpression of translocated genes.[80]

IgH translocations are considered as primary genetic events that occurs during early MM pathogenesis. Most patients with IgH translocations have a more aggressive disease with decreased OS and shorter time to relapse.[52, 75, 77, 81, 82] However, the most common t(11;14)(q13;q32) translocation observed in 15-20% cases, has a neutral prognosis due to heterogeneous outcomes despite its general association with favourable outcomes (Table 1.4).[79, 83, 84] The second most common IgH translocation of t(4;14)(p16;q32), observed in 15-20% cases, is associated with poor overall survival and is characterised by simultaneous overexpression of MMSET and FGFR3. Although occurring less commonly in MM, t(14;16)(q32:q23) is associated with poor survival in most cases owing to the dysregulation of MAF oncogene which is capable of initiating oncogenic transformation.[42] Despite this, the role of t(14;16)(q32,q23) in determining high-risk MM remains debateable, as a retrospective study was unable to confirm poor prognostic value of t(14;16)(q32,q23).[85] Taken together, t(4;14)(p16;q32) is likely the major IgH translocation contributing to the poor prognostic features of non-hyperdiploid MM.

1.2.8 Genomic complexity and clonal heterogeneity of MM

Gene expression profiling (GEP) using microarray technology has been used proficiently[2, 80, 86-92] to molecularly stratify MM patients and identified progression-associated molecular signature in the past 10 years (reviewed in [93]). Additionally, genome-wide profiling technology such as array comparative genomic hybridisation (aCGH) and SNP-array have identified chromosomal copy number alterations

underlying differential gene expression with prognostic significance.[94-98] As part of the ongoing efforts to apply GEP signature in clinical setting, muliple GEP signatures are integrated to improve the power of GEP in risk strattification[99] and to predict the likely response to therapy in MM patients.[100]

Table 1.4 Prevalence and cytogenetic risk of IgH translocations in MM patients.

Gene dysregulated	IgH Translocation	Prevalence %	Cytogenetic risk
FGFR3/MMSET	t(4;14)	15-20%	Intermediate
Cyclin D family			
CCND1	t(11;14)	15-20%	Standard
CCND2	t(11;14)	<1%	Standard
CCND3	t(6;14)	2%	Standard
MAF family			
c-MAF	t(14;16)	5%	High
MAFA	t(8;14)	<1%	Standard
MAFB	t(14;20)	2%	High

Adapted from references [42, 79, 101, 102].

The next generation sequencing (NGS) technology in whole-genome, whole-exome, or transcriptome sequencing has provided a better insight into genomic landscape and clonal heterogeneity in MM (reviewed in [103, 104]). Recent whole-exome sequencing (WES) studies performed in MM have identified recurrent mutations in genes affecting MAPK pathway (*KRAS, NRAS* and *BRAF*), DNA-repair pathway (*TP53*) and potential tumour suppressors (*FAM46C* and *DIS3*), which act as "driver mutations" in promoting MM progression.[105-108] Genomic assessment in WES analysis also further confirm the notion of intraclonal heterogeneity in MM[109, 110], in which most MM patients harbour at least three detectable genetically heterogeneous subclones, to as high as 10 subclones.[107] Moreover, longitudinal WES studies at two different time points, for example at different stage of diagnosis or at pre- or post-treatments, have provided a unique framework to dissect the evolution of clonal mutation composition over time.[105, 107, 109, 111] Currently, four clonal evolution models are identified: no change, linear, branching and differential clonal response. Better understanding in the dynamic of subclonal evolution may be used to predict the clonal evolution under selective pressure from and develop a comprehensive treatment strategy.

1.2.9 Treatments

The treatment paradigm for newly diagnosed MM are divided into four phases: 1) induction therapy, 2) autologous stem cell transplant for eligible patients or prolonged induction therapy for transplant ineligible patients, 3) maintenance/consolidation therapy and 4) treatment for relapsed disease. In the past 20 years, immunomodulatory drugs (IMiDs), such as thalidomide[112, 113] or its analogues lenalidomide[114, 115] and pomalidomide[116, 117], and the proteasome inhibitors (PI) bortezomib[118, 119] and carfilzomib[120, 121], have emerged as the active agents in frontline MM treatment and revolutionised MM treatment. In combination with alkylating chemotherapeutic agents, such as cyclophosphamide, vincristine, doxorubicin or melphalan, the IMiDs and PIs are used in triplet with dexamethasone as treatment regimens for newly diagnosed MM patients.[122, 123] Significant benefits have been observed with bortezomib-based regimens, with patients receiving bortezomib-based induction therapy displaying a better overall survival post-stem cell transplant compared with non-bortezomib based induction regimens.[124] Moreover, the second generation PI, carfilzomib has shown encouraging preliminary results in relapsed or refractory MM patients, with prolonged progression free survival, compared to bortezomib.[120, 125]

Although the overall survival of MM has improved with the emergence of novel agents, MM remains an incurable disease with essentially all patients eventually relapsing. Treatment selection in relation to relapsed MM is influenced by patient-related factors, disease-related factors and responses to prior treatments.[126, 127] Notably, patients who become refractory to IMiDs and bortezomib have poor overall outcome, with median survival of only 9 months.[128] Therefore, novel classes of drug targeting different pathogenic mechanisms of MM PC are being actively explored to overcome resistance to prior treatment in order to prolong OS. Histone deacetylase inhibitors, elotuzumab, daratumamab and kinesin spindle protein inhibitors are among the novel agents that have showed promising results for relapsed myeloma and are currently undergoing phase III clinical trials. Additionally, comprehensive understanding of the genomic landscape and subclonal evolution in MM have provided a better prospect in the effort to develop personalised therapies suited for each patient in the future.

Of interest, the t(4;14) translocation confers a negative impact on treatment responses to lenalidomide[129] and carfilzomib[115] in the relapsed setting. Although the mechanisms underlying the poor responses in t(4;14) patients remain to be elucidated,

studies have reported that the t(4;14) translocation is associated with aggressive relapses and metastatic extramedullary progression.[130] Therefore, potent agents that target pro-metastatic molecules represent attractive novel targets for treating high risk MM.

1.3 Tumour invasion and metastasis

Invasion and metastasis is one of the hallmarks of cancer [131] and is activated in tumour cells during cancer progression, resulting in movement of tumour cells from a primary site to distant organs. The invasion-metastasis cascade involves a multi-steps process, whereby primary tumour cells invade into the surrounding tissues, intravasate into blood vessels, survive in the vasculature during transport, extravasate into a distant site tissue and lastly, colonise and form a tumour at this secondary site (Figure 1.4).[132] Metastasis is an inefficient process and only about 0.02 % metastasised tumour cells will complete the whole invasion-metastasis cascade and survive to develop macroscopic metastases at secondary sites.[133, 134] Despite this, more than 90% of cancer related deaths are associated with metastasis and resistance to existing therapeutic agents.[135, 136]

Oncogenic transformation alone is not sufficient to establish metastasis.[137-139] Additional intrinsic genetic changes or pleotropic signals conferring tumour aggressiveness are required to initiate metastasis. Local invasion of tumour cells into surrounding stroma requires proteolytic degradation of basement membrane and extracellular matrix by tumour-secreted proteases belonging to the matrix metalloproteinase (MMP) family which allow the primary tumour to "break free" from their primary site. Moreover, tumour motility requires active remodelling of cytoskeleton, to enable dynamic cell elongation and directional movement, by the means of collective or single-cell migration.[140] During collective invasion, cancer cells of epithelial origin (e.g. carcinomas) migrate as epithelial sheets or detached cluster, keeping cell-cell junctions intact. Alternatively, individual tumour cells can detach from tumour mass and migrate using amoeboid or mesenchymal strategies, before entry into blood circulation.[132, 141, 142] Circulation pattern, structure of vasculature wall of distal organs and interactions of tumour cells with the endothelial cells are among the few factors influencing tumour extravasation to secondary sites (reviewed in [132, 139]).

Individual motile tumour cells that detach from primary tumours are characterised by the loss of cell-cell adhesion and organisation. These tumour cells may undergo epithelial-to-mesenchymal transition (EMT) to establish single cell invasion

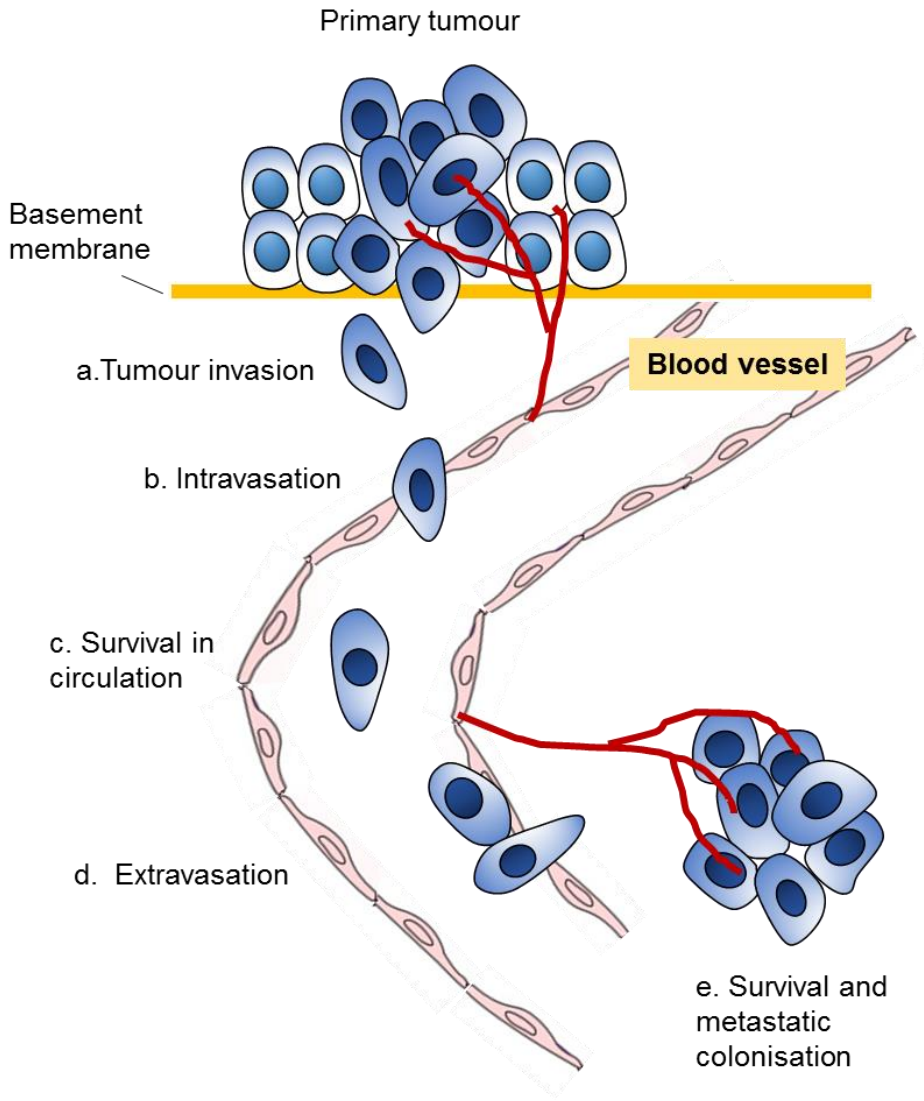


Figure 1.4 The invasion-metastasis cascade in cancers. During metastatic progression, tumour cells dissociate from primary tumour site, invade into basement membranes and intravasate into circulatory system. Metastatic tumour cells need to evade immune responses while in circulation before extravasate from blood vessels at secondary organs. Disseminated tumour cells need to adapt to survive in the new microenvironments to form metastases.[132]

and subsequently initiate the entire metastasis cascade. Interestingly, the histopathology of metastases at secondary sites have been shown to possess an epithelial phenotype similar to that of the primary tumour.[143] As the invasive, metastatic tumour cells were shown to have arrested proliferation[144-146], it has been suggested that the reverse mesenchymal-to-epithelial (MET) process is required for efficient colonisation and the formation of a macrometastasis at secondary sites.

EMT was first described as a biological process that occurs during gastrulation and embryonic development. EMT is fundamentally required to establish the three germinal layers (ectoderm, mesoderm and endoderm) and cells of neural crest origin. A similar EMT program is also reactivated as a physiological response during wound healing [144, 147], organ fibrosis [148, 149], as well as during cancer progression.[144, 150] EMT involves a morphological and functional transition from an epithelial to mesenchymal phenotype. Epithelial cells which are adhered to each other via adherent junctions, tight junctions, gap junctions and desmosomes, form highly organised epithelia layers with apical-basal polarity, separated from other tissues by basal lamina. In contrast, the spindle shape-like mesenchymal cells are loosely organised, lack polarity and are embedded within a mesh of ECM. During EMT, cell-cell junction molecules such as E-cadherin and occludin are downregulated, resulting in a loss of apical-basal polarity, thereby releasing single cells from primary tumour mass. N-cadherin, which facilitates transient cell adhesion, and fibronectin are upregulated[151, 152] to facilitate directional migration of mesenchymal cell. Additionally, epithelial intermediate filament, cytokeratin is switched to the mesenchymal intermediate filament of vimentin.[153, 154] Changes in epithelial- or mesenchymal-defining molecules serve as useful indicators that EMT has occurred within a tumour.

1.3.1 EMT-inducing transcription factors

Three main groups of transcription factors: SNAIL, ZEB and TWIST1 family members orchestrate the EMT program (reviewed in [155, 156]) by inhibiting and activating of a range of epithelial and mesenchymal markers, respectively. Both the SNAIL (SNAI1 and SNAI2) and the ZEB (ZEB1 and ZEB2) families directly repress E-cadherin expression and other cellular junction proteins such as the claudins, by binding to the E-box promoter region.[157-160] Snail and ZEB are highly elevated at the invasive front of various carcinomas to increase tumour invasiveness.[155, 158, 161]

TWIST1 is a basic/helix-loop-helix (bHLH) transcription factor that induces EMT program via repression of E-cadherin[162] and indirect upregulation of N-cadherin.[163] In addition, TWIST1 also promotes invadopodia-mediated ECM degradation[164], facilitating the tumour invasion aspect of cancer metastasis. TWIST1 regulates target gene expression through recruitment of coregulators or protein complexes. For example, TWIST1 interacts with several components in the Mi2/nucleosome remodeling and deacetylase (Mi2/NuRD) complex and recruits them to the E-cadherin promoter to repress E-cadherin transcription.[165]

1.3.2 EMT-activating signalling molecules

The expression of the EMT-inducing transcription factors Snail, ZEB and TWIST1 occurs in response to activation of various growth factor signalling pathways by factors derived from the surrounding tumour microenvironment, including transforming growth factors beta (TGFβ), bone morphogenic protein (BMP), Notch, Wnt, and Sonic Hedgehog (reviewed in [156, 166]).

Among these growth factors, TGFβ plays a major role in inducing EMT during cancer progression. In fact, TGFβ alone is sufficient to stimulate epithelial cells to undergo EMT *in vitro*.[167] Binding of TGFβ ligands to the TGFβ receptors induce oligomerisation of SMAD2/3 complex with SMAD4, which subsequently translocates into the nucleus and binds to target genes promoter regions to induce gene transcription. TGFβ can also activate SMAD independent responses, most commonly through receptor tyrosine kinase signalling such as PI3 kinase/Akt, Rho GTPase and Erk MAP kinase (reviewed in [168]). While TGFβ signalling induces Snail and TWIST1 transcription via a SMAD-dependent pathway, ZEB expression is induced by an indirect mechanism mediated by Ets-1.[169] BMP signalling is similar to the SMAD-dependent TGFβ signalling; with the exception that it uses BMP receptors that make use of SMAD1/5/8 rather than SMAD2/3 complexes.[170-173]

In Notch signalling, binding of the JAGGED family ligand to the Notch receptor initiates the proteolytic release of the Notch intracellular domain, which translocates to nucleus and binds to DNA-bound CSL (C protein binding factor 1/Suppressor of Hairless/ Lag-1) complex to activate transcription of a number of oncogenes, including NF-κB, Akt and p21.[174] Notch signalling regulates SNAIL expression, both directly and indirectly, via hypoxia-inducible factor 1α (HIF1α) to induce EMT in cancer.[175-177] Wnt

signals are transduced by Frizzled and LRP receptors and Wnt signalling has been shown to increase transcription of SNAI1, SNAI2 and TWIST1, which leads to EMT induction.[178-181]

Recent studies suggested that small non-coding micro RNA (miRNA) also plays a role in regulating the EMT program (reviewed in [182]). These 21-25 nucleotide-long mature miRNAs are incorporated into the RNA-induced silencing complex (RISC), which consists of endonucleolytically active Argonaute protein.[183, 184] Association of RISC with target mRNA via complementary binding of the miRNA seed sequences (2-8 nucleotides)[185] to the 3'UTR of the target mRNA can result in the inhibition of mRNA translation and/or mRNA for degradation.[186, 187] The miR-205 and miR-200 family (miR-200a/b/c, miR-141, and miR-429) inhibit EMT induction by targeting E-cadherin repressor, ZEB1 and ZEB2 mRNA, thereby maintaining an epithelial phenotype of the cell.[188-191] Loss of miR-200 was found to promote metastatic phenotype in EMT and has been linked to tumour progression[190, 192, 193] and serve as prognostic markers for advanced cancers.[194] Notably, ZEB1 and Snai1 can bind to the promoter region of miR-200, leading to the direct repression of mir-200 expression, forming a feed forward loop to maintain a mesenchymal state in response to EMT-inducing signals.[195-197]

1.4 Metastasis in MM

Metastasis is not usually used to describe dissemination of hematopoietic malignancies. Due to their hematopoietic origin, leukaemia and lymphoma cells are inherently motile and can circulate throughout the vasculature and are capable of efficient amoeboid migration to sites of infection or inflammation. Although rare, metastasis of haematological malignancies, such as leukaemia and lymphoma, to the heart, kidney, central nervous system, and other tissues have been reported.[198, 199] As leukemic and lymphoma cells are present in peripheral blood and lymphatic systems, the initial invasion and intravasation steps are not necessary in their metastasis. Despite being derived from terminally differentiated B-lymphocytes, studies suggest that the spread of MM PC and MM disease progression is dependent upon an "EMT-like" process resembling the invasion-metastasis cascade of cells of carcinoma origin (reviewed in [200]).

While solitary plasmacytomas, without evidence of distant metastasis, may resemble a primary tumour of MM, the majority of these patients remain undetected at

this stage. Diagnosis of symptomatic MM is most often associated with multiple myelomatous tumour lesions throughout the axial skeleton and appendicular due to continuous trafficking and micrometastasis formation (i.e. analogous to MGUS) of MM PCs. Similar to extravasation processes which result in the dissemination and metastasis of carcinomas into the peripheral blood circulation, MM PCs egress from BM and circulate via peripheral blood and home to distant BM niches, as evidenced by the presence of circulating malignant PCs in MM patients.[201, 202] In MM cases with extramedullary involvement, in addition to new BM regions, other organs such as liver, lungs, skin and brain also serve as secondary tumour sites for MM homing and colonisation. Notably, unlike MM PC, their normal PC counterparts are seemingly unable to invade and survive at these extramedullary sites. Following "metastases" MM PC can exist as solitatry tumour cells, small micro-metastases or larger vascularised metastases.[133]

1.4.1 Mechanisms regulating MM metastasis and BM homing

Although the regulation of MM metastasis has not been fully elucidated, MM PCs are likely to utilise mechanisms similar to those used by other cancer cell types. For example, the chemokine CXCL12, also known as stromal derived factor-1 (SDF-1), and its cognate receptor CXCR4, have been identified as a critical factor in regulating BM homing and migration of MM PCs.[203, 204] BM stromal cells are a major source of CXCL12, and stimulate chemotaxis of CXCR4-expressing MM PC. Binding of CXCL12 to the CXCR4 receptor on MM PC enhances the adhesion of MM PC to BM stromal cell via upregulation of VLA-4 and VLA-5, to retain MM PC within the BM niche.[205] Notably, elevated expression of CXCR4 on the MM PC surface was found to correlate with greater metastatic capabilities. For example, a study using matched patient samples showed that BM-resident PCs expressed lower surface CXCR4 when compared to MM PC from peripheral blood, and was attributed to CXCR4 internalisation induced by high CXCL12 level within BM.[206] Another study also suggested that surface CXCR4 was highly expressed on EMD-prone MM PC and was responsible for mediating dissemination and extramedullary tumour establishment.[207] In carcinomas, CXCR4 expression has been positively correlated with distant metastasis[208] and has been shown to drive the migration of breast[209], prostate[210] and colon cancers.[211] Therefore, disruption of CXCL12/CXCR4 axis is one likely mechanism leading to egress of MM PC from BM into peripheral blood. Additionally, MM also express other chemokine receptors

such as CCR1, CCR2 and CCR5 which drive specific homing and migration of MM PC towards regions enriched in MIP-1α (ligand of CCR1 and CCR5) and MCP-1, -2, -3 (ligands of CCR2).[212]

In addition to chemotaxis stimulated by chemokines, homing and colonisation of MM cells also requires proteolytic degradation of the subendothelial basement membrane and surrounding extracellular matrix, which involves the action of MMP-9.[213, 214] Both CXCL12 stimulation and interaction with BM endothelial cells upregulate MMP-9 production by MM PC, thereby increasing invasiveness of MM PC upon arrival at distant BM sites.[213-215] Inhibition of MMP-9 activity using MMP-9 antibodies or tissue inhibitors of metalloproteinases-1 (TIMP-1) abrogated CXCL12-induced migration[216, 217], suggesting that in addition to downstream signalling activated by CXCL12, MMP-9 also plays a crucial role in mediating MM PC migration. Consistent with these i*n vitro* studies, the treatment of 5T2MM bearing mice with SC-694, a broad MMP inhibitor, significantly reduced tumour burden compared with untreated mice[218], further strengthening the role of MMP-9 in disease progression. Similarly, MMP-9 action has also been observed in breast cancer and overexpression of MMP-9 in breast cancer cells, compared with normal breast tissue, has been correlated with a higher incidence of metastasis.[219]

1.4.2 Extramedullary MM is highly metastatic

Extramedullary MM disease (EMD) is an aggressive entity and is associated with adverse prognosis and poor response to therapy.[220, 221] Unsurprisingly, higher EMD incidence was reported at relapse in 7-24% of patients compared with to 3-5% in newly diagnosed MM patients. Interestingly, autopsy studies have shown that liver infiltration is evident in 40% of MM cases, suggesting MM PCs metastasis to distant organs occurs more commonly than anticipated. As EMD detection relies on availability of imaging technology such as MRI and PET-CT, it is likely most EMD occurs without being detected.[222]

Extraosseous and skeletal EMD highlight the BM-independent pathogenesis capabilities acquired by MM PCs and relapsed patients with extraosseous EMD as a result of haematogenous spread, have significantly poorer OS compared with skeletal EMD relapse.[223, 224] These finding suggest that the metastatic dissemination of MM PCs confers an aggressive disease phenotype. Indeed, higher number of circulating PCs have

been reported in relapsed MM patients (368/150, 000 events analysed) than in newly diagnosed MM (40/150,000 events analysed).[225, 226] Increased numbers of circulating MM PCs strongly predicts poor response to treatment and shorter time to progression[202, 227], reflecting increased metastatic property of the resistant MM subclones.

1.4.3 Cytogenetics of metastatic extramedullary MM

Although the precise mechanisms leading to extramedullary myeloma spread are not well understood, recent studies have used cytogenetics to identify genetic alterations that may be associated with extramedullary spread. For example, studies at multiple centres have evaluated the cytogenetics of MM PC at diagnosis of MM patients with extramedullary relapsed (refer to Table 1.5). Although the cytogenetic frequency observed were inconsistent across different studies, most likely due to the small sample sizes, high risk cytogenetic 13q deletion, t(4;14) and 17p deletion are the most frequent cytogenetic abnormalities observed in extramedullary relapse MM.[130, 228-230]

Previous studies[230-232] have suggested del 13q and del 17p as the prognostic markers of extramedullary progression. However, in the study by Billecke et al.[230], cytogenetic data from BM-resident PCs at diagnosis was lacking in majority of the extramedullary relapse MM patients, and as such, prevents a comprehensive evaluation of extramedullary relapse-promoting cytogenetics. Moreover, the incidence of del 17p was similar between extramedullary and medullary MM[231], suggesting that additional genetic factors are responsible for driving extramedullary metastasis. Furthermore, a study by Minnema et al. only investigated the incidence of deletion 13q in extramedullary relapse MM patients.[232] Due to the close association of del 13p with other high risk genetic features, specifically t(4;14)[75, 81, 233], the prognostic significance of del 13p in relation to extramedullary relapse maybe secondary to t(4;14).

Assuming extramedullary progression at relapse is indicative of an aggressive or advanced disease stage, early transforming events such as IgH translocation may also drive metastatic behaviour of resistant MM PCs at relapse. Indeed, high risk cytogenetic t(4;14) was found in 8-53% BM-resident PCs of extramedullary relapse MM patients[130, 228, 234], which is significantly higher when compared to 15-20% of medullary MM, indicating further clonal selection during extramedullary relapse. In addition, t(4; 14) frequency was significantly higher in extraosseous EMD (37%) derived from haematogenous spread compared to skeletal EMD (18%), emphasising that t(4;14)

mediates extramedullary metastasis of MM in advanced disease stages. However, the mechanisms on how t(4;14) might drive metastasis and extramedullary progression in MM remains to be fully understood.

Table 1.5 Cytogenetic distribution of MM patients with extramedullary relapse in various studies.

Authors	EMD definition	BM PCs cytogenetic at first diagnosis	EM tumour cytogenetic at relapse
2009 Kathodriou et al.[234]	Extraosseous and skeletal	33% t(4;14) 22% del 13q 0% del 17p	NA
2012 Rasche et al.[130]	Extraosseous	53% t(4;14) 58% del 13q 21% del 17p	NA
2013 Billecke et al.[230]	Extraosseous	8% t(4;14) 23% del 13q 8% del 17p	28.5% t(4;14) 28.5% del 13q 38% del 17p
2015 Besse et al.[228]	Extraosseous and skeletal	31% 14q32 disruption 48% del 13q 16% del 17p	60% t(4;14) 50% del 13q 11% del 17p

1.5 A role for EMT in MM metastasis

As discussed earlier, EMT and the reverse MET form the basis of local invasion and secondary colonisation of carcinoma metastasis. The sections that follow will examine common features between MM PCs during metastatic progression and carcinomas undergoing EMT from the aspects of cell plasticity, signalling pathways and gene expression signatures.

1.5.1 EMT-driven cell plasticity in carcinomas and MM PCs

The ability of a carcinoma cell to adopt a mesenchymal phenotype during EMT and vice versa during MET, reveals a high degree of cellular plasticity. Notably, this feature is shared by embryonic cells during morphogenesis, thereby allowing them to differentiate into different cell types (reviewed in [235, 236]). Ectopic expression of EMT-TFs or TGFβ1 stimulation in non-transformed epithelial cells is sufficient to reprogram differentiated epithelial cells into mesenchymal-appearing cells with stem cell-like properties.[237, 238] These putative cancer stem cell subpopulations, characterised by $CD44^{high}/CD24^{low}$ expression, are enriched in tumour initiation properties compared with the non-tumorigenic $CD44^{high}/CD24^{high}$ breast cancer cells.[239, 240]

Similarly, MM PC clones also display plasticity and are able to dedifferentiated into less mature phenotype. Histological and cytopathological studies dating back to the 1980's examined a large series of MM cases and showed that the morphology of MM PC is heterogeneous and could be divided into 4-6 histologic subtypes.[241-243] In keeping with the findings from metastatic carcinoma, the less mature, plasmablastic MM cells are associated with poor prognosis[244, 245] and possess an aggressive disease phenotype compared with the remaining subtypes.[246-248] In two small series studies, soft tissue tumour masses from relapsed MM patients showed histological progression to a less differentiated, immature morphology, including plasmablastic, anaplastic and pleomorphic phenotypes compared with BM MM PCs recovered at diagnosis.[249] These findings strengthen the association between dedifferentiation into immature plasmablast with MM PCs metastasis and aggressive relapse.

Several environmental factors can influence MM PCs dedifferentiation into an immature phenotype. For example, Balleari et al. reported that refractory MM patients showed a dedifferentiated plasmablast phenotype after treatment with thalidomide.[250] In these studies, following an initial, transient response to thalidomide, patients showed evidence of BM-resident MM PC with high expression of CD45 which had capacity for increased proliferation leading to increased bone marrow plasmacytosis.[250] Other studies also proposed that hypoxia had the capacity to induce MM PC dedifferentiation into less mature phenotype.[251, 252] MM cells cultured under hypoxic condition were shown to express lower plasma cells marker such as CD138, but higher B–cell markers including CD20 and CD45. Further to this, Yaccoby at el. also reported similar findings in long term co-culture of MM PCs with osteoclasts,[253] with MM PC acquiring an immature plasmablastic phenotype of $CD38^{low}/CD138^{intermediate}/CD45^{intermediate/high}$ compared with a pre-coculture phenotype of $CD38^{high}/CD138^{high}/CD45^{low/intermediate}$. These dedifferentiated MM PC showed evidence of enhanced tumour initiation ability and increased resistance to drug-induced apoptosis.[250, 252, 253] Given the established role of EMT in regulating carcinoma plasticity during metastasis, similar processes maybe involved in regulating MM PC plasticity.

1.5.2 EMT-activating signalling in MM metastasis

The EMT process is not limited to epithelial cancers as studies have identified activation of EMT program in non-epithelial cancers such as sarcomas and neuroectodermal

tumours.[131, 254] In addition, the key EMT transcription factors SNAIL and ZEB are involved in haematopoiesis.[255, 256] Ectopic expression of SNAI1, SNAI2 and ZEB, respectively, in mouse models resulted in development of both epithelial and non-epithelial cancers, with higher incidence of lymphoma and leukaemia.[257-259]

Signalling pathways known to activate EMT programs in epithelial cancer metastasis including TGFβ (reviewed in [260]), Notch (reviewed in [261]) and Wnt (review in [262]) were also found to be involved in development and pathogenesis of haematological malignancies including MM. Interestingly, these signalling pathways have been implicated in regulating MM PC migration, which is an important element in initiation of metastasis. High levels of TGFβ1 have been shown to induce the production of pro-survival cytokines, such as IL-6 and VEGF, by BM stromal cells.[263, 264] While VEGF is capable of inducing MM PC cell migration directly[265], IL-6 mediates cell migration indirectly via upregulation of other chemokines factors such as hepatocyte growth factor (HGF) and monocyte chemoattractant protein-1 (MCP-1).[266, 267] Notch signalling had also been shown to regulate cell migration via the CXCL12/CXCR4 axis. Inhibition of Notch signalling decreased CXCL2-induced MM PC migration by reducing CXCR4 surface expression on MM PC.[268] Studies by Qiang et al. showed that activation of Wnt signalling promoted MM PC cell invasion and migration via RhoA and PKC, independent of canonical Wnt/β-catenin pathway.[269] However, further studies are required to determine whether these signalling molecules are also capable of activating similar EMT program in MM PCs.

1.5.3 EMT-like expression profile in MM

Interestingly, there is emerging evidence to suggest that an EMT-like process can be activated in MM PCs and can contribute to tumour dissemination and aggressiveness.[270-272] Azab et al.[270] recently showed that tumour progression induces hypoxia and promotes metastasis of MM PC by activating EMT-related protein expression. In animal model studies, they showed that as the tumour burden increases, both the level of hypoxia in the BM MM PC and number of circulating cells increase proportionally, suggesting that hypoxia may drive the egress of MM PC into circulation. Hypoxic condition also increased protein level of the key EMT-promoting factor Snail1 and resulted in the downregulation of E-cadherin in two human myeloma cell lines, thereby contributing to decreased MM PC adhesion and increased migration and homing.

In another study by Sun et al.[271], the pro-inflammatory cytokine IL-17, was shown to stimulate an EMT expression profile in MM PCs. Previously known to be an EMT-inducing factor in lung cancer [273], IL-17 triggered repression of E-cadherin expression while inducing the expression of mesenchymal marker vimentin and EMT transcription factors SNAIL and SLUG in MM PCs. Biologically, IL-17 stimulation induced MM PCs proliferation, migration and apoptosis resistance, which contributed to the metastatic properties of the MM PC.[271]

Heparanase has also been shown to induce mesenchymal features and promote metastasis in MM PCs. Previously reported to play role in solid tumour metastasis[274, 275], ectopic heparanase expression in MM PCs induced a spreading morphology and a phenotype resembling mesenchymal cells. High heparanase expression reduced E-cadherin protein expression, while increasing vimentin and fibronectin protein expression[272], thus contributing to increase cell motility and angiogenic potential.[276] Of note, conditioned media from heparanse^{high} MM PCs was also shown to stimulate vimentin expression in endothelial cells *in vitro.*

1.5.4 Emerging role of MM oncogene MMSET in EMT

Previous studies have characterised MMSET, a histone methyltransferase, as the key oncogenic factor in t(4;14) patients. Excessive MMSET expression alters chromatin structure and gene transcription in MM cells, which subsequently regulates cell growth, adhesion and clonogenicity of MM PCs.[8, 277, 278] MMSET is also highly expressed in many solid tumours, generally associated with more advanced, aggressive tumour phenotype.[279] Recently a study by Ezponda et al. identified MMSET as a novel driver of EMT via TWIST1 activation in prostate cancer.[1] Overexpression of MMSET in non-transformed prostate epithelial cell lines resulted in increased cell migration and invasion, accompanied by a gene expression profile consistent with EMT activation. Conversely, knockdown of MMSET was associated with a loss of invasive, metastatic characteristics. MMSET-induced EMT, migration and invasion in prostate cancer cell line was found to require the expression of TWIST1 as subsequent knockdown of TWIST1 was able to reverse the phenotype stimulated by MMSET.

In the same study, Ezponda et al. showed that expression of the key EMT-TF TWIST1 was also elevated in t(4;14)-positive MM cell lines and patients samples compared with t(4;14)-negative samples.[1] This is consistent with increased expression

of mesenchymal markers N-cadherin and vimentin in t(4;14) MM patients.[5, 280] Studies from our group[6, 281] and others[5] have also shown that high N-cadherin expression in MM PC is associated with poor survival outcomes. In a preclinical mouse model of MM, knockdown of N-cadherin and vimentin in MM PCs significantly reduced tumour progression[272, 281], demonstrating an important role of mesenchymal markers in facilitating MM PCs homing and dissemination. In conjunction with the novel role of MMSET as an EMT driver[1], upregulation of mesenchymal markers in t(4;14) MM PCs may highlight EMT activation as part of global transcription changes driven by MMSET, which promotes an aggressive MM phenotype and disease progression.

1.6 Summary and objectives

Advances in treatment have seen significant improvements in the progression free and OS of myeloma patients. Despite this, MM remains largely incurable, as essentially all patients relapse. MM patients harbouring the t(4;14) translocation often display aggressive disease which is associated with increased circulating tumour cells and extramedullary disease. Although the mechanisms underlying MM PC metastasis remain to be fully elucidated, emerging evidence suggests that MM PC metastasis may resemble carcinoma metastasis and may rely on the activation of an EMT-like expression profile.[200, 270, 271] Whether MMSET is likely to activate an EMT-like process in t(4;14) MM PCs and leads to metastatic phenotype warrants further investigation.

Furthermore, although previous studies have shown that MMSET induces activation of TWIST1 to promote EMT in prostate cancer[1], the role of TWIST1 in MM pathogenesis remains largely unknown. High expression of TWIST1 in solid tumours confers resistance to chemotherapy and stem-cell characteristic.[282, 283] Whether poor prognosis associated with MMSET-high t(4;4) patients is mediated by biological changes resulting from increased TWIST1 expression remains to be determined.

The studies outlined in this thesis addressed the following aims:

1. To investigate whether BM PCs from t(4;14) MM patients exhibit an EMT-like expression profile using publicly available microarray datasets.
2. To investigate the role of MMSET-driven TWIST1 in MM pathogenesis.

1.7 References

1. Ezponda T, Popovic R, Shah MY, Martinez-Garcia E, Zheng Y, Min DJ, *et al.* The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. *Oncogene* 2012 Jul 16.

2. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A, *et al.* Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* 2010 Oct 7; **116**(14)**:** 2543-2553.

3. Wu SP, Pfeiffer RM, Ahn IE, Mailankody S, Sonneveld P, van Duin M, *et al.* Impact of Genes Highly Correlated with MMSET Myeloma on the Survival of Non-MMSET Myeloma Patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2016 Aug 15; **22**(16)**:** 4039-4044.

4. Zhan F, Hardin J, Kordsmeier B, Bumm K, Zheng M, Tian E, *et al.* Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood* 2002 Mar 01; **99**(5)**:** 1745-1757.

5. Groen RW, de Rooij MF, Kocemba KA, Reijmers RM, de Haan-Kramer A, Overdijk MB, *et al.* N-cadherin-mediated interaction with multiple myeloma cells inhibits osteoblast differentiation. *Haematologica* 2011 Nov; **96**(11)**:** 1653-1661.

6. Vandyke K, Chow AW, Williams SA, To LB, Zannettino AC. Circulating N-cadherin levels are a negative prognostic indicator in patients with multiple myeloma. *Br J Haematol* 2013 May; **161**(4)**:** 499-507.

7. Xie Z, Bi C, Chooi JY, Chan ZL, Mustafa N, Chng WJ. MMSET regulates expression of IRF4 in t(4;14) myeloma and its silencing potentiates the effect of bortezomib. *Leukemia* 2015 Dec; **29**(12)**:** 2347-2354.

8. Min DJ, Ezponda T, Kim MK, Will CM, Martinez-Garcia E, Popovic R, *et al.* MMSET stimulates myeloma cell growth through microRNA-mediated modulation of c-MYC. *Leukemia* 2013 Mar; **27**(3)**:** 686-694.

9. Popovic R, Martinez-Garcia E, Giannopoulou EG, Zhang Q, Zhang Q, Ezponda T, *et al.* Histone methyltransferase MMSET/NSD2 alters EZH2 binding and reprograms the myeloma epigenome through global and focal changes in H3K36 and H3K27 methylation. *PLoS Genet* 2014 Sep; **10**(9)**:** e1004566.

10. Angtuaco EJ, Fassas AB, Walker R, Sethi R, Barlogie B. Multiple myeloma: clinical review and diagnostic imaging. *Radiology* 2004 Apr; **231**(1)**:** 11-23.

11. Gebregziabher M, Bernstein L, Wang Y, Cozen W. Risk patterns of multiple myeloma in Los Angeles County, 1972-1999 (United States). *Cancer causes & control : CCC* 2006 Sep; **17**(7)**:** 931-938.

12. Kaya H, Peressini B, Jawed I, Martincic D, Elaimy AL, Lamoreaux WT, *et al.* Impact of age, race and decade of treatment on overall survival in a critical population analysis of 40,000 multiple myeloma patients. *Int J Hematol* 2012 Jan; **95**(1)**:** 64-70.

13. Bianchi G, Munshi NC. Pathogenesis beyond the cancer clone(s) in multiple myeloma. *Blood* 2015 May 14; **125**(20)**:** 3049-3058.

14. Hideshima T, Mitsiades C, Tonon G, Richardson PG, Anderson KC. Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. *Nature reviews Cancer* 2007 Aug; **7**(8)**:** 585-598.

15. Mitsiades CS, McMillin DW, Klippel S, Hideshima T, Chauhan D, Richardson PG, *et al.* The role of the bone marrow microenvironment in the pathophysiology of myeloma and its significance in the development of more effective therapies. *Hematol Oncol Clin North Am* 2007 Dec; **21**(6)**:** 1007-1034, vii-viii.

16. Dispenzieri A, Kyle RA. Multiple myeloma: clinical features and indications for therapy. *Best practice & research Clinical haematology* 2005; **18**(4)**:** 553-568.

17. Kyle RA, Gertz MA, Witzig TE, Lust JA, Lacy MQ, Dispenzieri A, *et al.* Review of 1027 patients with newly diagnosed multiple myeloma. *Mayo Clin Proc* 2003 Jan; **78**(1)**:** 21-33.

18. Bataille R, Harousseau JL. Multiple myeloma. *The New England journal of medicine* 1997 Jun 5; **336**(23)**:** 1657-1664.

19. McGrath MA, Penny R. Paraproteinemia: blood hyperviscosity and clinical manifestations. *Journal of Clinical Investigation* 1976; **58**(5)**:** 1155-1162.

20. Dimopoulos MA, Kastritis E, Rosinol L, Blade J, Ludwig H. Pathogenesis and treatment of renal failure in multiple myeloma. *Leukemia* 2008 06/05/online; **22**(8)**:** 1485-1493.

21. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015 Mar 1; **136**(5)**:** E359-386.

22. (AIHW) AIoHaW. Australian Cancer Incidence and Mortality (ACIM) books: Multiple Myeloma. Canberra; 2016.

23. Ferlay J SI, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013.

24. Brenner H, Gondos A, Pulte D. Expected long-term survival of patients diagnosed with multiple myeloma in 2006-2010. *Haematologica* 2009 Feb; **94**(2)**:** 270-275.

25. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. *Nature reviews Cancer* 2012 Apr 12; **12**(5)**:** 335-348.

26. Landgren O, Kyle RA, Pfeiffer RM, Katzmann JA, Caporaso NE, Hayes RB*, et al.* Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. *Blood* 2009 May 28; **113**(22)**:** 5412-5417.

27. Weiss BM, Abadie J, Verma P, Howard RS, Kuehl WM. A monoclonal gammopathy precedes multiple myeloma in most patients. *Blood* 2009 May 28; **113**(22)**:** 5418-5422.

28. Kyle RA, Therneau TM, Rajkumar SV, Larson DR, Plevak MF, Offord JR*, et al.* Prevalence of monoclonal gammopathy of undetermined significance. *The New England journal of medicine* 2006 Mar 30; **354**(13)**:** 1362-1369.

29. Dispenzieri A, Katzmann JA, Kyle RA, Larson DR, Melton LJ, 3rd, Colby CL*, et al.* Prevalence and risk of progression of light-chain monoclonal gammopathy of undetermined significance: a retrospective population-based cohort study. *Lancet* 2010 May 15; **375**(9727)**:** 1721-1728.

30. Blade J, Rosinol L. Smoldering multiple myeloma and monoclonal gammopathy of undetermined significance. *Curr Treat Options Oncol* 2006 May; **7**(3)**:** 237-245.

31. Kyle RA, Therneau TM, Rajkumar SV, Offord JR, Larson DR, Plevak MF*, et al.* A long-term study of prognosis in monoclonal gammopathy of undetermined significance. *The New England journal of medicine* 2002 Feb 21; **346**(8)**:** 564-569.

32. Bergsagel D. The incidence and epidemiology of plasma cell neoplasms. *Stem Cells* 1995 Aug; **13 Suppl 2:** 1-9.

33. Rajkumar SV, Kyle RA, Therneau TM, Melton LJ, 3rd, Bradwell AR, Clark RJ*, et al.* Serum free light chain ratio is an independent risk factor for progression in monoclonal gammopathy of undetermined significance. *Blood* 2005 Aug 1; **106**(3)**:** 812-817.

34. The International Myeloma Working G. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. *British Journal of Haematology* 2003; **121**(5)**:** 749-757.

35. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV*, et al.* International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol* 2014 Nov; **15**(12)**:** e538-548.

36. Kyle RA, Remstein ED, Therneau TM, Dispenzieri A, Kurtin PJ, Hodnefield JM*, et al.* Clinical course and prognosis of smoldering (asymptomatic) multiple

myeloma. *The New England journal of medicine* 2007 Jun 21; **356**(25)**:** 2582-2590.

37. Rajkumar SV, Merlini G, San Miguel JF. Haematological cancer: Redefining myeloma. *Nat Rev Clin Oncol* 2012 Sep; **9**(9)**:** 494-496.

38. Rajkumar SV. Multiple myeloma: 2012 update on diagnosis, risk-stratification, and management. *Am J Hematol* 2012 Jan; **87**(1)**:** 78-88.

39. Landgren O, Waxman AJ. Multiple myeloma precursor disease. *JAMA* 2010 Dec 1; **304**(21)**:** 2397-2404.

40. Blade J, Fernandez de Larrea C, Rosinol L, Cibeira MT, Jimenez R, Powles R. Soft-tissue plasmacytomas in multiple myeloma: incidence, mechanisms of extramedullary spread, and treatment approach. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011 Oct 1; **29**(28)**:** 3805-3812.

41. Short KD, Rajkumar SV, Larson D, Buadi F, Hayman S, Dispenzieri A*, et al.* Incidence of extramedullary disease in patients with multiple myeloma in the era of novel therapy, and the activity of pomalidomide on extramedullary myeloma. *Leukemia* 2011 Jun; **25**(6)**:** 906-908.

42. Kuehl WM, Bergsagel PL. Multiple myeloma: evolving genetic events and host interactions. *Nature reviews Cancer* 2002 Mar; **2**(3)**:** 175-187.

43. Drexler HG, Matsuo Y. Malignant hematopoietic cell lines: in vitro models for the study of multiple myeloma and plasma cell leukemia. *Leuk Res* 2000 Aug; **24**(8)**:** 681-703.

44. Mateos MV, Hernandez MT, Giraldo P, de la Rubia J, de Arriba F, Lopez Corral L*, et al.* Lenalidomide plus dexamethasone for high-risk smoldering multiple myeloma. *The New England journal of medicine* 2013 Aug 1; **369**(5)**:** 438-447.

45. Kumar S, Perez WS, Zhang MJ, Ballen K, Bashey A, To LB*, et al.* Comparable outcomes in nonsecretory and secretory multiple myeloma after autologous stem cell transplantation. *Biol Blood Marrow Transplant* 2008 Oct; **14**(10)**:** 1134-1140.

46. Durie BG, Salmon SE. A clinical staging system for multiple myeloma. Correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. *Cancer* 1975 Sep; **36**(3)**:** 842-854.

47. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J*, et al.* International staging system for multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 May 20; **23**(15)**:** 3412-3420.

48. Durie BG, Kyle RA, Belch A, Bensinger W, Blade J, Boccadoro M*, et al.* Myeloma management guidelines: a consensus report from the Scientific

Advisors of the International Myeloma Foundation. *Hematol J* 2003; **4**(6)**:** 379-398.

49. Jacobson JL, Hussein MA, Barlogie B, Durie BG, Crowley JJ, Southwest Oncology G. A new staging system for multiple myeloma patients based on the Southwest Oncology Group (SWOG) experience. *Br J Haematol* 2003 Aug; **122**(3)**:** 441-450.

50. Terpos E, Katodritou E, Roussou M, Pouli A, Michalis E, Delimpasi S, *et al.* High serum lactate dehydrogenase adds prognostic value to the international myeloma staging system even in the era of novel agents. *Eur J Haematol* 2010 Aug; **85**(2)**:** 114-119.

51. Dispenzieri A, Kyle R, Merlini G, Miguel JS, Ludwig H, Hajek R, *et al.* International Myeloma Working Group guidelines for serum-free light chain analysis in multiple myeloma and related disorders. *Leukemia* 2009 Feb; **23**(2)**:** 215-224.

52. Avet-Loiseau H, Durie BG, Cavo M, Attal M, Gutierrez N, Haessler J, *et al.* Combining fluorescent in situ hybridization data with ISS staging improves risk assessment in myeloma: an International Myeloma Working Group collaborative project. *Leukemia* 2013 Mar; **27**(3)**:** 711-717.

53. Palumbo A, Avet-Loiseau H, Oliva S, Lokhorst HM, Goldschmidt H, Rosinol L, *et al.* Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2015 Sep 10; **33**(26)**:** 2863-2869.

54. Bakkus MH, Heirman C, Van Riet I, Van Camp B, Thielemans K. Evidence that multiple myeloma Ig heavy chain VDJ genes contain somatic mutations but show no intraclonal variation. *Blood* 1992 Nov 1; **80**(9)**:** 2326-2335.

55. Vescio RA, Cao J, Hong CH, Lee JC, Wu CH, Der Danielian M, *et al.* Myeloma Ig heavy chain V region sequences reveal prior antigenic selection and marked somatic mutation but no intraclonal diversity. *J Immunol* 1995 Sep 1; **155**(5)**:** 2487-2497.

56. Shapiro-Shelef M, Calame K. Regulation of plasma-cell development. *Nat Rev Immunol* 2005 Mar; **5**(3)**:** 230-242.

57. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol* 2015 Mar; **15**(3)**:** 137-148.

58. Kosmas C, Stamatopoulos K, Stavroyianni N, Zoi K, Belessi C, Viniou N, *et al.* Origin and diversification of the clonogenic cell in multiple myeloma: lessons from the immunoglobulin repertoire. *Leukemia* 2000 Oct; **14**(10)**:** 1718-1726.

59. González D, van der Burg M, García-Sanz R, Fenton JA, Langerak AW, González M*, et al.* Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. *Blood* 2007 November 1, 2007; **110**(9)**:** 3112-3121.

60. Kuppers R, Sousa AB, Baur AS, Strickler JG, Rajewsky K, Hansmann ML. Common germinal-center B-cell origin of the malignant cells in two composite lymphomas, involving classical Hodgkin's disease and either follicular lymphoma or B-CLL. *Mol Med* 2001 May; **7**(5)**:** 285-292.

61. Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. *Cell* 2010 Apr 2; **141**(1)**:** 27-38.

62. Kuehl WM, Bergsagel PL. Molecular pathogenesis of multiple myeloma and its premalignant precursor. *J Clin Invest* 2012 Oct; **122**(10)**:** 3456-3463.

63. Gourzones-Dmitriev C, Kassambara A, Sahota S, Reme T, Moreaux J, Bourquard P*, et al.* DNA repair pathways in human multiple myeloma: role in oncogenesis and potential targets for treatment. *Cell Cycle* 2013 Sep 1; **12**(17)**:** 2760-2773.

64. Bergsagel PL, Kuehl WM. Chromosome translocations in multiple myeloma. *Oncogene* 2001 Sep 10; **20**(40)**:** 5611-5622.

65. Kunkel EJ, Butcher EC. Plasma-cell homing. *Nat Rev Immunol* 2003 Oct; **3**(10)**:** 822-829.

66. Seong C, Delasalle K, Hayes K, Weber D, Dimopoulos M, Swantkowski J*, et al.* Prognostic value of cytogenetics in multiple myeloma. *Br J Haematol* 1998 Apr; **101**(1)**:** 189-194.

67. Calasanz MJ, Cigudosa JC, Odero MD, Ferreira C, Ardanaz MT, Fraile A*, et al.* Cytogenetic analysis of 280 patients with multiple myeloma and related disorders: primary breakpoints and clinical correlations. *Genes, chromosomes & cancer* 1997 Feb; **18**(2)**:** 84-93.

68. Smadja NV, Louvet C, Isnard F, Dutel JL, Grange MJ, Varette C*, et al.* Cytogenetic study in multiple myeloma at diagnosis: comparison of two techniques. *Br J Haematol* 1995 Jul; **90**(3)**:** 619-624.

69. Lai JL, Zandecki M, Mary JY, Bernardi F, Izydorczyk V, Flactif M*, et al.* Improved cytogenetics in multiple myeloma: a study of 151 patients including 117 patients at diagnosis. *Blood* 1995 May 1; **85**(9)**:** 2490-2497.

70. Dewald GW, Kyle RA, Hicks GA, Greipp PR. The clinical significance of cytogenetic studies in 100 patients with multiple myeloma, plasma cell leukemia, or amyloidosis. *Blood* 1985 Aug; **66**(2)**:** 380-390.

71. Sawyer JR, Waldron JA, Jagannath S, Barlogie B. Cytogenetic findings in 200 patients with multiple myeloma. *Cancer Genet Cytogenet* 1995 Jul 1; **82**(1)**:** 41-49.

72. Fonseca R, Monge J, Dimopoulos MA. Staging and prognostication of multiple myeloma. *Expert Rev Hematol* 2014 Feb; **7**(1)**:** 21-31.

73. Smadja NV, Leroux D, Soulier J, Dumont S, Arnould C, Taviaux S, *et al.* Further cytogenetic characterization of multiple myeloma confirms that 14q32 translocations are a very rare event in hyperdiploid cases. *Genes, chromosomes & cancer* 2003 Nov; **38**(3)**:** 234-239.

74. Smadja NV, Bastard C, Brigaudeau C, Leroux D, Fruchart C, Groupe Francais de Cytogenetique H. Hypodiploidy is a major prognostic factor in multiple myeloma. *Blood* 2001 Oct 1; **98**(7)**:** 2229-2238.

75. Avet-Loiseau H, Facon T, Grosbois B, Magrangeas F, Rapp MJ, Harousseau JL, *et al.* Oncogenesis of multiple myeloma: 14q32 and 13q chromosomal abnormalities are not randomly distributed, but correlate with natural history, immunological features, and clinical presentation. *Blood* 2002 Mar 15; **99**(6)**:** 2185-2191.

76. Fonseca R, Debes-Marun CS, Picken EB, Dewald GW, Bryant SC, Winkler JM, *et al.* The recurrent IgH translocations are highly associated with nonhyperdiploid variant multiple myeloma. *Blood* 2003 Oct 1; **102**(7)**:** 2562-2567.

77. Avet-Loiseau H, Attal M, Moreau P, Charbonnel C, Garban F, Hulin C, *et al.* Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myelome. *Blood* 2007 Apr 15; **109**(8)**:** 3489-3495.

78. Chesi M, Bergsagel PL. Molecular pathogenesis of multiple myeloma: basic and clinical updates. *Int J Hematol* 2013 Mar; **97**(3)**:** 313-323.

79. Fonseca R, Bergsagel PL, Drach J, Shaughnessy J, Gutierrez N, Stewart AK, *et al.* International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. *Leukemia* 2009 Dec; **23**(12)**:** 2210-2221.

80. Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S, *et al.* The molecular classification of multiple myeloma. *Blood* 2006 Sep 15; **108**(6)**:** 2020-2028.

81. Fonseca R, Blood E, Rue M, Harrington D, Oken MM, Kyle RA, *et al.* Clinical and biologic implications of recurrent genomic aberrations in myeloma. *Blood* 2003 Jun 1; **101**(11)**:** 4569-4575.

82. Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ, *et al.* In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood* 2003 Feb 15; **101**(4)**:** 1520-1529.

83. Fonseca R, Blood EA, Oken MM, Kyle RA, Dewald GW, Bailey RJ, *et al.* Myeloma and the t(11;14)(q13;q32); evidence for a biologically defined unique subset of patients. *Blood* 2002 May 15; **99**(10)**:** 3735-3741.

84. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 Sep 10; **23**(26)**:** 6333-6338.

85. Avet-Loiseau H, Malard F, Campion L, Magrangeas F, Sebban C, Lioure B*, et al.* Translocation t(14;16) and multiple myeloma: is it really an independent prognostic factor? *Blood* 2011 Feb 10; **117**(6)**:** 2009-2011.

86. Shaughnessy JD, Jr., Zhan F, Burington BE, Huang Y, Colla S, Hanamura I*, et al.* A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* 2007 Mar 15; **109**(6)**:** 2276-2284.

87. Decaux O, Lode L, Magrangeas F, Charbonnel C, Gouraud W, Jezequel P*, et al.* Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myelome. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008 Oct 10; **26**(29)**:** 4798-4805.

88. Hose D, Reme T, Hielscher T, Moreaux J, Messner T, Seckinger A*, et al.* Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. *Haematologica* 2011 Jan; **96**(1)**:** 87-95.

89. Chng WJ, Braggio E, Mulligan G, Bryant B, Remstein E, Valdez R*, et al.* The centrosome index is a powerful prognostic marker in myeloma and identifies a cohort of patients that might benefit from aurora kinase inhibition. *Blood* 2008 Feb 1; **111**(3)**:** 1603-1609.

90. Moreaux J, Klein B, Bataille R, Descamps G, Maiga S, Hose D*, et al.* A high-risk signature for patients with multiple myeloma established from the molecular classification of human myeloma cell lines. *Haematologica* 2011 Apr; **96**(4)**:** 574-582.

91. Kuiper R, Broyl A, de Knegt Y, van Vliet MH, van Beers EH, van der Holt B*, et al.* A gene expression signature for high-risk multiple myeloma. *Leukemia* 2012 Nov; **26**(11)**:** 2406-2413.

92. Chung TH, Mulligan G, Fonseca R, Chng WJ. A novel measure of chromosome instability can account for prognostic difference in multiple myeloma. *PloS one* 2013; **8**(6)**:** e66361.

93. Chng WJ, Dispenzieri A, Chim CS, Fonseca R, Goldschmidt H, Lentzsch S*, et al.* IMWG consensus on risk stratification in multiple myeloma. *Leukemia* 2014 Feb; **28**(2)**:** 269-277.

94. Agnelli L, Bicciato S, Fabris S, Baldini L, Morabito F, Intini D*, et al.* Integrative genomic analysis reveals distinct transcriptional and genetic features associated

with chromosome 13 deletion in multiple myeloma. *Haematologica* 2007 Jan; **92**(1)**:** 56-65.

95. Avet-Loiseau H, Li C, Magrangeas F, Gouraud W, Charbonnel C, Harousseau JL*, et al.* Prognostic significance of copy-number alterations in multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009 Sep 20; **27**(27)**:** 4585-4590.

96. Dickens NJ, Walker BA, Leone PE, Johnson DC, Brito JL, Zeisig A*, et al.* Homozygous deletion mapping in myeloma samples identifies genes and an expression signature relevant to pathogenesis and outcome. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2010 Mar 15; **16**(6)**:** 1856-1864.

97. Walker BA, Leone PE, Chiecchio L, Dickens NJ, Jenner MW, Boyd KD*, et al.* A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. *Blood* 2010 Oct 14; **116**(15)**:** e56-65.

98. Carrasco DR, Tonon G, Huang Y, Zhang Y, Sinha R, Feng B*, et al.* High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. *Cancer Cell* 2006 Apr; **9**(4)**:** 313-325.

99. Chng WJ, Chung TH, Kumar S, Usmani S, Munshi N, Avet-Loiseau H*, et al.* Gene signature combinations improve prognostic stratification of multiple myeloma patients. *Leukemia* 2016 May; **30**(5)**:** 1071-1078.

100. van Laar R, Flinchum R, Brown N, Ramsey J, Riccitelli S, Heuck C*, et al.* Translating a gene expression signature for multiple myeloma prognosis into a robust high-throughput assay for clinical use. *BMC Medical Genomics* 2014; **7**(1)**:** 25.

101. Bergsagel PL, Chesi M. V. Molecular classification and risk stratification of myeloma. *Hematol Oncol* 2013 Jun; **31 Suppl 1:** 38-41.

102. Chng WJ, Glebov O, Bergsagel PL, Kuehl WM. Genetic events in the pathogenesis of multiple myeloma. *Best practice & research Clinical haematology* 2007 Dec; **20**(4)**:** 571-596.

103. Corre J, Munshi N, Avet-Loiseau H. Genetics of multiple myeloma: another heterogeneity level? *Blood* 2015 Mar 19; **125**(12)**:** 1870-1876.

104. Manier S, Salem KZ, Park J, Landau DA, Getz G, Ghobrial IM. Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol* 2016 Aug 17.

105. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I*, et al.* Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun* 2014; **5:** 2997.

106. Chapman MALMSKJJCKCACHCLBJ-PGJAMKCAK. Initial genome sequencing and analysis of multiple myeloma. *Nature* 2011; **471**(7339)**:** 467-472.

107. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D*, et al.* Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell* 2014 Jan 13; **25**(1)**:** 91-101.

108. Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM*, et al.* APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nat Commun* 2015; **6:** 6997.

109. Walker BA, Wardell CP, Melchor L, Brioli A, Johnson DC, Kaiser MF*, et al.* Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. *Leukemia* 2014 Feb; **28**(2)**:** 384-390.

110. Walker BA, Wardell CP, Melchor L, Hulkki S, Potter NE, Johnson DC*, et al.* Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. *Blood* 2012 Aug 2; **120**(5)**:** 1077-1086.

111. Egan JB, Shi CX, Tembe W, Christoforides A, Kurdoglu A, Sinari S*, et al.* Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood* 2012 Aug 2; **120**(5)**:** 1060-1066.

112. Attal M, Harousseau JL, Leyvraz S, Doyen C, Hulin C, Benboubker L*, et al.* Maintenance therapy with thalidomide improves survival in patients with multiple myeloma. *Blood* 2006 Nov 15; **108**(10)**:** 3289-3294.

113. Lokhorst HM, van der Holt B, Zweegman S, Vellenga E, Croockewit S, van Oers MH*, et al.* A randomized phase 3 study on the effect of thalidomide combined with adriamycin, dexamethasone, and high-dose melphalan, followed by thalidomide maintenance in patients with multiple myeloma. *Blood* 2010 Feb 11; **115**(6)**:** 1113-1120.

114. Kapoor P, Kumar S, Fonseca R, Lacy MQ, Witzig TE, Hayman SR*, et al.* Impact of risk stratification on outcome among patients with multiple myeloma receiving initial therapy with lenalidomide and dexamethasone. *Blood* 2009 Jul 16; **114**(3)**:** 518-521.

115. Stewart AK, Rajkumar SV, Dimopoulos MA, Masszi T, Spicka I, Oriol A*, et al.* Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. *The New England journal of medicine* 2015 Jan 8; **372**(2)**:** 142-152.

116. Leleu X, Karlin L, Macro M, Hulin C, Garderet L, Roussel M*, et al.* Pomalidomide plus low-dose dexamethasone in multiple myeloma with deletion 17p and/or translocation (4;14): IFM 2010-02 trial results. *Blood* 2015 Feb 26; **125**(9)**:** 1411-1417.

117. Dimopoulos MA, Weisel KC, Song KW, Delforge M, Karlin L, Goldschmidt H, *et al.* Cytogenetics and long-term survival of patients with refractory or relapsed and refractory multiple myeloma treated with pomalidomide and low-dose dexamethasone. *Haematologica* 2015 Oct; **100**(10)**:** 1327-1333.

118. Bensinger WI, Jagannath S, Vescio R, Camacho E, Wolf J, Irwin D, *et al.* Phase 2 study of two sequential three-drug combinations containing bortezomib, cyclophosphamide and dexamethasone, followed by bortezomib, thalidomide and dexamethasone as frontline therapy for multiple myeloma. *Br J Haematol* 2010 Feb; **148**(4)**:** 562-568.

119. Jagannath S, Durie BG, Wolf JL, Camacho ES, Irwin D, Lutzky J, *et al.* Extended follow-up of a phase 2 trial of bortezomib alone and in combination with dexamethasone for the frontline treatment of multiple myeloma. *Br J Haematol* 2009 Sep; **146**(6)**:** 619-626.

120. Jakubowiak AJ, Dytfeld D, Griffith KA, Lebovic D, Vesole DH, Jagannath S, *et al.* A phase 1/2 study of carfilzomib in combination with lenalidomide and low-dose dexamethasone as a frontline treatment for multiple myeloma. *Blood* 2012 Aug 30; **120**(9)**:** 1801-1809.

121. Jakubowiak AJ, Siegel DS, Martin T, Wang M, Vij R, Lonial S, *et al.* Treatment outcomes in patients with relapsed and refractory multiple myeloma and high-risk cytogenetics receiving single-agent carfilzomib in the PX-171-003-A1 study. *Leukemia* 2013 Dec; **27**(12)**:** 2351-2356.

122. Ludwig H, Miguel JS, Dimopoulos MA, Palumbo A, Garcia Sanz R, Powles R, *et al.* International Myeloma Working Group recommendations for global myeloma care. *Leukemia* 2014 May; **28**(5)**:** 981-992.

123. Mateos MV, Ocio EM, Paiva B, Rosinol L, Martinez-Lopez J, Blade J, *et al.* Treatment for patients with newly diagnosed multiple myeloma in 2015. *Blood Rev* 2015 Nov; **29**(6)**:** 387-403.

124. Sonneveld P, Goldschmidt H, Rosinol L, Blade J, Lahuerta JJ, Cavo M, *et al.* Bortezomib-based versus nonbortezomib-based induction treatment before autologous stem-cell transplantation in patients with previously untreated multiple myeloma: a meta-analysis of phase III randomized, controlled trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013 Sep 10; **31**(26)**:** 3279-3287.

125. Dimopoulos MA, Moreau P, Palumbo A, Joshua D, Pour L, Hajek R, *et al.* Carfilzomib and dexamethasone versus bortezomib and dexamethasone for patients with relapsed or refractory multiple myeloma (ENDEAVOR): a randomised, phase 3, open-label, multicentre study. *Lancet Oncol* 2016 Jan; **17**(1)**:** 27-38.

126. Mohty B, El-Cheikh J, Yakoub-Agha I, Avet-Loiseau H, Moreau P, Mohty M. Treatment strategies in relapsed and refractory multiple myeloma: a focus on drug

sequencing and 'retreatment' approaches in the era of novel agents. *Leukemia* 2012 Jan; **26**(1)**:** 73-85.

127. Jakubowiak A. Management strategies for relapsed/refractory multiple myeloma: current clinical perspectives. *Semin Hematol* 2012 Jul; **49 Suppl 1:** S16-32.

128. Kumar SK, Lee JH, Lahuerta JJ, Morgan G, Richardson PG, Crowley J*, et al.* Risk of progression and survival in multiple myeloma relapsing after therapy with IMiDs and bortezomib: a multicenter international myeloma working group study. *Leukemia* 2012 Jan; **26**(1)**:** 149-157.

129. Avet-Loiseau H, Leleu X, Roussel M, Moreau P, Guerin-Charbonnel C, Caillot D*, et al.* Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010 Oct 20; **28**(30)**:** 4630-4634.

130. Rasche L, Bernard C, Topp MS, Kapp M, Duell J, Wesemeier C*, et al.* Features of extramedullary myeloma relapse: high proliferation, minimal marrow involvement, adverse cytogenetics: a retrospective single-center study of 24 cases. *Ann Hematol* 2012 Jul; **91**(7)**:** 1031-1037.

131. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000 Jan 7; **100**(1)**:** 57-70.

132. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. *Cell* 2011 Oct 14; **147**(2)**:** 275-292.

133. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. *Nature reviews Cancer* 2002 Aug; **2**(8)**:** 563-572.

134. Luzzi KJ, MacDonald IC, Schmidt EE, Kerkvliet N, Morris VL, Chambers AF*, et al.* Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol* 1998 Sep; **153**(3)**:** 865-873.

135. Gupta GP, Massague J. Cancer metastasis: building a framework. *Cell* 2006 Nov 17; **127**(4)**:** 679-695.

136. Steeg PS. Tumor metastasis: mechanistic insights and clinical challenges. *Nat Med* 2006 Aug; **12**(8)**:** 895-904.

137. Minna JD, Kurie JM, Jacks T. A big step in the study of small cell lung cancer. *Cancer Cell* 2003 Sep; **4**(3)**:** 163-166.

138. Klein CA. The systemic progression of human cancer: a focus on the individual disseminated cancer cell--the unit of selection. *Adv Cancer Res* 2003; **89:** 35-67.

139. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. *Nature reviews Cancer* 2009 Apr; **9**(4)**:** 274-284.

140. Friedl P, Wolf K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nature reviews Cancer* 2003 May; **3**(5)**:** 362-374.

141. Pankova K, Rosel D, Novotny M, Brabek J. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cellular and molecular life sciences : CMLS* 2010 Jan; **67**(1)**:** 63-71.

142. Wan L, Pantel K, Kang Y. Tumor metastasis: moving new biological insights into the clinic. *Nat Med* 2013 Nov; **19**(11)**:** 1450-1464.

143. Brabletz T, Jung A, Reu S, Porzner M, Hlubek F, Kunz-Schughart LA*, et al.* Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proceedings of the National Academy of Sciences of the United States of America* 2001 Aug 28; **98**(18)**:** 10356-10361.

144. Thiery JP, Acloque H, Huang RYJ, Nieto MA. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* 2009; **139**(5)**:** 871-890.

145. Tsai JH, Yang J. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes & development* 2013 Oct 15; **27**(20)**:** 2192-2206.

146. Ocana OH, Corcoles R, Fabra A, Moreno-Bueno G, Acloque H, Vega S*, et al.* Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell* 2012 Dec 11; **22**(6)**:** 709-724.

147. Arnoux V, Nassour M, L'Helgoualc'h A, Hipskind RA, Savagner P. Erk5 controls Slug expression and keratinocyte activation during wound healing. *Molecular biology of the cell* 2008 Nov; **19**(11)**:** 4738-4749.

148. Iwano M, Plieth D, Danoff TM, Xue C, Okada H, Neilson EG. Evidence that fibroblasts derive from epithelium during tissue fibrosis. *J Clin Invest* 2002 Aug; **110**(3)**:** 341-350.

149. Zeisberg M, Yang C, Martino M, Duncan MB, Rieder F, Tanjore H*, et al.* Fibroblasts derive from hepatocytes in liver fibrosis via epithelial to mesenchymal transition. *The Journal of biological chemistry* 2007 Aug 10; **282**(32)**:** 23337-23347.

150. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation* 2009; **119**(6)**:** 1420-1428.

151. Bednarz-Knoll N, Alix-Panabieres C, Pantel K. Plasticity of disseminating cancer cells in patients with epithelial malignancies. *Cancer metastasis reviews* 2012 Dec; **31**(3-4)**:** 673-687.

152. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest* 2009 Jun; **119**(6)**:** 1420-1428.

153. Gilles C, Thompson EW. The Epithelial to Mesenchymal Transition and Metastatic Progression in Carcinoma. *The Breast Journal* 1996; **2**(1)**:** 83-96.

154. Thompson EW, Newgreen DF. Carcinoma Invasion and Metastasis: A Role for Epithelial-Mesenchymal Transition? *Cancer research* 2005; **65**(14)**:** 5991-5995.

155. Sanchez-Tillo E, Liu Y, de Barrios O, Siles L, Fanlo L, Cuatrecasas M, *et al.* EMT-activating transcription factors in cancer: beyond EMT and tumor invasiveness. *Cellular and molecular life sciences : CMLS* 2012 Oct; **69**(20)**:** 3429-3456.

156. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol* 2014; **15**(3)**:** 178-196.

157. Batlle E, Sancho E, Franci C, Dominguez D, Monfar M, Baulida J, *et al.* The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature cell biology* 2000 Feb; **2**(2)**:** 84-89.

158. Cano A, Perez-Moreno MA, Rodrigo I, Locascio A, Blanco MJ, del Barrio MG, *et al.* The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature cell biology* 2000 Feb; **2**(2)**:** 76-83.

159. Hajra KM, Chen DY, Fearon ER. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer research* 2002 Mar 15; **62**(6)**:** 1613-1618.

160. Eger A, Aigner K, Sonderegger S, Dampier B, Oehler S, Schreiber M, *et al.* DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene* 2005 Mar 31; **24**(14)**:** 2375-2385.

161. Spaderna S, Schmalhofer O, Wahlbuhl M, Dimmler A, Bauer K, Sultan A, *et al.* The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. *Cancer research* 2008 Jan 15; **68**(2)**:** 537-544.

162. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 2004 Jun 25; **117**(7)**:** 927-939.

163. Yang Z, Zhang X, Gang H, Li X, Li Z, Wang T, *et al.* Up-regulation of gastric cancer cell invasion by Twist is accompanied by N-cadherin and fibronectin expression. *Biochemical and biophysical research communications* 2007 Jul 6; **358**(3)**:** 925-930.

164. Eckert Mark A, Lwin Thinzar M, Chang Andrew T, Kim J, Danis E, Ohno-Machado L, *et al.* Twist1-Induced Invadopodia Formation Promotes Tumor Metastasis. *Cancer Cell* 2011; **19**(3)**:** 372-386.

165. Fu J, Qin L, He T, Qin J, Hong J, Wong J, *et al.* The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis. *Cell research* 2011 Feb; **21**(2)**:** 275-289.

166. Gonzalez DM, Medici D. Signaling mechanisms of the epithelial-mesenchymal transition. *Sci Signal* 2014 Sep 23; **7**(344)**:** re8.

167. Miettinen PJ, Ebner R, Lopez AR, Derynck R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *The Journal of cell biology* 1994 Dec; **127**(6 Pt 2)**:** 2021-2036.

168. Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. *Cell research* 2009 Feb; **19**(2)**:** 156-172.

169. Shirakihara T, Saitoh M, Miyazono K. Differential regulation of epithelial and mesenchymal markers by deltaEF1 proteins in epithelial mesenchymal transition induced by TGF-beta. *Molecular biology of the cell* 2007 Sep; **18**(9)**:** 3533-3544.

170. Derynck R, Zhang Y, Feng XH. Smads: transcriptional activators of TGF-beta responses. *Cell* 1998 Dec 11; **95**(6)**:** 737-740.

171. Hoodless PA, Haerry T, Abdollah S, Stapleton M, O'Connor MB, Attisano L*, et al.* MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell* 1996 May 17; **85**(4)**:** 489-500.

172. Chen D, Zhao M, Mundy GR. Bone morphogenetic proteins. *Growth Factors* 2004 Dec; **22**(4)**:** 233-241.

173. Yamashita H, ten Dijke P, Huylebroeck D, Sampath TK, Andries M, Smith JC*, et al.* Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects. *The Journal of cell biology* 1995 Jul; **130**(1)**:** 217-226.

174. Miele L, Golde T, Osborne B. Notch signaling in cancer. *Curr Mol Med* 2006 Dec; **6**(8)**:** 905-918.

175. Sahlgren C, Gustafsson MV, Jin S, Poellinger L, Lendahl U. Notch signaling mediates hypoxia-induced tumor cell migration and invasion. *Proceedings of the National Academy of Sciences of the United States of America* 2008 Apr 29; **105**(17)**:** 6392-6397.

176. Leong KG, Niessen K, Kulic I, Raouf A, Eaves C, Pollet I*, et al.* Jagged1-mediated Notch activation induces epithelial-to-mesenchymal transition through Slug-induced repression of E-cadherin. *The Journal of experimental medicine* 2007 Nov 26; **204**(12)**:** 2935-2948.

177. Timmerman LA, Grego-Bessa J, Raya A, Bertran E, Perez-Pomares JM, Diez J*, et al.* Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes & development* 2004 Jan 1; **18**(1)**:** 99-115.

178. Zhou BP, Hung MC. Wnt, hedgehog and snail: sister pathways that control by GSK-3beta and beta-Trcp in the regulation of metastasis. *Cell Cycle* 2005 Jun; **4**(6)**:** 772-776.

179. Zhou BP, Deng J, Xia W, Xu J, Li YM, Gunduz M, *et al.* Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nature cell biology* 2004 Oct; **6**(10)**:** 931-940.

180. Bachelder RE, Yoon SO, Franci C, de Herreros AG, Mercurio AM. Glycogen synthase kinase-3 is an endogenous inhibitor of Snail transcription: implications for the epithelial-mesenchymal transition. *The Journal of cell biology* 2005 Jan 3; **168**(1)**:** 29-33.

181. Howe LR, Watanabe O, Leonard J, Brown AM. Twist is up-regulated in response to Wnt1 and inhibits mouse mammary cell differentiation. *Cancer research* 2003 Apr 15; **63**(8)**:** 1906-1913.

182. Zhang J, Ma L. MicroRNA control of epithelial-mesenchymal transition and metastasis. *Cancer metastasis reviews* 2012 Dec; **31**(3-4)**:** 653-662.

183. Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* 2000 Mar 16; **404**(6775)**:** 293-296.

184. Hammond SM, Caudy AA, Hannon GJ. Post-transcriptional gene silencing by double-stranded RNA. *Nat Rev Genet* 2001 Feb; **2**(2)**:** 110-119.

185. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell* 2009 Jan 23; **136**(2)**:** 215-233.

186. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development* 2006 Mar 01; **20**(5)**:** 515-524.

187. Ameres SL, Martinez J, Schroeder R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* 2007 Jul 13; **130**(1)**:** 101-112.

188. Korpal M, Lee ES, Hu G, Kang Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *The Journal of biological chemistry* 2008 May 30; **283**(22)**:** 14910-14914.

189. Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes & development* 2008 Apr 1; **22**(7)**:** 894-907.

190. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nature cell biology* 2008 May; **10**(5)**:** 593-601.

191. Gregory PA, Bracken CP, Bert AG, Goodall GJ. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle* 2008 Oct; **7**(20)**:** 3112-3118.

192. Shinozaki A, Sakatani T, Ushiku T, Hino R, Isogai M, Ishikawa S, *et al.* Downregulation of microRNA-200 in EBV-associated gastric carcinoma. *Cancer research* 2010 Jun 1; **70**(11)**:** 4719-4727.

193. Wiklund ED, Bramsen JB, Hulf T, Dyrskjot L, Ramanathan R, Hansen TB, *et al.* Coordinated epigenetic repression of the miR-200 family and miR-205 in invasive bladder cancer. *Int J Cancer* 2011 Mar 15; **128**(6)**:** 1327-1334.

194. Hu X, Macdonald DM, Huettner PC, Feng Z, El Naqa IM, Schwarz JK, *et al.* A miR-200 microRNA cluster as prognostic marker in advanced ovarian cancer. *Gynecol Oncol* 2009 Sep; **114**(3)**:** 457-464.

195. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, *et al.* A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep* 2008 Jun; **9**(6)**:** 582-589.

196. Bracken CP, Gregory PA, Kolesnikoff N, Bert AG, Wang J, Shannon MF, *et al.* A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer research* 2008 Oct 1; **68**(19)**:** 7846-7854.

197. Gregory PA, Bracken CP, Smith E, Bert AG, Wright JA, Roslan S, *et al.* An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. *Molecular biology of the cell* 2011 May 15; **22**(10)**:** 1686-1698.

198. Trendowski M. The inherent metastasis of leukaemia and its exploitation by sonodynamic therapy. *Crit Rev Oncol Hematol* 2015 May; **94**(2)**:** 149-163.

199. Viadana E, Bross ID, Pickren JW. An autopsy study of the metastatic patterns of human leukemias. *Oncology* 1978; **35**(2)**:** 87-96.

200. Ghobrial IM. Myeloma as a model for the process of metastasis: implications for therapy. *Blood* 2012 Jul 5; **120**(1)**:** 20-30.

201. Rawstron AC, Owen RG, Davies FE, Johnson RJ, Jones RA, Richards SJ, *et al.* Circulating plasma cells in multiple myeloma: characterization and correlation with disease stage. *Br J Haematol* 1997 Apr; **97**(1)**:** 46-55.

202. Nowakowski GS, Witzig TE, Dingli D, Tracz MJ, Gertz MA, Lacy MQ, *et al.* Circulating plasma cells detected by flow cytometry as a predictor of survival in 302 patients with newly diagnosed multiple myeloma. *Blood* 2005 Oct 1; **106**(7)**:** 2276-2279.

203. Moller C, Stromberg T, Juremalm M, Nilsson K, Nilsson G. Expression and function of chemokine receptors in human multiple myeloma. *Leukemia* 2003 Jan; **17**(1)**:** 203-210.

204. Aggarwal R, Ghobrial IM, Roodman GD. Chemokines in multiple myeloma. *Exp Hematol* 2006 Oct; **34**(10)**:** 1289-1295.

205. Gazitt Y, Akay C. Mobilization of myeloma cells involves SDF-1/CXCR4 signaling and downregulation of VLA-4. *Stem Cells* 2004; **22**(1)**:** 65-73.

206. Alsayed Y, Ngo H, Runnels J, Leleu X, Singha UK, Pitsillides CM, *et al.* Mechanisms of regulation of CXCR4/SDF-1 (CXCL12)-dependent migration and homing in multiple myeloma. *Blood* 2007 Apr 1; **109**(7)**:** 2708-2717.

207. Roccaro AM, Mishima Y, Sacco A, Moschetta M, Tai YT, Shi J, *et al.* CXCR4 Regulates Extra-Medullary Myeloma through Epithelial-Mesenchymal-Transition-like Transcriptional Activation. *Cell Rep* 2015 Jul 28; **12**(4)**:** 622-635.

208. Zhang Z, Ni C, Chen W, Wu P, Wang Z, Yin J, *et al.* Expression of CXCR4 and breast cancer prognosis: a systematic review and meta-analysis. *BMC Cancer* 2014; **14:** 49.

209. Sobolik T, Su YJ, Wells S, Ayers GD, Cook RS, Richmond A. CXCR4 drives the metastatic phenotype in breast cancer through induction of CXCR2 and activation of MEK and PI3K pathways. *Molecular biology of the cell* 2014 Mar; **25**(5)**:** 566-582.

210. Singh S, Singh UP, Grizzle WE, Lillard JW, Jr. CXCL12-CXCR4 interactions modulate prostate cancer cell migration, metalloproteinase expression and invasion. *Lab Invest* 2004 Dec; **84**(12)**:** 1666-1676.

211. Gassmann P, Haier J, Schluter K, Domikowsky B, Wendel C, Wiesner U, *et al.* CXCR4 regulates the early extravasation of metastatic tumor cells in vivo. *Neoplasia* 2009 Jul; **11**(7)**:** 651-661.

212. Vande Broek I, Leleu X, Schots R, Facon T, Vanderkerken K, Van Camp B, *et al.* Clinical significance of chemokine receptor (CCR1, CCR2 and CXCR4) expression in human myeloma cells: the association with disease activity and survival. *Haematologica* 2006 Feb; **91**(2)**:** 200-206.

213. Vande Broek I, Asosingh K, Allegaert V, Leleu X, Facon T, Vanderkerken K, *et al.* Bone marrow endothelial cells increase the invasiveness of human multiple myeloma cells through upregulation of MMP-9: evidence for a role of hepatocyte growth factor. *Leukemia* 2004 May; **18**(5)**:** 976-982.

214. Van Valckenborgh E, Bakkus M, Munaut C, Noel A, St Pierre Y, Asosingh K, *et al.* Upregulation of matrix metalloproteinase-9 in murine 5T33 multiple myeloma cells by interaction with bone marrow endothelial cells. *Int J Cancer* 2002 Oct 20; **101**(6)**:** 512-518.

215. Asosingh K, Menu E, Van Valckenborgh E, Vande Broek I, Van Riet I, Van Camp B, *et al.* Mechanisms involved in the differential bone marrow homing of CD45 subsets in 5T murine models of myeloma. *Clin Exp Metastasis* 2002; **19**(7)**:** 583-591.

216. Parmo-Cabanas M, Molina-Ortiz I, Matias-Roman S, Garcia-Bernal D, Carvajal-Vergara X, Valle I*, et al.* Role of metalloproteinases MMP-9 and MT1-MMP in CXCL12-promoted myeloma cell invasion across basement membranes. *J Pathol* 2006 Jan; **208**(1)**:** 108-118.

217. Menu E, Asosingh K, Indraccolo S, De Raeve H, Van Riet I, Van Valckenborgh E*, et al.* The involvement of stromal derived factor 1alpha in homing and progression of multiple myeloma in the 5TMM model. *Haematologica* 2006 May; **91**(5)**:** 605-612.

218. Van Valckenborgh E, Croucher PI, De Raeve H, Carron C, De Leenheer E, Blacher S*, et al.* Multifunctional role of matrix metalloproteinases in multiple myeloma: a study in the 5T2MM mouse model. *Am J Pathol* 2004 Sep; **165**(3)**:** 869-878.

219. Yousef EM, Tahir MR, St-Pierre Y, Gaboury LA. MMP-9 expression varies according to molecular subtypes of breast cancer. *BMC Cancer* 2014; **14:** 609.

220. Varettoni M, Corso A, Pica G, Mangiacavalli S, Pascutto C, Lazzarino M. Incidence, presenting features and outcome of extramedullary disease in multiple myeloma: a longitudinal study on 1003 consecutive patients. *Ann Oncol* 2010 Feb; **21**(2)**:** 325-330.

221. Usmani SZ, Rodriguez-Otero P, Bhutani M, Mateos MV, Miguel JS. Defining and treating high-risk multiple myeloma. *Leukemia* 2015 Nov; **29**(11)**:** 2119-2125.

222. Thomas FB, Clausen KP, Greenberger NJ. Liver disease in multiple myeloma. *Arch Intern Med* 1973 Aug; **132**(2)**:** 195-202.

223. Pour L, Sevcikova S, Greslikova H, Kupska R, Majkova P, Zahradova L*, et al.* Soft-tissue extramedullary multiple myeloma prognosis is significantly worse in comparison to bone-related extramedullary relapse. *Haematologica* 2014 Feb; **99**(2)**:** 360-364.

224. Sevcikova S, Paszekova H, Besse L, Sedlarikova L, Kubaczkova V, Almasi M*, et al.* Extramedullary relapse of multiple myeloma defined as the highest risk group based on deregulated gene expression data. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2015 Jun; **159**(2)**:** 288-293.

225. Gonsalves WI, Morice WG, Rajkumar V, Gupta V, Timm MM, Dispenzieri A*, et al.* Quantification of clonal circulating plasma cells in relapsed multiple myeloma. *Br J Haematol* 2014 Nov; **167**(4)**:** 500-505.

226. Gonsalves WI, Rajkumar SV, Gupta V, Morice WG, Timm MM, Singh PP*, et al.* Quantification of clonal circulating plasma cells in newly diagnosed multiple myeloma: implications for redefining high-risk myeloma. *Leukemia* 2014 Oct; **28**(10)**:** 2060-2065.

227. Witzig TE, Gertz MA, Lust JA, Kyle RA, O'Fallon WM, Greipp PR. Peripheral blood monoclonal plasma cells as a predictor of survival in patients with multiple myeloma. *Blood* 1996 Sep 1; **88**(5)**:** 1780-1787.

228. Besse L, Sedlarikova L, Greslikova H, Kupska R, Almasi M, Penka M*, et al.* Cytogenetics in multiple myeloma patients progressing into extramedullary disease. *Eur J Haematol* 2016 Jul; **97**(1)**:** 93-100.

229. Qu X, Chen L, Qiu H, Lu H, Wu H, Qiu H*, et al.* Extramedullary manifestation in multiple myeloma bears high incidence of poor cytogenetic aberration and novel agents resistance. *Biomed Res Int* 2015; **2015:** 787809.

230. Billecke L, Murga Penas EM, May AM, Engelhardt M, Nagler A, Leiba M*, et al.* Cytogenetics of extramedullary manifestations in multiple myeloma. *Br J Haematol* 2013 Apr; **161**(1)**:** 87-94.

231. Billecke L, Penas EM, May AM, Engelhardt M, Nagler A, Leiba M*, et al.* Similar incidences of TP53 deletions in extramedullary organ infiltrations, soft tissue and osteolyses of patients with multiple myeloma. *Anticancer Res* 2012 May; **32**(5)**:** 2031-2034.

232. Minnema MC, van de Donk NW, Zweegman S, Hegenbart U, Schonland S, Raymakers R*, et al.* Extramedullary relapses after allogeneic non-myeloablative stem cell transplantation in multiple myeloma patients do not negatively affect treatment outcome. *Bone Marrow Transplant* 2008 May; **41**(9)**:** 779-784.

233. Chiecchio L, Protheroe RK, Ibrahim AH, Cheung KL, Rudduck C, Dagrada GP*, et al.* Deletion of chromosome 13 detected by conventional cytogenetics is a critical prognostic factor in myeloma. *Leukemia* 2006 Sep; **20**(9)**:** 1610-1617.

234. Katodritou E, Gastari V, Verrou E, Hadjiaggelidou C, Varthaliti M, Georgiadou S*, et al.* Extramedullary (EMP) relapse in unusual locations in multiple myeloma: Is there an association with precedent thalidomide administration and a correlation of special biological features with treatment and outcome? *Leuk Res* 2009 Aug; **33**(8)**:** 1137-1140.

235. Nieto MA. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science* 2013 Nov 8; **342**(6159)**:** 1234850.

236. Fabregat I, Malfettone A, Soukupova J. New Insights into the Crossroads between EMT and Stemness in the Context of Cancer. *J Clin Med* 2016; **5**(3).

237. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY*, et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008 May 16; **133**(4)**:** 704-715.

238. Morel AP, Lievre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS one* 2008; **3**(8)**:** e2888.

239. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 2003 Apr 1; **100**(7)**:** 3983-3988.

240. Ponti D, Costa A, Zaffaroni N, Pratesi G, Petrangolini G, Coradini D*, et al.* Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer research* 2005 Jul 1; **65**(13)**:** 5506-5511.

241. Greipp PR, Raymond NM, Kyle RA, O'Fallon WM. Multiple myeloma: significance of plasmablastic subtype in morphological classification. *Blood* 1985 Feb; **65**(2)**:** 305-310.

242. Bartl R, Frisch B, Burkhardt R, Fateh-Moghadam A, Mahl G, Gierster P*, et al.* Bone marrow histology in myeloma: its importance in diagnosis, prognosis, classification and staging. *Br J Haematol* 1982 Jul; **51**(3)**:** 361-375.

243. Bartl R, Frisch B, Fateh-Moghadam A, Kettner G, Jaeger K, Sommerfeld W. Histologic classification and staging of multiple myeloma. A retrospective and prospective study of 674 cases. *Am J Clin Pathol* 1987 Mar; **87**(3)**:** 342-355.

244. Rajkumar SV, Fonseca R, Lacy MQ, Witzig TE, Therneau TM, Kyle RA*, et al.* Plasmablastic morphology is an independent predictor of poor survival after autologous stem-cell transplantation for multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 1999 May; **17**(5)**:** 1551-1557.

245. Goasguen JE, Zandecki M, Mathiot C, Scheiff JM, Bizet M, Ly-Sunnaram B*, et al.* Mature plasma cells as indicator of better prognosis in multiple myeloma. New methodology for the assessment of plasma cell morphology. *Leuk Res* 1999 Dec; **23**(12)**:** 1133-1140.

246. Srija M, Zachariah PP, Unni VN, Mathew A, Rajesh R, Kurian G*, et al.* Plasmablastic myeloma presenting as rapidly progressive renal failure in a young adult. *Indian J Nephrol* 2014 Jan; **24**(1)**:** 41-44.

247. Lee CK, Ma ES, Shek TW, Lam CC, Au WY, Wan TS*, et al.* Plasmablastic transformation of multiple myeloma. *Hum Pathol* 2003 Jul; **34**(7)**:** 710-714.

248. Fassas ABT, Muwalla F, Berryman T, Benramdane R, Joseph L, Anaissie E*, et al.* Myeloma of the central nervous system: association with high-risk chromosomal abnormalities, plasmablastic morphology and extramedullary manifestations. *British Journal of Haematology* 2002; **117**(1)**:** 103-108.

249. Cerny J, Fadare O, Hutchinson L, Wang SA. Clinicopathological features of extramedullary recurrence/relapse of multiple myeloma. *Eur J Haematol* 2008 Jul; **81**(1)**:** 65-69.

250. Balleari E, Ghio R, Falcone A, Musto P. Possible multiple myeloma dedifferentiation following thalidomide therapy: a report of four cases. *Leukemia & lymphoma* 2004 Apr; **45**(4)**:** 735-738.

251. Kawano Y, Kikukawa Y, Fujiwara S, Wada N, Okuno Y, Mitsuya H*, et al.* Hypoxia reduces CD138 expression and induces an immature and stem cell-like transcriptional program in myeloma cells. *Int J Oncol* 2013 Dec; **43**(6)**:** 1809-1816.

252. Muz B, de la Puente P, Azab F, Luderer M, Azab AK. Hypoxia promotes stem cell-like phenotype in multiple myeloma cells. *Blood cancer journal* 2014; **4:** e262.

253. Yaccoby S. The phenotypic plasticity of myeloma plasma cells as expressed by dedifferentiation into an immature, resilient, and apoptosis-resistant phenotype. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2005 Nov 1; **11**(21)**:** 7599-7606.

254. Cates JM, Dupont WD, Barnes JW, Edmunds HS, Fasig JH, Olson SJ*, et al.* Markers of epithelial-mesenchymal transition and epithelial differentiation in sarcomatoid carcinoma: utility in the differential diagnosis with sarcoma. *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* 2008 May; **16**(3)**:** 251-262.

255. Pioli PD, Weis JH. Snail transcription factors in hematopoietic cell development: a model of functional redundancy. *Exp Hematol* 2014 Jun; **42**(6)**:** 425-430.

256. Goossens S, Janzen V, Bartunkova S, Yokomizo T, Drogat B, Crisan M*, et al.* The EMT regulator Zeb2/Sip1 is essential for murine embryonic hematopoietic stem/progenitor cell differentiation and mobilization. *Blood* 2011 May 26; **117**(21)**:** 5620-5630.

257. Caudell D, Harper DP, Novak RL, Pierce RM, Slape C, Wolff L*, et al.* Retroviral insertional mutagenesis identifies Zeb2 activation as a novel leukemogenic collaborating event in CALM-AF10 transgenic mice. *Blood* 2010 Feb 11; **115**(6)**:** 1194-1203.

258. Perez-Mancera PA, Gonzalez-Herrero I, Perez-Caro M, Gutierrez-Cianca N, Flores T, Gutierrez-Adan A*, et al.* SLUG in cancer development. *Oncogene* 2005 Apr 28; **24**(19)**:** 3073-3082.

259. Perez-Mancera PA, Perez-Caro M, Gonzalez-Herrero I, Flores T, Orfao A, de Herreros AG*, et al.* Cancer development induced by graded expression of Snail in mice. *Human molecular genetics* 2005 Nov 15; **14**(22)**:** 3449-3461.

260. Dong M, Blobe GC. Role of transforming growth factor-beta in hematologic malignancies. *Blood* 2006 Jun 15; **107**(12)**:** 4589-4596.

261. Gu Y, Masiero M, Banham AH. Notch signaling: its roles and therapeutic potential in hematological malignancies. *Oncotarget* 2016 Feb 26.

262. Kim Y, Thanendrarajan S, Schmidt-Wolf IG. Wnt/ss-catenin: a new therapeutic approach to acute myeloid leukemia. *Leuk Res Treatment* 2011; **2011:** 428960.

263. Urashima M, Ogata A, Chauhan D, Hatziyanni M, Vidriales MB, Dedera DA, *et al.* Transforming growth factor-beta1: differential effects on multiple myeloma versus normal B cells. *Blood* 1996 Mar 1; **87**(5)**:** 1928-1938.

264. Hayashi T, Hideshima T, Nguyen AN, Munoz O, Podar K, Hamasaki M, *et al.* Transforming growth factor beta receptor I kinase inhibitor down-regulates cytokine secretion and multiple myeloma cell growth in the bone marrow microenvironment. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004 Nov 15; **10**(22)**:** 7540-7546.

265. Podar K, Tai YT, Davies FE, Lentzsch S, Sattler M, Hideshima T, *et al.* Vascular endothelial growth factor triggers signaling cascades mediating multiple myeloma cell growth and migration. *Blood* 2001 Jul 15; **98**(2)**:** 428-435.

266. Hov H, Tian E, Holien T, Holt RU, Vatsveen TK, Fagerli UM, *et al.* c-Met signaling promotes IL-6-induced myeloma cell proliferation. *Eur J Haematol* 2009 Apr; **82**(4)**:** 277-287.

267. Arendt BK, Velazquez-Dones A, Tschumper RC, Howell KG, Ansell SM, Witzig TE, *et al.* Interleukin 6 induces monocyte chemoattractant protein-1 expression in myeloma cells. *Leukemia* 2002 Oct; **16**(10)**:** 2142-2147.

268. Mirandola L, Apicella L, Colombo M, Yu Y, Berta DG, Platonova N, *et al.* Anti-Notch treatment prevents multiple myeloma cells localization to the bone marrow via the chemokine system CXCR4/SDF-1. *Leukemia* 2013 Jul; **27**(7)**:** 1558-1566.

269. Qiang YW, Walsh K, Yao L, Kedei N, Blumberg PM, Rubin JS, *et al.* Wnts induce migration and invasion of myeloma plasma cells. *Blood* 2005 Sep 1; **106**(5)**:** 1786-1793.

270. Azab AK, Hu J, Quang P, Azab F, Pitsillides C, Awwad R, *et al.* Hypoxia promotes dissemination of multiple myeloma through acquisition of epithelial to mesenchymal transition-like features. *Blood* 2012 June 14, 2012; **119**(24)**:** 5782-5794.

271. Sun Y, Pan J, Mao S, Jin J. IL-17/miR-192/IL-17Rs regulatory feedback loop facilitates multiple myeloma progression. *PloS one* 2014; **9**(12)**:** e114647.

272. Li J, Pan Q, Rowan PD, Trotter TN, Peker D, Regal KM, *et al.* Heparanase promotes myeloma progression by inducing mesenchymal features and motility of myeloma cells. *Oncotarget* 2016 Mar 8; **7**(10)**:** 11299-11309.

273. Gu K, Li MM, Shen J, Liu F, Cao JY, Jin S, *et al.* Interleukin-17-induced EMT promotes lung cancer cell migration and invasion via NF-kappaB/ZEB1 signal pathway. *Am J Cancer Res* 2015; **5**(3)**:** 1169-1179.

274. Ilan N, Elkin M, Vlodavsky I. Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis. *Int J Biochem Cell Biol* 2006; **38**(12)**:** 2018-2039.

275. Sanderson RD, Yang Y, Suva LJ, Kelly T. Heparan sulfate proteoglycans and heparanase--partners in osteolytic tumor growth and metastasis. *Matrix Biol* 2004 Oct; **23**(6)**:** 341-352.

276. Purushothaman A, Uyama T, Kobayashi F, Yamada S, Sugahara K, Rapraeger AC*, et al.* Heparanase-enhanced shedding of syndecan-1 by myeloma cells promotes endothelial invasion and angiogenesis. *Blood* 2010 Mar 25; **115**(12)**:** 2449-2457.

277. Lauring J, Abukhdeir AM, Konishi H, Garay JP, Gustin JP, Wang Q*, et al.* The multiple myeloma–associated MMSET gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. *Blood* 2008 January 15, 2008; **111**(2)**:** 856-864.

278. Huang Z, Wu H, Chuai S, Xu F, Yan F, Englund N*, et al.* NSD2 is recruited through its PHD domain to oncogenic gene loci to drive multiple myeloma. *Cancer research* 2013 Oct 15; **73**(20)**:** 6277-6288.

279. Yang S, Zhang Y, Meng F, Liu Y, Xia B, Xiao M*, et al.* Overexpression of multiple myeloma SET domain (MMSET) is associated with advanced tumor aggressiveness and poor prognosis in serous ovarian carcinoma. *Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals* 2013 May; **18**(3)**:** 257-263.

280. Dring AM, Davies FE, Fenton JA, Roddam PL, Scott K, Gonzalez D*, et al.* A global expression-based analysis of the consequences of the t(4;14) translocation in myeloma. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004 Sep 1; **10**(17)**:** 5692-5701.

281. Mrozik KM, Cheong CM, Hewett D, Chow AW, Blaschuk OW, Zannettino AC*, et al.* Therapeutic targeting of N-cadherin is an effective treatment for multiple myeloma. *Br J Haematol* 2015 Nov; **171**(3)**:** 387-399.

282. Kwok WK, Ling MT, Lee TW, Lau TC, Zhou C, Zhang X*, et al.* Up-regulation of TWIST in prostate cancer and its implication as a therapeutic target. *Cancer research* 2005 Jun 15; **65**(12)**:** 5153-5162.

283. Zhuo WL, Wang Y, Zhuo XL, Zhang YS, Chen ZT. Short interfering RNA directed against TWIST, a novel zinc finger transcription factor, increases A549 cell sensitivity to cisplatin via MAPK/mitochondrial pathway. *Biochemical and biophysical research communications* 2008 May 16; **369**(4)**:** 1098-1102.

Chapter 2:
MMSET promotes the acquisition of an epithelial-to-mesenchymal transition-like gene expression signature in t(4;14) multiple myeloma

Statement of Authorship

Title of Paper	MMSET promotes the acquisition of an epithelial-to-mesenchymal transition-like gene expression signature in t(4;14) multiple myeloma
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Chee Man Cheong, Kate Vandyke, Krzysztof Mrozik, Chung H. Kok & Andrew C.W. Zannettino (2017) MMSET promotes the acquisition of an epithelial-to-mesenchymal transition-like gene expression signature in t(4;14) multiple myeloma

Principal Author

Name of Principal Author (Candidate)	Chee Man Cheong
Contribution to the Paper	Performed experiments, analysed data and wrote manuscript
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 23/12/2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Kate Vandyke
Contribution to the Paper	Designed experiments, data interpretation, bioinformatic analysis, co-supervised development of work and manuscript evaluation
Signature	Date 23/12/2016

Name of Co-Author	Chung Hoow Kok
Contribution to the Paper	Bioinformatic analysis, statistical analysis and data interpretation
Signature	Date 23/12/2016

Name of Co-Author	Krzysztof Mrozik		
Contribution to the Paper	Data interpretation and performed experiments		
Signature		Date	23/12/16

Name of Co-Author	Andrew Zannettino		
Contribution to the Paper	Designed experiments, data interpretation, manuscript revision, provided funding and supervised development of work		
Signature		Date	23/12/2016

Please cut and paste additional co-author panels here as required.

Chapter 2: MMSET promotes the acquisition of an epithelial-to-mesenchymal transition-like gene expression signature in t(4;14) multiple myeloma

Chee Man Cheong[1,2], Kate Vandyke[1,2,3], Chung How Kok[2], Krzysztof Mrozik[1,2], Andrew Zannettino[1,2,3,4]

1. Myeloma Research Group, Discipline of Physiology, Adelaide Medical School, The University of Adelaide, Adelaide, Australia.
2. Cancer Theme, South Australian Health & Medical Research Institute (SAHMRI), Adelaide, Australia.
3. SA Pathology, Adelaide, Australia.
4. Centre for Cancer Biology, University of South Australia, Adelaide, Australia.

Keywords: t(4;14) multiple myeloma, MMSET, EMT, gene expression profile, TWIST1

2.1 Abstract

The t(4;14) translocation is a recurrent chromosomal translocation, present in 15-20% of multiple myeloma (MM) patients and results in overexpression of fibroblast growth factor receptor 3 (FGFR3) and MM SET domain (MMSET/WHSC1). The t(4;14) translocation confers intermediate- to high-risk prognostic features and is associated with rapid disease onset, rapid relapse and poor survival. While studies show that t(4;14) MM is associated with a highly aggressive, migratory phenotype, the precise mechanisms by which the t(4;14) translocation mediates these features remains to be determined. Using previously published gene expression profile (GEP) data from newly diagnosed MM patients and transcriptomic data from human MM cell lines, we have identified 17 mesenchymal markers overexpressed in MMSET-high MM plasma cells. Notably, enrichment of epithelial-to-mesenchymal transition (EMT)-markers is more prominent in MMSET-high human MM cell lines compared to MMSET-high BM-resident MM cells from patients. Among the mesenchymal markers upregulated, TWIST1 expression correlates strongly with MMSET expression. TWIST1 is a key transcription factor implicated in activating EMT and invasion in a number of solid tumours. Importantly, *TWIST1* expression was reduced in MM plasma cells when MMSET was knocked down and increased when MMSET was re-expressed. These findings suggests a critical interplay between MMSET and TWIST1, which may play a role in promoting aggressive disease phenotype in t(4;14)-positive MM.

2.2 Introduction

Multiple myeloma (MM) is characterized by uncontrolled clonal proliferation of malignant plasma cells within the bone marrow (BM). The presence of multiple tumour lesions within BM of MM patients at diagnosis suggest that MM cells continuously disseminate to new distant BM sites during disease progression. The interaction between clonal MM PCs and cells of the BM microenvironment plays a crucial role in supporting the pathogenesis of MM.[1,2] Recent studies suggest that BM microenvironment-derived factors promote metastasis and dissemination of MM plasma cells via activation of an epithelial-to-mesenchymal transition (EMT)-like program in MM PCs.[3-5] However, whether intrinsic genetic abnormalities within the MM PCs contribute to acquisition of EMT-like features in MM plasma cells remains to be explored.

MM is a genetically complex and heterogeneous with respect to clinical response to therapy and at the intrinsic molecular level. The majority of the chromosomal abnormalities in MM involve recurrent translocations with the immunoglobulin heavy chain (IGH) locus at 14q32. The reciprocal chromosomal translocation of t(4;14) is seen in approximately 15% of MM patients[6-8] and is associated with poor prognosis and high risk of relapse.[9-11] The poor prognosis of t(4;14) MM is associated with acquisition of a highly aggressive, migratory phenotype, as evidenced by an increased incidence of circulating tumour cells and extramedullary dissemination.[12-14] The t(4;14) translocation results in the simultaneous dysregulation of two genes, namely *FGFR3* and *MMSET* [Wolf-Hirschhorn syndrome candidate 1 (*WHSC1*)]. While the MMSET transcript is overexpressed in all cases of t(4;14) MM, approximately 30% of them lack FGFR3 expression.[9,15,16] High throughput DNA microarrays also confirm consistent overexpression of MMSET and FGFR3 in 100% and 75% of MM cases, respectively. For example, MMSET expression is elevated in the MS subgroup, the molecular subgroup of MM patients characterised by UAMS criteria that overexpresses MMSET and harbour the t(4;14) translocation.[17] Notably, the poor prognosis associated with the t(4;14) translocation is independent of FGFR3 expression [15], suggesting that MMSET is the key oncogenic factor in this subgroup of MM patients.

The MMSET protein, also known as WHSC1 protein or nuclear receptor-binding SET domain 2 (NSD2), is encoded by *WHSC1* gene. MMSET is a histone

methyltransferase (HMT) that catalyses the methylation of lysine residues on both histone H3 and histone H4.[18,19] Overexpression of MMSET results in a global increase in H3K36me2 and a decrease in H3K27me3 across the genome, leading to changes in the chromatin structure allowing it to be more accessible to active transcription.[20,21] H3K36me2 has been suggested as the principal chromatin regulatory activity of MMSET that leads to oncogenic transformation in primary plasma cells[18], by modulating the efficiency of DNA repair and chemotherapy response.[22] In addition, MMSET regulates gene expression indirectly through recruitment of other histone modifying enzymes such as histone deacetylase 1 (HDAC1), HDAC 2 and histone demethylase LSD1 [23] or competes with other HMT such as EZH2.[21]

Elevated MMSET expression is also observed in a range of more aggressive, metastatic advanced solid tumours such as neuroblastoma, hepatocellular and ovarian carcinomas.[24] Recently, MMSET was identified as the novel driver of EMT in prostate cancer mediated by TWIST1 upregulation.[25] TWIST1 is a key transcription factor implicated in activating EMT and invasion.[26] Interestingly, elevation of mesenchymal markers N-cadherin and vimentin have been reported in t(4;14) MM patients.[27,28] However, whether MMSET activates an EMT-like programme as part of the global transcription changes that leads to an aggressive phenotype remains to be determined.

We hypothesise that MMSET overexpression promotes acquisition of an EMT-like phenotype in MM cells leading to the aggressive disease features of t(4;14) MM. In this study, we took advantage of the previously published GEPs from MM patients and human MM cell lines (HMCL) to determine whether the expression of EMT markers are regulated by MMSET.

2.3 Materials and Methods

Microarray datasets

Unprocessed CEL files used in meta-analysis were assembled from the following datasets: GSE19784 (n = 328)[29], GSE26760 (n = 304)[30], MTAB-363 (n = 226)[31] and MTAB-317 (n = 156) and downloaded from NCBI Gene Expression Omnibus(GEO) and ArrayExpress. All gene expression data were derived from CD138+ purified plasma cells of newly diagnosed MM patients, which were hybridized to Affymetrix Human Genome U133 Plus 2.0 cDNA microarray. All raw CEL files were normalised using the RMA function in the R statistical programming language, and gene expression levels were log2 transformed. One patient in E-MTAB-363 (V0681) failed

quality control (normalised unscaled standard error [NUSE] >1.05) and was excluded, and the remaining 165 files were re-normalised. The final dataset included GEPs of 1012 MM patients. For GSE19784 and GSE26760, UAMS molecular classification status[17] and EMT-status[32] were available for 320 and 225 patients respectively. Survival analysis of *MMSET* and *TWIST1* expressing MM patients were performed using GEPs of MM patients enrolled in total therapy 2 (TT2) trial, who were classified as MS subgroup (n = 68; GSE4851[17]). MAS5-normalised data in GSE4851 were downloaded from the Gene Expression Omnibus (GEO) and were $\log_2$ normalised prior to analysis.

Classification of MM patients

For meta-analysis, patients were classified into MMSET-low or MMSET-high groups based on the relationships of *WHSC1* (209053_s_at) and *FGFR3* (409053_s_at) expression levels in scatterplots (supplementary Figure S2.1). Samples that were clustered in Quadrant I and IV were classified as MMSET-high, while those clustered in Quadrant II and III were classified as MMSET-low MM. Expression thresholds were assigned to MMSET and FGFR3, where 30% of the MMSET-high population expressed low levels of FGFR3, corresponding to the frequency of FGFR3 overexpression observed t(4;14) MM.[9,15,16]

Human MM cell lines (HMCL) transcriptome data

Pre-processed expression file of 66 cell lines was downloaded from https://sites.google.com/site/jonathankeatslab/data-repository. HMCL were stratified based on associated genetic information. HMCLs with t(4;14) translocation were classified as MMSET-high while non-t(4;14) HMCLs were classified as MMSET-low. Differentially expressed genes were determined using quasi-likelihood F-test from edgeR package. Significantly regulated genes were determined using cut-off of 1-fold or greater changes in mean expression and a FDR < 0.05.

Statistical analysis

Differentially expressed gene probes and their corresponding p-values of patients GEPs were determined using LIMMA package. The p-values for each gene probes across four studies were then combined using Fisher's method. The significantly dysregulated gene probes were determined using cut-off of 1-fold or greater changes in mean expression and combined p-value < 0.01. Survival curves were compared using

the log-rank (Mantel-Cox) test with hazard ratios calculated using the Mantel-Haenszel calculation. KMS11 microarray and reverse transcription-quantitative PCR (RT-qPCR) data were analysed by two-way ANOVA with Sidak's multiple comparison tests or compared between groups using unpaired two-tailed *t* tests. A *p* value of 0.05 was considered statistically significant.

Cell culture

Human MM cell lines LP-1 and RPMI-8226 were purchased from the American Type Culture Collection and grown in RPMI-1640 supplemented with 10% (v/v) fetal calf serum and additives (2mM L-glutamine, 1mM sodium pyruvate, 15 mM HEPES, 50 U/ml penicillin and 50 μg/ml streptomycin; from Sigma-Aldrich, Sydney, Australia).

Generation of MMSET-overexpressing RPMI-8226 cells

The empty vector pRetroX-DsRed or pRetroX-MMSET-DsRed (from Prof. Jonathan Licht, Northwestern University, Chicago[25]) harboring full-length MMSET cDNA were transfected into the HEK293T cells together with pGP (which encodes for Pol and GAG proteins) and pVSVG (which encodes for the viral envelope protein). Retroviral supernatant collected from transiently transfected HEK293T cells was used to transduce RPMI-8226 cells as previously described.[33] Stable MMSET-overexpressing RPMI-8226 (RPMI8226-MMSET) or empty vector control (RPMI8226-EV) cell lines were generated from the top 10% DsRed-positive cells sorted using BD FACSFusion. Overexpression of MMSET was validated by RT-qPCR and immunoblotting (supplementary Figure S2.2).

Immunoblotting

Nuclear proteins were extracted using nuclear complex Co-IP kit (Active Motif, Carlsbad, USA). Protein were resolved in 10% SDS-polyacrylamide gels, transferred to PVDF membranes (GE Healthcare), blotted with primary antibodies, followed by alkaline phosphatase-conjugated secondary antibodies (Chemicon, Merck Milipore, Billeria, USA) and detected using enhanced chemifluorescence substrate (ECF, GE Healthcare). Primary antibodies were: MMSET (ab73259, Abcam, Cambridge, UK), TWIST1 (ab50887, Abcam), and HSP90 (sc-7947, Santa-Cruz, Dallas, USA).

Reverse transcription-quantitative polymerase chain reaction (RT-qPCR)

Total RNA was isolated from cell lines using TRIzol (Life Technologies, Scorsby, Australia). RNA (1 μg) was reverse transcribed to cDNA using Superscript III (Invitrogen, Life Technologies). qPCR was performed using RT2 SYBR Green qPCR Mastermix (Qiagen, Chadstone, Ausrtalia) on a Bio-Rad CFX Connect (Bio-Rad, Hercules, USA) using the following primers: human *GAPDH* (F: 5'-ACCCAGAAGACTGTGGATGG-3'; R: 5'-CAGTGAGCTTCCCGTTCAG-3'), human MMSET (F: 5'- AGAGGATACAGGACCCTACA-3'; R: 5'-GTGTTTCGTCTGCACTTTCG-3') and human TWIST1 (F: 5'-TCTTACGAGGAGCTGCAGACGCA-3'; R: 5'-ATCTTGGAGTCCAGCTCGTCGCT-3'). Gene expression was represented relative to GAPDH expression, calculated using the $2^{-\Delta CT}$ method.

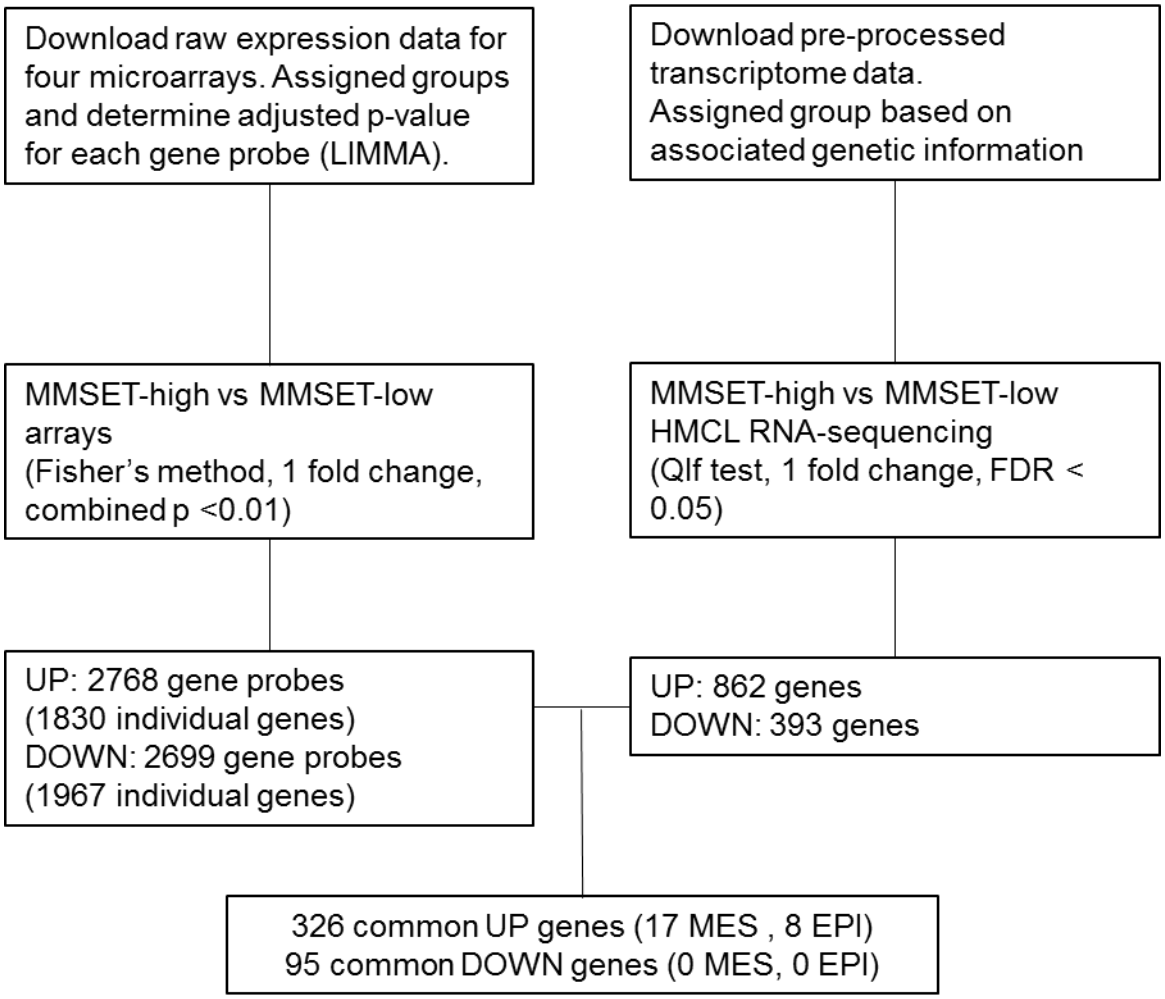
2.4 Results

2.4.1 Genes differentially regulated by MMSET in MM

To identify potential MMSET-driven EMT markers that were differentially expressed in MM patients, we interrogated four publicly available microarrays examining the GEP of CD138+-selected BMMNC from newly diagnosed MGUS, MM and healthy, age-matched controls. GEP of MM patients from each dataset were classified into MMSET-high or MMSET-low based on quadrant analysis performed using MMSET and FGFR3 expression levels as detailed in Materials and Methods section (supplementary Figure S2.1). Among the 1012 MM patients included in this meta-analysis, 136 (13.4%) were classified as MMSET-high. Linear Model for Microarray Analysis (LIMMA) was subsequently used to compute p-values for each gene probes between the MMSET-high and MMSET-low groups. In order to integrate the data from the four microarrays and to minimise variability among the independent datasets, the adjusted p-values calculated from the four individual microarrays were combined using Fisher's method.[34,35] A total of 2768 gene probes were significantly upregulated (combined p-value < 0.01, median $\log_2$ fold change > 0) and 2699 gene probes were significantly downregulated (combined p-value < 0.01, median $\log_2$ fold change < 0) in MMSET-high patient group, which correspond to 1830 and 1967 individual genes respectively (Figure 2.1).

To validate the MMSET-high signature identified in our meta-analysis, we compared our list with existing MMSET-regulated genes characterized by Wu et al.

Figure 2.1 Identification of mesenchymal markers overexpressed in MMSET-high and t(4;14)-positive human MM cell lines. (A) Gene expression profiles from four microarrays were compared between MMSET MMSET-high and MMSET MMSET-low patients. The Transcriptome transcriptomes from of 66 human MM cell lines were classified into t(4;14)-positive or negative based on their genetic information. 1830 and 1967 unique genes were upregulated and downregulated in MMSET-high patients, respectively. Among the 1830 uniquely upregulated genes, 326 were also upregulated in t(4;14)-positive cell lines and 17 of them were mesenchymal markers.


Download raw expression data for four microarrays. Assigned groups and determine adjusted p-value for each gene probe (LIMMA).
Download pre-processed transcriptome data. Assigned group based on associated genetic information
MMSET-high vs MMSET-low arrays (Fisher's method, 1 fold change, combined p <0.01)
MMSET-high vs MMSET-low HMCL RNA-sequencing (Qlf test, 1 fold change, FDR < 0.05)
UP: 2768 gene probes (1830 individual genes) DOWN: 2699 gene probes (1967 individual genes)
UP: 862 genes DOWN: 393 genes
326 common UP genes (17 MES , 8 EPI) 95 common DOWN genes (0 MES, 0 EPI)

(Wu-list), which was used to define signatures highly correlated with MMSET and associated with inferior survival in non-t(4;14) MM patients.[36] Among the differentially regulated gene probes in MMSET-high patients, 43 of them overlapped with the 71 genes in the Wu-list (supplementary Table S2.1), confirming our approach in identifying genes involved in the pathogenesis of MMSET MM

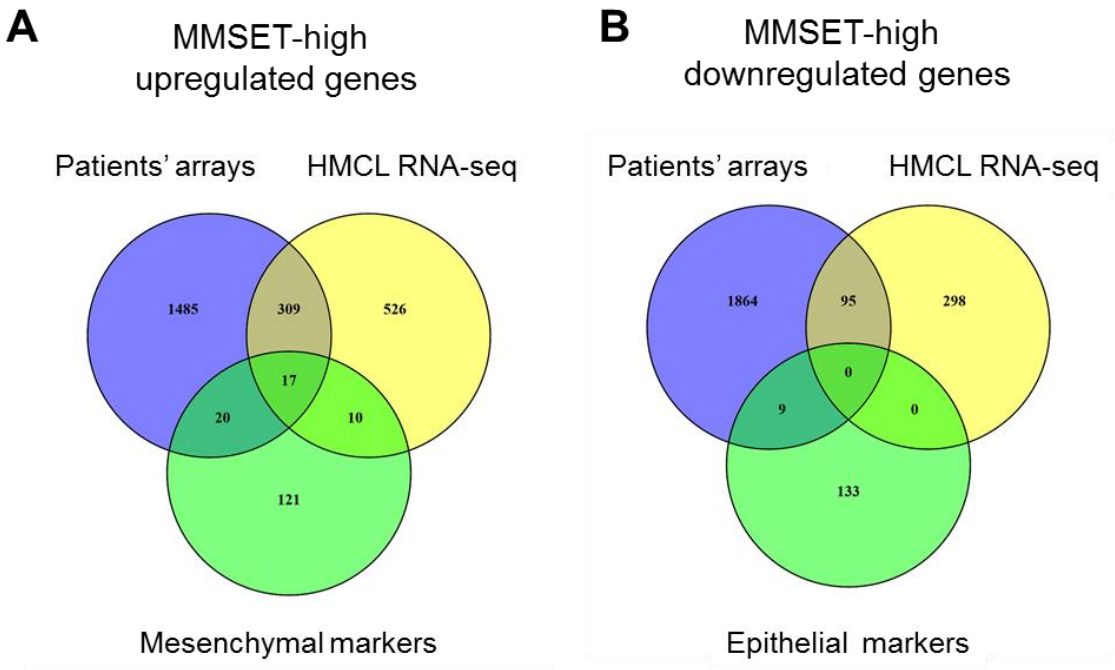
2.4.2 Mesenchymal markers driven by MMSET in MM

Among the differentially regulated genes, 37 of the upregulated and 9 of the downregulated genes corresponded to the generic EMT-signature[32], previously established using data from various epithelial cancers, with a typical EMT requiring a concurrent upregulation of mesenchymal markers and downregulation of epithelial markers (Figure 2.2A and B).

Evidence of an EMT associated with the t(4;14) translocation was also evaluated in 66 HMCL. Genetic information associated with the cell lines was used to stratify these HMCL into MMSET-high [t(4;14)-positive] or MMSET-low [t(4;14)-negative]. 862 genes were significantly upregulated (FDR < 0.05, log2 fold change > 0) in MMSET-high cell lines and 393 genes were downregulated (FDR < 0.05, log2 fold change < 0; Figure 2.1). While 27 of the upregulated genes corresponded to known mesenchymal markers, none of the downregulated genes corresponded to epithelial markers (Figure 2.2A and B).

The upregulated gene profile from MMSET-high patients' microarrays and HMCL shared 17 common mesenchymal markers (Figure 2.2A), which correlated positively with MMSET (Figure 2.3Figure 2). Examples of the upregulated mesenchymal genes included *TWIST1*, *MAP1B*, *UCHL1*, *CDH2* and *PDGFC* (Table 2.1). Conversely, no epithelial markers were shared between the downregulated profile from MMSET-high patients' microarrays and HMCLs. Only 3 genes (*UCHL1*, *CDH2* and *MAP1B*) in our 17-mesenchymal signature were common in Wu-list, most likely due to the exclusion of candidate genes with low fold change, as a cutoff of > than 2-fold difference was applied in Wu-list. As we hypothesized that MMSET would drive the acquisition of an EMT-like phenotype, thereby downregulating epithelial markers, it was interesting to find that the epithelial markers, DSG2 and FGFR3 were upregulated in the MMSET-high patients and HMCLs (supplementary Table S2.2).

Figure 2.2 Overlapped EMT-signature in MMSET-high MM patients and HMCL. Venn diagrams showing the upregulated (A) and downregulated (B) genes that overlapped between MMSET-high patients and HMCL. While 17 of the upregulated genes were previously recognised mesenchymal markers, none of the common downregulated genes corresponded to epithelial markers. (C) Gene set enrichment analysis (GSEA) of generic mesenchymal-signature between MMSET-high [t(4;14)-positive] and MMSET-low [non-t(4;14)] HMCL. The horizontal bar, in graded color from red to blue, represents the unique gene list from microarray, rank-ordered by signal-to-noise ratio. The vertical black lines represent the individual genes in the mesenchymal-signature gene set within the ranked gene list. The green curves show the enrichment score and reflect the degree to which each gene (black vertical lines) is represented at the top or bottom of the ranked gene list. Genes on the left side (red) correlated most strongly with increased EMT-signature. The heatmaps on the right show the top 50 genes for each phenotype. NES, normalized enrichment score; FDR, false discovery rate.


A
MMSET-high upregulated genes
Patients' arrays
HMCL RNA-seq
1485
309
526
17
20
10
121
Mesenchymal markers
B
MMSET-high downregulated genes
Patients' arrays
HMCL RNA-seq
1864
95
298
0
9
0
133
Epithelial markers

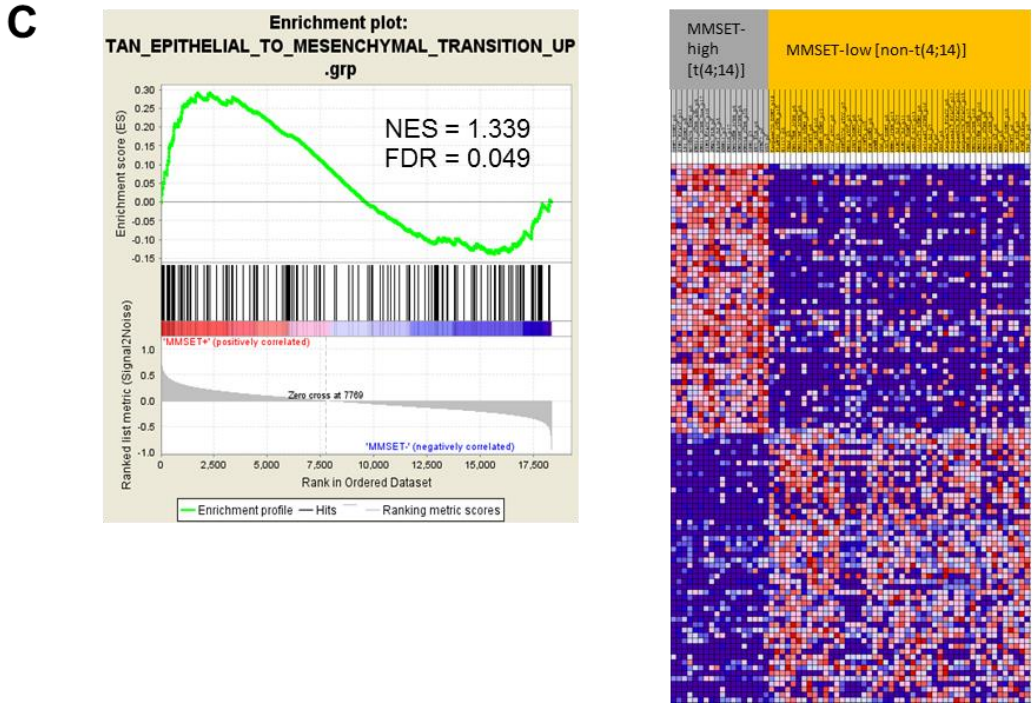
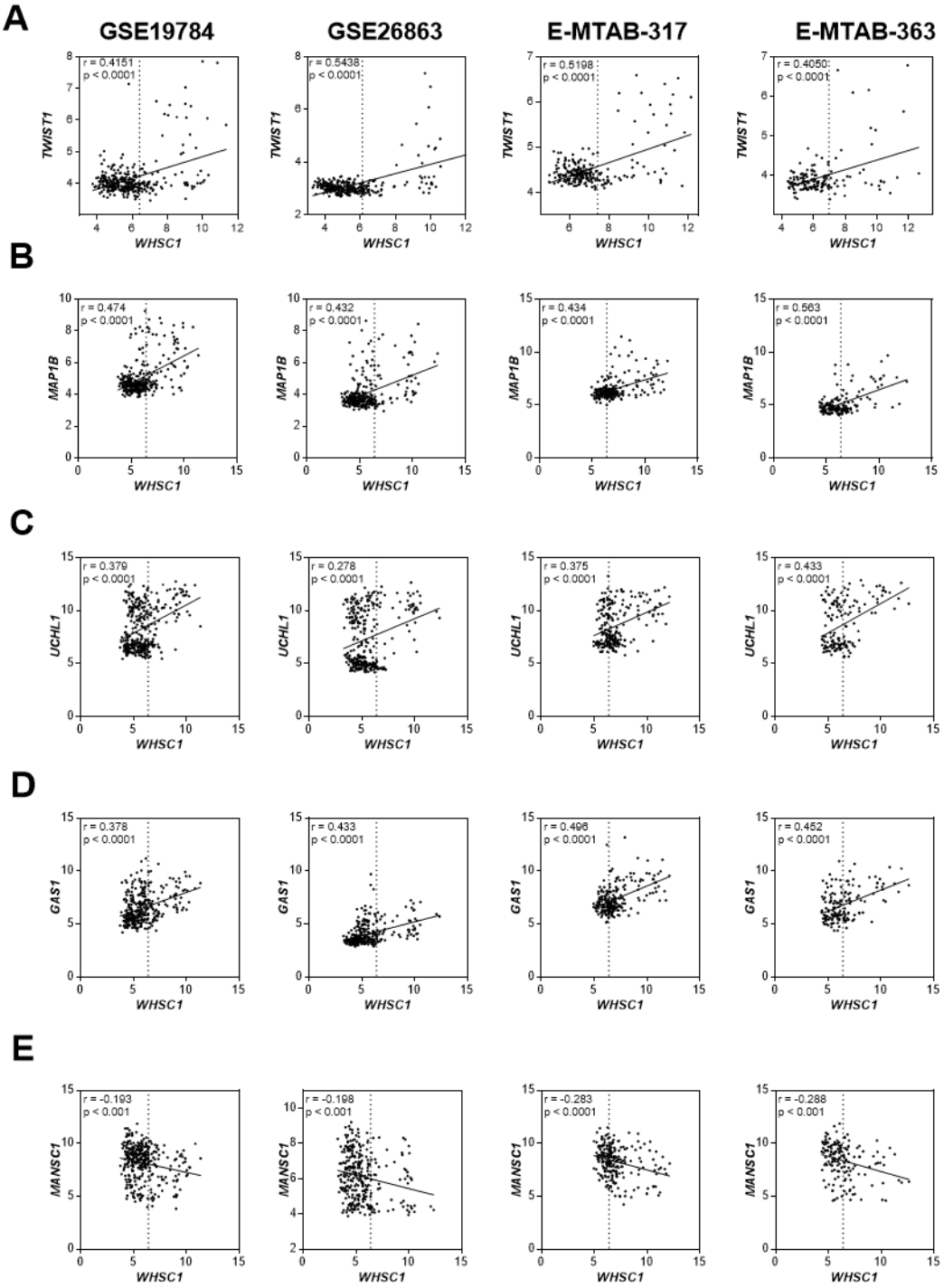

C
Enrichment plot:
TAN_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_UP
.grp
NES = 1.339
FDR = 0.049
Enrichment score (ES)
Ranked list metric (Signal2Noise)
'MMSET+' (positively correlated)
Zero cross at 7709
'MMSET-' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
MMSET-high [t(4;14)]
MMSET-low [non-t(4;14)]

Figure 2.3 Mesenchymal genes positively correlate with MMSET expression in MM patients. *MMSET* expression in 1012 newly diagnosed MM patients from E-GEOD-19784, E-GEOD-26760, E-MTAB-317 and E-MTAB363 were plotted against expression of mesenchymal genes *TWIST1* (A), *MAP1B* (B), *UCHL1* (C), *GAS1* (D) and the epithelial markers *MANSC1* (E). Pearson correlation coefficient *r* and the respective *p*-values are shown.


A
GSE19784
GSE26863
E-MTAB-317
E-MTAB-363
TWIST1
r = 0.4151
p < 0.0001
r = 0.5438
p < 0.0001
r = 0.5198
p < 0.0001
r = 0.4050
p < 0.0001
WHSC1
B
MAP1B
r = 0.474
p < 0.0001
r = 0.432
p < 0.0001
r = 0.434
p < 0.0001
r = 0.563
p < 0.0001
C
UCHL1
r = 0.379
p < 0.0001
r = 0.278
p < 0.0001
r = 0.375
p < 0.0001
r = 0.433
p < 0.0001
D
GAS1
r = 0.378
p < 0.0001
r = 0.433
p < 0.0001
r = 0.496
p < 0.0001
r = 0.452
p < 0.0001
E
MANSC1
r = -0.193
p < 0.001
r = -0.198
p < 0.001
r = -0.283
p < 0.0001
r = -0.288
p < 0.001

2.4.3 Genes upregulated in MMSET-high patients lack enrichment for canonical EMT-marker

The cohort of MMSET-regulated EMT-related genes was further investigated using Gene Set Enrichment analysis (GSEA). Mesenchymal markers were found to be enriched in t(4;14) HMCL (NES = 1.339, FDR < 0.05; Figure 2.2C). In relation to the MM patients GEP data, although enrichment for mesenchymal markers was observed in the MMSET-high patients, significance was not reached in 3 out of 4 datasets, with the exception of data from E-MTAB-317 (supplementary Figure S2.3). When assessed individually, the canonical EMT markers correlated weakly with WHSC1 (supplementary Figure S2.4). Mesenchymal markers CDH2 and VIM correlate positively with WHSC1 in 3 and 2 datasets, respectively. The mean relative expression of CDH1 was generally low across all datasets (4.61, 95% CI: 4.56-4.66) and correlated negatively with MMSET expression only in GSE19784.

To validate the lack of EMT enrichment in MMSET-high MM, EMT status32 associated with GEPs in GSE19784 and GSE26760 were plotted based on their UAMS classification and further analysed. The degree of EMT status (with an EMT score, -1 reflecting an epithelial state and +1 reflecting a mesenchymal state) for MM patients GEPs were previously computed using EMT signature established from epithelial cancers.[32] An estimation of an EMT phenotype requires an EMT score greater than 0 (derived from differences between the fold change of mesenchymal markers and epithelial markers). As shown in supplementary Figure S2.5, the EMT score for MS patient subtype (MMSET-high) was not significantly different from other UAMS subtypes in both GSE19784 and GSE26760 datasets ($p > 0.05$, Kruskal-Wallis test with multiple comparisons). While an EMT score for the MS subgroup of -0.089 is insufficient to suggest a EMT phenotype in MMSET-MM patients, the differential expression of 37 mesenchymal and 9 epithelial genes represents a MMSET-driven EMT signature in MMSET-high patients.

2.4.4 The EMT-related transcription factor (TF), TWIST, is elevated in MMSET-high MM patients and is associated with poor prognosis

Among the upregulated mesenchymal genes, TWIST1 was upregulated by the greatest fold change in t(4;14)-positive HMCL compared with t(4;14)-negative HMCL (log2 fold change 5.96, $p < 0.01$; Table 2.1). TWIST1 expression was consistently

Table 2.1 List of 17 mesenchymal genes differentially upregulated in MMSET-high patients (Fisher's method) and HMCL.

Gene symbol	Probe ID	median log2 FC MMSET-high vs MMSET-low patients	combined p-value	Log2 fold change MMSET-high vs MMSET-low HMCL	adjusted p-value
TWIST1	213943_at	0.78	2.54E-56	5.95	2.11E-06
MAP1B	212233_at	1.40	2.52E-39	5.13	5.40E-05
MAP1B	226084_at	1.44	1.12E-36		
MAP1B	214577_at	0.24	3.92E-07		
UCHL1	201387_s_at	2.69	3.06E-26	3.92	9.29E-03
UCHL1	1555834_at	0.09	1.38E-04		
GAS1	204457_s_at	1.56	2.70E-25	3.72	9.81E-08
CDH2	203440_at	1.04	3.14E-06	3.15	8.82E-02
CDH2	203441_s_at	0.35	3.44E-03		
PDGFC	218718_at	1.64	1.59E-11	2.93	4.02E-04
PDGFC	222719_s_at	0.24	1.64E-06		
NAP1L3	204749_at	1.92	2.95E-79	2.74	1.17E-03
AKT3	212607_at	0.92	4.81E-15	2.60	2.50E-05
AKT3	224229_s_at	0.26	3.00E-06		
AKT3	212609_s_at	0.32	2.64E-04		
AKT3	222880_at	0.25	3.52E-04		
DPYSL3	201431_s_at	0.14	4.58E-05	2.57	1.23E-02
GEM	204472_at	0.62	1.47E-04	2.51	6.95E-03
MYH10	212372_at	1.63	1.62E-75	2.45	1.21E-04
MYH10	213067_at	0.30	2.07E-14		
HEG1	213069_at	0.64	6.48E-11	2.17	6.91E-03
HEG1	212822_at	0.24	1.56E-03		
MAF	206363_at	1.32	3.24E-10	1.93	3.75E-01
MAF	209348_s_at	1.51	1.13E-06		
MAF	229327_s_at	0.90	3.60E-04		
JAM3	212813_at	1.44	1.52E-88	1.84	4.04E-05
JAM3	231721_at	0.05	4.51E-07		
TRPC1	205802_at	0.55	2.76E-20	1.75	2.03E-02
TRPC1	205803_s_at	0.52	4.41E-13		
TUBA1A	209118_s_at	1.30	5.75E-09	1.74	4.05E-03
COL5A2	221729_at	0.24	1.99E-06	0.99	4.80E-02
COL5A2	221730_at	0.19	8.10E-06		

upregulated in MMSET-high patients (p-value < 0.01) across all four microarray datasets examined as shown in Figure 2.4A. Elevated TWIST1 expression (greater than 2 standard deviations of the mean TWIST1 expression in MMSET-low/neg MM patient cohort) was observed in 45% (18/40; GSE19784), 58% (20/34; GSE26863), 50% (17/34; E-MTAB-317) and 33% (9/27; E-MTAB-363) of MMSET-high MM patients. From these data, TWIST1 represents a putative candidate gene responsible for modulating the EMT phenotype in MM.

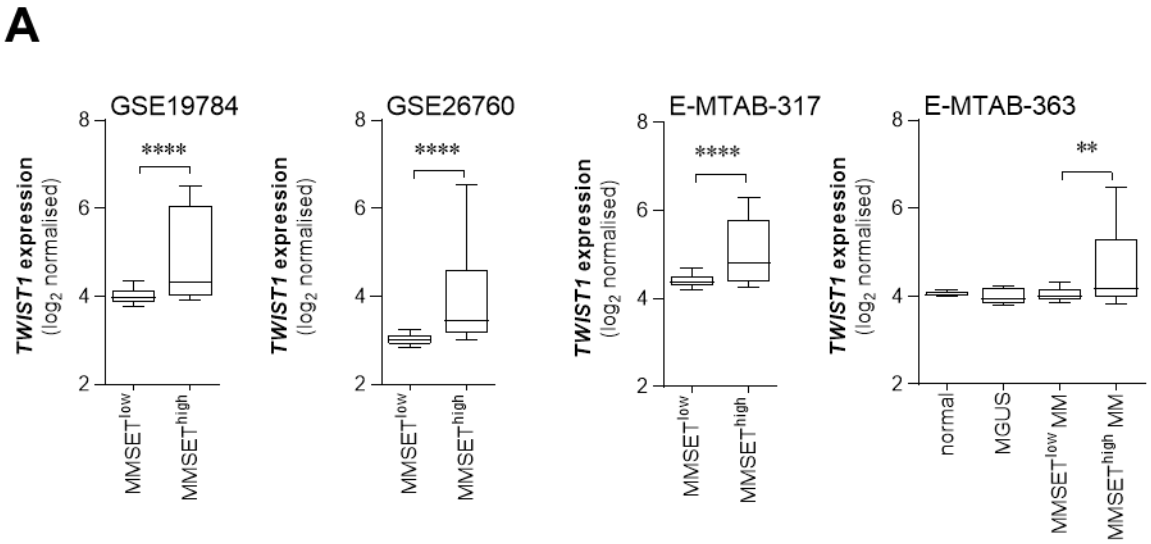
In order to determine whether the increased expression of TWIST1 in MMSET-high MM patient cohort was related to prognostic significance, we assessed the overall survival of newly diagnosed MM patients enrolled in the total therapy 2 (TT2) trial in the GSE4581 microarray datasets, with GEP-defined subgroups (UAMS classification). The MM patients of the MS subgroup (n = 68) were stratified based on median TWIST1 expression levels, where the third and fourth quartile with the highest TWIST1 expression was classified as TWIST1-high (n = 17 patients). The overall survival for patients of MS subgroup patients expressing high *TWIST1* was significantly poorer relative to patients who are not classified in MS subgroup. (p = 0.0375, log-rank [Mantel-Cox] test; Figure 2.4B).

2.4.5 Modulation of MMSET affects TWIST1 expression level in MM

To determine whether modulation of MMSET affects TWIST1 expression, we interrogated publicly available microarray data (GSE50074 and GSE24746) from the t(4;14)-positive KMS11 MM cells in which MMSET expression and function was genetically knocked down. As shown in Figure 2.5A, the knock down of MMSET on the translocated allele in KMS11 cells (MMSET TKO cells) resulted in 51% decrease in TWIST1 expression compared with KMS11 parental cells. TWIST1 expression was restored to basal level when MMSET was re-expressed in MMSET TKO cells. Similarly, in the GSE24746 dataset, ectopic expression of wild type MMSET in MMSET TKO cells increased TWIST1 expression by 70% (Figure 2.5B). The HMT-defective MMSET-Y1118A mutant resulted in a 50% increase in TWIST1 expression. Similarly, MMSET depletion using inducible shRNA knockdown system also led to a 41% reduction of TWIST1 expression (Figure 2.5C).

To validate the effect of MMSET modulation on TWIST1 expression in KMS11, we evaluated the effect of overexpressing wild type or mutant MMSET

Figure 2.4 TWIST1 expression is upregulated in MMSET-high MM patients. (A) *In silico* analyses of publicly available microarray performed in CD138+ plasma cells isolated by MACS from MM patients and normal age-matched controls, segregated into MMSET-low or MMSET-high (GSE19784 [MMSET-low, n = 288; MMSET-high, n = 40], GSE26760 [MMSET-low, n = 270; MMSET-high, n = 34], E-MTAB-317 [MMSET-low, n = 192; MMSET-high, n = 34], E-MTAB-363 [normal, n = 5; MGUS, n = 5, MMSET-low, n = 128; MMSET-high, n = 27]). $^{**}p < 0.01$, $^{****}p < 0.0001$ compared with MMSET-low, Mann-Whitney test. (B) Kaplan-Meier plot of overall survival for MM patients of MS subgroup (MS +) in GSE4581 with respect to TWIST1 expression. Solid black line represents non-MS subgroup patients (MS -; n = 346); solid gray line represents MS subgroup patients expressing low TWIST1 (MS+/TWIST1-low; n = 11); dashed line represents MS subgroup patients expressing high


A
GSE19784
GSE26760
E-MTAB-317
E-MTAB-363
TWIST1 expression (log2 normalised)
****
**
MMSET low
MMSET high
normal
MGUS
MMSET low MM
MMSET high MM

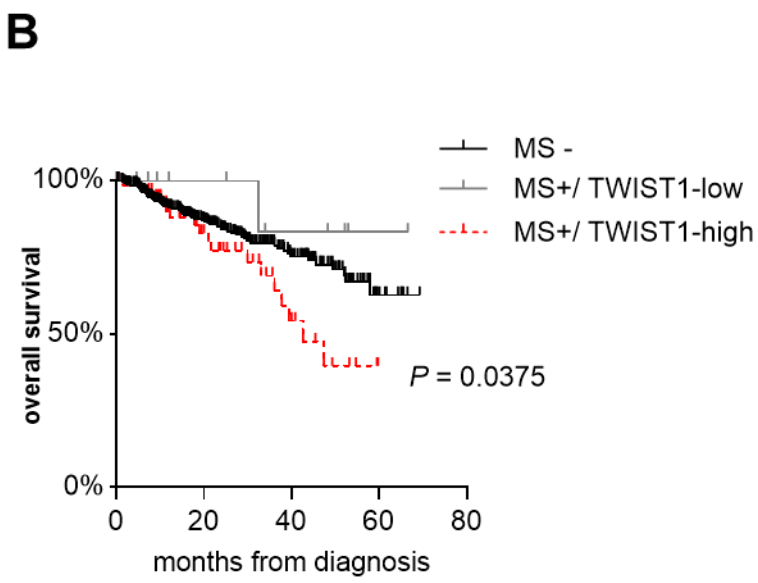
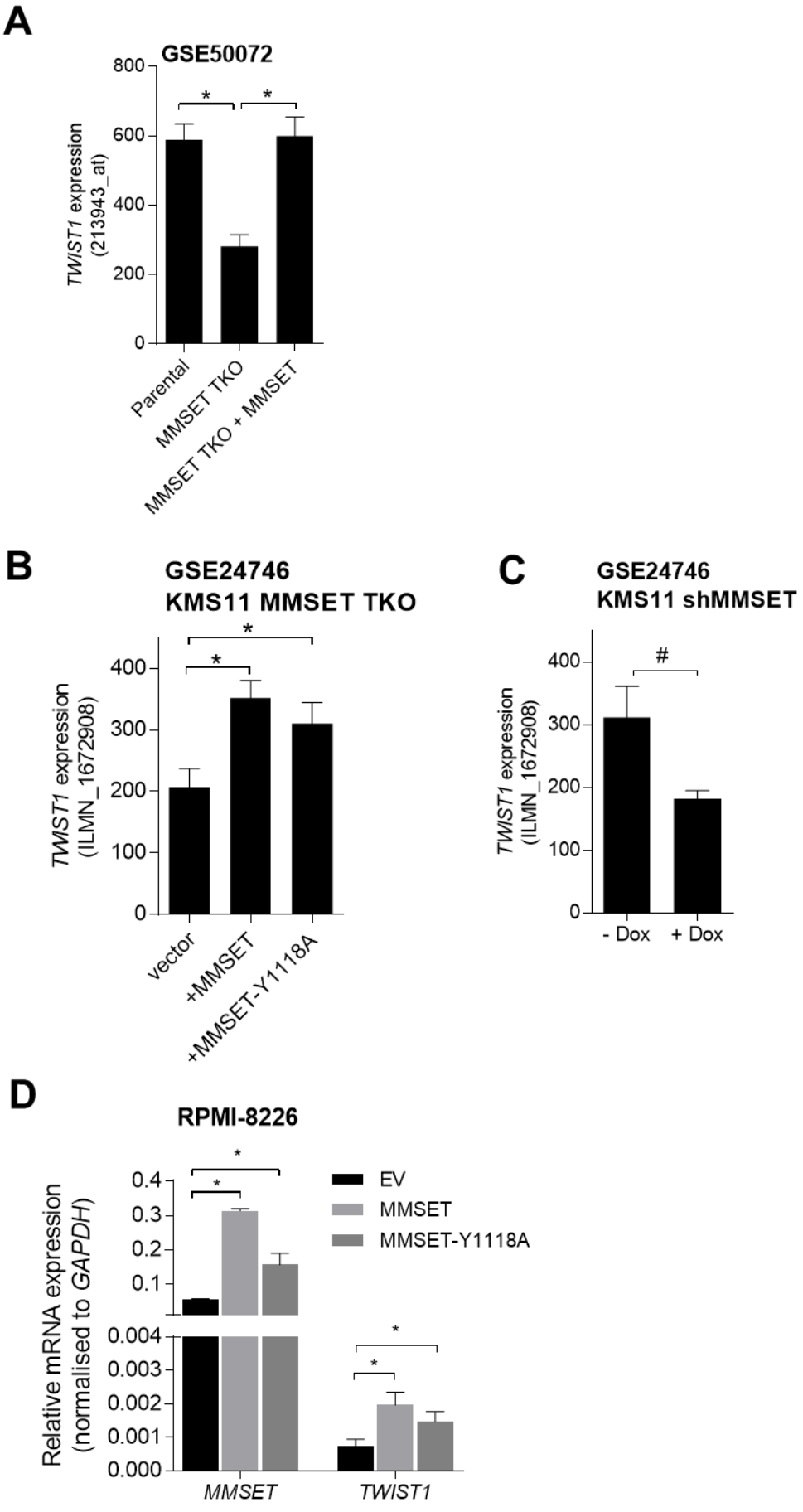

B
MS -
MS+/ TWIST1-low
MS+/ TWIST1-high
overall survival
100%
50%
0%
P = 0.0375
0
20
40
60
80
months from diagnosis

Figure 2.5 Modulation of MMSET level regulates TWIST1 expression. TWIST1 expression level in t(4;14)-positive KMS11 cells with presence or absence of MMSET using shRNA knockdown, knockout or repletion systems were extracted from two independent publicly available microarray GSE50071 (A) and GSE24746 (B and C). Relative *TWIST1* expression is represented in mean ± SD of biological triplicates. (*$p < 0.05$, one-way ANOVA; #$p < 0.05$, student's t-test). (D) Overexpression of *MMSET* in t(4;14)-negative RPMI-8226 cells increases *TWIST1* expression. Relative *TWIST1* levels were measured by RT-qPCR in t(4;14)-negative RPMI-8226 transduced with empty vector, wild-type or mutant MMSET retroviruses, normalised to *GAPDH*. (*$p < 0.05$, one-way ANOVA).


A
GSE50072
TWIST1 expression (213943_at)
800
600
400
200
0
*
*
Parental
MMSET TKO
MMSET TKO + MMSET
B
GSE24746
KMS11 MMSET TKO
TWIST1 expression (ILMN_1672908)
400
300
200
100
0
*
*
vector
+MMSET
+MMSET-Y1118A
C
GSE24746
KMS11 shMMSET
TWIST1 expression (ILMN_1672908)
400
300
200
100
0
#
- Dox
+ Dox
D
RPMI-8226
Relative mRNA expression (normalised to GAPDH)
0.4
0.3
0.2
0.1
0.004
0.003
0.002
0.001
0.000
*
*
*
*
EV
MMSET
MMSET-Y1118A
MMSET
TWIST1

(Y1118A) in MMSET-low [non-t(4;14)] RPMI-8226 cells. Overexpression of MMSET was validated by RT-qPCR (Figure 2.5D) and Western blot (supplementary Figure S2.2). As demonstrated in Figure 2.5D, overexpression of MMSET resulted in 158% increase in TWIST1 mRNA levels. The mutant MMSET-Y1118A yielded a 100% increase in TWIST1 expression, in agreement with its less potent stimulation on TWIST1 level observed in KMS11 cells.

2.5 Discussion

GEP studies published over the past decade have provided novel insights into the underlying biology of MM pathogenesis[31,37-41] and have served as a framework to molecularly stratify MM patients into different risk groups[42] with prognostic significance. MMSET overexpression was consistently detected in GEPs associated with t(4;14) translocation and identified as high-risk subgroup. In an effort to characterise EMT-markers influenced by MMSET, we analysed four independent GEP microarray studies comprised of 1012 MM patients and transcriptomic analysis from 66 HMCLs.

In this study, we identified 17 mesenchymal markers which were significantly upregulated in both MMSET-high patients and t(4;14)-positive HMCL. Among these genes, *TWIST1, MAP1B, UCHL1, GAS1, CDH2, MYH10, MAF* and *JAM3* were previously shown to be upregulated upon MMSET-overexpression in MM.[21,36,43,44] Conversely no epithelial markers were significantly downregulated in MMSET-high patients. As MM PC are derived from haematopoietic cells and not of epithelial origin, it was not surprising that no epithelial gene changes were detected. Unlike epithelial cells, cells of haematopoietic origin, like MM PC, do not rely on the expression of cell surface proteins involved in governing cell-cell junction to maintain cell polarity and lateral adhesion to form epithelial sheets.[45] The haematopoietic origin of MM PCs may account for the lack of significant enrichment of EMT-associated genes in MMSET-high patients. Despite this, the upregulation of 17 mesenchymal genes in both MMSET-high patients and HMCL appears to represent a MMSET-driven EMT signature in MM.

Moreover, the non-significant enrichment of EMT-related markers in MMSET-high MM was further complicated by upregulation of epithelial-associated markers such as FGFR3 and DSG2 (supplementary Table S2.2), consistent with previous MMSET studies.[28,46] In epithelial cancers, FGFR3 is associated with E-cadherin

expression[47,48], a canonical epithelial marker, and exerts a tumour suppressive effect in epithelial-like pancreatic cancer cell lines.[49] In contrast, other studies have shown that FGFR3 acts as oncogenic promoting molecule in mesenchymal-like pancreatic[49] and bladder carcinomas[47,50], suggesting a dichotomous role of FGFR3 in oncogenesis depending on the cellular context. In t(4;14) MM, FGFR3 is translocated and juxtaposed under the influence of potent IGH enhancer leading to FGFR3 overexpression.[51] FGFR3 has been considered as an attractive therapeutic target in t(4;14) MM due to anti-tumour effects of FGFR3 inhibition seen in mouse models, indicating an oncogenic promoting role of FGFR3 overexpression in MM. Similarly, while it is generally considered that DSG2 desmosomal proteins are repressed upon EMT activation thereby destabilizing the cell-cell junction[52], other studies have shown that DSG2 expression can promote tumour growth in prostate and skin cancers.[53,54] Although the functions of DSG2 in MM biology remain to be investigated, overexpression of DSG2 in MS subtype MM has been implicated in both stimulating cell proliferation and adhesion of MM PCs.[46]

Genes upregulated in t(4;14)-positive HMCL showed significant enrichment for mesenchymal markers. In contrast to patient-derived GEPs that were derived from primary BM-resident MM PC, HMCL are generally established from advanced stage or extramedullary disease.[55] The transcriptome of highly proliferative HMCL may reflect additional genetic changes leading to extramedullary progression downstream of MMSET and/or gained from *in vitro* culture, independent of the BM microenvironment. In order to delineate whether MMSET drives expression of mesenchymal markers leading to an aggressive disease phenotype, longitudinal GEP studies of BM-resident MM PC and PC recovered from peripheral blood and/or extramedullary sites may provide more insights.

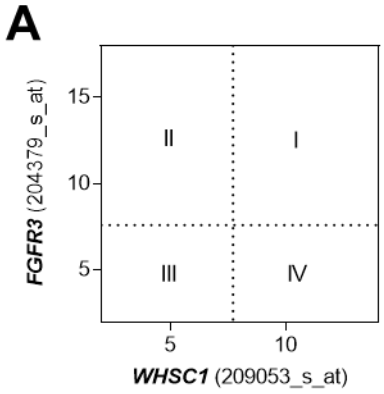
TWIST1 is upregulated in both MMSET-high patients and in t(4;14)-positive HMCL which show the highest fold change in MMSET expression. TWIST1 was not shortlisted as a MMSET-regulated gene in previous GEP studies performed on patient-derived MM PCs, possibly due to low expression changes of *TWIST1*. Interestingly, *TWIST1* overexpression was identified in MMSET-high HMCL from previous GEPs[44,56], consistent with our analysis of RNA-seq data of HMCL, confirming a potential role for MMSET on TWIST1 regulation. TWIST1 is a transcription factor that promotes tumour invasion and metastasis in a variety of solid tumour via activation of EMT.[26,57,58] In prostate cancer, MMSET-driven EMT requires

upregulation of TWIST1, leading to invasive phenotype.[25] Increased TWIST1 expression is also detected in hematological malignancies, including cutaneous T-cell lymphoma[59,60] and diffuse large B-cell lymphoma.[61,62] In MM, *TWIST1* expression was found to be highly expressed in extramedullary tumour compared with BM PC[63], suggesting a role for TWIST1 in mediating establishment of MM PC growth which is independent of BM niche. However, how TWIST1 expression affects the pathogenesis of MM remains to be determined.

This work provides a comprehensive evaluation of EMT-like process in MM and identified 17 mesenchymal markers that are upregulated in MMSET-high patients. Collectively, upregulation of mesenchymal markers in t(4;14) MM PCs may suggest activation of EMT-like program as part of the global transcription changes driven by MMSET, leading to an aggressive metastatic MM phenotype.

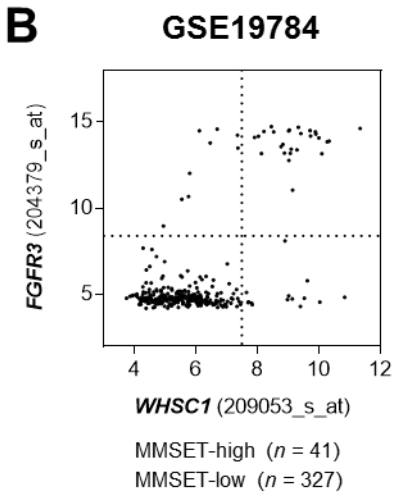
2.6 Supplementary Figures and Tables

supplementary Figure S2.1 Quadrant analysis to classify GEPs into MMSET-low or MMSET-high MM for meta-analysis. (A) Quadrant numbering. *FGFR3* (204379_s_at) expressions from each sample in dataset GSE19784 (B), GSE26760 (C), E-MTAB-317 (D) and E-MTAB-363 (E) were plotted against *WHSC1* (209053_s_at) expression. Samples scattered in Quadrant I and IV were classified as MMSET-high while samples scattered in Quadrant II and III were classified as MMSET-low MM.

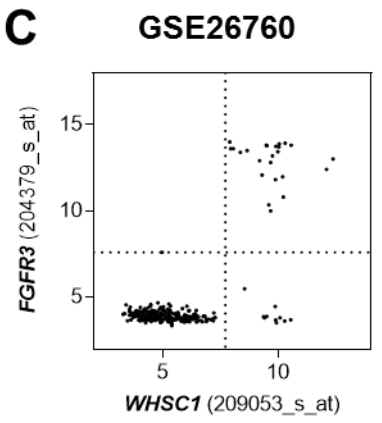


Quadrant I & IV: MMSET-high

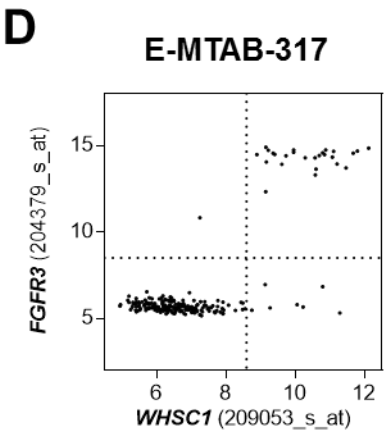
Quadrant II & III: MMSET-low



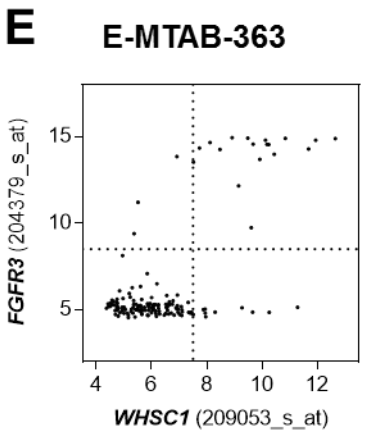
MMSET-high (n = 41)
MMSET-low (n = 327)



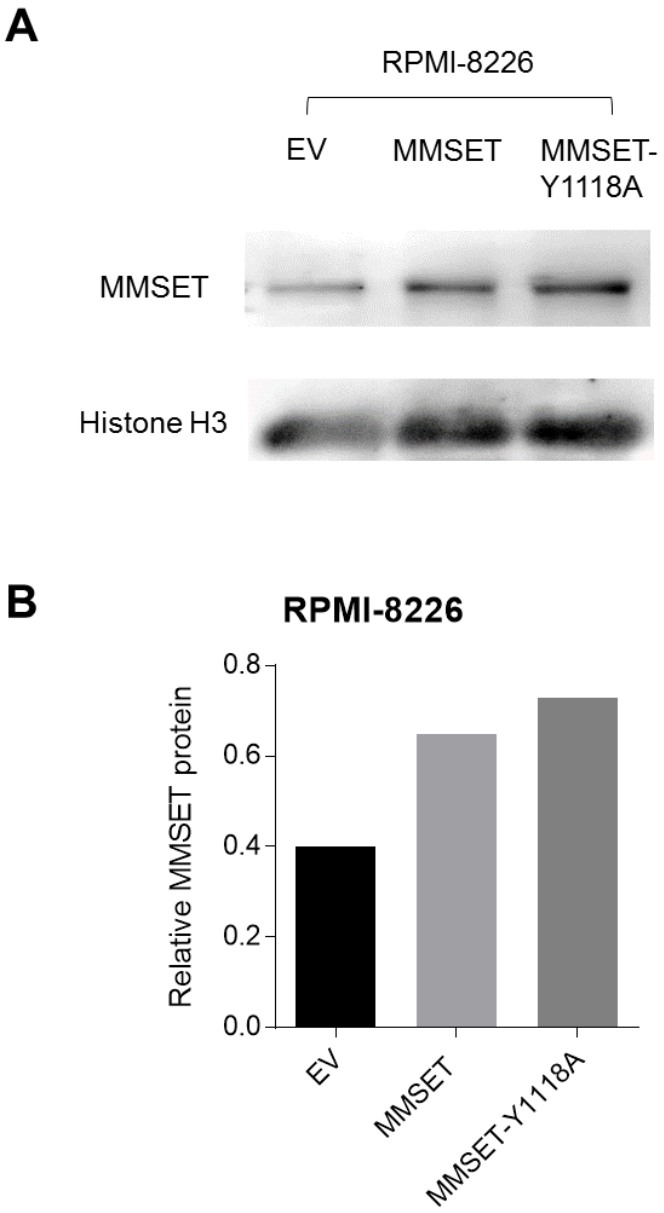
MMSET-high (n = 34)
MMSET-low (n = 270)



MMSET-high (n = 34)
MMSET-low (n = 192)



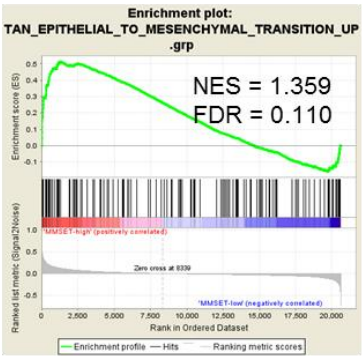
MMSET-high (n = 27)
MMSET-low (n = 128)

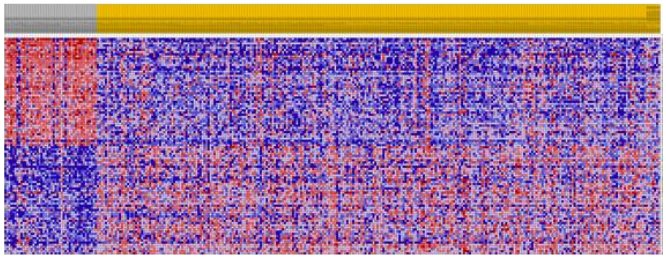


supplementary Figure S2.2 Overexpression of MMSET in t(4;14)-negative (MMSET-low) RPMI-8226 HMCL. (A) Nuclear protein extracts (30 μg) from RPMI-8226 cell lines transduced with empty vector, wild-type or mutant MMSET construct were resolved on 10% SDS-PAGE gels under reducing condition and immunoblotted with the indicated antibodies. (B) Protein levels were quantified by densitometry relative to the level of loading control Histone H3.

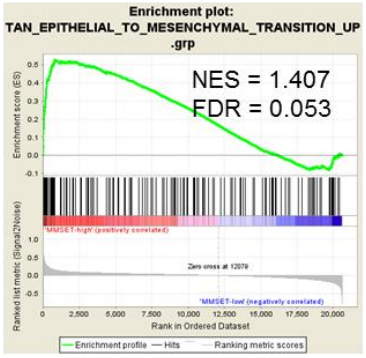
supplementary Figure S2.3 Gene deregulated by MMSET lack enrichment for canonical EMT-signature in 3 out of 4 microarrays analysed. Gene set enrichment analysis demonstrated enrichment plot of a generic mesenchymal-signature between MMSET-high and MMSET-low MM patients in GSE19784 (A), GSE26760 (B), E-MTAB-317 (C) and E-MTAB363 (D). The horizontal bar in graded color from red to blue represents the unique gene list from microarray, rank-ordered by signal-to-noise ratio. The vertical black lines represent the individual genes in the mesenchymal-signature gene set within the ranked gene list. The green curves show the enrichment score and reflect the degree to which each gene (black vertical lines) is represented at the top or bottom of the ranked gene list. Genes on the left side (red) correlated most strongly with increased EMT-signature. The heatmaps on the right show the top 50 feature (genes) for each phenotype. NES, normalized enrichment score; FDR, false discovery rate.

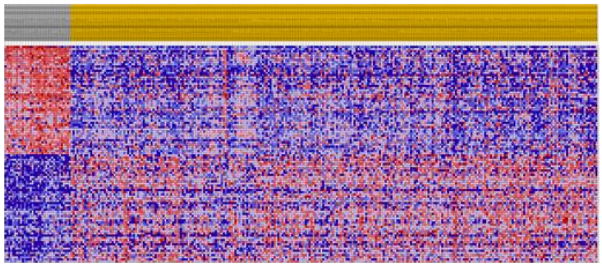
A. GSE19784



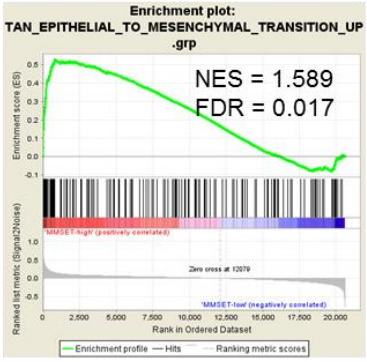


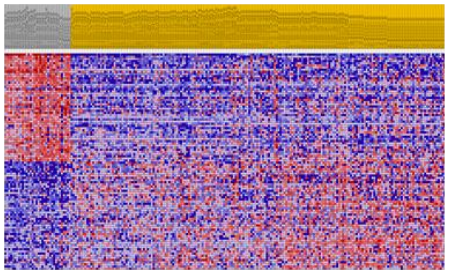
B. GSE26760



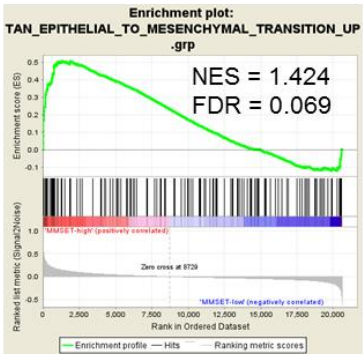


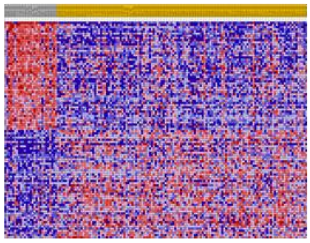
C. E-MTAB-317



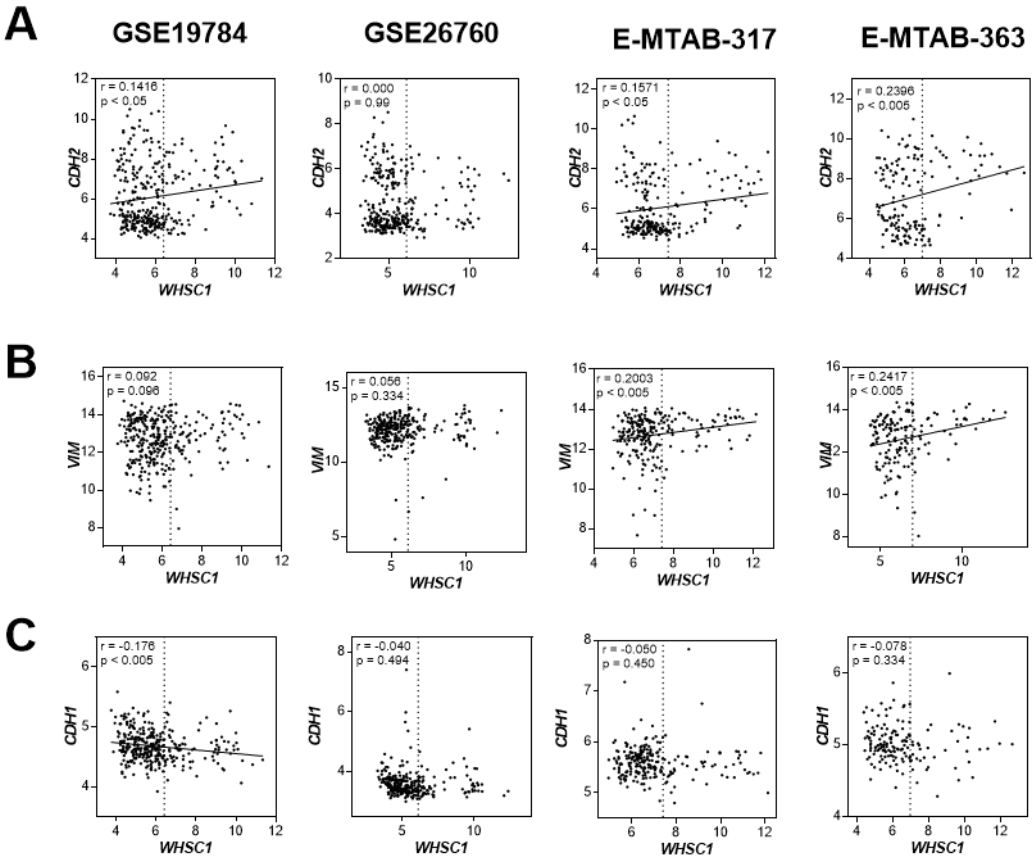


D. E-MTAB-363

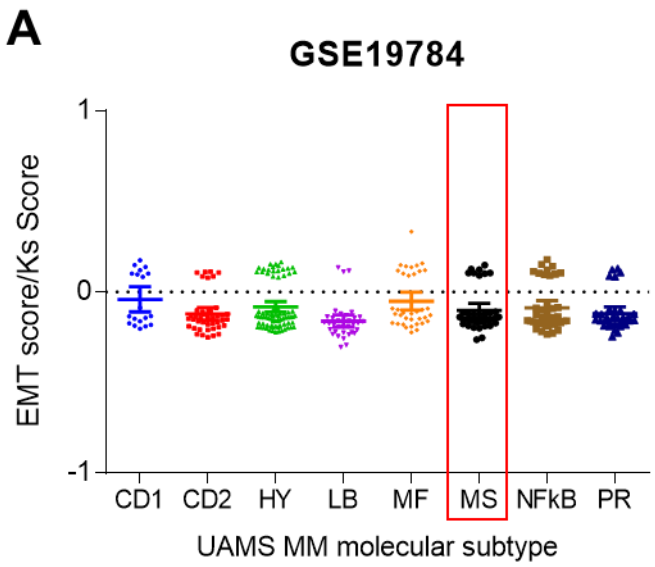


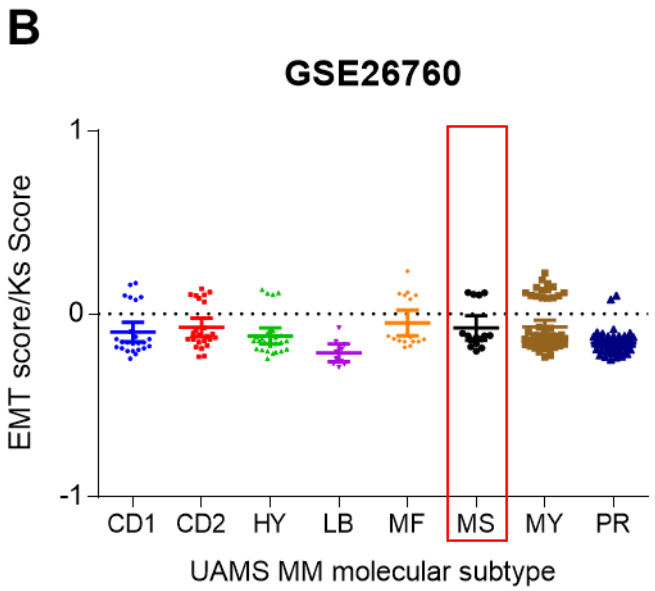


supplementary Figure S2.4 Weak correlation of canonical EMT markers with MMSET expression in MM patients. *MMSET* expression in 1012 newly diagnosed MM patients from GSE19784, GSE26760, E-MTAB-317 and E-MTAB363 were plotted against expression of canonical mesenchymal genes *CDH2* (A), *VIM* (B) and the epithelial markers *CDH1* (C). Pearson correlation coefficient *r* and the respective *p*-values are shown.


A
GSE19784
GSE26760
E-MTAB-317
E-MTAB-363
r = 0.1416
p < 0.05
r = 0.000
p = 0.99
r = 0.1571
p < 0.05
r = 0.2396
p < 0.005
CDH2
WHSC1
B
r = 0.092
p = 0.096
r = 0.056
p = 0.334
r = 0.2003
p < 0.005
r = 0.2417
p < 0.005
VIM
WHSC1
C
r = -0.176
p < 0.005
r = -0.040
p = 0.494
r = -0.050
p = 0.450
r = -0.078
p = 0.334
CDH1
WHSC1

supplementary Figure S2.5 EMT-score of MS subgroup is not significantly different from other MM subgroups. EMT score for GEPs in GSE19784 (n = 260) and GSE26760 (n = 224) was obtained from Tan et al. MM patients from were stratified into subgroups based on the UAMS criteria; namely, patients characterised by increased proliferation-related genes (PR), chromosomal translocations involving cyclin D1 and cyclin D3 (CD1 and CD2), MAF (MF) or MMSET (MS), as well as patients exhibiting hyperdiploidy (HY) and decreased prevalence of lytic bone disease (LB). Box and whiskers plots show the median and interquartile ranges for each cohort. Kruskal-Wallis test with Dunn's multiple comparison tests.


A
GSE19784
EMT score/Ks Score
1
0
-1
CD1
CD2
HY
LB
MF
MS
NFkB
PR
UAMS MM molecular subtype


B
GSE26760
EMT score/Ks Score
1
0
-1
CD1
CD2
HY
LB
MF
MS
MY
PR
UAMS MM molecular subtype

supplementary Table S2.1 List of 43 gene probes overlapped with Wu-list.[36]

Gene	ProbeID	median log2 FC MMSET-high vs MMSET-low patients	combined p-value (Fisher's method)	Wu-list linear regression coefficient	Wu-list p-value (SAM method)
CCND2	200953_s_at	4.10	5.95E-32	3.7	6.00E-36
KLF4	221841_s_at	2.93	1.28E-33	3.7	3.90E-35
UCHL1	201387_s_at	2.69	3.06E-26	3.3	2.50E-28
CCND2	200951_s_at	2.63	2.08E-22	3.7	5.20E-36
APP	200602_at	2.61	5.80E-20	2.6	1.20E-17
AZGP1	209309_at	2.56	3.36E-24	3.1	6.40E-25
FAM171A1	212771_at	2.37	2.83E-41	2.5	1.40E-16
CTHRC1	225681_at	2.27	3.92E-15	2.3	7.80E-15
IGFBP7	201163_s_at	2.20	4.24E-26	2.5	8.10E-17
GNAI1	227692_at	2.15	1.10E-74	2.4	8.70E-16
MYBL1	213906_at	2.15	2.85E-37	2.5	5.30E-17
SELL	204563_at	2.14	6.16E-23	2.2	4.10E-13
GNB4	225710_at	2.13	1.91E-17	2.2	7.90E-14
SOX4	201417_at	2.10	1.12E-18	2.4	2.70E-16
PKP2	207717_s_at	2.03	2.85E-21	2.7	9.30E-20
TMEM47	209656_s_at	2.02	7.26E-25	2.6	7.50E-18
MPPED2	205413_at	1.95	2.49E-23	2.4	6.90E-16
ETV1	221911_at	1.88	2.42E-06	2.6	1.10E-17
DSEL	232235_at	1.86	3.19E-36	2.8	1.80E-21
MYADM	225673_at	1.82	9.21E-15	2.1	4.10E-12
ROBO1	213194_at	1.82	1.65E-20	2.2	8.60E-14
MAGED4	221261_x_at	1.81	3.46E-35	2.6	8.10E-18
CDC42BPA	214464_at	1.81	1.22E-121	2.4	2.20E-15
PTP4A3	206574_s_at	1.69	1.82E-14	2.5	4.70E-17
PPIC	204518_s_at	1.52	2.02E-29	2.4	1.00E-15
AOC1	203559_s_at	1.51	2.92E-34	2.3	3.30E-14
KLF4	220266_s_at	1.50	1.81E-45	2.1	2.20E-12
MAGED4	223313_s_at	1.48	3.31E-37	2.9	5.00E-23
SGCB	205120_s_at	1.47	6.56E-24	2.5	3.90E-17
GABRB2	242344_at	1.46	2.88E-17	2.1	3.60E-12
ACKR3	212977_at	1.45	2.31E-17	2.3	4.00E-14
MAP1B	226084_at	1.44	1.12E-36	2.2	3.40E-13
PXDN	212012_at	1.40	5.49E-12	2.2	1.30E-13
RNF130	217865_at	1.39	7.68E-13	2.5	1.40E-16
SEPT9	208657_s_at	-0.65	3.04E-16	-2.4	1.20E-15
FSTL5	232010_at	-1.14	1.08E-04	-2.3	7.90E-15
SEPT9	41220_at	-1.33	1.05E-51	-2.2	5.20E-13
NOL4	206045_s_at	-1.54	5.56E-15	-2.6	4.70E-18
LAMP5	219463_at	-1.69	1.76E-05	-2.2	3.50E-13
NOL4	238605_at	-1.78	6.56E-14	-2.1	1.40E-12
EDNRB	204273_at	-1.85	1.02E-03	-2.1	7.20E-13
CCND1	208711_s_at	-1.96	4.84E-13	-2.4	6.00E-16
CCND1	208712_at	-2.94	2.48E-23	-3.6	1.70E-33

supplementary Table S2.2 List of 8 epithelial genes differentially upregulated in MMSET-high patients (Fisher's method) and t(4;14)-positive MM cell lines

Gene symbol	Probe ID	median log2 FC MMSET-high vs MMSET-low patients	combined p-value	Log2 FC t(4;14) vs non-t(4;14) HMCL	adjusted p-value
FGFR3	204379_s_at	6.41	1.84E-183	7.78	1.05E-05
FGFR3	204380_s_at	1.33	1.72E-106		
DSG2	217901_at	3.26	4.22E-72	5.81	5.13E-10
DSG2	1553105_s_at	1.36	1.59E-63		
DSP	200606_at	0.88	4.05E-34	3.11	3.12E-02
CXADR	203917_at	1.21	2.95E-18	2.13	8.53E-04
CXADR	226374_at	0.87	7.35E-15		
CXADR	239155_at	0.61	2.10E-09		
CXADR	1555716_a_at	0.02	1.94E-04		
CDS1	226185_at	0.72	3.18E-11	2.04	3.94E-02
CDS1	205709_s_at	0.71	1.29E-13		
CDS1	226187_at	0.28	2.16E-07		
EXPH5	214734_at	0.26	8.74E-04	1.84	1.78E-03
JUP	201015_s_at	0.48	5.34E-10	1.70	1.48E-02
C1orf106	219010_at	0.94	5.87E-05	1.69	3.72E-02

2.7 References

1. Ghobrial IM. Myeloma as a model for the process of metastasis: implications for therapy. *Blood*. 2012;120(1):20-30.
2. Manier S, Salem KZ, Park J, Landau DA, Getz G, Ghobrial IM. Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol*. 2016.
3. Azab AK, Hu J, Quang P, et al. Hypoxia promotes dissemination of multiple myeloma through acquisition of epithelial to mesenchymal transition-like features. *Blood*. 2012;119(24):5782-5794.
4. Li J, Pan Q, Rowan PD, et al. Heparanase promotes myeloma progression by inducing mesenchymal features and motility of myeloma cells. *Oncotarget*. 2016;7(10):11299-11309.
5. Sun Y, Pan J, Mao S, Jin J. IL-17/miR-192/IL-17Rs regulatory feedback loop facilitates multiple myeloma progression. *PLoS One*. 2014;9(12):e114647.
6. Bergsagel PL, Kuehl WM. Chromosome translocations in multiple myeloma. *Oncogene*. 2001;20(40):5611-5622.
7. Chng WJ, Glebov O, Bergsagel PL, Kuehl WM. Genetic events in the pathogenesis of multiple myeloma. *Best Pract Res Clin Haematol*. 2007;20(4):571-596.
8. Fonseca R, Bergsagel PL, Drach J, et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. *Leukemia*. 2009;23(12):2210-2221.
9. Keats JJ, Reiman T, Maxwell CA, et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood*. 2003;101(4):1520-1529.
10. Chang H, Sloan S, Li D, et al. The t(4;14) is associated with poor prognosis in myeloma patients undergoing autologous stem cell transplant. *Br J Haematol*. 2004;125(1):64-68.
11. Cavo M, Terragna C, Renzulli M, et al. Poor outcome with front-line autologous transplantation in t(4;14) multiple myeloma: low complete remission rate and short duration of remission. *J Clin Oncol*. 2006;24(3):e4-5.
12. Gonsalves WI, Rajkumar SV, Gupta V, et al. Quantification of clonal circulating plasma cells in newly diagnosed multiple myeloma: implications for redefining high-risk myeloma. *Leukemia*. 2014;28(10):2060-2065.
13. Besse L, Sedlarikova L, Greslikova H, et al. Cytogenetics in multiple myeloma patients progressing into extramedullary disease. *Eur J Haematol*. 2016;97(1):93-100.
14. Rasche L, Bernard C, Topp MS, et al. Features of extramedullary myeloma relapse: high proliferation, minimal marrow involvement, adverse cytogenetics: a retrospective single-center study of 24 cases. *Ann Hematol*. 2012;91(7):1031-1037.
15. Keats JJ, Maxwell CA, Taylor BJ, et al. Overexpression of transcripts originating from the MMSET locus characterizes all t(4;14)(p16;q32)-positive multiple myeloma patients. *Blood*. 2005;105(10):4060-4069.
16. Santra M, Zhan F, Tian E, Barlogie B, Shaughnessy J, Jr. A subset of multiple myeloma harboring the t(4;14)(p16;q32) translocation lacks FGFR3 expression but maintains an IGH/MMSET fusion transcript. *Blood*. 2003;101(6):2374-2376.
17. Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. *Blood*. 2006;108(6):2020-2028.
18. Kuo AJ, Cheung P, Chen K, et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol Cell*. 2011;44(4):609-620.

19. Li Y, Trojer P, Xu CF, et al. The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J Biol Chem*. 2009;284(49):34283-34295.
20. Martinez-Garcia E, Popovic R, Min D-J, et al. The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood*. 2011;117(1):211-220.
21. Popovic R, Martinez-Garcia E, Giannopoulou EG, et al. Histone methyltransferase MMSET/NSD2 alters EZH2 binding and reprograms the myeloma epigenome through global and focal changes in H3K36 and H3K27 methylation. *PLoS Genet*. 2014;10(9):e1004566.
22. Shah MY, Martinez-Garcia E, Phillip JM, et al. MMSET/WHSC1 enhances DNA damage repair leading to an increase in resistance to chemotherapeutic agents. *Oncogene*. 2016.
23. Marango J, Shimoyama M, Nishio H, et al. The MMSET protein is a histone methyltransferase with characteristics of a transcriptional corepressor. *Blood*. 2008;111(6):3145-3154.
24. Hudlebusch HR, Santoni-Rugiu E, Simon R, et al. The histone methyltransferase and putative oncoprotein MMSET is overexpressed in a large variety of human tumors. *Clin Cancer Res*. 2011;17(9):2919-2933.
25. Ezponda T, Popovic R, Shah MY, et al. The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. *Oncogene*. 2012.
26. Yang J, Mani SA, Donaher JL, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*. 2004;117(7):927-939.
27. Groen RW, de Rooij MF, Kocemba KA, et al. N-cadherin-mediated interaction with multiple myeloma cells inhibits osteoblast differentiation. *Haematologica*. 2011;96(11):1653-1661.
28. Dring AM, Davies FE, Fenton JA, et al. A global expression-based analysis of the consequences of the t(4;14) translocation in myeloma. *Clin Cancer Res*. 2004;10(17):5692-5701.
29. Broyl A, Hose D, Lokhorst H, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood*. 2010;116(14):2543-2553.
30. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467-472.
31. Hose D, Reme T, Hielscher T, et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. *Haematologica*. 2011;96(1):87-95.
32. Tan TZ, Miow QH, Miki Y, et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med*. 2014;6(10):1279-1293.
33. Isenmann S, Arthur A, Zannettino AC, et al. TWIST family of basic helix-loop-helix transcription factors mediate human mesenchymal stem cell growth and commitment. *Stem Cells*. 2009;27(10):2457-2468.
34. Fisher RA. Statistical methods for research workers: Genesis Publishing Pvt Ltd; 1925.
35. Campain A, Yang YH. Comparison study of microarray meta-analysis methods. *BMC Bioinformatics*. 2010;11:408.

36. Wu SP, Pfeiffer RM, Ahn IE, et al. Impact of Genes Highly Correlated with MMSET Myeloma on the Survival of Non-MMSET Myeloma Patients. *Clin Cancer Res*. 2016;22(16):4039-4044.
37. Chng WJ, Braggio E, Mulligan G, et al. The centrosome index is a powerful prognostic marker in myeloma and identifies a cohort of patients that might benefit from aurora kinase inhibition. *Blood*. 2008;111(3):1603-1609.
38. Chung TH, Mulligan G, Fonseca R, Chng WJ. A novel measure of chromosome instability can account for prognostic difference in multiple myeloma. *PLoS One*. 2013;8(6):e66361.
39. Decaux O, Lode L, Magrangeas F, et al. Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myelome. *J Clin Oncol*. 2008;26(29):4798-4805.
40. Dickens NJ, Walker BA, Leone PE, et al. Homozygous deletion mapping in myeloma samples identifies genes and an expression signature relevant to pathogenesis and outcome. *Clin Cancer Res*. 2010;16(6):1856-1864.
41. Moreaux J, Hose D, Kassambara A, et al. Osteoclast-gene expression profiling reveals osteoclast-derived CCR2 chemokines promoting myeloma cell migration. *Blood*. 2011;117(4):1280-1290.
42. Agnelli L, Tassone P, Neri A. Molecular profiling of multiple myeloma: from gene expression analysis to next-generation sequencing. *Expert Opin Biol Ther*. 2013;13 Suppl 1:S55-68.
43. Annunziata CM, Hernandez L, Davis RE, et al. A mechanistic rationale for MEK inhibitor therapy in myeloma based on blockade of MAF oncogene expression. *Blood*. 2011;117(8):2396-2404.
44. Moreaux J, Klein B, Bataille R, et al. A high-risk signature for patients with multiple myeloma established from the molecular classification of human myeloma cell lines. *Haematologica*. 2011;96(4):574-582.
45. Chiba H, Osanai M, Murata M, Kojima T, Sawada N. Transmembrane proteins of tight junctions. *Biochim Biophys Acta*. 2008;1778(3):588-600.
46. Brito JLR, Walker B, Jenner M, et al. MMSET deregulation affects cell cycle progression and adhesion regulons in t(4;14) myeloma plasma cells. *Haematologica*. 2009;94(1):78-86.
47. Cheng T, Roth B, Choi W, Black PC, Dinney C, McConkey DJ. Fibroblast growth factor receptors-1 and -3 play distinct roles in the regulation of bladder cancer growth and metastasis: implications for therapeutic targeting. *PLoS One*. 2013;8(2):e57284.
48. Hanze J, Henrici M, Hegele A, Hofmann R, Olbert PJ. Epithelial mesenchymal transition status is associated with anti-cancer responses towards receptor tyrosine-kinase inhibition by dovitinib in human bladder cancer cells. *BMC Cancer*. 2013;13:589.
49. Lafitte M, Moranvillier I, Garcia S, et al. FGFR3 has tumor suppressor properties in cells with epithelial phenotype. *Mol Cancer*. 2013;12:83.
50. Gust KM, McConkey DJ, Awrey S, et al. Fibroblast growth factor receptor 3 is a rational therapeutic target in bladder cancer. *Mol Cancer Ther*. 2013;12(7):1245-1254.
51. Chesi M, Nardini E, Lim RS, Smith KD, Kuehl WM, Bergsagel PL. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood*. 1998;92(9):3025-3034.

52. Kurrey NK, K A, Bapat SA. Snail and Slug are major determinants of ovarian cancer invasiveness at the transcription level. *Gynecol Oncol*. 2005;97(1):155-165.
53. Barber AG, Castillo-Martin M, Bonal DM, Rybicki BA, Christiano AM, Cordon-Cardo C. Characterization of desmoglein expression in the normal prostatic gland. Desmoglein 2 is an independent prognostic factor for aggressive prostate cancer. *PLoS One*. 2014;9(6):e98786.
54. Brennan D, Mahoney MG. Increased expression of Dsg2 in malignant skin carcinomas: A tissue-microarray based study. *Cell Adh Migr*. 2009;3(2):148-154.
55. Drexler HG, Matsuo Y. Malignant hematopoietic cell lines: in vitro models for the study of multiple myeloma and plasma cell leukemia. *Leuk Res*. 2000;24(8):681-703.
56. Lombardi L, Poretti G, Mattioli M, et al. Molecular characterization of human multiple myeloma cell lines by integrative genomics: insights into the biology of the disease. *Genes Chromosomes Cancer*. 2007;46(3):226-238.
57. Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer*. 2005;5(9):744-749.
58. Kwok WK, Ling MT, Lee TW, et al. Up-regulation of TWIST in prostate cancer and its implication as a therapeutic target. *Cancer Res*. 2005;65(12):5153-5162.
59. Goswami M, Duvic M, Dougherty A, Ni X. Increased Twist expression in advanced stage of mycosis fungoides and Sezary syndrome. *J Cutan Pathol*. 2012;39(5):500-507.
60. van Doorn R, Dijkman R, Vermeer MH, et al. Aberrant expression of the tyrosine kinase receptor EphA4 and the transcription factor twist in Sezary syndrome identified by gene expression analysis. *Cancer Res*. 2004;64(16):5578-5586.
61. Merindol N, Riquet A, Szablewski V, Eliaou JF, Puisieux A, Bonnefoy N. The emerging role of Twist proteins in hematopoietic cells and hematological malignancies. *Blood Cancer J*. 2014;4:e206.
62. Lemma S, Karihtala P, Haapasaari KM, et al. Biological roles and prognostic values of the epithelial-mesenchymal transition-mediating transcription factors Twist, ZEB1 and Slug in diffuse large B-cell lymphoma. *Histopathology*. 2013;62(2):326-333.
63. Yang JZ, Lian WG, Sun LX, Qi DW, Ding Y, Zhang XH. High nuclear expression of Twist1 in the skeletal extramedullary disease of myeloma patients predicts inferior survival. *Pathol Res Pract*. 2016;212(3):210-216.

Chapter 3:

TWIST1 promotes migration and motility of multiple myeloma plasma cells *in vitro*

Statement of Authorship

Title of Paper	TWIST1 promotes motility of myeloma cells in vitro
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Chee Man Cheong, Krzysztof Mrozik, Duncan Hewett, Kate Vandyke & Andrew C.W. Zannettino (2017) TWIST1 promotes motility of myeloma cells in vitro

Principal Author

Name of Principal Author (Candidate)	Chee Man Cheong
Contribution to the Paper	Design experiments, performed experiments, analysed data and wrote manuscript
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 23/12/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Krzysztof Mrozik
Contribution to the Paper	Data interpretation and performed experiments
Signature	Date: 23/12/16

Name of Co-Author	Duncan Hewett
Contribution to the Paper	Data interpretation and manuscript evaluation
Signature	Date: 23/12/16

Name of Co-Author	Kate Vandyke		
Contribution to the Paper	Design experiments, data interpretation, manuscript evaluation and co-supervised development of work		
Signature		Date	23/12/16

Name of Co-Author	Andrew Zannettino		
Contribution to the Paper	Designed experiments, data interpretation, manuscript revision, provided funding and supervised development of work		
Signature		Date	23/12/2016

Please cut and paste additional co-author panels here as required.

Chapter 3: TWIST1 promotes migration and motility of multiple myeloma plasma cells *in vitro*

Chee Man Cheong[1,2], Krzysztof Mrozik[1,2], Duncan Hewett[1,2], Kate Vandyke[1,2,3], Andrew Zannettino[1,2,3,4]

1. Myeloma Research Group, Discipline of Physiology, Adelaide Medical School, The University of Adelaide, Adelaide, Australia.
2. Cancer Theme, South Australian Health & Medical Research Institute (SAHMRI), Adelaide, Australia.
3. SA Pathology, Adelaide, Australia.
4. Centre for Cancer Biology, University of South Australia, Adelaide, Australia.

Keywords: multiple myeloma, TWIST1, RNA-seq, migration

3.1 Abstract

TWIST1 is a key transcription factor that orchestrates the epithelial-to-mesenchymal transition during cancer metastasis. While overexpression of TWIST1 was found to associate with MMSET expression in t(4;14) multiple myeloma (MM; Chapter 2), a functional role for TWIST1 in MM PC biology and MM disease development remains to be determined. In this study, we characterised the transcriptomic changes mediated by TWIST1 overexpression in MM PCs using RNA-sequencing. Furthermore, we assessed the consequences of TWIST1 overexpression on the growth, adhesion and migration of MM PCs *in vitro*. TWIST1 was found to upregulate the expressions of *LGR6*, *RAPGEF3* and *PREX1*, which have been previously shown to regulate cell motility. While TWIST1 overexpression did not affect cell proliferation *in* vitro, TWIST1 enhanced MM cell adhesion to fibronectin and increased cell migration by enhancing dynamic F-actin polymerisation. Taken together, these results suggest that TWIST1 expression may contribute to MM PC dissemination and disease pathogenesis by regulating cytoskeletal remodelling and cell adhesion to extracellular matrix components

3.2 Introduction

TWIST1 is a transcription factor belonging to the basic helix-loop-helix (bHLH) family of proteins.[2, 3] TWIST1 typically form homodimers or heterodimers with class A bHLH transcription factor E proteins (E12 and E47), mediated by the conserved bHLH domain.[4] Dimerised TWIST1 transcription factors regulate gene transcription by binding to the conserved E-box (CANNTG) motif within the promoter regions of genes[5, 6]. The nature of the dimerised partner dictates the DNA binding affinity and preferred target sites, allowing TWIST1 to act as either a transcription activator or repressor, thereby mediating different biological outcomes.[7, 8]

TWIST1 was first described in *Drosophila* as a transcription factor essential for mesodermal and myogenic tissue development during embryogenesis [9] In rodents, *TWIST1* expression is first detected at embryonic day 7.0 - 7.5 (E7.0-E7.5) following mesoderm differentiation.[10] In contrast to its myogenesis-promoting role in Drosophila, TWIST1 was found to repress muscle cell differentiation in mouse and human embryoid bodies.[11-14] In adult human tissues, TWIST1 is expressed in mesoderm-derived tissues such as heart and skeletal muscle.[15] Additionally, TWIST1 is expressed in bone marrow-derived mesenchymal stem cells (MSC) and plays an important role in maintaining an immature stem cells phenotype. TWIST1 overexpression in MSC was found to promote adipogenesis, while inhibiting osteogenesis and chondrogenesis.[16] Furthermore, TWIST1 has been implicated in the self-renewal of haematopoietic stem cells (HSC), with murine HSCs, which express high levels of TWIST1, exhibiting superior hematopoietic reconstitution and engraftment potential.[17] Recent studies also show that TWIST1 can regulate inflammation by mediating the suppression of pro-inflammatory cytokine TNFα in macrophages[18, 19] and inhibiting pathogenic inflammation in T helper 1 lymphocytes.[20-22]

In addition to its roles in mesodermal development, blood cell production and inflammation, TWIST1 has also been implicated in the development and progression of cancer. To this end, TWIST1 is expressed at high levels in various tumours including carcinomas, sarcomas, neuroblastomas and gliomas (reviewed in[23, 24]) and is associated with invasive, metastatic phenotype. Other roles for TWIST1 in oncogenesis include overcoming oncogene-induced senescence and apoptosis[25-27], chemoresistance[28-30], and by promoting a stem cell like phenotype.[31, 32]

Recently, TWIST1 has emerged as a novel risk factor in heamatological malignancies. Dysregulation of TWIST1 has been detected in several haematological malignancies including in myelodysplastic syndrome (MDS)[33, 34], acute myeloid leukaemia (AML)[35], chronic myeloid leukaemia (CML)[36, 37], anaplastic large cell lymphoma (ALCL)[38] and diffuse large B-cell lymphoma (DBCL)[39] (reviewed in [40, 41]). Notably, TWIST1 upregulation in haematological malignancies has been shown to inhibit apoptosis.[33, 42] For example, in MDS, TWIST1 interacts with p53 to inhibit TNFα-induced apoptosis. In addition, TWIST1 has also been implicated in the increased resistance towards therapeutic agents, with TWIST1 knockdown in CML and ALCL tumour cells resulting in increased sensitivity of Imatinib[36] and Crizotinib[38], respectively. The mechanism(s) leading to TWIST1 re-expression in cancer cells appears to be context-dependent. For example, deregulation of multiple transcription factors and signalling pathways such as STAT3, HIF-1α, WNT1, MAPK, TGFβ, and NF-κB (reviewed in[23, 43]), deregulation of microRNA[44, 45] and post-translational modifications of TWIST1 protein[46], have been shown to activate TWIST1 in different cancer types. Additionally, epigenetic regulation of TWIST1 has been reported in prostate cancer.

To date, the role of TWIST1 in multiple myeloma (MM) disease progression has not been determined. However, as shown in Chapter 2 and in studies by other groups[47, 48], ectopic expression of the histone methyltransferase MMSET, results in increased TWIST1 expression in MM PCs, leading to increased cell invasion and migration.[47] This close association between TWIST1 and *MMSET* in MM, highlights a potential role for TWIST1 in pathogenesis of MM as part of the mechanisms regulated by *MMSET* overexpression. In this study, we utilised RNA sequencing to determine the transcriptomic changes resulting from *TWIST1* overexpression in MM PCs. In addition, we investigated the functional consequences of TWIST1 overexpression in relation to the growth, adhesion and migration of MM PC *in vitro*.

3.3 Materials and Methods

Cell cultures

Human MM cell lines WL2 and KMS11 were kindly provided by Professor Doug Joshua (Royal Prince Alfred Hospital, Sydney, Australia) and Professor Andrew Spencer (Monash University, Melbourne, Australia), respectively. All cell lines were cultured in RPMI-1640 medium supplemented with 10% (v/v) fetal calf serum and

additives (2mM L-glutamine, 1mM sodium pyruvate, 15 mM HEPES, 50 U/ml penicillin and 50 μg/ml streptomycin; from Sigma-Aldrich, Sydney, Australia).

Generation of TWIST1-overexpressing WL2 cells

The WL2 cell line overexpressing TWIST1 were generated by retroviral infection with pruf-IRES-GFP harbouring full length cDNA encoding human *TWIST1* (kindly provided by Prof. Stan Gronthos, The University of Adelaide) as previously described.[16] Briefly, empty vector pruf-IRES-GFP or pruf-IRES-GFP-hTWIST1 were transfected into HEK293T cells together with packaging vector pEQPAM3[49] (which encodes for Pol, GAG and viral envelope proteins). Retroviral supernatant collected from transiently transfected HEK293T cells was used to transduce WL2 cells as previously described.[50] Stable TWIST1-overexpressing WL2 (WL2-TWIST1) or empty vector control (WL2-EV) cell lines were generated from the top 30% GFP-positive cells sorted using BD FACSFusion.

Library preparation for whole transcriptome RNA-sequencing

Total RNA was extracted from stable TWIST1 overexpressing WL2 cells and empty vector control using RNA kit (Qiagen) according to the manufacturer's instructions. RNA quality was determined using Bioanalyzer 2200 (Agilent) and were stored at -80°C. All samples used for cDNA library construction had a RIN score > 8. Library construction and RNA-sequencing with NextSeq500 (Illumina) were performed at David Gunn Genomics Facility (SAHMRI, Adelaide). RNA-seq libraries were prepared using NEXTflex™ Rapid Directional mRNA-Seq Kit Bundle with RNA-Seq Barcodes and poly(A) beads (BIOO Scientific) according to manufacturer's instruction.

RNA-sequencing analysis

Raw fastq files containing single-end reads (1 × 75 bp) obtained from RNA-sequencing were analysed according workflow described in supplementary Figure S3.1. Briefly, read quality was assessed using FastQC and over-represented adapter sequences were trimmed from the raw fastq files using Trimmomatic version 0.33 and repeated quality assessment. Filtered reads were mapped to the reference genome hg19 using STAR version 2.5.0b. BAM files from individual lanes of the same sample were combined. The number of reads mapped were counted using featureCounts (part of Rsubread version 1.12.6). Transcripts expressed at levels below 5 counts per million

reads, in at least three libraries, were filtered out from downstream analysis. Differential expression was determined using quasi-likelihood F-test from edgeR R package to account for variability due to relatively small sample size.[51] Significantly regulated genes were identified using a cut-off of 1-fold or greater changes in mean expression and FDR < 0.05. Reads per kilobase of transcript per million mapper reads (RPKM) values were calculated with edgeR.

GO enrichment analysis

To determine the biological implications of the differentially expressed genes, gene ontology (GO) annotation enrichment analysis was performed using DAVID Bioinformatics Resources 6.8.[1] Functional Annotation Clustering module with high stringency was used to reduce redundancy of highly similar annotations into functional annotation groups.

Microarray analysis

Microarray analysis was performed as previously described in Chapter 2. A total of 135 MMSET-positive GEPs from four microarray datasets (GSE19784, GSE26760, MTAB-363 and MTAB-317) were used. Patients were segregated into TWIST1-low or TWIST1-high based on the median of TWIST1 expression level (213943_at). Differentially expressed gene probes and their corresponding p-values of patients GEPs were determined using Linear Models for Microarray Data (LIMMA) package. The p-values for each gene probes, across four studies, were then combined using Fisher's method. The significantly deregulated gene probes were identified using a cut-off of 1-fold or greater changes in mean expression and combined p-value < 0.05.

Reverse transcription-quantitative polymerase chain reaction (RT-qPCR)

Total RNA was isolated from cell lines using TRIzol (Life Technologies, Mulgrave, Australia) and cDNA was synthesized using Superscript III First-Strand Synthesis System (Life Technologies). qPCR were performed using RT2 SYBR Green qPCR Mastermix (Qiagen, Chadstone, Ausrtalia) on the Bio-Rad CFX Connect (Bio-Rad) using the following primers: *GAPDH* (F: 5'-ACCCAGAAGACTGTGGATGG-3'; R: 5'-CAGTGAGCTTCCCGTTCAG-3'), *TWIST1* (F: 5'-TCTTACGAGGAGCTGCAGACGCA-3'; R: 5'-ATCTTGGAGTCCAGCTCGTCGCT-3')., *HECW* (F: 5'-GCAAGCCCAGAACCTCATTA-3' and R: 5'- CCCAGGCTCAATTCTCCATA-3'),

UNC13C (F: 5'- CCAGGAACGGGAGATCATAA-3' and R: 5'- TGGTGACCATGTGTTGCTTT-3'), *RAPGEF3* (F: 5'- ACTCACCAACAGCGAGGAGT-3'; R: 5'- TGCCGATAGAGCCTAAGGTG-3'), *LGR6* (F: 5'- CCTCAGCATGAACAACCTCA-3'; R: 5'- TTCCTCCCAGCTGATTGTTC-3'), *PREX1* (F: 5'- CCTCAACGAGATCTTGGGCA-3'; R: 5'- CGGCTCCGGGTGTAAACAATA-3'), *CDH1* (F: 5'- ATAGAGAACGCATTGCCACATA-3'; R: 5'- ATGACAGACCCCTTAAAGACCT-3') and *CDH2* (F: 5'- GGCAGTAAAATTGAGCCTGAAG-3'; R: 5'-AGTTTTCTGGCAAGTTGATTGG-3'). Gene expression was represented relative to GAPDH expression, calculated using the $2^{-\Delta CT}$ method.

Immunoblotting

Whole cell lysates were prepared in reducing buffer as previously described.[52] Proteins (30μg) were resolved in 12% SDS-polyacrylamide gels, transferred to PVDF membranes (GE Healthcare). Membranes were incubated in 2.5% enhanced chemiluminescence Blocking Agent (GE Healthcare) for 1 hour at room temperature and then probed overnight at 4°C with antibodies against TWIST1 (1:200 dilution; ab50887, Abcam) and HSP90 (1:5000 dilution; sc-7947, Santa Cruz). Proteins of interest were detected by incubating the membrane in alkaline phosphatase-conjugated anti-mouse IgG (1:5000 dilution; Millipore). Proteins were visualised using enhanced chemifluorescence substrate (ECF, GE Healthcare) on a ChemiDoc (BioRad).

Proliferation assay

For WST-1 assays, WL2 cells were seeded at 1×10^5 cells/well in phenol red-free RPMI-1640 media in 96-well plates in triplicates. Relative cell proliferation was measured at 24-hour intervals by adding WST-1 reagent (Roche, Basel, Switzerland) according to the manufacturer's instruction. For BrdU assays, WL2 cells were seeded at 1×10^5 cells/well in triplicate in 96-well plates in triplicates. BrdU (Roche) was added to the cells incubated for 2 hours at 37°C. BrdU incorporation was measured using a BrdU Cell Proliferation ELISA kit (Roche) according to manufacturer's protocol. Absorbance was measured at 450 nm with a microplate reader.

Cell cycle analysis

Cell cycle distributions were assessed by PI staining. WL2 cells were cultured at 5×10^5 cells/mL for 24 and 48 hours. Cells were fixed in 70% ice cold ethanol, washed twice in PBS and stained with 40 μg/mL PI containing 20 μg/mL RNaseA for 30 min, followed by analysis on a BD LSRFortessa flow cytometer.

Adhesion assay

Adhesion assays examining the adhesion of PC to BM stromal cells and BM endothelial cells were performed as previously described.[53] BMSCs and BMECs were seeded at 1×10^4 cells per well in a 96-well plate and allowed to adhere overnight. Empty wells (plastic) were used as controls for adhesion. For adhesion to fibronectin, 96-well plates were coated with 10 μg/ml plasma fibronectin or BSA overnight at 37°C. WL2 cells were added at 1×10^5 cells per well in serum-free culture media and incubated for indicated period of time at 37˚C. Non-adherent cells were gently aspirated, followed by three washes with HBSS with 5% FCS. Images were taken of each well at × 10 magnification and adhered cells was counted using ImageJ software (http://fiji.sc).

Migration assay

Migration assays were performed as previously described.[54] Briefly, WL2 or KMS11 cells were washed once in serum free RPMI-1640 media, seeded (1×10^5 cells) on 8 μm transwells in serum-free RPMI-1640 media in quadruplicate and allowed to migrate towards 10% FCS. Migrated cells were enumerated using microscopy after 18 hours as previously described.[53] Cell migration was represented as the percentage of cells migrated to 10% FCS relative to basal migration (0% FCS).

Actin polymerisation

WL2 cells (1×10^5 cells) were resuspended in 100μL serum-free RPMI-1640 media and stimulated with or without CXCL12 (200 ng/mL) for indicated durations. Cells were then fixed in 2% paraformaldehyde for 15 min at room temperature, permeabilised in 0.2% saponin and stained with 0.5 unit of phalloidin-A680 for 30 min on ice in the dark. The mean fluorescence intensity (MFI) for each sample was determined using flow cytometry. Data was represented as percentage change in MFI in relation to unstimulated samples (mean ± SD of a representative experiment in triplicate).

RhoA and Rac1 activation assays

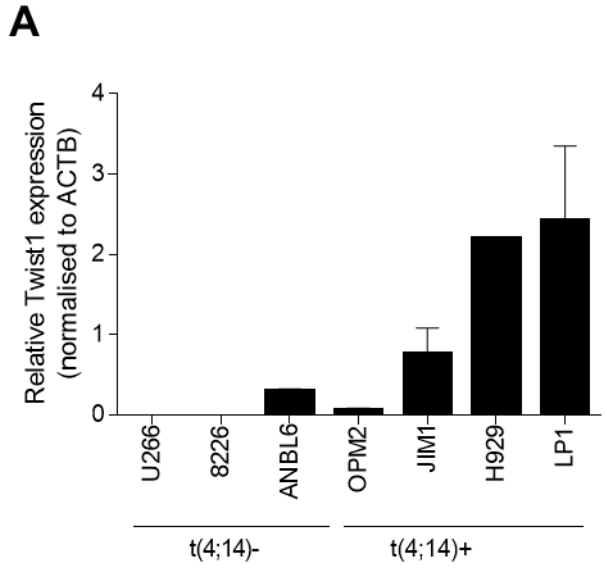
Activated RhoA or Rac1 proteins were detected using the RhoA or Rac1 pull-down activation assay biochem kit (Cytoskeleton). Briefly, WL2 cells were serum-starved overnight and stimulated with CXCL12 (200 ng/mL) for 1 minute. Cells were then pelleted, washed with ice-cold PBS and lysed in lysis buffer. Protein concentrations were determined using RcDc Reagent (BioRad) and adjusted to the same concentration. Equal amounts of protein lysates (350µg) were added to Rhothekin or PAK1 agarose beads and incubated for 1 hour with rotation at 4°C to pull down RhoA and Rac1 GTPases respectively. Beads were pelleted by centrifugation (1 min, 5000 ×*g* at 4°C), washed twice, resuspended and boiled in 20µL of 2X Laemmli buffer. Protein lysates were resolved in 12% SDS-PAGE and transferred to PVDF membranes (GE Healthcare). Activated RhoA and Rac1 were detected by immunoblotting with respective antibodies included in the kit (1:200 dilution) and visualised according to immunoblotting method above.

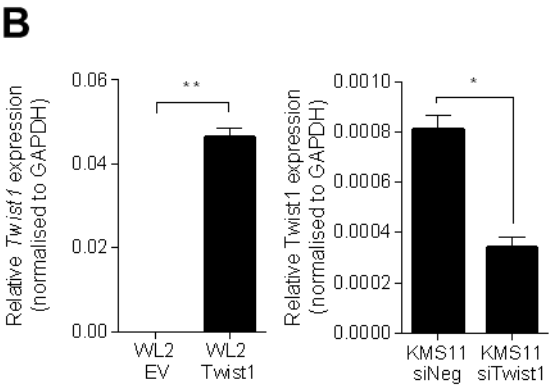
3.4 Results

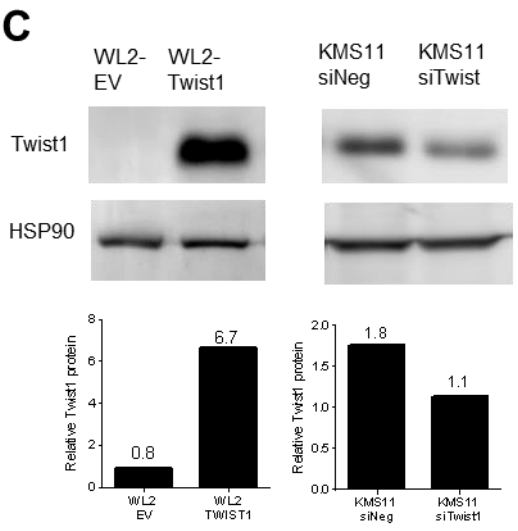
3.4.1 Generation and characterization of TWIST1 overexpressing t(4;14) negative MM cells

RT-qPCR was utilized to examine *TWIST1* expression in t(4;14)-positive HMCL. As anticipated, TWIST1 was expressed in t(4;14)-positive cell lines and was absent in t(4;14)-negative cell lines (Figure 3.1A). To further examine the contribution of TWIST1 expression to the biology of t(4;14) MM cells, *TWIST1* was stably overexpressed in t(4;14)-negative HMCL, WL2. RT-qPCR and Western blotting was utilized to confirm *TWIST1* overexpression in the WL2-TWIST1 cells compared to empty vector control (WL2-EV) (Figure 3.1B). In separate experiments, TWIST1 siRNA transfection, was used to knockdown *TWIST1* in the t(4;14)-positive KMS11 HMCL. As shown in Figure 3.1C, TWIST1 mRNA and protein expression was reduced by 60% in the KMS11-siTWIST1 HMCL compared with the KMS11-siNeg HMCL.

Figure 3.1 TWIST1 overexpression in WL2 cells and transient knockdown in KMS11 cells. (A) *TWIST1* expression in a panel of t(4;14)+ and t(4;14)- MM cell lines was analysed using qualitative real-time PCR (qPCR). Data were normalised to reference gene *GAPDH* (mean ± SD of triplicates). (B) *TWIST1* was overexpressed in t(4;14)-negative WL2 cells by retroviral transduction *TWIST1* cDNA-harbouring vector or empty vector (EV) as control. Complementary, *TWIST1* was knockdown in t(4;14)-positive KMS11 cells by transfecting the cells with *TWIST* siRNAs (siTwist) or negative control siRNA (siNeg). Overexpression and knockdown of *TWIST1* mRNA levels was confirmed by qPCR, normalised to *GAPDH* (mean ± range of two independent experiments). (C) Total protein extracts (30 μg) from (B) were resolved on 12% SDS-PAGE gels under reducing condition and immunoblotted with the indicated antibodies. Protein levels were quantified by densitometry relative to the level of loading control HSP90.


A
Relative Twist1 expression (normalised to ACTB)
4
3
2
1
0
U266
8226
ANBL6
OPM2
JIM1
H929
LP1
t(4;14)-
t(4;14)+


B
Relative Twist1 expression (normalised to GAPDH)
0.06
0.04
0.02
0.00
**
WL2 EV
WL2 Twist1
Relative Twist1 expression (normalised to GAPDH)
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
*
KMS11 siNeg
KMS11 siTwist1


C
WL2-EV
WL2-Twist1
KMS11 siNeg
KMS11 siTwist
Twist1
HSP90
Relative Twist1 protein
8
6
4
2
0
0.8
6.7
WL2 EV
WL2 TWIST1
Relative Twist1 protein
2.0
1.5
1.0
0.5
0.0
1.8
1.1
KMS11 siNeg
KMS11 siTwist1

3.4.2 Alignment of sequencing reads to reference genome

To investigate the transcriptome-wide effects of *TWIST1* overexpression in HMCL, RNA-Seq was performed. Briefly, polyA mRNA was isolated from WL2-EV cells and WL2-TWIST1 cells from three biological replicates and RNAseq was performed as described in the Materials and Methods. Data was analysed according workflow described in supplementary Figure S3.1. Using the Illumina NextSeq500, over 28 million single-end reads were obtained per sample after removal of low quality reads and adapters sequence. The average length of each read was 75 bp and all samples displayed high quality metrics with a Phred score of more than 30 (supplementary Figure S3.2) and an even nucleotide distributions along each read (supplementary Figure S3.3). Approximately 90% of the reads were aligned to the hg19 reference genome using STAR aligner, and more than 70% were uniquely mapped to reference genes (Table 3.1). The percentage of mapped reads was very similar between the two groups (WL2-EV, 90.1% ± 0.001; WL2-TWIST1, 90.1% ± 0.002; mean ± SD of triplicates), suggesting there were no detectable bias in the sequence data. The multidimensional scaling plots showed clear separation between WL2-EV and WL2-TWIST1 replicate libraries (supplementary Figure S3.4), indicating a likelihood of significant differences between transcriptomes in response to TWIST1 overexpression.

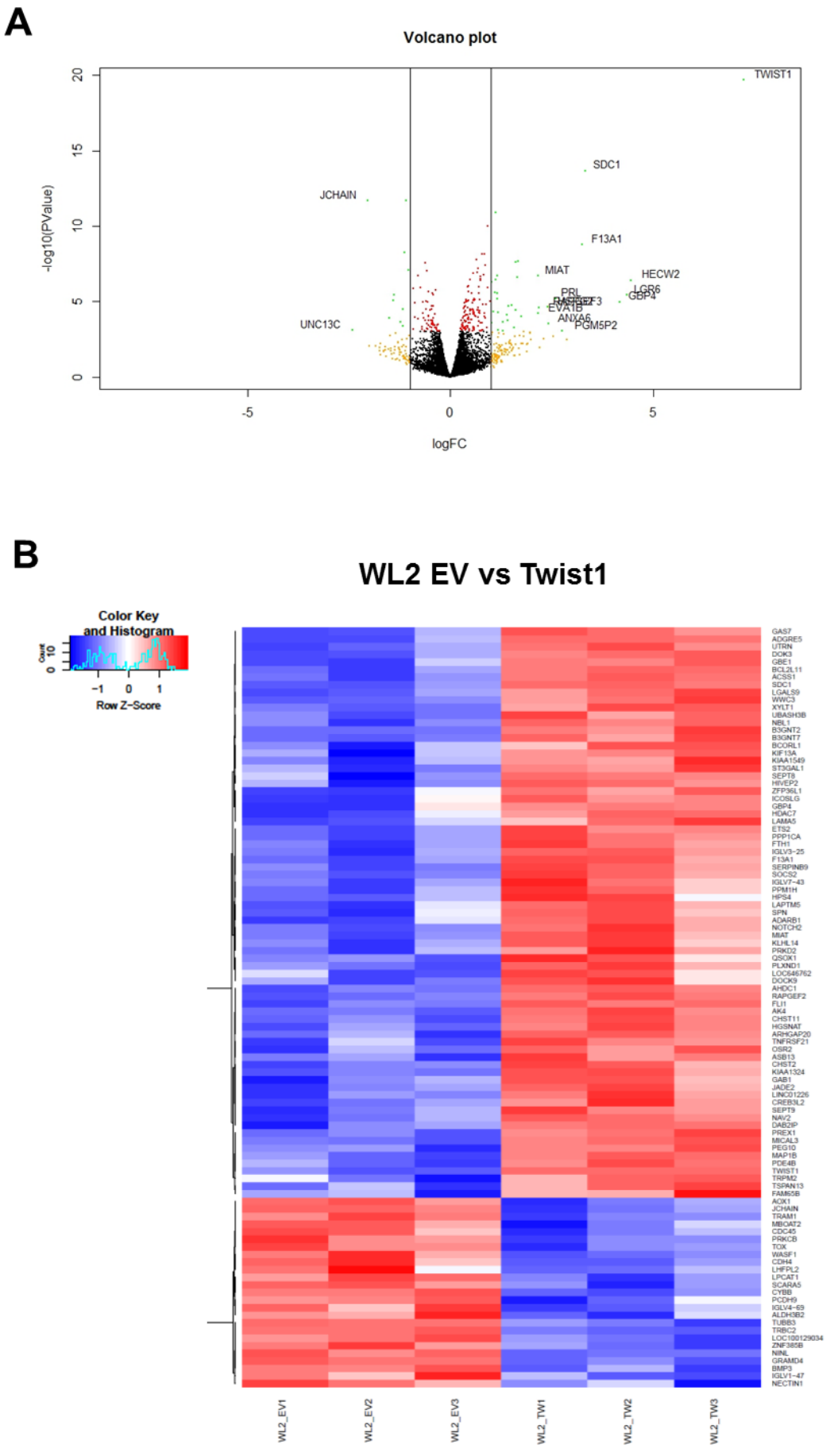
3.4.3 Differential expressed genes (DEGs) induced by TWIST1 expression in MM cells

Compared to the empty vector control, 225 genes were differentially regulated in WL2-TWIST1 HMCL, of which 159 (70.7%) genes were upregulated and 66 (29.3%) were downregulated (Table S3.1). Of these, 13 transcripts did not code for protein species. As anticipated, the most highly upregulated gene in the WL2-TWIST1 HMCL was *TWIST1* which exhibited a log2 fold change of 7.2. The overall changes in gene expression that were associated with *TWIST1* overexpression are represented in volcano plot and heatmap as shown in Figure 3.2A and B, of which *HECW, LGR6, GBP4, SDC1, F13A1, JCHAIN* and *UNC13C* were examples of the most differentially regulated genes. RT-qPCR was utilised to validate the differential expression of genes that exhibited the most statistically significant difference (fold change) in expression in RNA-seq analysis. Consistent with RNA-seq analysis, RT-qPCR revealed

Table 3.1 Raw and mapped reads of each WL2-EV or WL2-TWIST1 sample from RNA-seq analysis.

	Total reads	Mapped to genome		Unique gene mapping	
		Reads	Frequency	Reads	Frequency
WL2-EV_A	32566277	29388730	90.2%	23513167	72.2%
WL2-EV_B	31239910	28160648	90.1%	22128166	70.8%
WL2-EV_C	31117593	28016186	90.0%	22389381	72.0%
WL2-TWIST1_A	28885281	26035561	90.1%	20757236	71.9%
WL2-TWIST1_B	30678225	27704801	90.3%	21969783	71.6%
WL2-TWIST1_C	30529126	27455559	89.9%	21842845	71.5%

Figure 3.2 Differentially expressed genes in TWIST1-overexpressing WL2 cells identified from RNA-Seq by edgeR. (A) Genes differentially expressed between WL2-TWIST1 and WL2-EV cells were plotted against significance of change in volcano plot. Red dots indicate differentially expressed genes (fold change >1, FDR < 0.05) and green dots indicate genes with greater than 2 fold changes in expression. Genes with greater than 4 fold difference in expression are labelled. (B) Heatmap showed log2 cpm values and clustering for 53 genes differentially expressed in WL2-TWIST1 (39 upregulated, 14 downregulated; FDR <0.05, absolute fold change > 2) by RNA sequencing. Each column represents a biological replicate of WL2-TWIST1 cells or control. Upregulated genes were coloured in red while downregulated genes were coloured in blue


A
Volcano plot
-log10(PValue)
logFC
TWIST1
SDC1
JCHAIN
F13A1
MIAT
HECW2
LGR6
GBP4
PRL
EVA1B
ANXA6
PGM5P2
UNC13C
B
WL2 EV vs Twist1
Color Key and Histogram
Count
Row Z-Score
WL2_EV1
WL2_EV2
WL2_EV3
WL2_TW1
WL2_TW2
WL2_TW3

that *HECW* was significantly upregulated ($p < 0.0001$) while *LGR6* was significantly downregulated ($p < 0.0001$) in WL2-TWIST1 HMCL (supplementary Figure S3.5).

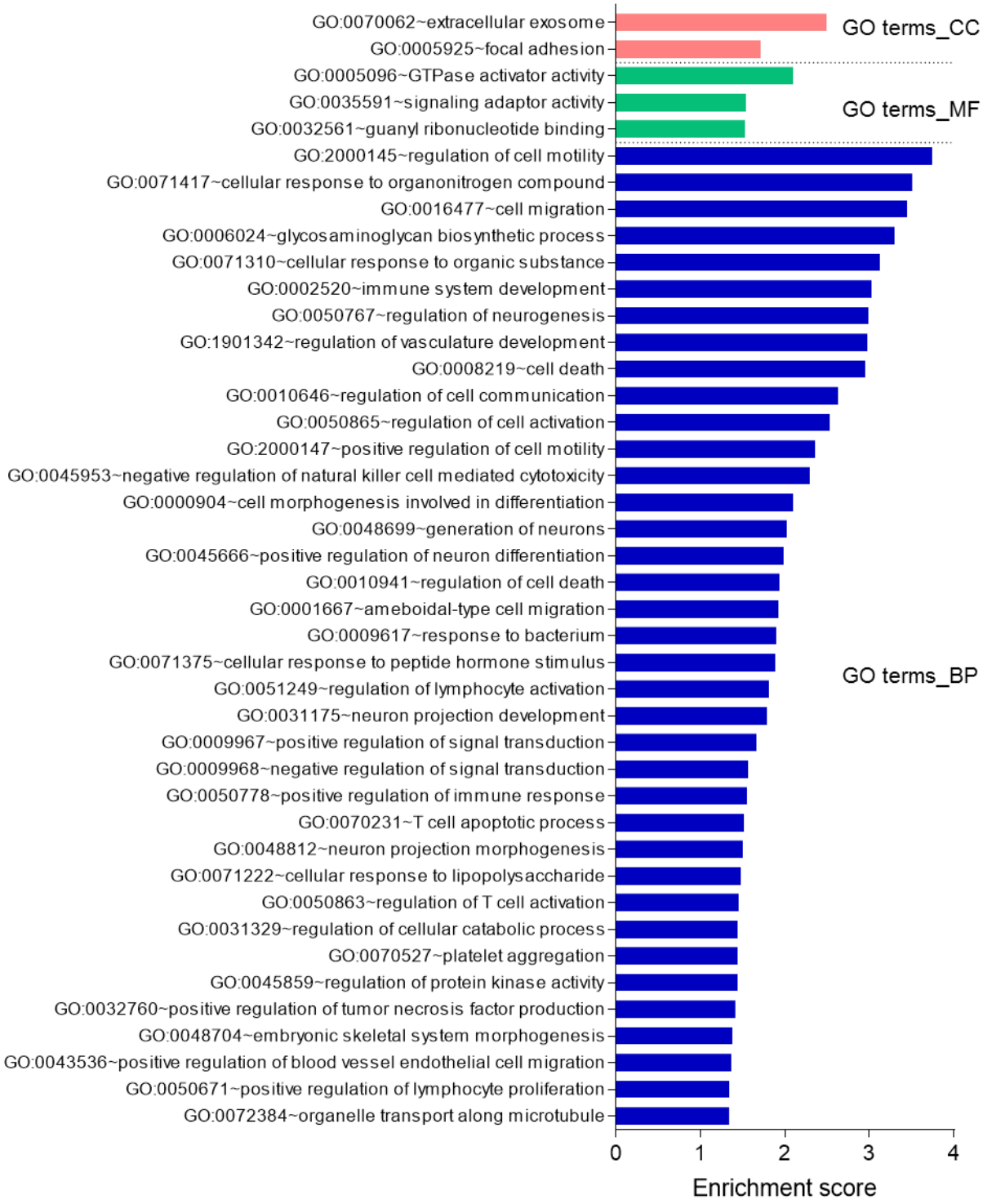
3.4.4 Pathways enrichment of DEGs induced by TWIST1 expression

To identify the potential biological significance of the DEGs, gene ontology (GO) analysis was performed using DAVID Bioinformatics Resources 6.8. Genes were categorized into GO terms associated with cellular components (CC), molecular functions (MF) and biological processes (BP). While 53 GO terms were significantly enriched in the upregulated DEG list, no GO terms were significantly enriched in the downregulated DEG list. Redundant GO terms were reduced by passing the GO terms through Functional Annotation Clustering module. Among the upregulated DEGs, 42 functional clusters were significantly enriched (EASE > 1.3, Figure 3.3). A larger proportion of the upregulated DEGs were found to be associated with the membrane-bounded vesicles (47 of 160 genes, 29.3%, Table S3.2), as indicated by the top CC annotation cluster, suggesting that TWIST1 may alter the content of exosomes which could serve to modulate the microenvironment of MM PCs. In addition, genes involved in molecular functions such as GTPase activator activity, adaptor and GTP binding (20 of 160 genes, 12.5%, Table S3.3) were significantly enriched in WL2-TWIST1 HMCL transcriptome. Among the biological processes, GO term clusters related to cell motility, response to organic compounds, cell death, regulation of cell communication and differentiation were significantly enriched (supplementary Table S3.4), suggesting a role of TWIST1 in regulating these processes.

3.4.5 Correlation of transcriptome analysis with patients GEP

To determine whether expression of putative TWIST1 targets identified from RNA-seq in WL2-TWIST1 HMCL were clinically relevant, the list of significant DEGs were compared with GEPs of MMSET-positive patients that harboured high TWIST1 expression levels (as previously described in Chapter 2). A total of 135 MMSET-positive GEPs, from four microarray datasets, were stratified based on *TWIST1* expression. Differentially expressed genes were compared between the top 50% and bottom 50% of MMSET-patients and were combined using Fisher's method as previously described (Section 2.3, Chapter 2). Using this approach, 12 genes were

Figure 3.3 Gene Ontology analysis associated with cellular components (CC), molecular functions (MF) and biological processes (BP). DAVID Functional Annotation Clustering analysis of DEGs upregulated upon *TWIST1* overexpression in WL2 cells. Significance was determined by enrichment score less than 1.3.[1]


GO:0070062~extracellular exosome
GO:0005925~focal adhesion
GO:0005096~GTPase activator activity
GO:0035591~signaling adaptor activity
GO:0032561~guanyl ribonucleotide binding
GO:2000145~regulation of cell motility
GO:0071417~cellular response to organonitrogen compound
GO:0016477~cell migration
GO:0006024~glycosaminoglycan biosynthetic process
GO:0071310~cellular response to organic substance
GO:0002520~immune system development
GO:0050767~regulation of neurogenesis
GO:1901342~regulation of vasculature development
GO:0008219~cell death
GO:0010646~regulation of cell communication
GO:0050865~regulation of cell activation
GO:2000147~positive regulation of cell motility
GO:0045953~negative regulation of natural killer cell mediated cytotoxicity
GO:0000904~cell morphogenesis involved in differentiation
GO:0048699~generation of neurons
GO:0045666~positive regulation of neuron differentiation
GO:0010941~regulation of cell death
GO:0001667~ameboidal-type cell migration
GO:0009617~response to bacterium
GO:0071375~cellular response to peptide hormone stimulus
GO:0051249~regulation of lymphocyte activation
GO:0031175~neuron projection development
GO:0009967~positive regulation of signal transduction
GO:0009968~negative regulation of signal transduction
GO:0050778~positive regulation of immune response
GO:0070231~T cell apoptotic process
GO:0048812~neuron projection morphogenesis
GO:0071222~cellular response to lipopolysaccharide
GO:0050863~regulation of T cell activation
GO:0031329~regulation of cellular catabolic process
GO:0070527~platelet aggregation
GO:0045859~regulation of protein kinase activity
GO:0032760~positive regulation of tumor necrosis factor production
GO:0048704~embryonic skeletal system morphogenesis
GO:0043536~positive regulation of blood vessel endothelial cell migration
GO:0050671~positive regulation of lymphocyte proliferation
GO:0072384~organelle transport along microtubule
GO terms_CC
GO terms_MF
GO terms_BP
0
1
2
3
4
Enrichment score

identified as being differentially expressed (8 upregulated and 4 downregulated) in both WL2-TWIST1 HMCL and TWIST1-high MM patients (Table 3.2 and 3.3). Of these, genes involved in cell migration (*STAT5A* and *PLXNB1*) and regulation of organelle organisation (*KIAA1324*) were upregulated, while genes involved in the regulation of EMT (*LDLRAD4*) were found to be downregulated. Consistent with the GO functional analysis described in the previous section, increased *TWIST1* expression in MM cells activated genes related to cell motility.

3.4.6 TWIST1 expression does not affect MM cell growth

To determine whether TWIST1 expression confers growth advantage to MM cells, cell proliferation was assessed using the WST-1 assay. Over a three-day time course, the basal proliferation of WL2-TWIST1 cells was not statistically different to that of vector control cells (p = 0.317, two-way ANOVA with Sidak's multiple comparison tests, Figure 3.4A). Similarly, both *TWIST1* overexpression in WL2 cells and transient knockdown of TWIST1 in KMS11 did not affect relative BrDU incorporation over 2 hours (WL2: $p = 0.520$, KMS11: $p = 0.954$, unpaired student's T-test; Figure 3.4B). In addition, *TWIST1* overexpression in WL2 cells, did not affect cell cycle distribution, as determined by propidium iodide staining at 24 and 48 hours after culture (Figure 3.4C and D).

3.4.7 TWIST1 expression modulates MM cell adhesion to fibronectin

Given that interactions of MM PCs and stromal cells are crucial in mediating MM cell survival and progression, the effect of TWIST1 expression on the modulation of MM cell adhesion to cells and components in BM microenvironment was assessed. As shown in Figure 3.5A and B, elevated TWIST1 did not affect the adhesion of WL2 cells to BM stromal cells ($p = 0.088$) or BM endothelial cells ($p = 0.103$). In addition to cells of the BM microenvironment, the effects of *TWIST1* expression on MM cell adhesion to fibronectin, a key extracellular matrix component involved in MM pathogenesis was evaluated.[55] As shown in Figure 3.5C, TWIST1 significantly enhanced the capacity of WL2-TWIST1 to adhere to fibronectin by 32.5% , compared with WL2-EV HMCL ($p = 0.036$, unpaired t-test).

Table 3.2 List of 8 commonly upregulated genes by TWIST1 in MMSET-patients (Fisher's method) and WL2-TWIST1 cell lines

Gene	Probe ID	median log2 FC TWIST-high vs TWIST-low MMSET-patients	combined p-value	Log2 FC WL2-TWIST1 vs WL2-EV	adjusted p-value
TWIST1	213943_at	1.60	5.68E-28	7.23	2.46E-16
STAT5A	203010_at	0.19	9.89E-03	1.41	1.54E-02
KIAA1324	226248_s_at	0.13	9.11E-03	0.75	1.09E-04
LOC646762	1568597_at	0.36	8.58E-03	0.70	6.06E-03
LOC646762	1560006_a_at	0.27	3.71E-04		
PLXNB1	215807_s_at	0.13	6.14E-03	0.45	2.04E-02
ETS2	201329_s_at	0.18	3.33E-03	0.44	3.70E-04
ACSS1	234484_s_at	0.18	2.02E-03	0.43	1.79E-03
AHDC1	205002_at	0.06	1.36E-03	0.34	3.29E-03

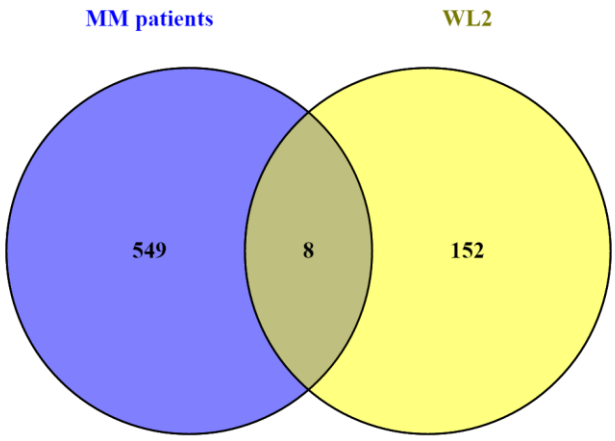


Table 3.3 List of 4 commonly downregulated genes by TWIST1 in MMSET-patients (Fisher's method) and WL2-TWIST1 cell lines

Gene	ProbeID	median log2 FC TWIST-high vs TWIST-low MMSET-patients	combined p-value	Log2 fold change WL2-TWIST1 vs WL2-EV	adjusted p-value
LDLRAD4	209574_s_at	-7.92E-02	8.12E-03	-0.54	2.79E-02
TLE3	212770_at	-4.39E-01	3.78E-04	-0.44	1.55E-02
TLE3	206472_s_at	-5.06E-01	-5.17E-01		
TRAM1	201399_s_at	-3.96E-01	7.56E-03	-0.37	5.21E-04
PLCL2	216218_s_at	-3.51E-01	2.16E-03	-0.31	2.75E-02

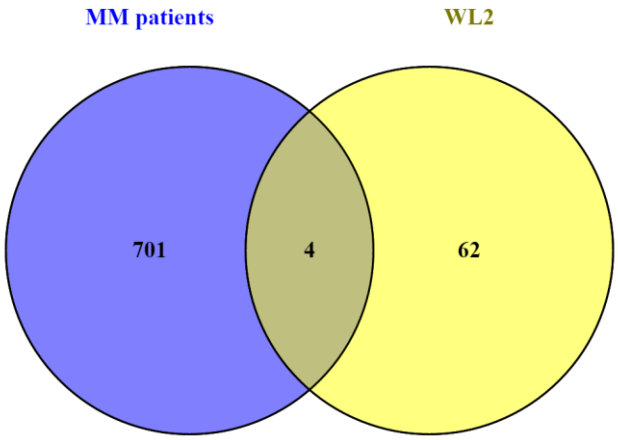


Figure 3.4 TWIST1 expression does not affect MM cell growth as determined by WST-1 and BrDU assays. (A) WL2 cells were seeded at 1×10^5 cells/mL in 96-well plate and cultured over three days. Cells were reincubated for additional two hours following addition of 10 µL WST-1 reagent per well. Absorbance values (450nm) are represented relative to time 0 (mean ± SEM of three independent experiments). (B) Growth of KMS11 cells was assessed at 48 hours post siRNAs transfection. BrDU substrate was added to KMS11 or WL2 cells seeded at 1×10^5 cells/mL in 96-well plates. BrDU incorporation was quantitated after 24 hours by immunoassay, measured by absorbance at 450nm. Graphs depict mean ± range of two independent experiments. (C & D) Cell cycle distributions were assessed by PI staining after culturing WL2 at 5×10^5 cells/mL for 24 and 48 hours, followed by analysis on flow cytometer. Representative FACS plots (PI histograms) are shown. Graphs depict mean ± SEM of three independent experiments.

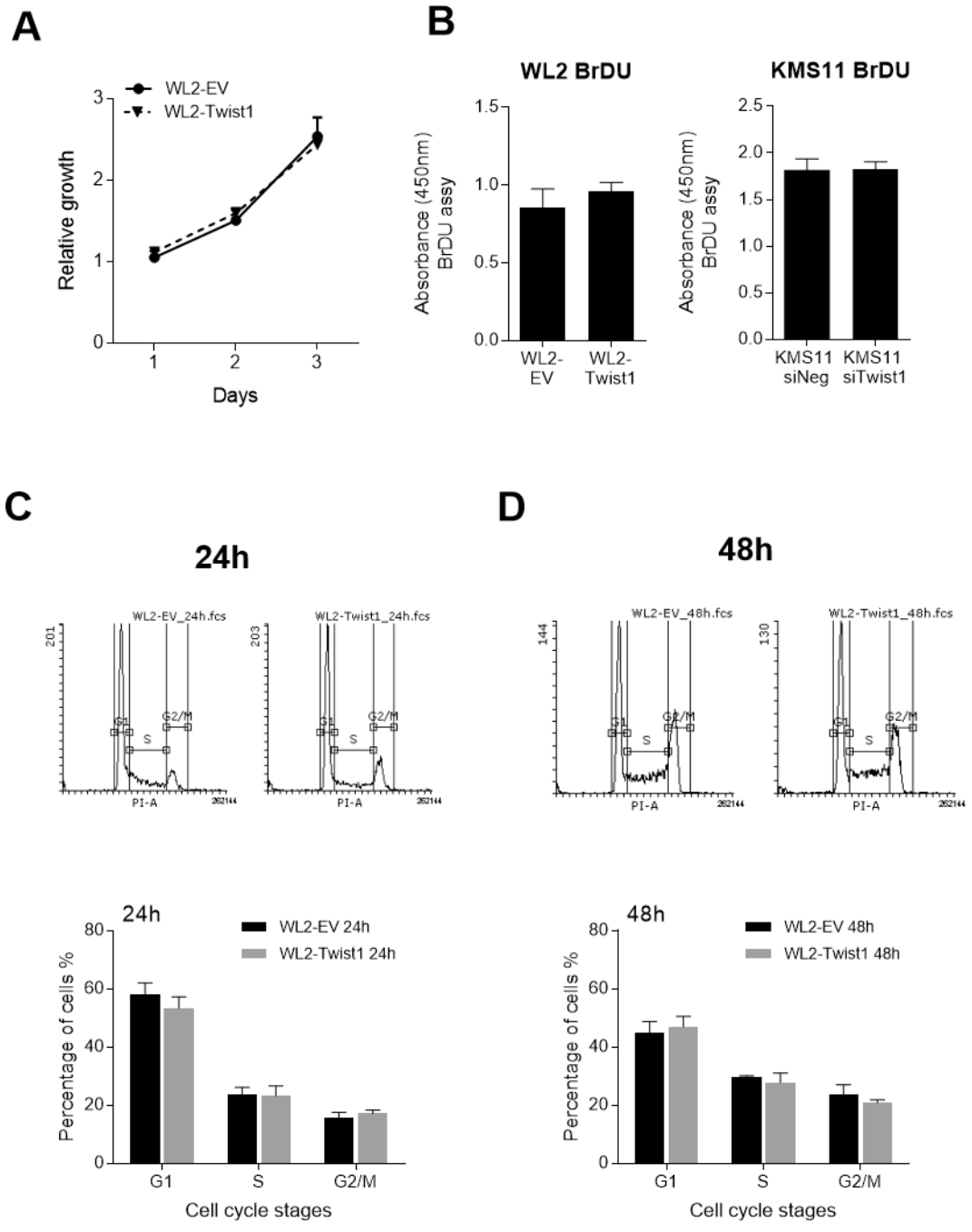
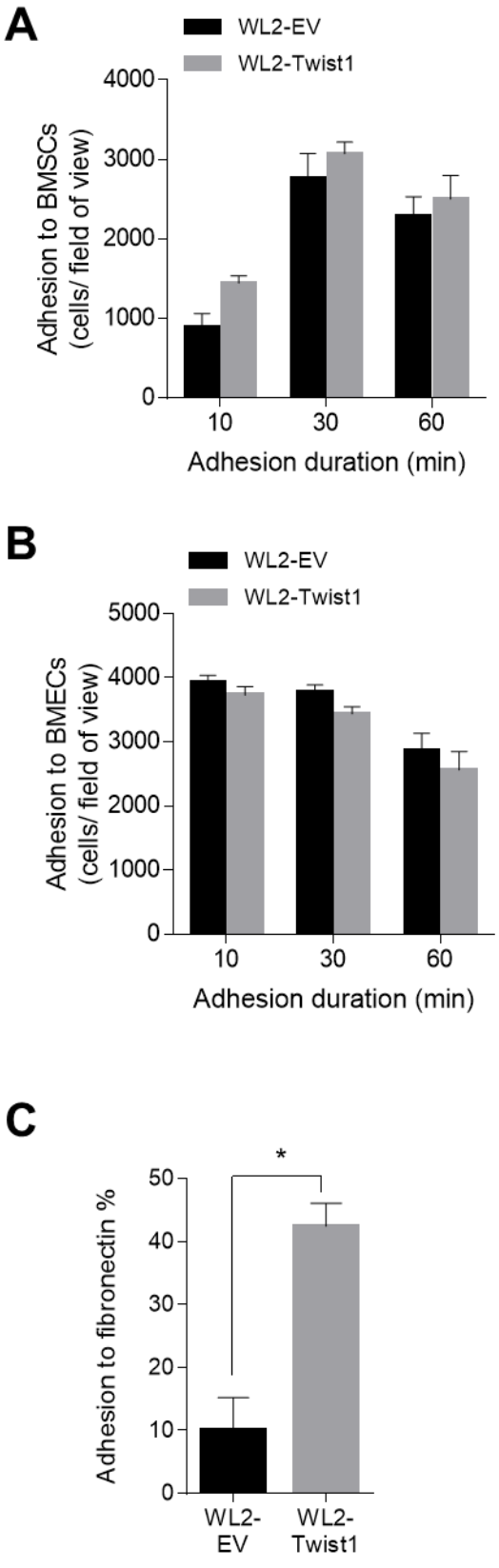

A
WL2-EV
WL2-Twist1
Relative growth
Days
B
WL2 BrDU
KMS11 BrDU
Absorbance (450nm) BrDU assy
WL2-EV
WL2-Twist1
KMS11 siNeg
KMS11 siTwist1
C
24h
WL2-EV_24h.fcs
WL2-Twist1_24h.fcs
G1
S
G2/M
PI-A
24h
WL2-EV 24h
WL2-Twist1 24h
Percentage of cells %
G1
S
G2/M
Cell cycle stages
D
48h
WL2-EV_48h.fcs
WL2-Twist1_48h.fcs
48h
WL2-EV 48h
WL2-Twist1 48h

Figure 3.5 TWIST1 expression modulates the MM cell adhesion. *TWIST1* expression did not affect adhesion rate of WL2 cells to monolayers of (A) BM stromal cells and (B) BM endothelial cells. Data represents mean ± SD of three experiments in quadruplicate (C) WL2 cells (1×10^5) were serum-starved overnight and allowed to adhere to fibronectin-coated 96-wells for 1 hour at 37°C in serum free media. Data is presented relative to BSA adhesion and the graph depicts mean ± range of two independent experiments performed in triplicate.

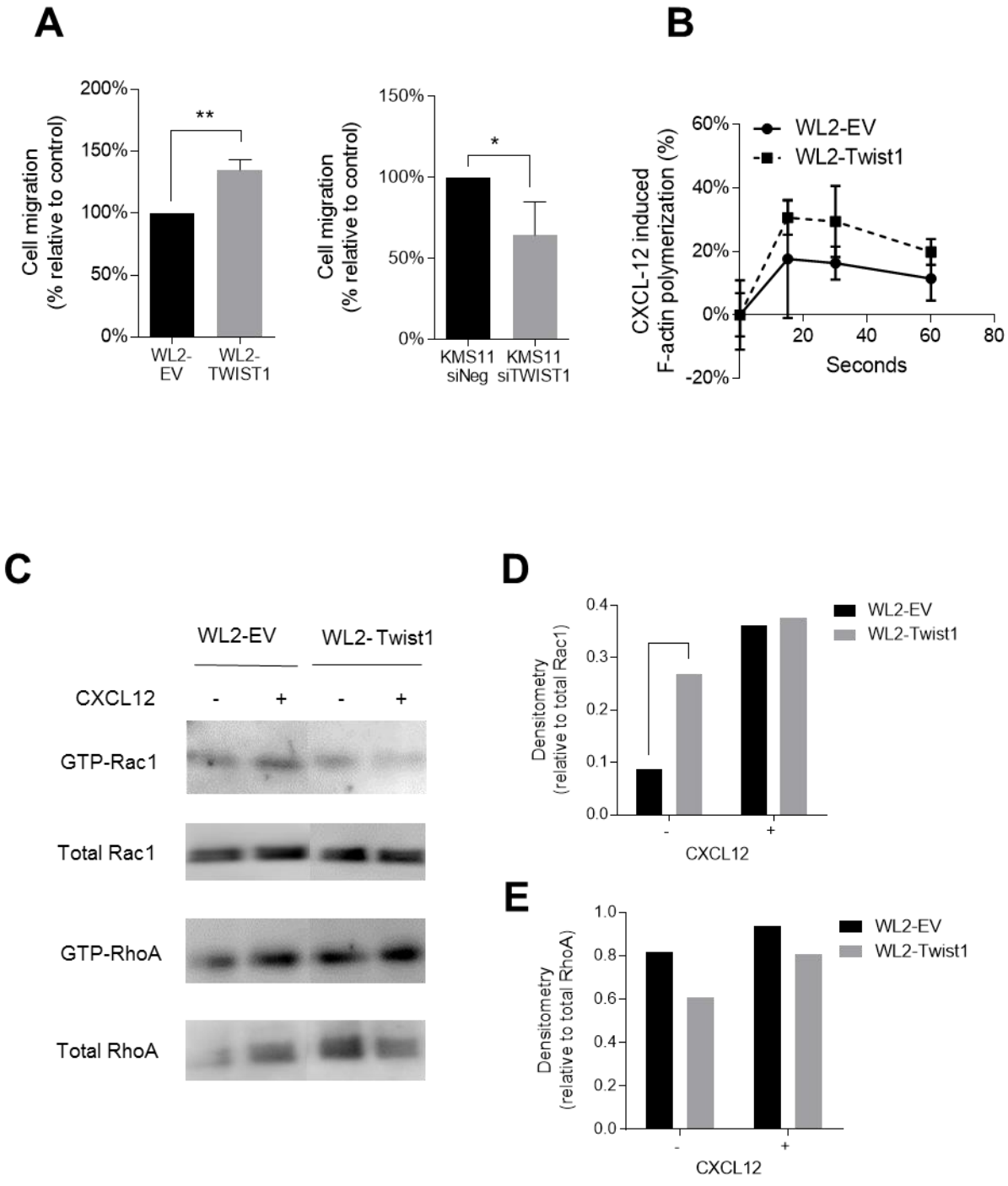

A
WL2-EV
WL2-Twist1
Adhesion to BMSCs (cells/ field of view)
4000
3000
2000
1000
0
10
30
60
Adhesion duration (min)
B
WL2-EV
WL2-Twist1
Adhesion to BMECs (cells/ field of view)
5000
4000
3000
2000
1000
0
10
30
60
Adhesion duration (min)
C
*
Adhesion to fibronectin %
50
40
30
20
10
0
WL2-EV
WL2-Twist1

3.4.8 TWIST1 expression increases MM cell motility

As described above (section 3.4.4), GO analysis of TWIST1-overexpression in HMCL identified changes in MM HMCL expression of genes related to cell migration. To confirm that these changes in gene expression led to changes in MM cell behavior *in vitro* trans-well migration assays were performed using FCS gradient as chemoattractant. As shown in Figure 3.6A, TWIST1 overexpression induced a 35% increase in WL2 migration to FCS, compared with the vector control cells ($p < 0.05$). In contrast, transient TWIST1 knock down in KMS11 cells, reduced cell migration by 36% ($p < 0.05$, two-tailed student's t-test). As cell motility is reliant on mechanical forces generated from remodeling of cytoskeletal proteins, the effect of TWIST1 expression on actin cytoskeleton polymerization was assessed. The degree of actin polymerization in response to the potent chemokine CXCL12 was determined using phalloidin labelling. As shown in Figure 3.6B, while statistical significance was not reached ($p = 0.211$, two-way ANOVA with Sidak's multiple comparison tests) TWIST1 overexpression in WL2 cells resulted in elevated actin polymerization compared with vector control cells throughout the course of the assay. CXCL12-induced actin polymerisation peaked at 15 seconds post-stimulation in both WL2-EV (17±15.8%) and WL2-TWIST1 cells (30±4.1%) and subsequently, at 60 seconds, reduced to 11±6.2% and 20±3.4% in WL2-EV and WL2-TWIST1 cells, respectively.

As detailed in Figure 3.3, RNA-seq analysis revealed that TWIST1-overexpression in WL2 cells led to an upregulation in the expression of GTPases. As Rac1 and RhoA activation has been implicated in modulating cell motility of MM PCs in response to CXCL12[56], we examined whether TWIST1 expression could affect activities of Rac1 and RhoA in MM PCs. As shown in Figure 3.6C, the levels of GTP-bound Rac1 and RhoA were elevated following CXCL12 stimulation in both WL2-EV and WL2-TWIST1 cells. While RhoA activity remained largely unchanged, irrespective of TWIST1 expression in WL2 cells (Figure 3.6E), the level of GTP-bound Rac1 was elevated in WL2-TWIST1 cells to levels comparable to WL2-EV cell exposed to CXCL12 (Figure 3.6D), suggesting that TWIST1 activates Rac1 in MM PCs leading to enhanced cell migration.

Figure 3.6 TWIST1 expression promotes MM cell motility. (A) The number of cells that migrated across 8 μm transwells to the lower chamber in response to 10% FCS was assessed. WL2 or KMS11 cells (1×10^5) at 48 hours post siRNA transfection were seeded onto transwells. Cells migrated into lower chamber after 18 hours was photographed using inverted microscope and quantitated using ImageJ software. Data is presented as percentage of migrated cells in relative to control. Graphs depict mean ± SEM of three independent experiments in quadruplicate. (B) *Twist* expression enhanced CXCL12-induced F-actin polymerisation in MM cells. CXCL12-stimulated (200 ng/mL) WL2 cells for indicated duration were fixed in 2% paraformaldehyde for 15 min, followed by permeabilisation in 0.2% saponin. F-actin polymerisation was quantitated by flow cytometer after staining the cells with phalloidin-A680 for 30 min on ice in the dark. Data is presented as percentage change in MFI (mean ± SD of a representative experiment in triplicate). (C) The expression of GTP-bound Rac1, total Rac1, GTP-bound RhoA and RhoA in response to CXCL12 stimulation for 1 min. Activated Rac1 and RhoA (GTP-bound) were pulled down with Rhothekin or PAK1 agarose beads, respectively, and resolved in 12% SDS-PAGE under reducing condition. Rac1 and RhoA proteins were detected by Western immunoblot. The levels of GTP-bound Rac1 (D) and GTP-bound RhoA (E) in WL2-EV ad WL2-TWIST1 cells were quantified by densitometry relative to total Rac1 and RhoA, respectively.


A
Cell migration (% relative to control)
200%
150%
100%
50%
0%
**
WL2-EV
WL2-TWIST1
Cell migration (% relative to control)
150%
100%
50%
0%
*
KMS11 siNeg
KMS11 siTWIST1
B
CXCL-12 induced F-actin polymerization (%)
60%
40%
20%
0%
-20%
WL2-EV
WL2-Twist1
20
40
60
80
Seconds
C
WL2-EV
WL2-Twist1
CXCL12
-
+
-
+
GTP-Rac1
Total Rac1
GTP-RhoA
Total RhoA
D
Densitometry (relative to total Rac1)
0.4
0.3
0.2
0.1
0.0
WL2-EV
WL2-Twist1
-
+
CXCL12
E
Densitometry (relative to total RhoA)
1.0
0.8
0.6
0.4
0.2
0.0
WL2-EV
WL2-Twist1
-
+
CXCL12

3.5 Discussion

The pathogenesis and progression of t(4;14) MM is driven by the overexpression of the histone methyltransferase, MMSET, which leads to altered gene expression. We, and others[47, 48, 57] have previously shown that TWIST1 expression is activated by MMSET in both prostate cancer and MM. In an effort to elucidate the transcriptome-wide effects of TWIST1 on MM cell, independent of aberrant MMSET expression, we overexpressed TWIST1 in the MMSET$^{low/negative}$ MM cell line, WL2, and performed RNA-sequencing analysis.

TWIST1 has previously been shown to play a role as an oncogenic transcription factor that promotes invasion and metastasis.[29, 58, 59] Furthermore, upregulation of TWIST1 has been associated with development of more aggressive epithelial cancers.[59-61] In hematological cancers, increased TWIST1 expression has previously been implicated in promoting motility of AML cells[62], which leads to aggressive phenotype. In this study, we demonstrated that ectopic expression of TWIST1 in MM cells, resulted in the upregulation of genes associated with regulation of cell motility. Importantly, *in vitro* transwell migration assays towards FCS also confirmed that TWIST1 promotes MM cell migration. Furthermore, GEP of MM patients further validated the effect of TWIST1 on cell motility genes in MM. Among these genes, *PREX1*[63], *MARCKS*[64], and *MICAL3*[65] have been shown to regulate cytoskeleton organisation, a fundamental mechanism that governs cell polarity and facilitates directional cell movement. Our *in vitro* migration assay also confirmed that TWIST1 promotes MM cell migration and elevation of actin polymerization, consistent with the GO analysis of the RNA sequencing data. Whether TWIST1 regulates the expression of these genes directly or indirectly through cooperation with other molecules to modulate cytoskeleton dynamic in MM cells remains to be determined and requires further investigation.

TWIST1 promotes carcinoma invasion and metastasis via activation of epithelial-to-mesenchymal transition (EMT), in which carcinoma cells lose their epithelial polarity and become more migratory. This is often characterised by the loss of E-cadherin and gain of N-cadherin, known as the “cadherin switch”. An EMT-like phenotype in MM plasma cells was recently reported, in which E-cadherin was downregulated in response to hypoxia, IL-17[66] and heparanase.[67] However, it should be noted that RNA-seq data from 66 HMCL (as described in section 2.4.2, Chapter 2)

shows that E-cadherin expression varies across MM cell lines (supplementary Figure S3.6A) owing to genetic heterogeneity among MM patients from which these cell lines were derived. Notably, our studies did not identify changes in E- and N-cadherin expression when TWIST1 was overexpressed in WL2 cells (data not shown). Furthermore, TWIST1 knockdown in the KMS11 HMCL was found to decrease the expression of both E- and N- cadherin (supplementary Figure S3.6B). These findings suggest that TWIST1 is capable of promoting MM cells migration independent of the cadherin switch, in agreement with previous findings in glioma and mammary epithelial cells.[68, 69]

Cell migration during metastasis also requires the control of adhesion to extracellular matrix. While strong adhesion immobilises tumour cells at their primary site, minimal adhesion to extracellular matrices enables tumour cells to migrate along extracellular substrates such as fibronectin.[70] In MM, interactions with the cellular components or extracellular matrix within the bone microenvironment is critical for MM PC survival, growth and drug resistance during disease progression.[71, 72] Previous studies have shown that adhesion of MM PCs to fibronectin induces cell-cycle arrest and inhibits apoptosis.[73, 74] The studies presented in this chapter show that TWIST1 expression did not affect the capacity of MM cells to adhere to either BM stromal cells or BM endothelial cells but did increase MM cell adhesion to fibronectin. These findings suggest that TWIST1 did not affect MM PCs retention within the BM microenvironment but rather promoted their interaction with fibronectin, thereby enhancing their migratory capacity.[55] TWIST1-enhanced adhesion to fibronectin has also been observed in glioma.[68] Although, none of the fibronectin receptors such as integrin α4β1, α4β7 and α5β1[75, 76] were differentially expressed in our transcriptome data, upregulation of potent integrin activators such as talin, kindlin[77, 78] and Ras GTPases[79] can trigger integrin conformational rearrangements into extended forms, thereby increasing their binding affinity to extracellular ligands including fibronectin. As seen in the RNA-seq analysis (supplementary Table S3.1), guanine-nucleotide exchange factor (GEF), that activates Ras by exchanging bound GDP for GTP, *RAPGEF2* and *RAPGEF3*, were upregulated in WL2-TWIST1 cells, suggesting TWIST1-enhanced adhesiveness to fibronectin may be regulated by Ras GTPase-mediated activation of integrins in MM PCs.

Previous overexpression and knockdown studies in solid tumours have demonstrated that TWIST1 mediates proliferation of tumour cells downstream of

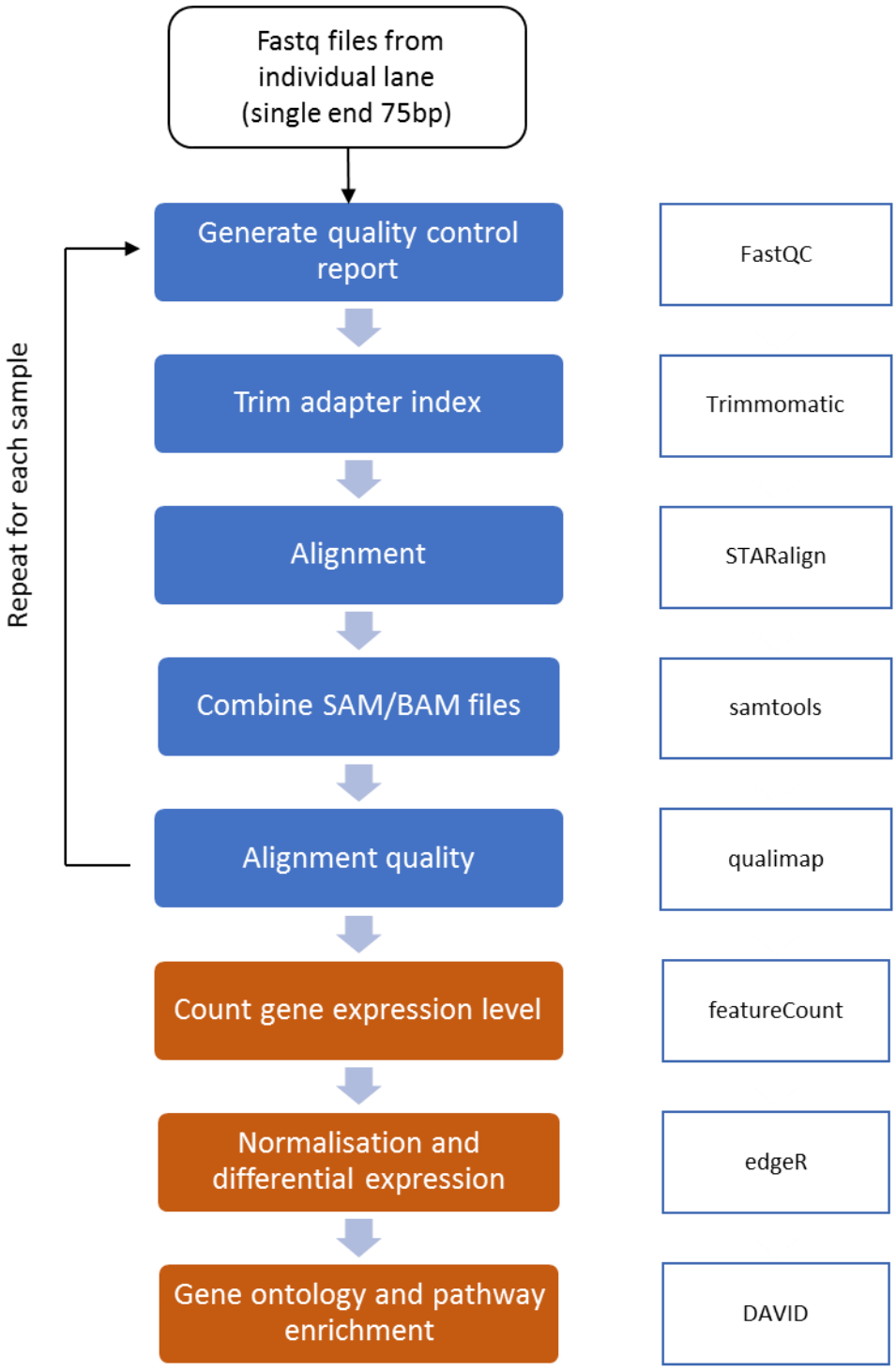
SIRT6 and p63.[80, 81] In the present study, TWIST1 overexpression in MMSET-low or TWIST1- knockdown in MMSET-high HMCL did not affect cell proliferation *in vitro,* suggesting that TWIST1 does not mediate the enhanced proliferation observed when MMSET is overexpressed in MM cells.[82] TWIST1 has also been reported to mediate angiogenesis in various cancers,[61, 83] including in MM.[84] Using immunohistochemistry, Yang et al. recently showed that patients with high TWIST1-expressing MM cells display elevated microvessel density.[84] In accord with this finding, our GO analysis of HMCL overexpressing TWIST1 showed enrichment of genes involved in angiogenesis including histone deacetlyase 7 (HDAC7), hypoxia-inducible factor 1α (HIF1A), laminin α5 (LAMA5) and perlecan/heparan sulfate proteoglycan 2 (HSPG2). Based on these findings, we postulate that TWIST1-induced HDAC7 may stabilise HIF-1α, an essential transcription factors in driving tumorigenic angiogenesis.[85, 86] Furthermore, extracellular secretion of laminin α5[87] and perlecan[88] by MM PC, may modify nearby basal membranes to activate angiogenesis. Clearly, additional studies are required to determine the precise mechanisms by which TWIST1 overexpression in cancer cells leads to enhanced angiogenesis.

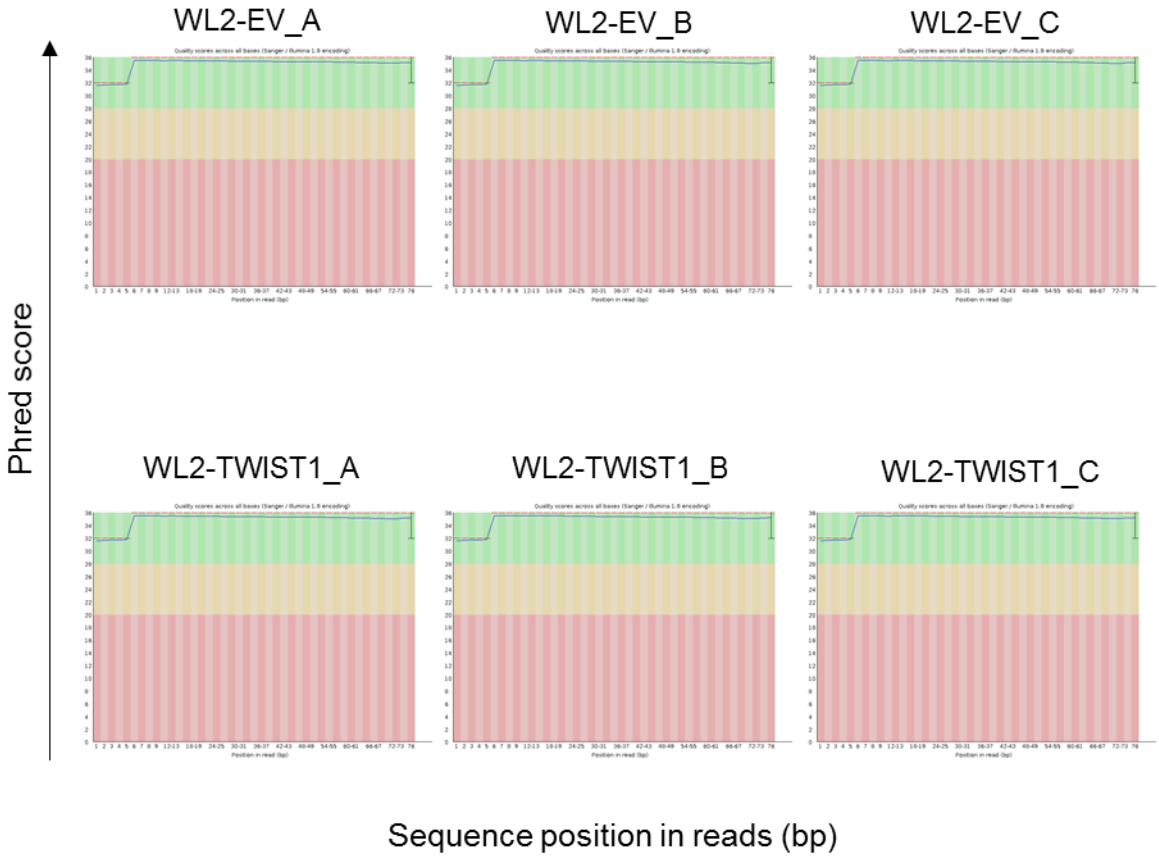
The activation of small GTPases have been implicated in wide range of biological activities including cell growth, cytoskeletal organization, cell migration and adhesion.[89] Consistent with enrichment of GTPases activity from GO analysis, we showed that the activity of Rac1 was increased in MM in which TWIST1 was overexpressed. These findings suggest that like epithelial cancers[90], TWIST1 activates Rac1-mediated F-actin polymerization leading to enhanced cell motility. In MM, activation RhoA and Rac1 have been shown to mediate CXCL12-induced chemotaxis, adhesion and drug resistance.[56, 91] Further studies using pharmacological or genetic inhibition Rac1 and RhoA are required to identify potential strategies to overcome TWIST1-induced cell motility in MM PCs.

In summary, transcriptomic analysis of the effects of TWIST1 overexpression in MM HMCL identified global gene changes and provides a framework to delineate mechanisms regulated by TWIST1 in MM pathogenesis. The *in vitro* data presented here suggest that TWIST1 contributes to pathogenesis of MM by promoting cell motility via regulation of actin remodeling and cell adhesion to fibronectin. Further studies are required to determine the effect of TWIST1 *in vivo* in order to determine therapeutic and prognostic implications of TWIST1.

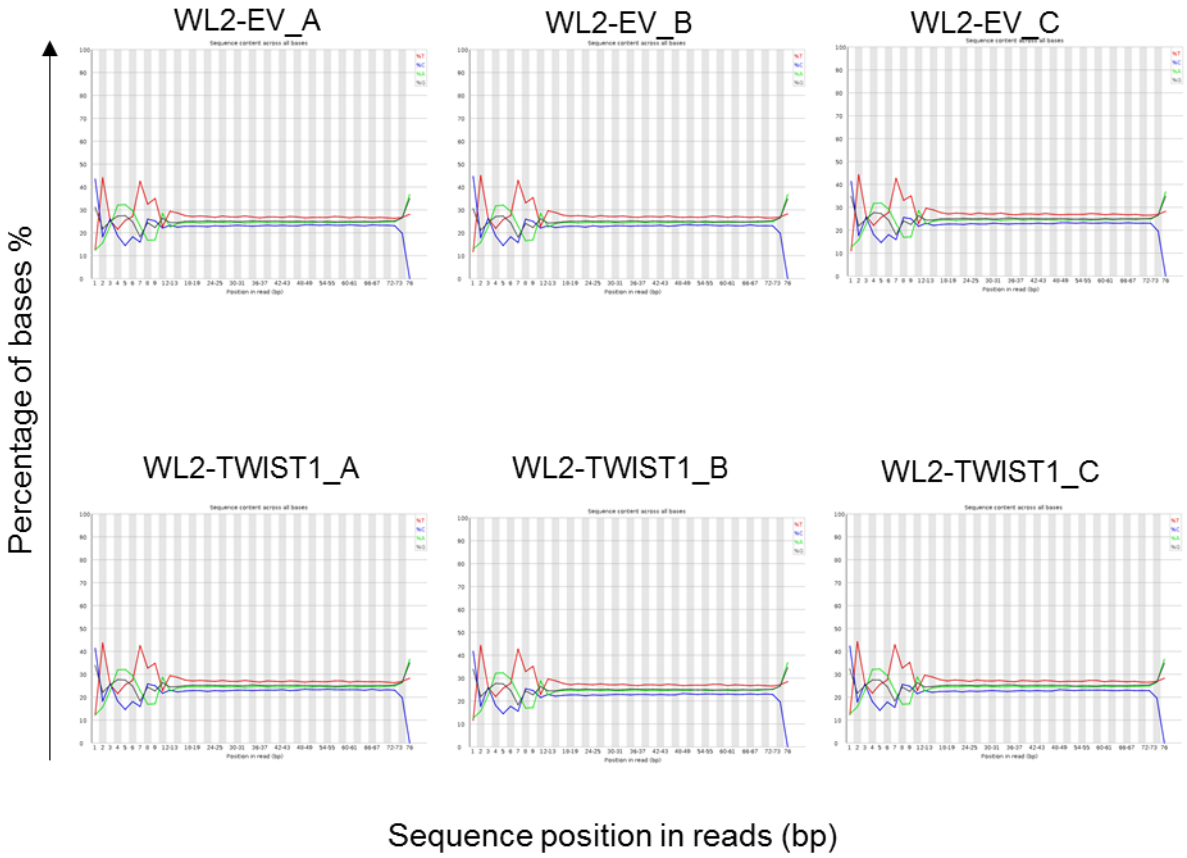
3.6 Supplementary Figures and Tables

supplementary Figure S3.1 RNA sequencing analysis workflow. Quality of raw RNA-Seq fastq files is performed using FastQC. Overrepresented adapter sequences were trimmed from the raw fastq files and repeated QC. Filtered reads are mapped to the reference genome hg19 using STARalign. BAM files from individual lanes of the same sample were combined. Gene expression level and differential expression were determined using featureCount and edgeR respectively. Unfilled boxes on the right represent the packages used to perform each task.

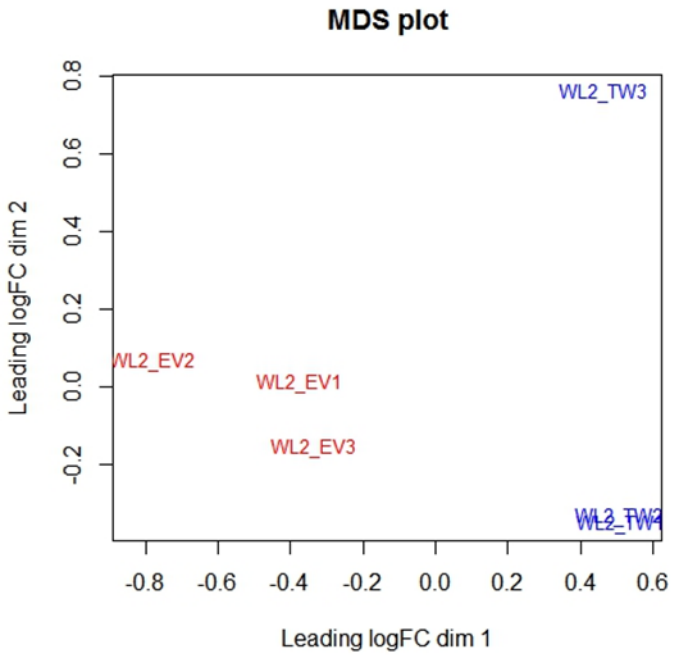

Fastq files from individual lane (single end 75bp)
Generate quality control report
FastQC
Trim adapter index
Trimmomatic
Alignment
STARalign
Combine SAM/BAM files
samtools
Alignment quality
qualimap
Count gene expression level
featureCount
Normalisation and differential expression
edgeR
Gene ontology and pathway enrichment
DAVID
Repeat for each sample



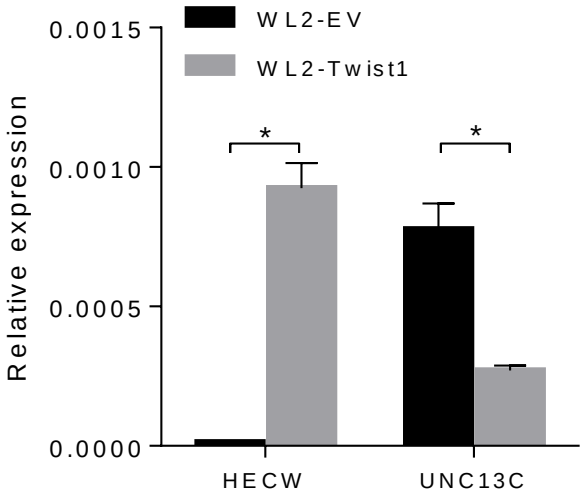
supplementary Figure S3.2 Quality assessment of RNA-Seq reads. The quality of reads was assessed using fastQC. Representative plot from first lane of each sample were shown. Phred score are plotted on the y-axis. Position of reads are shown on the x-axis. Blue line represents the mean quality while red line represents the media value. All sequence reads have Phred score more than 30.



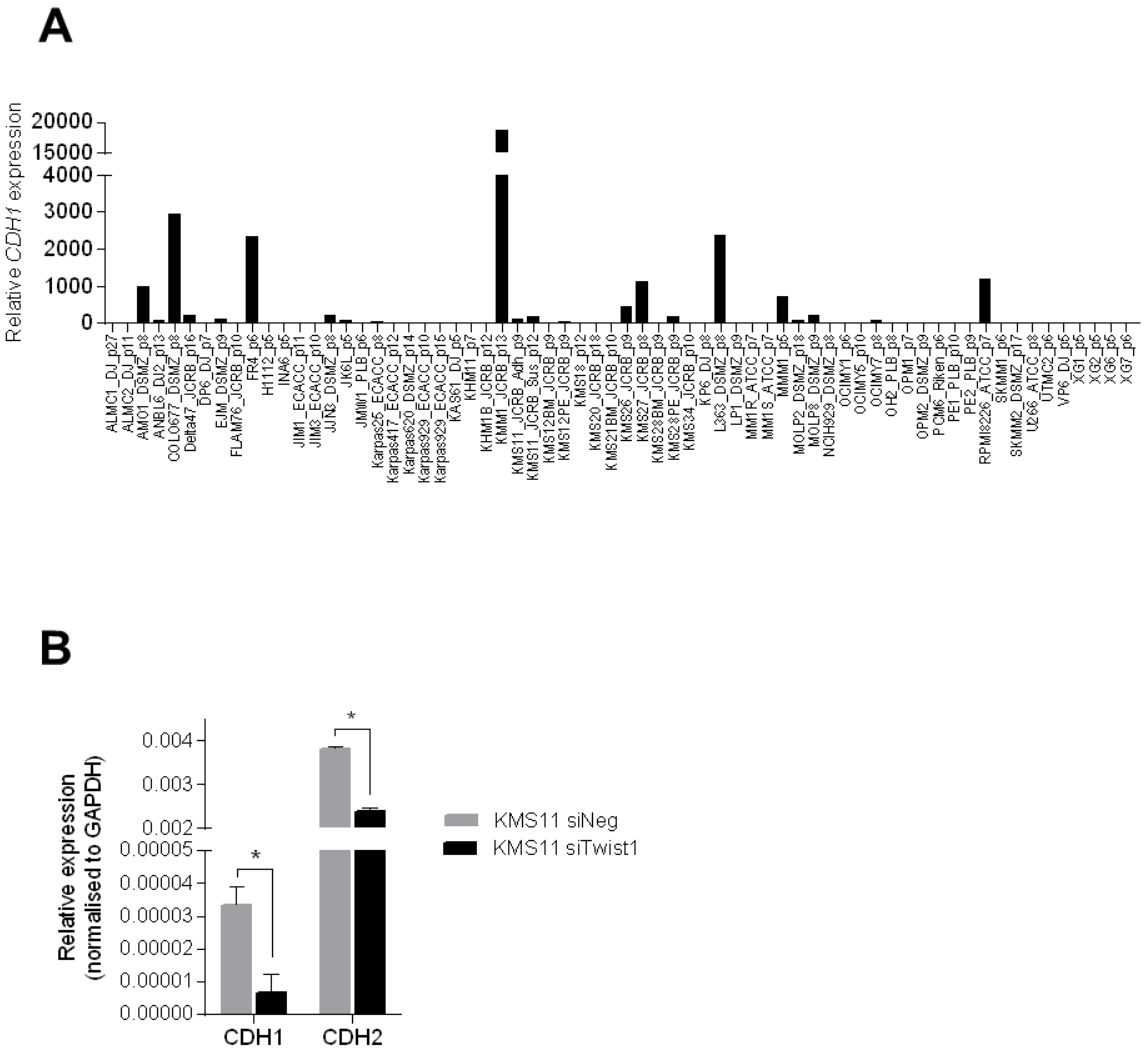
supplementary Figure S3.3 Per base sequence content by FastQC of the raw reads. Percentage of each nucleotide along the read position is indicated by colour (T = red; C = blue; C= green; G = black). Nucleotide composition along the reads position. Representative plot from first lane of each sample were shown. It is common in RNA-seq data to have changing composition at the beginning of the reads (1-12 bp) and becomes evenly distributed further down the reads.



supplementary Figure S3.4 Multidimensional scaling plot to evaluate libraries similarities. Graph was plotted using edgeR package in R to visualised similarity of WL2-EV and WL2-TWIST1 replicate samples. The root-mean-square of top 500 genes with largest absolute log2 fold change between two sample libraries (leading logFC) was calculated. Strong separation is observed amongst libraries from two groups in the first dimension.



supplementary Figure S3.5 Validation of RNA-seq data by qPCR. Relative mRNA levels of representative DEG with greatest fold change (HECW and UNC13C) were upregulated and downregulated respectively in WL2-TWIST1 cells compared with vector control cells. Data were normalised to reference gene GAPDH (mean ± SD of triplicates).



supplementary Figure S3.6 . Modulation of TWIST1 expression does not generate cadherin switch in KMS11 cells. (A) *CDH1* mRNA expression encode for E-cadherin is shown for a panel of human MM cell lines assessed by RNA-Seq. (B) *TWIST1* was knockdown in KMS11 cells by transfecting the cells with *TWIST1* siRNAs (siTwist) or negative control siRNA (siNeg). *CDH1* and *CDH2* mRNA level was assessed by qPCR 48 hours post transfection and normalised to reference gene *GAPDH*. Graph depict mean ± SD of triplicates from a representative experiment (* $p < 0.05$ unpaired student's t-test).

Table S3.1. List of genes differentially regulated in TWIST1-overexpressing WL2 cells.

GeneID	Symbol	chromosome	type_of_gene	log2 fold change	PValue	FDR
7291	TWIST1	7	protein-coding	7.233	2.01E-20	2.46E-16
57520	HECW2	2	protein-coding	4.459	3.79E-07	1.52E-04
59352	LGR6	1	protein-coding	4.347	3.81E-06	9.54E-04
115361	GBP4	1	protein-coding	4.181	1.04E-05	1.95E-03
6382	SDC1	2	protein-coding	3.335	2.07E-14	1.27E-10
2162	F13A1	6	protein-coding	3.253	1.66E-09	2.90E-06
595135	PGM5P2	9	pseudo	2.758	8.90E-04	4.89E-02
5617	PRL	6	protein-coding	2.591	6.28E-06	1.52E-03
309	ANXA6	5	protein-coding	2.421	3.02E-04	2.25E-02
3339	HSPG2	1	protein-coding	2.400	2.29E-05	3.51E-03
10411	RAPGEF3	12	protein-coding	2.185	2.39E-05	3.61E-03
55194	EVA1B	1	protein-coding	2.172	5.98E-05	6.85E-03
440823	MIAT	22	ncRNA	2.171	2.03E-07	1.03E-04
8938	BAIAP3	16	protein-coding	1.750	9.24E-05	9.44E-03
976	ADGRE5	19	protein-coding	1.667	2.15E-08	2.20E-05
93010	B3GNT7	2	protein-coding	1.661	2.34E-07	1.10E-04
87	ACTN1	14	protein-coding	1.607	3.82E-05	4.97E-03
153090	DAB2IP	9	protein-coding	1.607	2.43E-08	2.30E-05
27343	POLL	10	protein-coding	1.575	5.38E-04	3.51E-02
89797	NAV2	11	protein-coding	1.511	1.75E-05	2.90E-03
7226	TRPM2	21	protein-coding	1.443	1.97E-05	3.14E-03
6776	STAT5A	17	protein-coding	1.414	1.76E-04	1.54E-02
28793	IGLV3-25	22	other	1.391	7.41E-05	7.89E-03
126917	IFFO2	1	protein-coding	1.303	2.22E-04	1.81E-02
9053	MAP7	6	protein-coding	1.292	9.19E-04	4.98E-02
8835	SOCS2	12	protein-coding	1.180	5.57E-05	6.50E-03
114822	RHPN1	8	protein-coding	1.171	7.84E-04	4.53E-02
6482	ST3GAL1	8	protein-coding	1.167	1.84E-07	9.79E-05
5142	PDE4B	1	protein-coding	1.156	6.45E-06	1.52E-03
9750	FAM65B	6	protein-coding	1.156	2.67E-06	7.57E-04
6693	SPN	16	protein-coding	1.124	3.66E-07	1.52E-04
57580	PREX1	20	protein-coding	1.119	1.27E-11	3.12E-08
79754	ASB13	10	protein-coding	1.103	2.65E-06	7.57E-04
116988	AGAP3	7	protein-coding	1.055	4.43E-05	5.43E-03
7504	XK	X	protein-coding	1.031	2.57E-04	2.03E-02
23129	PLXND1	3	protein-coding	1.021	5.28E-07	2.02E-04
23348	DOCK9	13	protein-coding	0.984	9.65E-06	1.88E-03
23089	PEG10	7	protein-coding	0.924	9.37E-11	1.91E-07
27242	TNFRSF21	6	protein-coding	0.919	1.27E-06	4.22E-04
201294	UNC13D	17	protein-coding	0.912	5.45E-04	3.53E-02
50515	CHST11	12	protein-coding	0.887	1.65E-05	2.77E-03
9693	RAPGEF2	4	protein-coding	0.857	4.05E-08	3.31E-05
57685	CACHD1	1	protein-coding	0.855	1.26E-04	1.18E-02
58527	ABRACL	6	protein-coding	0.849	6.07E-05	6.88E-03
100132074	FOXO6	1	protein-coding	0.848	1.01E-04	1.01E-02
55841	WWC3	X	protein-coding	0.844	6.91E-09	8.47E-06
4681	NBL1	1	protein-coding	0.826	1.73E-07	9.79E-05
6821	SUOX	12	protein-coding	0.805	7.47E-04	4.40E-02
116039	OSR2	8	protein-coding	0.798	1.65E-05	2.77E-03
5790	PTPRCAP	11	protein-coding	0.791	2.59E-04	2.04E-02
56894	AGPAT3	21	protein-coding	0.787	7.54E-04	4.42E-02
79930	DOK3	5	protein-coding	0.786	6.85E-09	8.47E-06
23312	DMXL2	15	protein-coding	0.759	8.45E-04	4.72E-02
57535	KIAA1324	1	protein-coding	0.755	2.22E-07	1.09E-04
9783	RIMS3	1	protein-coding	0.754	2.66E-04	2.06E-02
64764	CREB3L2	7	protein-coding	0.749	3.04E-07	1.38E-04
57553	MICAL3	22	protein-coding	0.737	3.85E-07	1.52E-04
57670	KIAA1549	7	protein-coding	0.735	7.55E-06	1.71E-03
1280	COL2A1	12	protein-coding	0.731	8.29E-04	4.70E-02
57565	KLHL14	18	protein-coding	0.725	4.00E-05	5.01E-03
23576	DDAH1	1	protein-coding	0.721	6.30E-05	7.08E-03
10221	TRIB1	8	protein-coding	0.716	2.14E-04	1.76E-02
5768	QSOX1	1	protein-coding	0.712	1.02E-06	3.69E-04
23338	JADE2	5	protein-coding	0.704	2.84E-06	7.57E-04
646762	LOC646762	7	ncRNA	0.704	4.99E-05	6.06E-03
6319	SCD	10	protein-coding	0.694	4.04E-04	2.84E-02
1445	CSK	15	protein-coding	0.694	1.16E-04	1.12E-02
9435	CHST2	3	protein-coding	0.689	6.13E-08	4.70E-05
57569	ARHGAP20	11	protein-coding	0.685	3.66E-05	4.83E-03

GeneID	Symbol	chromosome	type_of_gene	log2 fold change	PValue	FDR
63971	KIF13A	6	protein-coding	0.679	2.49E-05	3.69E-03
10678	B3GNT2	2	protein-coding	0.671	1.19E-07	7.68E-05
5272	SERPINB9	6	protein-coding	0.661	3.52E-05	4.74E-03
3965	LGALS9	17	protein-coding	0.647	1.58E-08	1.76E-05
463	ZFHX3	16	protein-coding	0.621	9.36E-05	9.48E-03
23176	SEPT8	5	protein-coding	0.611	7.51E-06	1.71E-03
25960	ADGRA2	8	protein-coding	0.605	8.75E-05	9.02E-03
10743	RAI1	17	protein-coding	0.596	7.63E-04	4.45E-02
283131	NEAT1	11	ncRNA	0.593	6.12E-04	3.81E-02
138050	HGSNAT	8	protein-coding	0.592	5.29E-05	6.24E-03
3911	LAMA5	20	protein-coding	0.592	6.31E-06	1.52E-03
4208	MEF2C	5	protein-coding	0.591	6.37E-05	7.09E-03
84959	UBASH3B	11	protein-coding	0.586	8.90E-06	1.79E-03
55652	SLC48A1	12	protein-coding	0.579	5.26E-04	3.46E-02
126282	TNFAIP8L1	19	protein-coding	0.572	5.35E-04	3.51E-02
149086	LINC01225	1	pseudo	0.571	5.82E-04	3.64E-02
57060	PCBP4	3	protein-coding	0.570	4.61E-04	3.15E-02
347735	SERINC2	1	protein-coding	0.569	6.76E-05	7.40E-03
23308	ICOSLG	21	protein-coding	0.568	3.85E-05	4.97E-03
27075	TSPAN13	7	protein-coding	0.567	3.92E-05	5.00E-03
3097	HIVEP2	6	protein-coding	0.563	5.17E-05	6.16E-03
64131	XYLT1	16	protein-coding	0.561	1.39E-07	8.51E-05
2549	GAB1	4	protein-coding	0.560	8.52E-06	1.79E-03
51564	HDAC7	12	protein-coding	0.560	8.14E-06	1.78E-03
25865	PRKD2	19	protein-coding	0.556	3.19E-05	4.54E-03
205	AK4	1	protein-coding	0.554	1.28E-05	2.20E-03
2200	FBN1	15	protein-coding	0.546	8.51E-04	4.72E-02
284551	LINC01226	1	ncRNA	0.529	2.72E-06	7.57E-04
60673	ATG101	12	protein-coding	0.524	1.56E-04	1.40E-02
57460	PPM1H	12	protein-coding	0.519	3.59E-05	4.79E-03
1268	CNR1	6	protein-coding	0.516	1.60E-04	1.42E-02
9516	LITAF	16	protein-coding	0.497	7.03E-05	7.62E-03
81558	FAM117A	17	protein-coding	0.494	2.73E-04	2.09E-02
7402	UTRN	6	protein-coding	0.488	3.35E-05	4.66E-03
63035	BCORL1	X	protein-coding	0.486	3.99E-05	5.01E-03
10018	BCL2L11	2	protein-coding	0.480	3.42E-07	1.50E-04
55635	DEPDC1	1	protein-coding	0.480	2.68E-04	2.06E-02
91408	BTF3L4	1	protein-coding	0.479	8.37E-04	4.72E-02
1396	CRIP1	14	protein-coding	0.466	7.02E-04	4.22E-02
104	ADARB1	21	protein-coding	0.458	8.50E-06	1.79E-03
2313	FLI1	11	protein-coding	0.457	9.87E-07	3.66E-04
4131	MAP1B	5	protein-coding	0.457	2.20E-06	6.57E-04
80727	TTYH3	7	protein-coding	0.453	1.49E-04	1.35E-02
5364	PLXNB1	3	protein-coding	0.451	2.62E-04	2.04E-02
6790	AURKA	20	protein-coding	0.444	1.31E-04	1.21E-02
9415	FADS2	11	protein-coding	0.441	1.08E-04	1.06E-02
2114	ETS2	21	protein-coding	0.436	1.06E-06	3.70E-04
57584	ARHGAP21	10	protein-coding	0.430	4.21E-04	2.93E-02
89781	HPS4	22	protein-coding	0.429	2.50E-05	3.69E-03
5163	PDK1	2	protein-coding	0.429	5.56E-04	3.55E-02
84532	ACSS1	20	protein-coding	0.428	8.99E-06	1.79E-03
2632	GBE1	3	protein-coding	0.426	7.54E-05	7.96E-03
10226	PLIN3	19	protein-coding	0.423	1.36E-04	1.25E-02
6812	STXBP1	9	protein-coding	0.416	4.23E-04	2.93E-02
4853	NOTCH2	1	protein-coding	0.396	1.01E-05	1.93E-03
2495	FTH1	11	protein-coding	0.394	1.10E-05	1.99E-03
157285	SGK223	8	protein-coding	0.391	6.98E-04	4.22E-02
8522	GAS7	17	protein-coding	0.390	7.78E-06	1.73E-03
29760	BLNK	10	protein-coding	0.390	7.36E-05	7.89E-03
10957	PNRC1	6	protein-coding	0.389	1.98E-04	1.66E-02
3133	HLA-E	6	protein-coding	0.385	2.03E-04	1.68E-02
2781	GNAZ	22	protein-coding	0.380	6.62E-04	4.06E-02
59345	GNB4	3	protein-coding	0.379	2.35E-04	1.89E-02
58528	RRAGD	6	protein-coding	0.378	4.93E-04	3.33E-02
128239	IQGAP3	1	protein-coding	0.375	8.20E-04	4.67E-02
10491	CRTAP	3	protein-coding	0.374	5.62E-04	3.57E-02
55824	PAG1	8	protein-coding	0.369	8.05E-04	4.61E-02
677	ZFP36L1	14	protein-coding	0.368	4.19E-05	5.19E-03
57189	KIAA1147	7	protein-coding	0.368	5.56E-04	3.55E-02
149076	ZNF362	1	protein-coding	0.368	2.03E-04	1.68E-02

GeneID	Symbol	chromosome	type_of_gene	log2 fold change	PValue	FDR
28776	IGLV7-43	22	other	0.365	1.97E-05	3.14E-03
3995	FADS3	11	protein-coding	0.362	8.11E-05	8.49E-03
10801	SEPT9	17	protein-coding	0.357	1.15E-05	2.05E-03
5499	PPP1CA	11	protein-coding	0.354	5.07E-05	6.09E-03
10024	TROAP	12	protein-coding	0.345	5.57E-04	3.55E-02
5896	RAG1	11	protein-coding	0.344	1.78E-04	1.54E-02
27245	AHDC1	1	protein-coding	0.343	2.12E-05	3.29E-03
994	CDC25B	20	protein-coding	0.339	1.03E-04	1.01E-02
7805	LAPTM5	1	protein-coding	0.333	2.97E-05	4.34E-03
4082	MARCKS	6	protein-coding	0.325	6.22E-04	3.85E-02
5660	PSAP	10	protein-coding	0.318	3.61E-04	2.60E-02
22882	ZHX2	8	protein-coding	0.308	1.23E-04	1.18E-02
4033	LRMP	12	protein-coding	0.304	4.80E-04	3.27E-02
10659	CELF2	10	protein-coding	0.302	2.46E-04	1.96E-02
2037	EPB41L2	6	protein-coding	0.294	4.97E-04	3.33E-02
3107	HLA-C	6	protein-coding	0.293	7.91E-04	4.55E-02
3091	HIF1A	14	protein-coding	0.284	3.86E-04	2.75E-02
3833	KIFC1	6	protein-coding	0.283	3.45E-04	2.54E-02
7431	VIM	10	protein-coding	0.278	1.88E-04	1.59E-02
3106	HLA-B	6	protein-coding	0.267	7.12E-04	4.24E-02
71	ACTG1	17	protein-coding	0.254	5.71E-04	3.59E-02
4311	MME	3	protein-coding	-0.261	8.62E-04	4.76E-02
9734	HDAC9	7	protein-coding	-0.297	9.03E-04	4.93E-02
116236	ABHD15	17	protein-coding	-0.299	3.40E-04	2.52E-02
23228	PLCL2	3	protein-coding	-0.310	3.86E-04	2.75E-02
84916	UTP4	16	protein-coding	-0.317	9.06E-04	4.93E-02
54545	MTMR12	5	protein-coding	-0.332	6.68E-04	4.07E-02
4776	NFATC4	14	protein-coding	-0.343	5.07E-04	3.38E-02
23471	TRAM1	8	protein-coding	-0.367	1.66E-06	5.21E-04
5983	RFC3	13	protein-coding	-0.379	6.42E-05	7.09E-03
990	CDC6	17	protein-coding	-0.384	5.67E-04	3.58E-02
79623	GALNT14	2	protein-coding	-0.386	3.60E-04	2.60E-02
651	BMP3	4	protein-coding	-0.413	2.01E-05	3.16E-03
10381	TUBB3	16	protein-coding	-0.414	3.60E-06	9.31E-04
22806	IKZF3	17	protein-coding	-0.415	2.79E-04	2.12E-02
1439	CSF2RB	22	protein-coding	-0.415	3.59E-04	2.60E-02
8565	YARS	1	protein-coding	-0.425	1.30E-04	1.20E-02
199786	FAM129C	19	protein-coding	-0.439	2.40E-04	1.93E-02
7090	TLE3	15	protein-coding	-0.440	1.79E-04	1.55E-02
5579	PRKCB	16	protein-coding	-0.442	2.84E-06	7.57E-04
5101	PCDH9	13	protein-coding	-0.448	1.87E-05	3.06E-03
84679	SLC9A7	X	protein-coding	-0.448	3.47E-04	2.54E-02
5984	RFC4	3	protein-coding	-0.451	8.47E-04	4.72E-02
7975	MAFK	7	protein-coding	-0.457	5.13E-04	3.40E-02
28638	TRBC2	7	other	-0.458	3.12E-05	4.49E-03
22902	RUFY3	4	protein-coding	-0.465	2.82E-04	2.12E-02
57577	CCDC191	3	protein-coding	-0.477	1.87E-04	1.59E-02
89884	LHX4	1	protein-coding	-0.496	4.49E-04	3.09E-02
10184	LHFPL2	5	protein-coding	-0.506	3.46E-05	4.71E-03
833	CARS	11	protein-coding	-0.528	1.88E-04	1.59E-02
10129	FRY	13	protein-coding	-0.531	1.01E-04	1.01E-02
1002	CDH4	20	protein-coding	-0.536	2.16E-06	6.57E-04
753	LDLRAD4	18	protein-coding	-0.537	3.93E-04	2.79E-02
5026	P2RX5	17	protein-coding	-0.538	8.73E-05	9.02E-03
23504	RIMBP2	12	protein-coding	-0.547	6.80E-04	4.13E-02
129642	MBOAT2	2	protein-coding	-0.552	1.06E-05	1.95E-03
8936	WASF1	6	protein-coding	-0.574	9.04E-08	6.15E-05
8745	ADAM23	2	protein-coding	-0.579	2.80E-04	2.12E-02
285704	RGMB	5	protein-coding	-0.582	6.26E-04	3.86E-02
28822	IGLV1-47	22	other	-0.600	9.05E-06	1.79E-03
79888	LPCAT1	5	protein-coding	-0.601	3.40E-05	4.68E-03
22981	NINL	20	protein-coding	-0.608	2.66E-08	2.33E-05
171586	ABHD3	18	protein-coding	-0.621	1.72E-04	1.52E-02
113510	HELQ	4	protein-coding	-0.632	8.47E-04	4.72E-02
22822	PHLDA1	12	protein-coding	-0.663	1.28E-04	1.20E-02
8318	CDC45	22	protein-coding	-0.676	1.06E-05	1.95E-03
6474	SHOX2	3	protein-coding	-0.680	5.98E-05	6.85E-03
26084	ARHGEF26	3	protein-coding	-0.693	4.94E-04	3.33E-02
10347	ABCA7	19	protein-coding	-0.694	7.12E-04	4.24E-02
100129034	LOC100129034	9	ncRNA	-0.729	1.19E-06	4.04E-04

GeneID	Symbol	chromosome	type_of_gene	log2 fold change	PValue	FDR
81793	TLR10	4	protein-coding	-0.791	1.40E-04	1.27E-02
316	AOX1	2	protein-coding	-0.802	1.81E-07	9.79E-05
5818	NECTIN1	11	protein-coding	-0.866	1.24E-05	2.18E-03
24138	IFIT5	10	protein-coding	-0.872	1.11E-04	1.07E-02
286133	SCARA5	8	protein-coding	-0.908	1.35E-06	4.36E-04
8832	CD84	1	protein-coding	-0.910	7.22E-04	4.28E-02
23151	GRAMD4	22	protein-coding	-1.037	7.92E-08	5.71E-05
1536	CYBB	X	protein-coding	-1.097	1.91E-12	6.11E-09
151126	ZNF385B	2	protein-coding	-1.130	5.43E-09	8.32E-06
122618	PLD4	14	protein-coding	-1.165	4.12E-04	2.89E-02
9760	TOX	8	protein-coding	-1.172	3.31E-05	4.66E-03
28797	IGLV3-19	22	other	-1.219	2.32E-04	1.89E-02
28784	IGLV4-69	22	other	-1.391	3.65E-06	9.31E-04
222	ALDH3B2	11	protein-coding	-1.401	8.66E-06	1.79E-03
2334	AFF2	X	protein-coding	-1.501	1.26E-04	1.18E-02
3512	JCHAIN	4	protein-coding	-2.037	1.99E-12	6.11E-09
440279	UNC13C	15	protein-coding	-2.404	7.83E-04	4.53E-02

Table S3.2. Function annotation clustering (cellular components) of Twist1-upregulated genes in WL2 cells.

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
Annotation Cluster 1	Enrichment Score: 2.4928918815294936											
GOTERM_CC_FAT	GO:0031988~membrane-bounded vesicle	47	29.375	0.00109564	GNAZ, F13A1, HPS4, VIM, UTRN, KIAA1324, FTH1, ST3GAL1, ARHGAP21, ACTG1, ANXA6, SERINC2, TTYH3, JADE2, DMXL2, PDE4B, PLIN3, B3GNT2, RAPGEF3, CSK, RAPGEF2, QSOX1, DDAH1, SPN, DAB2IP, PSAP, ADGRE5, FBN1, STXBP1, HSPG2, ACTN1, HLA-C, HLA-B, AK4, HLA-E, TRPM2, LGALS9, EPB41L2, SERPINB9, PPP1CA, SDC1, UNC13D, GBE1, LAMA5, GNB4, MARCKS, ICOSLG	123	3608	14505	1.536186523	0.320145756	0.053633406	1.489373321
GOTERM_CC_FAT	GO:0070062~extracellular exosome	38	23.75	0.002248825	GNAZ, UTRN, VIM, KIAA1324, FTH1, ST3GAL1, ACTG1, ANXA6, SERINC2, TTYH3, JADE2, B3GNT2, RAPGEF3, CSK, QSOX1, DDAH1, SPN, DAB2IP, PSAP, ADGRE5, FBN1, STXBP1, HSPG2, ACTN1, HLA-C, HLA-B, AK4, HLA-E, LGALS9, EPB41L2, SERPINB9, SDC1, PPP1CA, GBE1, LAMA5, GNB4, MARCKS, ICOSLG	123	2799	14505	1.601007328	0.547278297	0.094311373	3.034740345
GOTERM_CC_FAT	GO:1903561~extracellular vesicle	38	23.75	0.002460425	GNAZ, UTRN, VIM, KIAA1324, FTH1, ST3GAL1, ACTG1, ANXA6, SERINC2, TTYH3, JADE2, B3GNT2, RAPGEF3, CSK, QSOX1, DDAH1, SPN, DAB2IP, PSAP, ADGRE5, FBN1, STXBP1, HSPG2, ACTN1, HLA-C, HLA-B, AK4, HLA-E, LGALS9, EPB41L2, SERPINB9, SDC1, PPP1CA, GBE1, LAMA5, GNB4, MARCKS, ICOSLG	123	2813	14505	1.593039286	0.579847198	0.091852561	3.315851485
GOTERM_CC_FAT	GO:0043230~extracellular organelle	38	23.75	0.002476184	GNAZ, UTRN, VIM, KIAA1324, FTH1, ST3GAL1, ACTG1, ANXA6, SERINC2, TTYH3, JADE2, B3GNT2, RAPGEF3, CSK, QSOX1, DDAH1, SPN, DAB2IP, PSAP, ADGRE5, FBN1, STXBP1, HSPG2, ACTN1, HLA-C, HLA-B, AK4, HLA-E, LGALS9, EPB41L2, SERPINB9, SDC1, PPP1CA, GBE1, LAMA5, GNB4, MARCKS, ICOSLG	123	2814	14505	1.592473174	0.582177193	0.083570176	3.336757548
GOTERM_CC_FAT	GO:0044421~extracellular region part	44	27.5	0.022862506	GNAZ, F13A1, VIM, UTRN, KIAA1324, COL2A1, FTH1, ST3GAL1, ACTG1, ANXA6, SERINC2, TTYH3, JADE2, DMXL2, B3GNT2, RAPGEF3, CSK, QSOX1, DDAH1, SPN, DAB2IP, PSAP, ADGRE5, CRTAP, FBN1, STXBP1, HSPG2, ACTN1, HLA-C, HLA-B, AK4, HLA-E, LGALS9, EPB41L2, SERPINB9, NBL1, PPP1CA, SDC1, GBE1, LAMA5, NAV2, GNB4, MARCKS, ICOSLG	123	3873	14505	1.3397316	0.999708661	0.287666649	27.13657941
Annotation Cluster 2	Enrichment Score: 1.7153826261241822											
GOTERM_CC_FAT	GO:0005925~focal adhesion	9	5.625	0.018415709	EPB41L2, ANXA6, ACTG1, SDC1, ADGRE5, VIM, HSPG2, ACTN1, MARCKS	123	395	14505	2.686940414	0.998559494	0.267696365	22.46425677
GOTERM_CC_FAT	GO:0005924~cell-substrate adherens junction	9	5.625	0.019176394	EPB41L2, ANXA6, ACTG1, SDC1, ADGRE5, VIM, HSPG2, ACTN1, MARCKS	123	398	14505	2.666687094	0.998903521	0.266408053	23.28271596
GOTERM_CC_FAT	GO:0030055~cell-substrate junction	9	5.625	0.020225382	EPB41L2, ANXA6, ACTG1, SDC1, ADGRE5, VIM, HSPG2, ACTN1, MARCKS	123	402	14505	2.640152894	0.999247655	0.268537543	24.39824352
Annotation Cluster 3	Enrichment Score: 0.958154515836375											
GOTERM_CC_FAT	GO:0031091~platelet alpha granule	4	2.5	0.025284848	F13A1, STXBP1, ACTN1, QSOX1	123	75	14505	6.289430894	0.999878394	0.283859252	29.57059465
GOTERM_CC_FAT	GO:0031093~platelet alpha granule lumen	3	1.875	0.078113997	F13A1, ACTN1, QSOX1	123	55	14505	6.432372506	1	0.494220209	67.15354083
GOTERM_CC_FAT	GO:0034774~secretory granule lumen	3	1.875	0.163242575	F13A1, ACTN1, QSOX1	123	86	14505	4.113726602	1	0.6079342	91.28005065
GOTERM_CC_FAT	GO:0060205~cytoplasmic membrane-bounded vesicle lumen	3	1.875	0.221003935	F13A1, ACTN1, QSOX1	123	105	14505	3.369337979	1	0.666761507	96.72437611
GOTERM_CC_FAT	GO:0031983~vesicle lumen	3	1.875	0.227198137	F13A1, ACTN1, QSOX1	123	107	14505	3.306359699	1	0.66924067	97.06346612
Annotation Cluster 4	Enrichment Score: 0.9255153446301186											
GOTERM_CC_FAT	GO:0042612~MHC class I protein complex	3	1.875	0.005151041	HLA-C, HLA-B, HLA-E	123	13	14505	27.21388368	0.837625898	0.140571542	6.825080361
GOTERM_CC_FAT	GO:0098553~lumenal side of endoplasmic reticulum membrane	3	1.875	0.02456712	HLA-C, HLA-B, HLA-E	123	29	14505	12.19932717	0.999842427	0.295470369	28.85738314
GOTERM_CC_FAT	GO:0071556~integral component of lumenal side of endoplasmic reticulum membra	3	1.875	0.02456712	HLA-C, HLA-B, HLA-E	123	29	14505	12.19932717	0.999842427	0.295470369	28.85738314
GOTERM_CC_FAT	GO:0042611~MHC protein complex	3	1.875	0.027833211	HLA-C, HLA-B, HLA-E	123	31	14505	11.4122738	0.99995161	0.290097931	32.04972629
GOTERM_CC_FAT	GO:0012507~ER to Golgi transport vesicle membrane	3	1.875	0.08057602	HLA-C, HLA-B, HLA-E	123	56	14505	6.317508711	1	0.489346149	68.3341643
GOTERM_CC_FAT	GO:0030670~phagocytic vesicle membrane	3	1.875	0.090657912	HLA-C, HLA-B, HLA-E	123	60	14505	5.896341463	1	0.474448214	72.77024701
GOTERM_CC_FAT	GO:0030134~ER to Golgi transport vesicle	3	1.875	0.139856546	HLA-C, HLA-B, HLA-E	123	78	14505	4.53564728	1	0.574860461	87.28308426
GOTERM_CC_FAT	GO:0030658~transport vesicle membrane	4	2.5	0.148057021	DMXL2, HLA-C, HLA-B, HLA-E	123	158	14505	2.985489349	1	0.580099641	88.8459337
GOTERM_CC_FAT	GO:0045335~phagocytic vesicle	3	1.875	0.172186458	HLA-C, HLA-B, HLA-E	123	89	14505	3.975061661	1	0.613350147	92.4728658
GOTERM_CC_FAT	GO:0031901~early endosome membrane	3	1.875	0.261436644	HLA-C, HLA-B, HLA-E	123	118	14505	2.998139727	1	0.698375801	98.42078136
GOTERM_CC_FAT	GO:0030662~coated vesicle membrane	3	1.875	0.332880756	HLA-C, HLA-B, HLA-E	123	141	14505	2.509081474	1	0.769825108	99.60768838
GOTERM_CC_FAT	GO:0030135~coated vesicle	4	2.5	0.337618541	UNC13D, HLA-C, HLA-B, HLA-E	123	244	14505	1.933226709	1	0.772254532	99.64415446
GOTERM_CC_FAT	GO:0030176~integral component of endoplasmic reticulum membrane	3	1.875	0.348209954	HLA-C, HLA-B, HLA-E	123	146	14505	2.423154026	1	0.781702036	99.71461126
GOTERM_CC_FAT	GO:0031227~intrinsic component of endoplasmic reticulum membrane	3	1.875	0.36342325	HLA-C, HLA-B, HLA-E	123	151	14505	2.342917138	1	0.789577168	99.79344494
GOTERM_CC_FAT	GO:0030666~endocytic vesicle membrane	3	1.875	0.375499242	HLA-C, HLA-B, HLA-E	123	155	14505	2.28245476	1	0.799903779	99.84107997
GOTERM_CC_FAT	GO:0005769~early endosome	4	2.5	0.494220753	KIFC1, HLA-C, HLA-B, HLA-E	123	314	14505	1.502252602	1	0.878124984	99.99113442
Annotation Cluster 5	Enrichment Score: 0.9062183808776271											
GOTERM_CC_FAT	GO:0044448~cell cortex part	4	2.5	0.09666522	EPB41L2, UTRN, RAPGEF3, SEPT9	123	130	14505	3.628517824	1	0.48453505	75.13199841
GOTERM_CC_FAT	GO:0030864~cortical actin cytoskeleton	3	1.875	0.109117228	EPB41L2, UTRN, RAPGEF3	123	67	14505	5.280305788	1	0.510085884	79.43517491
GOTERM_CC_FAT	GO:0030863~cortical cytoskeleton	3	1.875	0.181209637	EPB41L2, UTRN, RAPGEF3	123	92	14505	3.845440085	1	0.623720041	93.52148194
Annotation Cluster 6	Enrichment Score: 0.7151953737674561											
GOTERM_CC_FAT	GO:0030175~filopodium	4	2.5	0.047190039	FAM65B, UTRN, RAPGEF3, ADGRA2	123	96	14505	4.913617886	0.999999959	0.368641482	48.40231594
GOTERM_CC_FAT	GO:0098858~actin-based cell projection	4	2.5	0.208446661	FAM65B, UTRN, RAPGEF3, ADGRA2	123	187	14505	2.522499022	1	0.65177552	95.92283098
GOTERM_CC_FAT	GO:0031253~cell projection membrane	3	1.875	0.727057233	FAM65B, UTRN, RAPGEF3	123	303	14505	1.167592369	1	0.97341841	99.99999809
Annotation Cluster 7	Enrichment Score: 0.5887978265495467											

Table S3.2. Function annotation clustering (cellular components) of Twist1-upregulated genes in WL2 cells.

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0005802~trans-Golgi network	4	2.5	0.232660422	KIF13A, CHST2, KIAA1324, LGR6	123	198	14505	2.382360187	1	0.674735577	97.33518458
GOTERM_CC_FAT	GO:0098791~Golgi subcompartment	5	3.125	0.255038116	ST3GAL1, KIF13A, CHST2, KIAA1324, LGR6	123	305	14505	1.933226709	1	0.696151346	98.22285455
GOTERM_CC_FAT	GO:0031984~organelle subcompartment	5	3.125	0.288588431	ST3GAL1, KIF13A, CHST2, KIAA1324, LGR6	123	323	14505	1.825492713	1	0.732092896	99.0542513
Annotation Cluster 8	Enrichment Score: 0.5576211282970794											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0044432~endoplasmic reticulum part	14	8.75	0.19451712	SCD, CRTAP, FADS3, FADS2, HLA-C, COL2A1, HLA-B, HLA-E, NOTCH2, XYLT1, KLHL14, CREB3L2, LRMP, AGPAT3	123	1183	14505	1.395583778	1	0.637681118	94.82317353
GOTERM_CC_FAT	GO:0005789~endoplasmic reticulum membrane	12	7.5	0.223543123	NOTCH2, XYLT1, SCD, KLHL14, FADS3, CREB3L2, LRMP, FADS2, HLA-C, HLA-B, HLA-E, AGPAT3	123	1004	14505	1.409484015	1	0.666968561	96.86754411
GOTERM_CC_FAT	GO:0042175~nuclear outer membrane-endoplasmic reticulum membrane network	12	7.5	0.243607887	NOTCH2, XYLT1, SCD, KLHL14, FADS3, CREB3L2, LRMP, FADS2, HLA-C, HLA-B, HLA-E, AGPAT3	123	1025	14505	1.380606782	1	0.681060274	97.81101209
GOTERM_CC_FAT	GO:0005783~endoplasmic reticulum	15	9.375	0.555271006	GNAZ, SCD, CRTAP, FADS3, FADS2, HLA-C, COL2A1, HLA-B, HLA-E, NOTCH2, XYLT1, KLHL14, CREB3L2, LRMP, AGPAT3	123	1678	14505	1.054173086	1	0.914463458	99.99847595
Annotation Cluster 9	Enrichment Score: 0.12383793175691386											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_FAT	GO:0045121~membrane raft	3	1.875	0.716768749	CNR1, CSK, PAG1	123	297	14505	1.191180094	1	0.972152396	99.99999683
GOTERM_CC_FAT	GO:0098857~membrane microdomain	3	1.875	0.718505592	CNR1, CSK, PAG1	123	298	14505	1.187182845	1	0.971835187	99.99999709
GOTERM_CC_FAT	GO:0098589~membrane region	3	1.875	0.825423807	CNR1, CSK, PAG1	123	373	14505	0.948473158	1	0.991138063	100

Table S3.3. Function annotation clustering (molecular functions) of Twist1-upregulated genes in WL2 cells.

Annotation Cluster 1	Enrichment Score: 2.092628213089191												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0005096~GTPase activator activity	9	5.625	0.002955629	ARHGAP21, DAB2IP, ARHGAP20, PLXNB1, PREX1, IQGAP3, DEPDC1, RAPGEF2, AGAP3	134	280	15429	3.700986141	0.746015601	0.746015601	4.128779817	
GOTERM_MF_FAT	GO:0030695~GTPase regulator activity	9	5.625	0.005546517	ARHGAP21, DAB2IP, ARHGAP20, PLXNB1, PREX1, IQGAP3, DEPDC1, RAPGEF2, AGAP3	134	311	15429	3.332077554	0.923860323	0.474705616	7.617103981	
GOTERM_MF_FAT	GO:0060589~nucleoside-triphosphatase regulator activity	9	5.625	0.009439303	ARHGAP21, DAB2IP, ARHGAP20, PLXNB1, PREX1, IQGAP3, DEPDC1, RAPGEF2, AGAP3	134	341	15429	3.038932901	0.987613581	0.518986097	12.63701873	
GOTERM_MF_FAT	GO:0008047~enzyme activator activity	10	6.25	0.027534486	ARHGAP21, DAB2IP, ARHGAP20, PLXNB1, PSAP, PREX1, IQGAP3, DEPDC1, RAPGEF2, AGAP3	134	495	15429	2.326096789	0.999997569	0.630056597	32.81506692	
Annotation Cluster 2	Enrichment Score: 1.5466075739256726												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0035591~signaling adaptor activity	5	3.125	0.004600996	DAB2IP, SOCS2, GAB1, PAG1, BLNK	134	78	15429	7.380884041	0.881776332	0.656163312	6.357981992	
GOTERM_MF_FAT	GO:0005070~SH3/SH2 adaptor activity	4	2.5	0.012447426	SOCS2, GAB1, PAG1, BLNK	134	56	15429	8.224413646	0.996970406	0.563285339	16.34110784	
GOTERM_MF_FAT	GO:0060090~binding, bridging	5	3.125	0.069213372	DAB2IP, SOCS2, GAB1, PAG1, BLNK	134	179	15429	3.216251147	1	0.809944416	64.00187377	
GOTERM_MF_FAT	GO:0030674~protein binding, bridging	4	2.5	0.16422841	SOCS2, GAB1, PAG1, BLNK	134	162	15429	2.843007186	1	0.913100657	92.23462867	
Annotation Cluster 3	Enrichment Score: 1.5238489061797296												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0032561~guanyl ribonucleotide binding	9	5.625	0.023468877	GNAZ, ANXA6, AK4, RRAGD, RAPGEF2, GBP4, AGAP3, SEPT8, SEPT9	134	403	15429	2.571404763	0.999983225	0.666983364	28.70130874	
GOTERM_MF_FAT	GO:0019001~guanyl nucleotide binding	9	5.625	0.023774799	GNAZ, ANXA6, AK4, RRAGD, RAPGEF2, GBP4, AGAP3, SEPT8, SEPT9	134	404	15429	2.5650399	0.99998549	0.636796004	29.01881929	
GOTERM_MF_FAT	GO:0005525~GTP binding	8	5	0.048066679	GNAZ, ANXA6, AK4, RRAGD, GBP4, AGAP3, SEPT8, SEPT9	134	383	15429	2.405050466	1	0.738577766	50.42569009	
Annotation Cluster 4	Enrichment Score: 0.7581222570095613												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0042605~peptide antigen binding	3	1.875	0.030904092	HLA-C, HLA-B, HLA-E	134	32	15429	10.79454291	0.999999513	0.645897273	36.05615511	
GOTERM_MF_FAT	GO:0042277~peptide binding	4	2.5	0.389125648	CRIP1, HLA-C, HLA-B, HLA-E	134	260	15429	1.77141217	1	0.959805637	99.91067943	
GOTERM_MF_FAT	GO:0033218~amide binding	4	2.5	0.442106994	CRIP1, HLA-C, HLA-B, HLA-E	134	283	15429	1.62744581	1	0.971426627	99.97547004	
Annotation Cluster 5	Enrichment Score: 0.7300918186331518												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0008375~acetylglucosaminyltransferase activity	3	1.875	0.069096743	XYLT1, B3GNT7, B3GNT2	134	50	15429	6.908507463	1	0.82531606	63.93756777	
GOTERM_MF_FAT	GO:0016758~transferase activity, transferring hexosyl groups	4	2.5	0.264532288	GBE1, XYLT1, B3GNT7, B3GNT2	134	207	15429	2.224962146	1	0.941872848	98.74327495	
GOTERM_MF_FAT	GO:0008194~UDP-glycosyltransferase activity	3	1.875	0.353011035	XYLT1, B3GNT7, B3GNT2	134	144	15429	2.398787313	1	0.952857467	99.7975653	
Annotation Cluster 6	Enrichment Score: 0.7201214599917979												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0017137~Rab GTPase binding	4	2.5	0.103986788	UNC13D, DMXL2, HPS4, MICAL3	134	131	15429	3.515779879	1	0.879755405	79.07171217	
GOTERM_MF_FAT	GO:0017016~Ras GTPase binding	5	3.125	0.197321979	UNC13D, DMXL2, HPS4, MICAL3, IQGAP3	134	266	15429	2.164319381	1	0.926424243	95.63261365	
GOTERM_MF_FAT	GO:0031267~small GTPase binding	5	3.125	0.233457137	UNC13D, DMXL2, HPS4, MICAL3, IQGAP3	134	286	15429	2.012968375	1	0.939045118	97.73401166	
GOTERM_MF_FAT	GO:0051020~GTPase binding	5	3.125	0.27488648	UNC13D, DMXL2, HPS4, MICAL3, IQGAP3	134	308	15429	1.86918492	1	0.939670752	98.9731026	
Annotation Cluster 7	Enrichment Score: 0.60396335908126												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0032555~purine ribonucleotide binding	22	13.75	0.13050962	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, PDE4B, RAPGEF3, CSK, GBP4, RAPGEF2, AGAP3, SEPT8, SEPT9	134	1885	15429	1.343829922	1	0.909112644	86.35901425	
GOTERM_MF_FAT	GO:0017076~purine nucleotide binding	22	13.75	0.137180756	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, PDE4B, RAPGEF3, CSK, GBP4, RAPGEF2, AGAP3, SEPT8, SEPT9	134	1898	15429	1.334625607	1	0.912825693	87.77643545	
GOTERM_MF_FAT	GO:0032553~ribonucleotide binding	22	13.75	0.138748811	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, PDE4B, RAPGEF3, CSK, GBP4, RAPGEF2, AGAP3, SEPT8, SEPT9	134	1901	15429	1.332519412	1	0.907888415	88.08909587	
GOTERM_MF_FAT	GO:0097367~carbohydrate derivative binding	25	15.625	0.150764836	GNAZ, KIFC1, VIM, AURKA, RRAGD, TRIB1, ANXA6, ACTG1, KIF13A, SGK223, ACSS1, PDE4B, RAPGEF3, RAPGEF2, CSK, AGAP3, PDK1, FBN1, AK4, PRKD2, SDC1, NAV2, GBP4, SEPT8, SEPT9	134	2241	15429	1.284491199	1	0.899018715	90.24952997	
GOTERM_MF_FAT	GO:0000166~nucleotide binding	24	15	0.319284691	PDK1, GNAZ, KIFC1, MICAL3, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, PDE4B, CELF2, RAPGEF3, CSK, GBP4, RAPGEF2, AGAP3, SEPT8, SEPT9	134	2394	15429	1.15430367	1	0.956024465	99.58251016	
GOTERM_MF_FAT	GO:1901265~nucleoside phosphate binding	24	15	0.320037191	PDK1, GNAZ, KIFC1, MICAL3, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, PDE4B, CELF2, RAPGEF3, CSK, GBP4, RAPGEF2, AGAP3, SEPT8, SEPT9	134	2395	15429	1.153821706	1	0.953998768	99.58903641	
GOTERM_MF_FAT	GO:0035639~purine ribonucleoside triphosphate binding	19	11.875	0.321023362	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, CSK, GBP4, AGAP3, SEPT8, SEPT9	134	1841	15429	1.188318321	1	0.952083188	99.59744564	
GOTERM_MF_FAT	GO:0032550~purine ribonucleoside binding	19	11.875	0.329502699	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, CSK, GBP4, AGAP3, SEPT8, SEPT9	134	1851	15429	1.181898449	1	0.949466678	99.66342651	
GOTERM_MF_FAT	GO:0032549~ribonucleoside binding	19	11.875	0.332058664	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, CSK, GBP4, AGAP3, SEPT8, SEPT9	134	1854	15429	1.179985992	1	0.948481198	99.68124863	
GOTERM_MF_FAT	GO:0001883~purine nucleoside binding	19	11.875	0.332058664	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, CSK, GBP4, AGAP3, SEPT8, SEPT9	134	1854	15429	1.179985992	1	0.948481198	99.68124863	
GOTERM_MF_FAT	GO:0001882~nucleoside binding	19	11.875	0.338043395	PDK1, GNAZ, KIFC1, AURKA, AK4, RRAGD, TRIB1, KIF13A, ACTG1, ANXA6, PRKD2, SGK223, ACSS1, NAV2, CSK, GBP4, AGAP3, SEPT8, SEPT9	134	1861	15429	1.175547571	1	0.949439244	99.71960362	
GOTERM_MF_FAT	GO:0036094~small molecule binding	25	15.625	0.374798976	GNAZ, KIFC1, AURKA, RRAGD, TRIB1, ANXA6, ACTG1, KIF13A, SGK223, ACSS1, PDE4B, RAPGEF3, RAPGEF2, CSK, AGAP3, PDK1, MICAL3, AK4, LGALS9, PRKD2, NAV2, CELF2, GBP4, SEPT8, SEPT9	134	2579	15429	1.116147645	1	0.957217861	99.87573066	
Annotation Cluster 8	Enrichment Score: 0.47362067428695476												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0004721~phosphoprotein phosphatase activity	4	2.5	0.209698221	PPP1CA, UBASH3B, PPM1H, CDC25B	134	183	15429	2.51676046	1	0.925305426	96.49980214	
GOTERM_MF_FAT	GO:0042578~phosphoric ester hydrolase activity	5	3.125	0.411388256	PPP1CA, UBASH3B, PPM1H, PDE4B, CDC25B	134	378	15429	1.523039564	1	0.966896638	99.94736377	
GOTERM_MF_FAT	GO:0016791~phosphatase activity	4	2.5	0.439836576	PPP1CA, UBASH3B, PPM1H, CDC25B	134	282	15429	1.633216894	1	0.972057643	99.97400905	
Annotation Cluster 9	Enrichment Score: 0.3399523232995989												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0044212~transcription regulatory region DNA binding	9	5.625	0.454072112	MEF2C, RAI1, FLI1, OSR2, ETS2, CREB3L2, HIVEP2, ZFHX3, TWIST1	134	852	15429	1.216286525	1	0.965828956	99.9819875	
GOTERM_MF_FAT	GO:0000975~regulatory region DNA binding	9	5.625	0.458021933	MEF2C, RAI1, FLI1, OSR2, ETS2, CREB3L2, HIVEP2, ZFHX3, TWIST1	134	855	15429	1.212018853	1	0.963031505	99.98375752	
GOTERM_MF_FAT	GO:0001067~regulatory region nucleic acid binding	9	5.625	0.459337524	MEF2C, RAI1, FLI1, OSR2, ETS2, CREB3L2, HIVEP2, ZFHX3, TWIST1	134	856	15429	1.210602943	1	0.962096757	99.9843102	
Annotation Cluster 10	Enrichment Score: 0.3179277645576781												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0098641~cadherin binding involved in cell-cell adhesion	4	2.5	0.457899954	DAB2IP, PLIN3, DOCK9, SEPT9	134	290	15429	1.588162635	1	0.964394664	99.98370536	

Table S3.3. Function annotation clustering (molecular functions) of Twist1-upregulated genes in WL2 cells.

GOTERM_MF_FAT	GO:0098632~protein binding involved in cell-cell adhesion	4	2.5	0.480132708	DAB2IP, PLIN3, DOCK9, SEPT9	134	300	15429	1.535223881	1	0.965451856	99.99102618	
GOTERM_MF_FAT	GO:0098631~protein binding involved in cell adhesion	4	2.5	0.49109155	DAB2IP, PLIN3, DOCK9, SEPT9	134	305	15429	1.510056276	1	0.967833161	99.99337519	
GOTERM_MF_FAT	GO:0045296~cadherin binding	4	2.5	0.495444017	DAB2IP, PLIN3, DOCK9, SEPT9	134	307	15429	1.500218776	1	0.968021428	99.99413813	
Annotation Cluster 11	Enrichment Score: 0.2833385073383917												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0004672~protein kinase activity	7	4.375	0.495782659	PDK1, SGK223, PRKD2, STAT5A, AURKA, CSK, TRIB1	134	653	15429	1.234291788	1	0.966926192	99.99419392	
GOTERM_MF_FAT	GO:0016773~phosphotransferase activity, alcohol group as acceptor	8	5	0.518563861	PDK1, SGK223, PRKD2, STAT5A, GAB1, AURKA, CSK, TRIB1	134	785	15429	1.173419527	1	0.972689216	99.99699487	
GOTERM_MF_FAT	GO:0016301~kinase activity	9	5.625	0.549402838	PDK1, SGK223, PRKD2, STAT5A, GAB1, AURKA, AK4, CSK, TRIB1	134	926	15429	1.119088682	1	0.978608041	99.99882963	
Annotation Cluster 12	Enrichment Score: 0.26756979647920603												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0032559~adenyl ribonucleotide binding	15	9.375	0.451765414	ACTG1, KIF13A, PDK1, PRKD2, KIFC1, SGK223, ACSS1, NAV2, PDE4B, AURKA, RAPGEF3, AK4, CSK, RAPGEF2, TRIB1	134	1533	15429	1.126632006	1	0.967795303	99.9808725	
GOTERM_MF_FAT	GO:0030554~adenyl nucleotide binding	15	9.375	0.46277849	ACTG1, KIF13A, PDK1, PRKD2, KIFC1, SGK223, ACSS1, NAV2, PDE4B, AURKA, RAPGEF3, AK4, CSK, RAPGEF2, TRIB1	134	1544	15429	1.118605483	1	0.960536896	99.98567419	
GOTERM_MF_FAT	GO:0005524~ATP binding	12	7.5	0.753363311	ACTG1, KIF13A, PDK1, KIFC1, SGK223, PRKD2, ACSS1, NAV2, AURKA, AK4, CSK, TRIB1	134	1496	15429	0.923597254	1	0.99507043	99.99999978	
Annotation Cluster 13	Enrichment Score: 0.2268255834319366												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0016462~pyrophosphatase activity	8	5	0.568050435	GNAZ, KIF13A, KIFC1, NAV2, RRAGD, GBP4, TRPM2, SEPT9	134	822	15429	1.120601373	1	0.981809255	99.999359	
GOTERM_MF_FAT	GO:0016818~hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	8	5	0.570662115	GNAZ, KIF13A, KIFC1, NAV2, RRAGD, GBP4, TRPM2, SEPT9	134	824	15429	1.117881466	1	0.981585343	99.99941205	
GOTERM_MF_FAT	GO:0016817~hydrolase activity, acting on acid anhydrides	8	5	0.573266719	GNAZ, KIF13A, KIFC1, NAV2, RRAGD, GBP4, TRPM2, SEPT9	134	826	15429	1.115174732	1	0.979835898	99.99946087	
GOTERM_MF_FAT	GO:0017111~nucleoside-triphosphatase activity	7	4.375	0.66615455	GNAZ, KIF13A, KIFC1, NAV2, RRAGD, GBP4, SEPT9	134	777	15429	1.037313433	1	0.98927565	99.99998367	
Annotation Cluster 14	Enrichment Score: 0.2201555809393601												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0000982~transcription factor activity, RNA polymerase II core promoter proximal region sequence-	5	3.125	0.338964095	MEF2C, FLI1, HIF1A, ETS2, CREB3L2	134	341	15429	1.688296056	1	0.947585772	99.72510812	
GOTERM_MF_FAT	GO:0000978~RNA polymerase II core promoter proximal region sequence-specific DNA binding	4	2.5	0.595772639	MEF2C, FLI1, ETS2, CREB3L2	134	356	15429	1.293727989	1	0.982949965	99.99975082	
GOTERM_MF_FAT	GO:0000987~core promoter proximal region sequence-specific DNA binding	4	2.5	0.632863723	MEF2C, FLI1, ETS2, CREB3L2	134	376	15429	1.224912671	1	0.987433958	99.99993674	
GOTERM_MF_FAT	GO:0001159~core promoter proximal region DNA binding	4	2.5	0.636442007	MEF2C, FLI1, ETS2, CREB3L2	134	378	15429	1.218431651	1	0.986933383	99.99994498	
GOTERM_MF_FAT	GO:0000977~RNA polymerase II regulatory region sequence-specific DNA binding	5	3.125	0.764003125	MEF2C, FLI1, OSR2, ETS2, CREB3L2	134	599	15429	0.961116787	1	0.995640101	99.99999988	
GOTERM_MF_FAT	GO:0001012~RNA polymerase II regulatory region DNA binding	5	3.125	0.768542856	MEF2C, FLI1, OSR2, ETS2, CREB3L2	134	603	15429	0.954741219	1	0.995763422	99.99999991	
Annotation Cluster 15	Enrichment Score: 0.2053350586530742												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0043169~cation binding	37	23.125	0.572982177	GNAZ, RAI1, F13A1, UTRN, RAG1, COL2A1, FTH1, ZFP36L1, ANXA6, JADE2, PEG10, OSR2, PDE4B, CSK, RAPGEF2, DDAH1, AGAP3, POLL, SUOX, CRIP1, ADARB1, ADGRE5, SCD, MICAL3, FBN1, ZHX2, HSPG2, ACTN1, TRPM2, NOTCH2, PRKD2, PPP1CA, GBE1, ZNF362, HIVEP2, ZFHX3, HDAC7	134	4236	15429	1.005723859	1	0.980547967	99.99945572	
GOTERM_MF_FAT	GO:0046872~metal ion binding	36	22.5	0.616637893	GNAZ, RAI1, F13A1, UTRN, RAG1, COL2A1, FTH1, ZFP36L1, ANXA6, JADE2, PEG10, OSR2, PDE4B, CSK, RAPGEF2, DDAH1, AGAP3, POLL, SUOX, CRIP1, ADARB1, ADGRE5, SCD, MICAL3, FBN1, ZHX2, HSPG2, ACTN1, TRPM2, NOTCH2, PRKD2, PPP1CA, ZNF362, HIVEP2, ZFHX3, HDAC7	134	4186	15429	0.990230406	1	0.98541472	99.99988288	
GOTERM_MF_FAT	GO:0043167~ion binding	37	23.125	0.685209699	GNAZ, RAI1, F13A1, UTRN, RAG1, COL2A1, FTH1, ZFP36L1, ANXA6, JADE2, PEG10, OSR2, PDE4B, CSK, RAPGEF2, DDAH1, AGAP3, POLL, SUOX, CRIP1, ADARB1, ADGRE5, SCD, MICAL3, FBN1, ZHX2, HSPG2, ACTN1, TRPM2, NOTCH2, PRKD2, PPP1CA, GBE1, ZNF362, HIVEP2, ZFHX3, HDAC7	134	4415	15429	0.964948192	1	0.990853349	99.99999293	
Annotation Cluster 16	Enrichment Score: 0.17413081139049413												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:1990837~sequence-specific double-stranded DNA binding	7	4.375	0.597621032	MEF2C, CRIP1, FLI1, OSR2, ETS2, CREB3L2, ZFHX3	134	724	15429	1.113249361	1	0.982627327	99.99976657	
GOTERM_MF_FAT	GO:0003690~double-stranded DNA binding	7	4.375	0.701586393	MEF2C, CRIP1, FLI1, OSR2, ETS2, CREB3L2, ZFHX3	134	807	15429	0.998751595	1	0.991649391	99.9999967	
GOTERM_MF_FAT	GO:0000976~transcription regulatory region sequence-specific DNA binding	6	3.75	0.71630921	MEF2C, FLI1, OSR2, ETS2, CREB3L2, ZFHX3	134	691	15429	0.999784003	1	0.992869613	99.99999839	
Annotation Cluster 17	Enrichment Score: 0.08843832692567553												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0099600~transmembrane receptor activity	11	6.875	0.749578663	ANXA6, TNFRSF21, PLXNB1, ADGRE5, CNR1, HLA-C, PLXND1, LGR6, ADGRA2, SPN, TRPM2	134	1365	15429	0.927882565	1	0.9949991	99.99999973	
GOTERM_MF_FAT	GO:0038023~signaling receptor activity	11	6.875	0.789772714	TNFRSF21, NOTCH2, PLXNB1, ADGRE5, CNR1, HLA-C, PLXND1, LGR6, ADGRA2, SPN, TRPM2	134	1417	15429	0.893831829	1	0.996451544	99.99999998	
GOTERM_MF_FAT	GO:0004888~transmembrane signaling receptor activity	10	6.25	0.806692137	TNFRSF21, PLXNB1, ADGRE5, CNR1, HLA-C, PLXND1, LGR6, ADGRA2, SPN, TRPM2	134	1311	15429	0.878274531	1	0.997256883	99.99999999	
GOTERM_MF_FAT	GO:0060089~molecular transducer activity	12	7.5	0.869742316	ANXA6, TNFRSF21, NOTCH2, PLXNB1, ADGRE5, CNR1, HLA-C, PLXND1, LGR6, ADGRA2, SPN, TRPM2	134	1680	15429	0.822441365	1	0.999214573	100	
GOTERM_MF_FAT	GO:0004872~receptor activity	12	7.5	0.869742316	ANXA6, TNFRSF21, NOTCH2, PLXNB1, ADGRE5, CNR1, HLA-C, PLXND1, LGR6, ADGRA2, SPN, TRPM2	134	1680	15429	0.822441365	1	0.999214573	100	
Annotation Cluster 18	Enrichment Score: 0.05572404634120627												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0022836~gated channel activity	3	1.875	0.783861427	ANXA6, TTYH3, TRPM2	134	332	15429	1.040437871	1	0.99624462	99.99999997	
GOTERM_MF_FAT	GO:0005216~ion channel activity	3	1.875	0.885751856	ANXA6, TTYH3, TRPM2	134	426	15429	0.810857683	1	0.99941282	100	
GOTERM_MF_FAT	GO:0022838~substrate-specific channel activity	3	1.875	0.897894567	ANXA6, TTYH3, TRPM2	134	442	15429	0.781505369	1	0.999552267	100	
GOTERM_MF_FAT	GO:0015267~channel activity	3	1.875	0.918679896	ANXA6, TTYH3, TRPM2	134	474	15429	0.728745513	1	0.999779385	100	
GOTERM_MF_FAT	GO:0022803~passive transmembrane transporter activity	3	1.875	0.919261025	ANXA6, TTYH3, TRPM2	134	475	15429	0.727211312	1	0.999771143	100	

Table S3.4. Function annotation clustering (biological processes) of Twist1-upregulated genes in WL2 cells.

Annotation Cluster 1	Enrichment Score: 3.751488757480068												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0051270~regulation of cellular component movement	20	12.5	5.26E-05	DAB2IP, ADARB1, PLXNB1, STAT5A, ACTN1, LGR6, LGALS9, TRIB1, PRKD2, NBL1, HIF1A, LAMA5, PDE4B, GAB1, RAPGEF3, RAPGEF2, PLXND1, ADGRA2, HDAC7, TWIST1	141	815	16651	2.897968063	0.168083614	0.020239415	0.096279141	
GOTERM_BP_FAT	GO:2000145~regulation of cell motility	18	11.25	2.03E-04	DAB2IP, ADARB1, PLXNB1, STAT5A, LGR6, LGALS9, TRIB1, PRKD2, NBL1, HIF1A, LAMA5, GAB1, RAPGEF3, RAPGEF2, PLXND1, ADGRA2, HDAC7, TWIST1	141	755	16651	2.815443145	0.508674347	0.034908541	0.371291975	
GOTERM_BP_FAT	GO:0030334~regulation of cell migration	17	10.625	2.80E-04	DAB2IP, ADARB1, PLXNB1, STAT5A, LGR6, LGALS9, TRIB1, PRKD2, NBL1, HIF1A, LAMA5, GAB1, RAPGEF3, RAPGEF2, PLXND1, ADGRA2, HDAC7	141	703	16651	2.855714617	0.624888378	0.03846207	0.511936417	
GOTERM_BP_FAT	GO:0040012~regulation of locomotion	18	11.25	3.30E-04	DAB2IP, ADARB1, PLXNB1, STAT5A, LGR6, LGALS9, TRIB1, PRKD2, NBL1, HIF1A, LAMA5, GAB1, RAPGEF3, RAPGEF2, PLXND1, ADGRA2, HDAC7, TWIST1	141	787	16651	2.700965152	0.684471164	0.037720349	0.601973739	
Annotation Cluster 2	Enrichment Score: 3.504858541228309												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0071417~cellular response to organonitrogen compound	16	10	2.37E-05	MEF2C, SOCS2, STAT5A, FBN1, RRAGD, TRPM2, ZFP36L1, HIF1A, PDE4B, GAB1, GNB4, MARCKS, RAPGEF3, CSK, RAPGEF2, PRL	141	506	16651	3.734140667	0.079423858	0.016414896	0.043308345	
GOTERM_BP_FAT	GO:1901699~cellular response to nitrogen compound	16	10	1.17E-04	MEF2C, SOCS2, STAT5A, FBN1, RRAGD, TRPM2, ZFP36L1, HIF1A, PDE4B, GAB1, GNB4, MARCKS, RAPGEF3, CSK, RAPGEF2, PRL	141	583	16651	3.240952277	0.336929575	0.027019857	0.214837581	
GOTERM_BP_FAT	GO:1901701~cellular response to oxygen-containing compound	20	12.5	3.91E-04	MEF2C, DAB2IP, LITAF, SOCS2, PSAP, STAT5A, FBN1, RRAGD, TRPM2, TRIB1, ZFP36L1, HIF1A, PDE4B, GAB1, MARCKS, GNB4, RAPGEF3, CSK, RAPGEF2, PRL	141	952	16651	2.480928542	0.745247832	0.039421462	0.713232286	
GOTERM_BP_FAT	GO:0010243~response to organonitrogen compound	18	11.25	8.85E-04	MEF2C, SOCS2, STAT5A, FBN1, RRAGD, TRPM2, ZFP36L1, SDC1, HIF1A, CNR1, PDE4B, GAB1, GNB4, MARCKS, RAPGEF3, CSK, RAPGEF2, PRL	141	859	16651	2.474574592	0.954871145	0.057841173	1.608665573	
GOTERM_BP_FAT	GO:1901698~response to nitrogen compound	18	11.25	0.003113543	MEF2C, SOCS2, STAT5A, FBN1, RRAGD, TRPM2, ZFP36L1, SDC1, HIF1A, CNR1, PDE4B, GAB1, GNB4, MARCKS, RAPGEF3, CSK, RAPGEF2, PRL	141	967	16651	2.19820018	0.999981749	0.102401561	5.551380742	
Annotation Cluster 3	Enrichment Score: 3.447146836902513												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0016477~cell migration	24	15	1.50E-04	MEF2C, DAB2IP, ADARB1, PLXNB1, STAT5A, PREX1, LGR6, LGALS9, TRPM2, TRIB1, PRKD2, NBL1, SDC1, HIF1A, LAMA5, PDE4B, RAPGEF3, CSK, ADGRA2, RAPGEF2, PLXND1, SPN, HDAC7, TWIST1	141	1193	16651	2.375702235	0.408180026	0.030384857	0.274196279	
GOTERM_BP_FAT	GO:0048870~cell motility	25	15.625	3.23E-04	MEF2C, STAT5A, PREX1, LGR6, TRIB1, PDE4B, GAB1, RAPGEF3, PLXND1, RAPGEF2, CSK, SPN, TWIST1, DAB2IP, ADARB1, PLXNB1, LGALS9, TRPM2, PRKD2, NBL1, SDC1, HIF1A, LAMA5, ADGRA2, HDAC7	141	1341	16651	2.201569698	0.677549437	0.038276002	0.590682914	
GOTERM_BP_FAT	GO:0051674~localization of cell	25	15.625	3.23E-04	MEF2C, STAT5A, PREX1, LGR6, TRIB1, PDE4B, GAB1, RAPGEF3, PLXND1, RAPGEF2, CSK, SPN, TWIST1, DAB2IP, ADARB1, PLXNB1, LGALS9, TRPM2, PRKD2, NBL1, SDC1, HIF1A, LAMA5, ADGRA2, HDAC7	141	1341	16651	2.201569698	0.677549437	0.038276002	0.590682914	
GOTERM_BP_FAT	GO:0040011~locomotion	26	16.25	0.001037731	MEF2C, STAT5A, PREX1, LGR6, TRIB1, PDE4B, GAB1, B3GNT2, RAPGEF3, PLXND1, RAPGEF2, CSK, SPN, TWIST1, DAB2IP, ADARB1, PLXNB1, LGALS9, TRPM2, PRKD2, NBL1, SDC1, HIF1A, LAMA5, ADGRA2, HDAC7	141	1543	16651	1.989887986	0.973560753	0.060714976	1.88364829	
Annotation Cluster 4	Enrichment Score: 3.305034316063756												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0006024~glycosaminoglycan biosynthetic process	8	5	4.70E-05	ST3GAL1, SDC1, XYLT1, B3GNT7, CHST11, CHST2, HSPG2, B3GNT2	141	113	16651	8.360509634	0.151641523	0.023219327	0.086044021	
GOTERM_BP_FAT	GO:0006023~aminoglycan biosynthetic process	8	5	4.97E-05	ST3GAL1, SDC1, XYLT1, B3GNT7, CHST11, CHST2, HSPG2, B3GNT2	141	114	16651	8.28717183	0.15969844	0.021514496	0.091034517	
GOTERM_BP_FAT	GO:0030203~glycosaminoglycan metabolic process	9	5.625	7.10E-05	ST3GAL1, HGSNAT, SDC1, XYLT1, B3GNT7, CHST11, CHST2, HSPG2, B3GNT2	141	163	16651	6.520428143	0.21989168	0.024526466	0.129897949	
GOTERM_BP_FAT	GO:0006022~aminoglycan metabolic process	9	5.625	1.03E-04	ST3GAL1, HGSNAT, SDC1, XYLT1, B3GNT7, CHST11, CHST2, HSPG2, B3GNT2	141	172	16651	6.179242949	0.303602128	0.027449614	0.18921988	
GOTERM_BP_FAT	GO:1903510~mucopolysaccharide metabolic process	6	3.75	0.002758982	ST3GAL1, XYLT1, B3GNT7, CHST11, CHST2, B3GNT2	141	114	16651	6.215378873	0.999936662	0.095794054	4.934231161	
GOTERM_BP_FAT	GO:1901137~carbohydrate derivative biosynthetic process	9	5.625	0.312449403	ST3GAL1, SDC1, XYLT1, B3GNT7, CHST11, CHST2, HSPG2, B3GNT2, AK4	141	763	16651	1.392961713	1	0.861938037	99.89523954	
Annotation Cluster 5	Enrichment Score: 3.128974064961488												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0071310~cellular response to organic substance	35	21.875	3.43E-04	MEF2C, TNFRSF21, LITAF, STAT5A, IQGAP3, COL2A1, RRAGD, TRIB1, ZFP36L1, PEG10, PDE4B, GAB1, CHST11, CREB3L2, RAPGEF3, CSK, RAPGEF2, PRL, TWIST1, DAB2IP, SOCS2, PSAP, FBN1, HLA-C, HLA-B, HLA-E, TRPM2, LGALS9, BCL2L11, NBL1, PRKD2, HIF1A, GNB4, MARCKS, ADGRA2	141	2231	16651	1.852634222	0.698640353	0.037952991	0.625875984	
GOTERM_BP_FAT	GO:0070887~cellular response to chemical stimulus	39	24.375	6.99E-04	MEF2C, TNFRSF21, LITAF, PREX1, STAT5A, IQGAP3, COL2A1, RRAGD, TRIB1, ZFP36L1, PEG10, PDE4B, GAB1, CHST11, CREB3L2, RAPGEF3, CSK, RAPGEF2, PRL, TWIST1, PDK1, CRIP1, DAB2IP, SOCS2, PSAP, FBN1, HLA-C, HLA-B, HLA-E, TRPM2, LGALS9, BCL2L11, SERPINB9, NBL1, PRKD2, HIF1A, GNB4, MARCKS, ADGRA2	141	2702	16651	1.704513599	0.913310906	0.049670393	1.271881296	
GOTERM_BP_FAT	GO:0010033~response to organic substance	40	25	0.001713384	MEF2C, TNFRSF21, LITAF, STAT5A, IQGAP3, AURKA, COL2A1, RRAGD, TRIB1, ZFP36L1, PEG10, CNR1, PDE4B, GAB1, CHST11, CREB3L2, RAPGEF3, CSK, RAPGEF2, PRL, TWIST1, CRIP1, DAB2IP, SOCS2, PSAP, FBN1, HLA-C, HLA-B, HLA-E, TRPM2, LGALS9, BCL2L11, NBL1, PRKD2, SDC1, HIF1A, GNB4, MARCKS, ADGRA2, ZFHX3	141	2932	16651	1.611080472	0.997521922	0.072259668	3.091982945	
Annotation Cluster 6	Enrichment Score: 3.031791384661175												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0002520~immune system development	18	11.25	6.55E-04	MEF2C, POLL, PREX1, FBN1, RAG1, ACTN1, HLA-B, BCL2L11, TRPM2, LGALS9, TRIB1, ZFP36L1, NOTCH2, HIF1A, FLI1, UBASH3B, SPN, BLNK	141	836	16651	2.542654993	0.89895778	0.047600387	1.19267283	
GOTERM_BP_FAT	GO:0002521~leukocyte differentiation	13	8.125	7.58E-04	MEF2C, PREX1, FBN1, RAG1, HLA-B, TRPM2, LGALS9, TRIB1, ZFP36L1, NOTCH2, UBASH3B, SPN, BLNK	141	481	16651	3.191681043	0.929509665	0.050676339	1.378721559	
GOTERM_BP_FAT	GO:0048534~hematopoietic or lymphoid organ development	17	10.625	9.78E-04	MEF2C, PREX1, FBN1, RAG1, ACTN1, HLA-B, BCL2L11, TRPM2, LGALS9, TRIB1, ZFP36L1, NOTCH2, HIF1A, FLI1, UBASH3B, SPN, BLNK	141	789	16651	2.544445343	0.967414325	0.060354319	1.776238971	
GOTERM_BP_FAT	GO:0030097~hemopoiesis	16	10	0.001537417	MEF2C, PREX1, FBN1, RAG1, ACTN1, HLA-B, TRPM2, LGALS9, TRIB1, ZFP36L1, NOTCH2, HIF1A, FLI1, UBASH3B, SPN, BLNK	141	746	16651	2.532808549	0.995408562	0.070167521	2.778646984	
Annotation Cluster 7	Enrichment Score: 2.988604448894012												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0050767~regulation of neurogenesis	17	10.625	3.18E-04	MEF2C, TNFRSF21, DAB2IP, PLXNB1, PREX1, MAP1B, VIM, ZHX2, XK, FOXO6, NBL1, HIF1A, CNR1, MARCKS, RAPGEF2, PLXND1, ZFHX3	141	711	16651	2.823582807	0.671090389	0.03893506	0.580362202	
GOTERM_BP_FAT	GO:0045664~regulation of neuron differentiation	15	9.375	4.85E-04	MEF2C, NBL1, DAB2IP, PLXNB1, CNR1, PREX1, VIM, MAP1B, ZHX2, XK, MARCKS, PLXND1, RAPGEF2, ZFHX3, FOXO6	141	594	16651	2.982126227	0.816831586	0.041546021	0.884528471	
GOTERM_BP_FAT	GO:0051960~regulation of nervous system development	17	10.625	0.001115917	MEF2C, TNFRSF21, DAB2IP, PLXNB1, PREX1, MAP1B, VIM, ZHX2, XK, FOXO6, NBL1, HIF1A, CNR1, MARCKS, RAPGEF2, PLXND1, ZFHX3	141	799	16651	2.51259997	0.979894744	0.063038326	2.024201364	

Table S3.4. Function annotation clustering (biological processes) of Twist1-upregulated genes in WL2 cells.

					MEF2C, TNFRSF21, DAB2IP, PLXNB1, PREX1, MAP1B, VIM, STXBP1, ZHX2, XK, AURKA, FOXO6, NBL1,								
GOTERM_BP_FAT	GO:0048699~generation of neurons	22	13.75	0.006458733	HIF1A, CNR1, MARCKS, B3GNT2, PLXND1, RAPGEF2, ZFHX3, HDAC7, TWIST1	141	1397	16651	1.859719663	1	0.157819257	11.19052554	
Annotation Cluster 8	Enrichment Score: 2.983440211138212												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:1901342~regulation of vasculature development	10	6.25	2.67E-04	PRKD2, DAB2IP, HIF1A, RAPGEF3, PLXND1, RAPGEF2, ADGRA2, DDAH1, HDAC7, TWIST1	141	250	16651	4.723687943	0.606484531	0.038114432	0.486990481	
GOTERM_BP_FAT	GO:0045765~regulation of angiogenesis	9	5.625	6.55E-04	PRKD2, DAB2IP, HIF1A, RAPGEF3, PLXND1, ADGRA2, DDAH1, HDAC7, TWIST1	141	226	16651	4.702786669	0.898922683	0.048602432	1.192493209	
					ZFP36L1, PRKD2, DAB2IP, HIF1A, LAMA5, HSPG2, RAPGEF3, ADGRA2, PLXND1, RAPGEF2, DDAH1,								
GOTERM_BP_FAT	GO:0048514~blood vessel morphogenesis	13	8.125	9.71E-04	HDAC7, TWIST1	141	495	16651	3.101411276	0.966651948	0.061034894	1.764347849	
					MEF2C, DAB2IP, HSPG2, ZFP36L1, PRKD2, HIF1A, LAMA5, RAPGEF3, RAPGEF2, PLXND1, ADGRA2,								
GOTERM_BP_FAT	GO:0001568~blood vessel development	14	8.75	0.001282137	DDAH1, HDAC7, TWIST1	141	583	16651	2.835833242	0.988768971	0.065754742	2.322377585	
					MEF2C, DAB2IP, HSPG2, ZFP36L1, PRKD2, HIF1A, LAMA5, RAPGEF3, RAPGEF2, PLXND1, ADGRA2,								
GOTERM_BP_FAT	GO:0001944~vasculature development	14	8.75	0.00211754	DDAH1, HDAC7, TWIST1	141	617	16651	2.679563663	0.999399129	0.082631189	3.808035542	
GOTERM_BP_FAT	GO:0001525~angiogenesis	11	6.875	0.002731232	PRKD2, DAB2IP, HIF1A, LAMA5, HSPG2, RAPGEF3, PLXND1, ADGRA2, DDAH1, HDAC7, TWIST1	141	416	16651	3.122630251	0.999930185	0.095825784	4.885768933	
Annotation Cluster 9	Enrichment Score: 2.955164219096973 4												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
					MEF2C, TNFRSF21, LITAF, RAG1, AURKA, COL2A1, ZFP36L1, ANXA6, PEG10, PCBP4, CNR1, CHST11,								
					RAPGEF3, RAPGEF2, SPN, TWIST1, PDK1, DAB2IP, CRIP1, SOCS2, PSAP, STXBP1, ACTN1, LGALS9,								
GOTERM_BP_FAT	GO:0008219~cell death	32	20	4.64E-04	TRPM2, BCL2L11, NOTCH2, SERPINB9, PRKD2, PPP1CA, HIF1A, HDAC7	141	1991	16651	1.89801625	0.802833738	0.042935028	0.846315115	
					MEF2C, TNFRSF21, LITAF, RAG1, AURKA, COL2A1, ZFP36L1, ANXA6, PEG10, PCBP4, CNR1, CHST11,								
					RAPGEF3, RAPGEF2, SPN, TWIST1, PDK1, DAB2IP, CRIP1, SOCS2, STXBP1, LGALS9, BCL2L11, NOTCH2,								
GOTERM_BP_FAT	GO:0006915~apoptotic process	28	17.5	0.001620305	SERPINB9, PPP1CA, HIF1A, HDAC7	141	1774	16651	1.863912943	0.996566051	0.071039342	2.926360358	
					MEF2C, TNFRSF21, LITAF, RAG1, AURKA, COL2A1, ZFP36L1, ANXA6, PEG10, PCBP4, CNR1, CHST11,								
					RAPGEF3, RAPGEF2, SPN, TWIST1, PDK1, DAB2IP, CRIP1, SOCS2, STXBP1, ACTN1, LGALS9, BCL2L11,								
GOTERM_BP_FAT	GO:0012501~programmed cell death	29	18.125	0.001813207	NOTCH2, SERPINB9, PPP1CA, HIF1A, HDAC7	141	1880	16651	1.821634978	0.998253551	0.075402867	3.269309078	
Annotation Cluster 10	Enrichment Score: 2.6318146937230478												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
					MEF2C, TNFAIP8L1, LITAF, PREX1, IQGAP3, AURKA, COL2A1, RRAGD, LGR6, RIMS3, TRIB1, PEG10,								
					ARHGAP20, UBASH3B, CNR1, WWC3, GAB1, CHST11, RAPGEF3, CSK, RAPGEF2, PRL, PAG1, BLNK,								
					TWIST1, DAB2IP, SOCS2, PSAP, PLXNB1, FBN1, STXBP1, TRPM2, LGALS9, BCL2L11, NOTCH2, NBL1,								
GOTERM_BP_FAT	GO:0010646~regulation of cell communication	42	26.25	0.001142141	PRKD2, PPP1CA, HIF1A, MARCKS, ADGRA2, HDAC7	141	3069	16651	1.61612002	0.981659344	0.063449088	2.071300634	
					MEF2C, TNFAIP8L1, LITAF, PREX1, IQGAP3, AURKA, COL2A1, RRAGD, LGR6, RIMS3, TRIB1, PEG10,								
					ARHGAP20, UBASH3B, CNR1, WWC3, GAB1, CHST11, RAPGEF3, CSK, RAPGEF2, PRL, PAG1, BLNK,								
					TWIST1, DAB2IP, SOCS2, PSAP, PLXNB1, FBN1, STXBP1, TRPM2, LGALS9, BCL2L11, NOTCH2, NBL1,								
GOTERM_BP_FAT	GO:0023051~regulation of signaling	42	26.25	0.001574037	PRKD2, PPP1CA, HIF1A, MARCKS, ADGRA2, HDAC7	141	3119	16651	1.590212357	0.995961553	0.069957444	2.843933338	
					MEF2C, TNFAIP8L1, LITAF, PREX1, IQGAP3, AURKA, COL2A1, RRAGD, LGR6, TRIB1, PEG10, ARHGAP20,								
					UBASH3B, WWC3, GAB1, CHST11, CSK, RAPGEF2, PRL, PAG1, BLNK, TWIST1, DAB2IP, SOCS2, PLXNB1,								
GOTERM_BP_FAT	GO:0009966~regulation of signal transduction	36	22.5	0.007076545	PSAP, FBN1, LGALS9, BCL2L11, NOTCH2, NBL1, PRKD2, PPP1CA, HIF1A, ADGRA2, HDAC7	141	2765	16651	1.537547613	1	0.166992117	12.19654288	
Annotation Cluster 11	Enrichment Score: 2.53536775962131												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0050865~regulation of cell activation	13	8.125	0.001254074	MEF2C, ZFP36L1, TNFRSF21, UNC13D, UBASH3B, CNR1, RAG1, CSK, HLA-E, LGALS9, SPN, ICOSLG, PAG1	141	510	16651	3.010193297	0.987608668	0.065319131	2.272096678	
GOTERM_BP_FAT	GO:0002694~regulation of leukocyte activation	12	7.5	0.002260836	ZFP36L1, MEF2C, TNFRSF21, UNC13D, CNR1, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	475	16651	2.983381859	0.99963646	0.085140258	4.060712327	
GOTERM_BP_FAT	GO:0051249~regulation of lymphocyte activation	10	6.25	0.008735837	ZFP36L1, MEF2C, TNFRSF21, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	417	16651	2.831947208	1	0.188483262	14.84540588	
Annotation Cluster 12	Enrichment Score: 2.359872009530389												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:2000147~positive regulation of cell motility	11	6.875	0.003129328	PRKD2, DAB2IP, HIF1A, STAT5A, RAPGEF3, RAPGEF2, LGR6, ADGRA2, HDAC7, LGALS9, TWIST1	141	424	16651	3.063712699	0.999982733	0.10100041	5.578769264	
GOTERM_BP_FAT	GO:0051272~positive regulation of cellular component movement	11	6.875	0.003812087	PRKD2, DAB2IP, HIF1A, STAT5A, RAPGEF3, RAPGEF2, LGR6, ADGRA2, HDAC7, LGALS9, TWIST1	141	436	16651	2.979390331	0.999998429	0.111539716	6.756205704	
GOTERM_BP_FAT	GO:0040017~positive regulation of locomotion	11	6.875	0.003873934	PRKD2, DAB2IP, HIF1A, STAT5A, RAPGEF3, RAPGEF2, LGR6, ADGRA2, HDAC7, LGALS9, TWIST1	141	437	16651	2.972572504	0.999998736	0.112310697	6.86217252	
GOTERM_BP_FAT	GO:0030335~positive regulation of cell migration	10	6.25	0.007865852	PRKD2, DAB2IP, HIF1A, STAT5A, RAPGEF3, RAPGEF2, LGR6, ADGRA2, HDAC7, LGALS9	141	410	16651	2.880297526	1	0.1768259	13.46614719	
Annotation Cluster 13	Enrichment Score: 2.29611358887147												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0045953~negative regulation of natural killer cell mediated cytotoxicity	4	2.5	1.56E-04	SERPINB9, HLA-B, HLA-E, LGALS9	141	13	16651	36.3360611	0.42153043	0.029951693	0.286105956	
GOTERM_BP_FAT	GO:0002716~negative regulation of natural killer cell mediated immunity	4	2.5	1.98E-04	SERPINB9, HLA-B, HLA-E, LGALS9	141	14	16651	33.74062817	0.499617474	0.035785193	0.361765985	
GOTERM_BP_FAT	GO:0002707~negative regulation of lymphocyte mediated immunity	5	3.125	2.55E-04	SERPINB9, HLA-B, HLA-E, LGALS9, SPN	141	37	16651	15.95840521	0.590149487	0.038038636	0.465802418	
GOTERM_BP_FAT	GO:0001911~negative regulation of leukocyte mediated cytotoxicity	4	2.5	3.01E-04	SERPINB9, HLA-B, HLA-E, LGALS9	141	16	16651	29.52304965	0.6508331	0.039661481	0.549254321	
GOTERM_BP_FAT	GO:0031342~negative regulation of cell killing	4	2.5	5.11E-04	SERPINB9, HLA-B, HLA-E, LGALS9	141	19	16651	24.86151549	0.832629866	0.041667665	0.93131337	
GOTERM_BP_FAT	GO:0002704~negative regulation of leukocyte mediated immunity	5	3.125	6.99E-04	SERPINB9, HLA-B, HLA-E, LGALS9, SPN	141	48	16651	12.30127069	0.91345701	0.048714548	1.272753003	
GOTERM_BP_FAT	GO:0042270~protection from natural killer cell mediated cytotoxicity	3	1.875	0.001029824	SERPINB9, HLA-B, HLA-E	141	6	16651	59.04609929	0.97281827	0.061290578	1.869423349	
GOTERM_BP_FAT	GO:0042267~natural killer cell mediated cytotoxicity	5	3.125	0.001171038	SERPINB9, UNC13D, HLA-B, HLA-E, LGALS9	141	55	16651	10.73565442	0.983425064	0.06398786	2.123176612	
GOTERM_BP_FAT	GO:0002228~natural killer cell mediated immunity	5	3.125	0.001830548	SERPINB9, UNC13D, HLA-B, HLA-E, LGALS9	141	62	16651	9.523564402	0.998356552	0.075204292	3.300082892	
GOTERM_BP_FAT	GO:0031341~regulation of cell killing	5	3.125	0.002058175	SERPINB9, HLA-B, HLA-E, BCL2L11, LGALS9	141	64	16651	9.225953014	0.999260086	0.082241641	3.703171421	
GOTERM_BP_FAT	GO:0001906~cell killing	6	3.75	0.002758982	SERPINB9, UNC13D, HLA-B, HLA-E, BCL2L11, LGALS9	141	114	16651	6.215378873	0.999936662	0.095794054	4.934231161	
GOTERM_BP_FAT	GO:0042269~regulation of natural killer cell mediated cytotoxicity	4	2.5	0.003127145	SERPINB9, HLA-B, HLA-E, LGALS9	141	35	16651	13.49625127	0.9999826	0.101870857	5.574981117	
GOTERM_BP_FAT	GO:0002715~regulation of natural killer cell mediated immunity	4	2.5	0.003667312	SERPINB9, HLA-B, HLA-E, LGALS9	141	37	16651	12.76672417	0.999997388	0.111251056	6.507702799	
GOTERM_BP_FAT	GO:0045824~negative regulation of innate immune response	4	2.5	0.004578514	SERPINB9, HLA-B, HLA-E, LGALS9	141	40	16651	11.80921986	0.999999894	0.124274441	8.061387633	
GOTERM_BP_FAT	GO:0001909~leukocyte mediated cytotoxicity	5	3.125	0.005494083	SERPINB9, UNC13D, HLA-B, HLA-E, LGALS9	141	84	16651	7.029297535	0.999999996	0.140827711	9.597912519	
GOTERM_BP_FAT	GO:0001910~regulation of leukocyte mediated cytotoxicity	4	2.5	0.01002964	SERPINB9, HLA-B, HLA-E, LGALS9	141	53	16651	8.912618761	1	0.202358008	16.85819758	
GOTERM_BP_FAT	GO:0002703~regulation of leukocyte mediated immunity	6	3.75	0.011080732	SERPINB9, UNC13D, HLA-B, HLA-E, LGALS9, SPN	141	159	16651	4.45630938	1	0.210448517	18.4601992	
GOTERM_BP_FAT	GO:0002698~negative regulation of immune effector process	5	3.125	0.011144813	SERPINB9, HLA-B, HLA-E, LGALS9, SPN	141	103	16651	5.732630999	1	0.210403087	18.55691701	
GOTERM_BP_FAT	GO:0002706~regulation of lymphocyte mediated immunity	5	3.125	0.018073179	SERPINB9, HLA-B, HLA-E, LGALS9, SPN	141	119	16651	4.961857083	1	0.262056525	28.39764504	
GOTERM_BP_FAT	GO:0050777~negative regulation of immune response	5	3.125	0.021788568	SERPINB9, HLA-B, HLA-E, LGALS9, SPN	141	126	16651	4.686198356	1	0.283719977	33.20048411	
GOTERM_BP_FAT	GO:0002449~lymphocyte mediated immunity	7	4.375	0.026246079	SERPINB9, UNC13D, IGLV7-43, HLA-B, HLA-E, LGALS9, SPN	141	269	16651	3.07303119	1	0.308776768	38.56090817	
GOTERM_BP_FAT	GO:0031348~negative regulation of defense response	5	3.125	0.043162974	SERPINB9, HLA-B, HLA-E, LGALS9, SPN	141	156	16651	3.785006365	1	0.389354235	55.42995722	
GOTERM_BP_FAT	GO:0002443~leukocyte mediated immunity	7	4.375	0.066757344	SERPINB9, UNC13D, IGLV7-43, HLA-B, HLA-E, LGALS9, SPN	141	339	16651	2.438481977	1	0.478792224	71.78735313	

Table S3.4. Function annotation clustering (biological processes) of Twist1-upregulated genes in WL2 cells.

GOTERM_BP_FAT	GO:0002699~positive regulation of immune effector process	4	2.5	0.16005673	UNC13D, HLA-B, HLA-E, LGALS9	141	164	16651	2.880297526	1	0.686596541	95.90154233
GOTERM_BP_FAT	GO:0002697~regulation of immune effector process	6	3.75	0.303383325	SERPINB9, UNC13D, HLA-B, HLA-E, LGALS9, SPN	141	435	16651	1.628857911	1	0.855458294	99.86683305
GOTERM_BP_FAT	GO:0045088~regulation of innate immune response	5	3.125	0.364580027	SERPINB9, DAB2IP, HLA-B, HLA-E, LGALS9	141	363	16651	1.626614306	1	0.893320311	99.97528195

Annotation Cluster 14	Enrichment Score: 2.099146060630365											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0000904~cell morphogenesis involved in differentiation	16	10	0.001914865	MEF2C, DAB2IP, PLXNB1, PREX1, MAP1B, STXBP1, XK, ACTN1, UNC13D, HIF1A, LAMA5, MARCKS, B3GNT2, RAPGEF2, PLXND1, TWIST1	141	763	16651	2.476376379	0.998777107	0.077623362	3.449578959
GOTERM_BP_FAT	GO:0032989~cellular component morphogenesis	20	12.5	0.015529804	MEF2C, DAB2IP, PLXNB1, PREX1, MAP1B, STXBP1, XK, ACTN1, AURKA, ACTG1, NBL1, UNC13D, HIF1A, LAMA5, MARCKS, B3GNT2, RAPGEF2, PLXND1, SEPT9, TWIST1	141	1328	16651	1.778496967	1	0.244856039	24.92358051
GOTERM_BP_FAT	GO:0000902~cell morphogenesis	19	11.875	0.016953433	MEF2C, DAB2IP, PLXNB1, PREX1, MAP1B, STXBP1, XK, ACTN1, AURKA, NBL1, UNC13D, HIF1A, LAMA5, MARCKS, B3GNT2, RAPGEF2, PLXND1, SEPT9, TWIST1	141	1247	16651	1.799319786	1	0.254185033	26.88731073

Annotation Cluster 15	Enrichment Score: 2.0256821863342496											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0048699~generation of neurons	22	13.75	0.006458733	MEF2C, TNFRSF21, DAB2IP, PLXNB1, PREX1, MAP1B, VIM, STXBP1, ZHX2, XK, AURKA, FOXO6, NBL1, HIF1A, CNR1, MARCKS, B3GNT2, PLXND1, RAPGEF2, ZFHX3, HDAC7, TWIST1	141	1397	16651	1.859719663	1	0.157819257	11.19052554
GOTERM_BP_FAT	GO:0022008~neurogenesis	23	14.375	0.006476022	MEF2C, TNFRSF21, DAB2IP, PLXNB1, PREX1, MAP1B, VIM, STXBP1, ZHX2, XK, AURKA, FOXO6, NBL1, HIF1A, CNR1, MARCKS, B3GNT2, CSK, PLXND1, RAPGEF2, ZFHX3, HDAC7, TWIST1	141	1489	16651	1.824123954	1	0.157116894	11.21882548
GOTERM_BP_FAT	GO:0030182~neuron differentiation	19	11.875	0.020021556	MEF2C, DAB2IP, PLXNB1, PREX1, MAP1B, VIM, STXBP1, ZHX2, XK, AURKA, FOXO6, NBL1, HIF1A, CNR1, MARCKS, B3GNT2, RAPGEF2, PLXND1, HDAC7	141	1270	16651	1.76673368	1	0.275058785	30.95557866

Annotation Cluster 16	Enrichment Score: 1.986933984422298											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0045666~positive regulation of neuron differentiation	9	5.625	0.007566466	MEF2C, NBL1, DAB2IP, PLXNB1, CNR1, MAP1B, MARCKS, RAPGEF2, FOXO6	141	336	16651	3.163183891	1	0.172897509	12.98663924
GOTERM_BP_FAT	GO:0050769~positive regulation of neurogenesis	10	6.25	0.007985898	MEF2C, NBL1, DAB2IP, HIF1A, PLXNB1, CNR1, MAP1B, MARCKS, RAPGEF2, FOXO6	141	411	16651	2.873289503	1	0.178141975	13.65771399
GOTERM_BP_FAT	GO:0051962~positive regulation of nervous system development	10	6.25	0.018112617	MEF2C, NBL1, DAB2IP, HIF1A, PLXNB1, CNR1, MAP1B, MARCKS, RAPGEF2, FOXO6	141	471	16651	2.507265363	1	0.261485006	28.45029847

Annotation Cluster 17	Enrichment Score: 1.9404552909151223											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0010941~regulation of cell death	24	15	0.004933378	MEF2C, TNFRSF21, DAB2IP, SOCS2, PSAP, RAG1, STXBP1, ACTN1, COL2A1, AURKA, BCL2L11, LGALS9, TRPM2, ZFP36L1, NOTCH2, SERPINB9, PPP1CA, HIF1A, CNR1, CHST11, RAPGEF3, RAPGEF2, HDAC7, TWIST1	141	1546	16651	1.833255347	0.999999969	0.130253351	8.659835693
GOTERM_BP_FAT	GO:0042981~regulation of apoptotic process	21	13.125	0.016673763	MEF2C, TNFRSF21, DAB2IP, SOCS2, RAG1, STXBP1, ACTN1, COL2A1, AURKA, BCL2L11, LGALS9, ZFP36L1, NOTCH2, SERPINB9, PPP1CA, HIF1A, CNR1, CHST11, RAPGEF3, RAPGEF2, TWIST1	141	1432	16651	1.731799002	1	0.252674353	26.50541262
GOTERM_BP_FAT	GO:0043067~regulation of programmed cell death	21	13.125	0.018342396	MEF2C, TNFRSF21, DAB2IP, SOCS2, RAG1, STXBP1, ACTN1, COL2A1, AURKA, BCL2L11, LGALS9, ZFP36L1, NOTCH2, SERPINB9, PPP1CA, HIF1A, CNR1, CHST11, RAPGEF3, RAPGEF2, TWIST1	141	1446	16651	1.71503193	1	0.261172904	28.75634592

Annotation Cluster 18	Enrichment Score: 1.9239742007042109											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0001667~ameboidal-type cell migration	9	5.625	0.006018992	PRKD2, DAB2IP, HIF1A, LAMA5, STAT5A, PLXND1, ADGRA2, HDAC7, TWIST1	141	323	16651	3.290494697	0.999999999	0.149977952	10.46784241
GOTERM_BP_FAT	GO:0043542~endothelial cell migration	6	3.75	0.009999891	PRKD2, DAB2IP, STAT5A, PLXND1, ADGRA2, HDAC7	141	155	16651	4.571310913	1	0.204153505	16.81242591
GOTERM_BP_FAT	GO:0010631~epithelial cell migration	7	4.375	0.01345957	PRKD2, DAB2IP, HIF1A, STAT5A, PLXND1, ADGRA2, HDAC7	141	231	16651	3.578551472	1	0.231577701	21.97876605
GOTERM_BP_FAT	GO:0090132~epithelium migration	7	4.375	0.014264148	PRKD2, DAB2IP, HIF1A, STAT5A, PLXND1, ADGRA2, HDAC7	141	234	16651	3.532672607	1	0.235720408	23.13598249
GOTERM_BP_FAT	GO:0010632~regulation of epithelial cell migration	6	3.75	0.015487213	PRKD2, DAB2IP, HIF1A, STAT5A, ADGRA2, HDAC7	141	173	16651	4.095683188	1	0.245359846	24.86407009
GOTERM_BP_FAT	GO:0090130~tissue migration	7	4.375	0.015972963	PRKD2, DAB2IP, HIF1A, STAT5A, PLXND1, ADGRA2, HDAC7	141	240	16651	3.444355792	1	0.246569446	25.5401486

Annotation Cluster 19	Enrichment Score: 1.9031930458079356											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0009617~response to bacterium	13	8.125	0.003775768	MEF2C, TNFRSF21, SERPINB9, DAB2IP, HIF1A, LITAF, CNR1, PDE4B, HLA-B, HLA-E, LGALS9, SPN, TRIB1	141	583	16651	2.633273725	0.999998216	0.111465926	6.693922369
GOTERM_BP_FAT	GO:0051707~response to other organism	16	10	0.016585749	MEF2C, TNFRSF21, DAB2IP, ADARB1, LITAF, HLA-B, HLA-E, LGALS9, BCL2L11, TRIB1, SERPINB9, UNC13D, HIF1A, CNR1, PDE4B, SPN	141	971	16651	1.945906465	1	0.252592863	26.38483667
GOTERM_BP_FAT	GO:0043207~response to external biotic stimulus	16	10	0.016585749	MEF2C, TNFRSF21, DAB2IP, ADARB1, LITAF, HLA-B, HLA-E, LGALS9, BCL2L11, TRIB1, SERPINB9, UNC13D, HIF1A, CNR1, PDE4B, SPN	141	971	16651	1.945906465	1	0.252592863	26.38483667
GOTERM_BP_FAT	GO:0009607~response to biotic stimulus	16	10	0.023482936	MEF2C, TNFRSF21, DAB2IP, ADARB1, LITAF, HLA-B, HLA-E, LGALS9, BCL2L11, TRIB1, SERPINB9, UNC13D, HIF1A, CNR1, PDE4B, SPN	141	1014	16651	1.863387749	1	0.293827044	35.288148

Annotation Cluster 20	Enrichment Score: 1.8885577849857704											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0071375~cellular response to peptide hormone stimulus	10	6.25	0.00133097	ZFP36L1, HIF1A, SOCS2, STAT5A, GAB1, FBN1, GNB4, MARCKS, CSK, PRL	141	313	16651	3.772913693	0.990535098	0.064406117	2.409814218
GOTERM_BP_FAT	GO:1901653~cellular response to peptide	10	6.25	0.002160529	ZFP36L1, HIF1A, SOCS2, STAT5A, GAB1, FBN1, GNB4, MARCKS, CSK, PRL	141	336	16651	3.514648767	0.999483208	0.083311004	3.883905004
GOTERM_BP_FAT	GO:0043434~response to peptide hormone	10	6.25	0.011788322	ZFP36L1, HIF1A, SOCS2, STAT5A, GAB1, FBN1, GNB4, MARCKS, CSK, PRL	141	438	16651	2.696168917	1	0.21545044	19.52217601
GOTERM_BP_FAT	GO:1901652~response to peptide	10	6.25	0.020201653	ZFP36L1, HIF1A, SOCS2, STAT5A, GAB1, FBN1, GNB4, MARCKS, CSK, PRL	141	480	16651	2.460254137	1	0.275058013	31.18760648
GOTERM_BP_FAT	GO:0032870~cellular response to hormone stimulus	11	6.875	0.045314363	ZFP36L1, MEF2C, HIF1A, SOCS2, STAT5A, GAB1, FBN1, GNB4, MARCKS, CSK, PRL	141	644	16651	2.017102771	1	0.395837553	57.23008613
GOTERM_BP_FAT	GO:0009725~response to hormone	12	7.5	0.150262348	ZFP36L1, MEF2C, SDC1, HIF1A, SOCS2, STAT5A, GAB1, FBN1, GNB4, MARCKS, CSK, PRL	141	918	16651	1.54368887	1	0.672780863	94.93200762

Annotation Cluster 21	Enrichment Score: 1.8148423337195263											
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0051249~regulation of lymphocyte activation	10	6.25	0.008735837	ZFP36L1, MEF2C, TNFRSF21, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	417	16651	2.831947208	1	0.188483262	14.84540588
GOTERM_BP_FAT	GO:0070489~T cell aggregation	10	6.25	0.014041763	ZFP36L1, TNFRSF21, PREX1, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	451	16651	2.618452297	1	0.236913887	22.81776183
GOTERM_BP_FAT	GO:0042110~T cell activation	10	6.25	0.014041763	ZFP36L1, TNFRSF21, PREX1, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	451	16651	2.618452297	1	0.236913887	22.81776183
GOTERM_BP_FAT	GO:0071593~lymphocyte aggregation	10	6.25	0.014227468	ZFP36L1, TNFRSF21, PREX1, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	452	16651	2.612659261	1	0.236289803	23.08358085
GOTERM_BP_FAT	GO:0050863~regulation of T cell activation	8	5	0.014688433	TNFRSF21, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	305	16651	3.097500291	1	0.237443476	23.73967459
GOTERM_BP_FAT	GO:0070486~leukocyte aggregation	10	6.25	0.01557885	ZFP36L1, TNFRSF21, PREX1, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	459	16651	2.572814784	1	0.244445511	24.99205455
GOTERM_BP_FAT	GO:0022407~regulation of cell-cell adhesion	9	5.625	0.018331725	TNFRSF21, UBASH3B, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	394	16651	2.697537531	1	0.263147479	28.74215973
GOTERM_BP_FAT	GO:1903037~regulation of leukocyte cell-cell adhesion	8	5	0.01833223	TNFRSF21, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	319	16651	2.961559839	1	0.262096568	28.74283107
GOTERM_BP_FAT	GO:0007159~leukocyte cell-cell adhesion	10	6.25	0.024615808	ZFP36L1, TNFRSF21, PREX1, RAG1, HLA-E, CSK, ICOSLG, PAG1, LGALS9, SPN	141	497	16651	2.376100575	1	0.299494568	36.64940689

3.7 References

1. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocols* 2008 12//print; **4**(1)**:** 44-57.

2. Thisse B, el Messal M, Perrin-Schmitt F. The twist gene: isolation of a Drosophila zygotic gene necessary for the establishment of dorsoventral pattern. *Nucleic Acids Res* 1987 Apr 24; **15**(8)**:** 3439-3453.

3. Thisse B, Stoetzel C, Gorostiza-Thisse C, Perrin-Schmitt F. Sequence of the twist gene and nuclear localization of its protein in endomesodermal cells of early Drosophila embryos. *EMBO J* 1988 Jul; **7**(7)**:** 2175-2183.

4. Castanon I, Baylies MK. A Twist in fate: evolutionary comparison of Twist structure and function. *Gene* 2002 Apr 3; **287**(1-2)**:** 11-22.

5. Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 1989 Mar 10; **56**(5)**:** 777-783.

6. Murre C, McCaw PS, Vaessin H, Caudy M, Jan LY, Jan YN*, et al.* Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 1989 Aug 11; **58**(3)**:** 537-544.

7. Jones N. Transcriptional regulation by dimerization: two sides to an incestuous relationship. *Cell* 1990 Apr 6; **61**(1)**:** 9-11.

8. Kadesch T. Consequences of heteromeric interactions among helix-loop-helix proteins. *Cell Growth Differ* 1993 Jan; **4**(1)**:** 49-55.

9. Baylies MK, Bate M. twist: a myogenic switch in Drosophila. *Science* 1996 Jun 7; **272**(5267)**:** 1481-1484.

10. Fuchtbauer EM. Expression of M-twist during postimplantation development of the mouse. *Dev Dyn* 1995 Nov; **204**(3)**:** 316-322.

11. Hebrok M, Wertz K, Fuchtbauer EM. M-twist is an inhibitor of muscle differentiation. *Dev Biol* 1994 Oct; **165**(2)**:** 537-544.

12. Koutsoulidou A, Mastroyiannopoulos NP, Furling D, Uney JB, Phylactou LA. Endogenous TWIST expression and differentiation are opposite during human muscle development. *Muscle Nerve* 2011 Dec; **44**(6)**:** 984-986.

13. Rohwedel J, Horak V, Hebrok M, Fuchtbauer EM, Wobus AM. M-twist expression inhibits mouse embryonic stem cell-derived myogenic differentiation in vitro. *Exp Cell Res* 1995 Sep; **220**(1)**:** 92-100.

14. Spicer DB, Rhee J, Cheung WL, Lassar AB. Inhibition of myogenic bHLH and MEF2 transcription factors by the bHLH protein Twist. *Science* 1996 Jun 7; **272**(5267)**:** 1476-1480.

15. Wang SM, Coljee VW, Pignolo RJ, Rotenberg MO, Cristofalo VJ, Sierra F. Cloning of the human twist gene: its expression is retained in adult mesodermally-derived tissues. *Gene* 1997 Mar 10; **187**(1)**:** 83-92.

16. Isenmann S, Arthur A, Zannettino AC, Turner JL, Shi S, Glackin CA, *et al.* TWIST family of basic helix-loop-helix transcription factors mediate human mesenchymal stem cell growth and commitment. *Stem Cells* 2009 Oct; **27**(10)**:** 2457-2468.

17. Dong CY, Liu XY, Wang N, Wang LN, Yang BX, Ren Q, *et al.* Twist-1, a novel regulator of hematopoietic stem cell self-renewal and myeloid lineage development. *Stem Cells* 2014 Dec; **32**(12)**:** 3173-3182.

18. Sharif MN, Sosic D, Rothlin CV, Kelly E, Lemke G, Olson EN, *et al.* Twist mediates suppression of inflammation by type I IFNs and Axl. *The Journal of experimental medicine* 2006 Aug 7; **203**(8)**:** 1891-1901.

19. Sosic D, Richardson JA, Yu K, Ornitz DM, Olson EN. Twist regulates cytokine gene expression through a negative feedback loop that represses NF-kappaB activity. *Cell* 2003 Jan 24; **112**(2)**:** 169-180.

20. Niesner U, Albrecht I, Janke M, Doebis C, Loddenkemper C, Lexberg MH, *et al.* Autoregulation of Th1-mediated inflammation by twist1. *The Journal of experimental medicine* 2008 Aug 4; **205**(8)**:** 1889-1901.

21. Pham D, Vincentz JW, Firulli AB, Kaplan MH. Twist1 regulates Ifng expression in Th1 cells by interfering with Runx3 function. *J Immunol* 2012 Jul 15; **189**(2)**:** 832-840.

22. Pham D, Walline CC, Hollister K, Dent AL, Blum JS, Firulli AB, *et al.* The transcription factor Twist1 limits T helper 17 and T follicular helper cell development by repressing the gene encoding the interleukin-6 receptor alpha chain. *The Journal of biological chemistry* 2013 Sep 20; **288**(38)**:** 27423-27433.

23. Ansieau S, Morel AP, Hinkal G, Bastid J, Puisieux A. TWISTing an embryonic transcription factor into an oncoprotein. *Oncogene* 2010 Jun 3; **29**(22)**:** 3173-3184.

24. Qin Q, Xu Y, He T, Qin C, Xu J. Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. *Cell research* 2012 Jan; **22**(1)**:** 90-106.

25. Ansieau S, Bastid J, Doreau A, Morel AP, Bouchet BP, Thomas C, *et al.* Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer Cell* 2008 Jul 08; **14**(1)**:** 79-89.

26. Maestro R, Dei Tos AP, Hamamori Y, Krasnokutsky S, Sartorelli V, Kedes L*, et al.* Twist is a potential oncogene that inhibits apoptosis. *Genes & development* 1999 Sep 01; **13**(17)**:** 2207-2217.

27. Valsesia-Wittmann S, Magdeleine M, Dupasquier S, Garin E, Jallas AC, Combaret V*, et al.* Oncogenic cooperation between H-Twist and N-Myc overrides failsafe programs in cancer cells. *Cancer Cell* 2004 Dec; **6**(6)**:** 625-630.

28. Cheng GZ, Chan J, Wang Q, Zhang W, Sun CD, Wang LH. Twist transcriptionally up-regulates AKT2 in breast cancer cells leading to increased migration, invasion, and resistance to paclitaxel. *Cancer research* 2007 Mar 01; **67**(5)**:** 1979-1987.

29. Kwok WK, Ling MT, Lee TW, Lau TC, Zhou C, Zhang X*, et al.* Up-regulation of TWIST in prostate cancer and its implication as a therapeutic target. *Cancer research* 2005 Jun 15; **65**(12)**:** 5153-5162.

30. Wang X, Ling MT, Guan XY, Tsao SW, Cheung HW, Lee DT*, et al.* Identification of a novel function of TWIST, a bHLH protein, in the development of acquired taxol resistance in human cancer cells. *Oncogene* 2004 Jan 15; **23**(2)**:** 474-482.

31. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY*, et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008 May 16; **133**(4)**:** 704-715.

32. Morel AP, Lievre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS one* 2008; **3**(8)**:** e2888.

33. Li X, Marcondes AM, Gooley TA, Deeg HJ. The helix-loop-helix transcription factor TWIST is dysregulated in myelodysplastic syndromes. *Blood* 2010 Sep 30; **116**(13)**:** 2304-2314.

34. Stirewalt DL, Mhyre AJ, Marcondes M, Pogosova-Agadjanyan E, Abbasi N, Radich JP*, et al.* Tumour necrosis factor-induced gene expression in human marrow stroma: clues to the pathophysiology of MDS? *Br J Haematol* 2008 Feb; **140**(4)**:** 444-453.

35. Kawagoe H, Kandilci A, Kranenburg TA, Grosveld GC. Overexpression of N-Myc rapidly causes acute myeloid leukemia in mice. *Cancer research* 2007 Nov 15; **67**(22)**:** 10677-10685.

36. Cosset E, Hamdan G, Jeanpierre S, Voeltzel T, Sagorny K, Hayette S*, et al.* Deregulation of TWIST-1 in the CD34+ compartment represents a novel prognostic factor in chronic myeloid leukemia. *Blood* 2011 Feb 3; **117**(5)**:** 1673-1676.

37. Tipping AJ, Deininger MW, Goldman JM, Melo JV. Comparative gene expression profile of chronic myeloid leukemia cells innately resistant to imatinib mesylate. *Exp Hematol* 2003 Nov; **31**(11)**:** 1073-1080.

38. Zhang J, Wang P, Wu F, Li M, Sharon D, Ingham RJ*, et al.* Aberrant expression of the transcriptional factor Twist1 promotes invasiveness in ALK-positive anaplastic large cell lymphoma. *Cell Signal* 2012 Apr; **24**(4)**:** 852-858.

39. Lemma S, Karihtala P, Haapasaari KM, Jantunen E, Soini Y, Bloigu R*, et al.* Biological roles and prognostic values of the epithelial-mesenchymal transition-mediating transcription factors Twist, ZEB1 and Slug in diffuse large B-cell lymphoma. *Histopathology* 2013 Jan; **62**(2)**:** 326-333.

40. Merindol N, Riquet A, Szablewski V, Eliaou JF, Puisieux A, Bonnefoy N. The emerging role of Twist proteins in hematopoietic cells and hematological malignancies. *Blood cancer journal* 2014 Apr 25; **4:** e206.

41. Norozi F, Ahmadzadeh A, Shahjahani M, Shahrabi S, Saki N. Twist as a new prognostic marker in hematological malignancies. *Clin Transl Oncol* 2016 Feb; **18**(2)**:** 113-124.

42. Kim MS, Kim GM, Choi YJ, Kim HJ, Kim YJ, Jin W. TrkC promotes survival and growth of leukemia cells through Akt-mTOR-dependent up-regulation of PLK-1 and Twist-1. *Mol Cells* 2013 Aug; **36**(2)**:** 177-184.

43. Zhu QQ, Ma C, Wang Q, Song Y, Lv T. The role of TWIST1 in epithelial-mesenchymal transition and cancers. *Tumour Biol* 2016 Jan; **37**(1)**:** 185-197.

44. Tsukerman P, Yamin R, Seidel E, Khawaled S, Schmiedel D, Bar-Mag T*, et al.* MiR-520d-5p directly targets TWIST1 and downregulates the metastamiR miR-10b. *Oncotarget* 2014 Dec 15; **5**(23)**:** 12141-12150.

45. Yu J, Xie F, Bao X, Chen W, Xu Q. miR-300 inhibits epithelial to mesenchymal transition and metastasis by targeting Twist in human epithelial cancer. *Mol Cancer* 2014 May 24; **13:** 121.

46. Xue G, Hemmings BA. Phosphorylation of basic helix-loop-helix transcription factor Twist in development and disease. *Biochem Soc Trans* 2012 Feb; **40**(1)**:** 90-93.

47. Ezponda T, Popovic R, Shah MY, Martinez-Garcia E, Zheng Y, Min DJ*, et al.* The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. *Oncogene* 2012 Jul 16.

48. Moreaux J, Klein B, Bataille R, Descamps G, Maiga S, Hose D*, et al.* A high-risk signature for patients with multiple myeloma established from the molecular classification of human myeloma cell lines. *Haematologica* 2011 Apr; **96**(4)**:** 574-582.

49. Persons DA, Mehaffey MG, Kaleko M, Nienhuis AW, Vanin EF. An improved method for generating retroviral producer clones for vectors lacking a selectable marker gene. *Blood Cells Mol Dis* 1998 Jun; **24**(2)**:** 167-182.

50. Diamond P, Labrinidis A, Martin SK, Farrugia AN, Gronthos S, To LB*, et al.* Targeted disruption of the CXCL12/CXCR4 axis inhibits osteolysis in a murine model of myeloma-associated bone loss. *J Bone Miner Res* 2009 Jul; **24**(7)**:** 1150-1161.

51. Lund SP, Nettleton D, McCarthy DJ, Smyth GK. Detecting differential expression in RNA-sequence data using quasi-likelihood with shrunken dispersion estimates. *Stat Appl Genet Mol Biol* 2012 Oct 22; **11**(5).

52. Noll JE, Hewett DR, Williams SA, Vandyke K, Kok C, To LB*, et al.* SAMSN1 is a tumor suppressor gene in multiple myeloma. *Neoplasia* 2014 Jul; **16**(7)**:** 572-585.

53. Cheong CM, Chow AW, Fitter S, Hewett DR, Martin SK, Williams SA*, et al.* Tetraspanin 7 (TSPAN7) expression is upregulated in multiple myeloma patients and inhibits myeloma tumour development in vivo. *Exp Cell Res* 2015 Mar 1; **332**(1)**:** 24-38.

54. Mrozik KM, Cheong CM, Hewett D, Chow AW, Blaschuk OW, Zannettino AC*, et al.* Therapeutic targeting of N-cadherin is an effective treatment for multiple myeloma. *Br J Haematol* 2015 Nov; **171**(3)**:** 387-399.

55. Podar K, Tai YT, Lin BK, Narsimhan RP, Sattler M, Kijima T*, et al.* Vascular endothelial growth factor-induced migration of multiple myeloma cells is associated with beta 1 integrin- and phosphatidylinositol 3-kinase-dependent PKC alpha activation. *The Journal of biological chemistry* 2002 Mar 8; **277**(10)**:** 7875-7881.

56. Azab AK, Azab F, Blotta S, Pitsillides CM, Thompson B, Runnels JM*, et al.* RhoA and Rac1 GTPases play major and differential roles in stromal cell-derived factor-1-induced cell adhesion and chemotaxis in multiple myeloma. *Blood* 2009 Jul 16; **114**(3)**:** 619-629.

57. Lombardi L, Poretti G, Mattioli M, Fabris S, Agnelli L, Bicciato S*, et al.* Molecular characterization of human multiple myeloma cell lines by integrative genomics: insights into the biology of the disease. *Genes, chromosomes & cancer* 2007 Mar; **46**(3)**:** 226-238.

58. Brabletz T, Hlubek F, Spaderna S, Schmalhofer O, Hiendlmeyer E, Jung A*, et al.* Invasion and metastasis in colorectal cancer: epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. *Cells Tissues Organs* 2005; **179**(1-2)**:** 56-65.

59. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C*, et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 2004 Jun 25; **117**(7)**:** 927-939.

60. Gajula RP, Chettiar ST, Williams RD, Thiyagarajan S, Kato Y, Aziz K, *et al.* The twist box domain is required for Twist1-induced prostate cancer metastasis. *Mol Cancer Res* 2013 Nov; **11**(11)**:** 1387-1400.

61. Mironchik Y, Winnard PT, Jr., Vesuna F, Kato Y, Wildes F, Pathak AP, *et al.* Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer. *Cancer research* 2005 Dec 1; **65**(23)**:** 10801-10809.

62. Xu J, Zhang W, Yan XJ, Lin XQ, Li W, Mi JQ, *et al.* DNMT3A mutation leads to leukemic extramedullary infiltration mediated by TWIST1. *J Hematol Oncol* 2016 Oct 10; **9**(1)**:** 106.

63. Damoulakis G, Gambardella L, Rossman KL, Lawson CD, Anderson KE, Fukui Y, *et al.* P-Rex1 directly activates RhoG to regulate GPCR-driven Rac signalling and actin polarity in neutrophils. *J Cell Sci* 2014 Jun 1; **127**(Pt 11)**:** 2589-2600.

64. Eckert RE, Neuder LE, Park J, Adler KB, Jones SL. Myristoylated alanine-rich C-kinase substrate (MARCKS) protein regulation of human neutrophil migration. *Am J Respir Cell Mol Biol* 2010 May; **42**(5)**:** 586-594.

65. Giridharan SS, Caplan S. MICAL-family proteins: Complex regulators of the actin cytoskeleton. *Antioxid Redox Signal* 2014 May 1; **20**(13)**:** 2059-2073.

66. Sun Y, Pan J, Mao S, Jin J. IL-17/miR-192/IL-17Rs regulatory feedback loop facilitates multiple myeloma progression. *PloS one* 2014; **9**(12)**:** e114647.

67. Li J, Pan Q, Rowan PD, Trotter TN, Peker D, Regal KM, *et al.* Heparanase promotes myeloma progression by inducing mesenchymal features and motility of myeloma cells. *Oncotarget* 2016 Mar 8; **7**(10)**:** 11299-11309.

68. Mikheeva SA, Mikheev AM, Petit A, Beyer R, Oxford RG, Khorasani L, *et al.* TWIST1 promotes invasion through mesenchymal change in human glioblastoma. *Mol Cancer* 2010 Jul 20; **9:** 194.

69. Shamir ER, Pappalardo E, Jorgens DM, Coutinho K, Tsai WT, Aziz K, *et al.* Twist1-induced dissemination preserves epithelial identity and requires E-cadherin. *The Journal of cell biology* 2014 Mar 3; **204**(5)**:** 839-856.

70. Akiyama SK, Aota S, Yamada KM. Function and receptor specificity of a minimal 20 kilodalton cell adhesive fragment of fibronectin. *Cell Adhes Commun* 1995 Feb; **3**(1)**:** 13-25.

71. Mitsiades CS, McMillin DW, Klippel S, Hideshima T, Chauhan D, Richardson PG, *et al.* The role of the bone marrow microenvironment in the pathophysiology of myeloma and its significance in the development of more effective therapies. *Hematol Oncol Clin North Am* 2007 Dec; **21**(6)**:** 1007-1034, vii-viii.

72. Podar K, Chauhan D, Anderson KC. Bone marrow microenvironment and the identification of new targets for myeloma therapy. *Leukemia* 2009 Jan; **23**(1)**:** 10-24.

73. Damiano JS, Dalton WS. Integrin-mediated drug resistance in multiple myeloma. *Leukemia & lymphoma* 2000 Jun; **38**(1-2)**:** 71-81.

74. Hazlehurst LA, Damiano JS, Buyuksal I, Pledger WJ, Dalton WS. Adhesion to fibronectin via beta1 integrins regulates p27kip1 levels and contributes to cell adhesion mediated drug resistance (CAM-DR). *Oncogene* 2000 Sep 7; **19**(38)**:** 4319-4327.

75. Plow EF, Haas TA, Zhang L, Loftus J, Smith JW. Ligand binding to integrins. *The Journal of biological chemistry* 2000 Jul 21; **275**(29)**:** 21785-21788.

76. Ruoslahti E. Fibronectin and its receptors. *Annu Rev Biochem* 1988; **57:** 375-413.

77. Shattil SJ, Kim C, Ginsberg MH. The final steps of integrin activation: the end game. *Nat Rev Mol Cell Biol* 2010 04//print; **11**(4)**:** 288-300.

78. Kinashi T. Intracellular signalling controlling integrin activation in lymphocytes. *Nat Rev Immunol* 2005 07//print; **5**(7)**:** 546-559.

79. Kinbara K, Goldfinger LE, Hansen M, Chou FL, Ginsberg MH. Ras GTPases: integrins' friends or foes? *Nat Rev Mol Cell Biol* 2003 Oct; **4**(10)**:** 767-776.

80. Han Z, Liu L, Liu Y, Li S. Sirtuin SIRT6 suppresses cell proliferation through inhibition of Twist1 expression in non-small cell lung cancer. *Int J Clin Exp Pathol* 2014; **7**(8)**:** 4774-4781.

81. Qiang L, Zhao B, Ming M, Wang N, He TC, Hwang S, *et al.* Regulation of cell proliferation and migration by p62 through stabilization of Twist1. *Proceedings of the National Academy of Sciences of the United States of America* 2014 Jun 24; **111**(25)**:** 9241-9246.

82. Min DJ, Ezponda T, Kim MK, Will CM, Martinez-Garcia E, Popovic R, *et al.* MMSET stimulates myeloma cell growth through microRNA-mediated modulation of c-MYC. *Leukemia* 2013 Mar; **27**(3)**:** 686-694.

83. Hu L, Roth JM, Brooks P, Ibrahim S, Karpatkin S. Twist is required for thrombin-induced tumor angiogenesis and growth. *Cancer research* 2008 Jun 1; **68**(11)**:** 4296-4302.

84. Yang JZ, Lian WG, Sun LX, Qi DW, Ding Y, Zhang XH. High nuclear expression of Twist1 in the skeletal extramedullary disease of myeloma patients predicts inferior survival. *Pathol Res Pract* 2016 Mar; **212**(3)**:** 210-216.

85. Pugh CW, Ratcliffe PJ. Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat Med* 2003 Jun; **9**(6)**:** 677-684.

86. Kato H, Tamamizu-Kato S, Shibasaki F. Histone deacetylase 7 associates with hypoxia-inducible factor 1alpha and increases transcriptional activity. *The Journal of biological chemistry* 2004 Oct 1; **279**(40)**:** 41966-41974.

87. Gopal SK, Greening DW, Zhu HJ, Simpson RJ, Mathias RA. Transformed MDCK cells secrete elevated MMP1 that generates LAMA5 fragments promoting endothelial cell angiogenesis. *Sci Rep* 2016 Jun 21; **6:** 28321.

88. Gubbiotti MA, Neill T, Iozzo RV. A current view of perlecan in physiology and pathology: A mosaic of functions. *Matrix Biol* 2016 Sep 6.

89. Parri M, Chiarugi P. Rac and Rho GTPases in cancer cell motility control. *Cell Commun Signal* 2010 Sep 07; **8:** 23.

90. Yang WH, Lan HY, Huang CH, Tai SK, Tzeng CH, Kao SY*, et al.* RAC1 activation mediates Twist1-induced cancer cell migration. *Nature cell biology* 2012 Mar 11; **14**(4)**:** 366-374.

91. Kobune M, Chiba H, Kato J, Kato K, Nakamura K, Kawano Y*, et al.* Wnt3/RhoA/ROCK signaling pathway is involved in adhesion-mediated drug resistance of multiple myeloma in an autocrine mechanism. *Mol Cancer Ther* 2007 Jun; **6**(6)**:** 1774-1784.

Chapter 4:
TWIST1 promotes dissemination and multiple myeloma disease progression *in vivo*

Statement of Authorship

Title of Paper	TWIST1 promotes dissemination and progression of multiple myeloma in vivo
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Chee Man Cheong, Krzysztof Mrozik, Duncan Hewett, Jacqueline Noll, Kate Vandyke & Andrew C.W. Zannettino (2017) TWIST1 promotes dissemination and progression of multiple myeloma in vivo.

Principal Author

Name of Principal Author (Candidate)	Chee Man Cheong
Contribution to the Paper	Design experiments, performed experiments, analysed data and wrote manuscript
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 23/12/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Krzysztof Mrozik
Contribution to the Paper	Data interpretation and performed experiments
Signature	Date: 23/12/16

Name of Co-Author	Duncan Hewett
Contribution to the Paper	Data interpretation and manuscript evaluation
Signature	Date: 23/12/16

Name of Co-Author	Kate Vandyke		
Contribution to the Paper	Design experiments, data interpretation, manuscript evaluation and co-supervised development of work		
Signature		Date	23/12/16

Name of Co-Author	Jacqueline Noll		
Contribution to the Paper	Performed animal injections		
Signature		Date	23/12/16

Name of Co-Author	Andrew Zannettino		
Contribution to the Paper	Designed experiments, data interpretation, manuscript revision, provided funding and supervised development of work		
Signature		Date	23/12/2016

Please cut and paste additional co-author panels here as required.

Chapter 4: TWIST1 promotes dissemination and multiple myeloma disease progression *in vivo*

Chee Man Cheong[1,2], Krzysztof Mrozik[1,2], Duncan Hewett[1,2], Jacqueline Noll[1,2], Kate Vandyke[1,2,3], Andrew C.W Zannettino[1,2,3,4]

1. Myeloma Research Group, Discipline of Physiology, Adelaide Medical School, The University of Adelaide, Adelaide, Australia.
2. Cancer Theme, South Australian Health & Medical Research Institute (SAHMRI), Adelaide, Australia.
3. SA Pathology, Adelaide, Australia.
4. Centre for Cancer Biology, University of South Australia, Adelaide, Australia.

Keywords: multiple myeloma, TWIST1, migration, extramedullary infiltration, C57Bl/KaLwRij

4.1 Abstract

The expression of the TWIST1 transcription factor promotes invasion and metastasis of various cancers. We have previously demonstrated that TWIST1 expression is upregulated in newly-diagnosed MM patients exhibiting MMSET overexpression (Chapter 2) and promotes myeloma cell migration *in vitro* (Chapter 3). In this study, we investigated the effects of TWIST1 overexpression on 5TGM1 murine myeloma plasma cell dissemination and MM disease progression *in vivo*. Furthermore, we evaluated the impact of TWIST1 overexpression on the 5TGM1 PC transcriptome and whether TWIST1 affected cell migration as previously seen in HMCLs (Chapter 3). As detailed herein, we found that TWIST1 significantly increased tumour burden, splenic infiltration and the number of circulating tumour cells 4 weeks after intravenous inoculation of C57BL/KaLwRij mice with 5TGM1-TWIST1 cells. Furthermore, we found that TWIST1 significantly increased transendothelial migration and actin polymerisation of 5TGM1 cells *in vitro*. Finally, we demonstrated that TWIST1 overexpression significantly enhanced 5TGM1 cell proliferation in response to yet to be identified, spleen-derived soluble factors. These findings suggest that elevated TWIST1 levels enhances MM disease aggressiveness by promoting tumour cell dissemination to distant BM sites and extramedullary infiltration.

4.2 Introduction

Multiple myeloma (MM), a malignancy of antibody-secreting plasma cells (PCs), is characterised by the proliferation of clonal PCs in the bone marrow (BM) and excessive production of non-functional monoclonal immunoglobulin (Ig). Although the overall survival outcomes for MM patients have improved significantly since the introduction of novel therapies, the survival rate for high-risk MM patients remains universally poor.[1]

Murine models of MM have played an important role in our understanding of the mechanisms of MM pathogenesis (reviewed in [2, 3]). These *in vivo* experimental models enable the complex interactions between MM cells and the tumour microenvironments to be studied and overcome many of the limitations associated with *in vitro* systems.[4] In addition, murine models have served as excellent preclinical models of myeloma and have assisted in the evaluation of the efficacy and toxicity of novel therapeutic strategies for MM disease control.

To date, the most common preclinical models of MM include the systemic SCID/NOD model, SCID-Hu model, and the 5TMM model. In the NOD/SCID model, human myeloma cell lines are injected intravenously into immune compromised NOD/SCID mice to establish systemic disease, or subcutaneously to establish a localized plasmacytoma.[5] This model provides a platform to evaluate the efficacy of therapeutics against both medullary and extramedullary disease *in vivo*.[6, 7] The SCID/Hu model involves engraftment of human fetal bone into immune compromised SCID mice prior to injection of primary MM cells or human MM cell lines.[8, 9] Despite being a highly relevant model to investigate the interaction between human MM cells and human BM microenvironment, this model is restricted by the availability of human fetal bone tissues, is technically challenging and ethically fraught. Moreover, the allogeneic nature of fetal bone and MM cells is often associated with variable engraftment rates and results.[10]

In contrast to SCID-hu and NOD/SCID models, the 5TMM murine model is a syngeneic, immunocompetent model of MM.[11] This model was first described by Radl et al.[12] and was based on the observation that the C57BL/KaLwRij mice spontaneously developed monoclonal B cell proliferative disorders at low frequency (0.5% of the mice at 2 years of age). Subsequent transplantation of the BM cells from ageing mice into young C57BL/KaLwRij mice lead to the development of a myeloma-like disease resembling human MM, including intramedullary plasmacytosis, elevated serum paraprotein, extramedullary dissemination, osteolytic bone lesions and renal

dysfunction.[12] The two most commonly used 5T2 and 5T33 cell lines were established from serial transplantation from different donor mice into young syngeneic recipient mice.[13-15] Inoculation of 5T2 cell line into C57BL/KaLwRij results in relatively slow rates of disease progression and the formation of widespread bone disease within 3 months. In contrast, inoculation of C57BL/KaLwRij with the 5T33 cell line results in PC line derived from the IgG2bκ-secreting 5T33 subclone, which can be cultured *in vitro*.[14-16] The establishment of 5TGM1 cell line has enabled researchers to genetically modify the cells *in vitro*[17] and thus, unravel the role of specific molecular targets underlying disease pathogenesis.[18-20]

TWIST1 is a basic helix-loop-helix transcription factor that promotes invasion and metastasis in a variety of epithelial cancers.[21-23] As detailed in Chapter 2, we have shown that TWIST1 expression is upregulated in approximately 50% of the newly diagnosed, MMSET-high, MM patients. Furthermore, as detailed in Chapter 3, ectopic expression of TWIST1 in human MM cell line resulted in increased cell migration *in vitro*, demonstrating a potential role for TWIST1 in MM disease development. In this Chapter, we investigated the effect of TWIST1 overexpression in the murine MM 5TGM1 PC line on MM disease development *in vivo* using the systemic C57BL/KaLwRij murine model of MM.

4.3 Materials and Methods

Cell culture

All culture media were supplemented with additives (2mM L-glutamine, 1mM sodium pyruvate, 15 mM HEPES, 50 U/ml penicillin and 50 μg/ml streptomycin; from Sigma-Aldrich, Sydney, Australia). 5TGM1 cells were cultured in Iscove's Modified Eagle's Medium (IMDM) supplemented with 20% fetal calf serum (FCS) and additives, unless otherwise specified. The transformed BM endothelial cells (BMEC) were cultured in M199 medium supplemented with 20% FCS, additives, 15 μg/mL heparin and 15 μg/mL of the endothelial cell growth factor supplement, as previously described.[24] C57BL/KaLwRij bone marrow stromal cells (BMSC) were cultured in α-modified Eagle's medium (α-MEM) supplemented with 10% FCS, additives and 100 mM L-ascorbate-2-phosphate. All cells were maintained in a humidified incubation chamber at 37°C with 5% CO_2.

Generation of TWIST1-overexpressing 5TGM1 cells

The full-length murine *Twist1* coding sequence was amplified from pCMV-SPORT6-TWIST1 cDNA vector (accession BC083139, clone ID 6516673, Open Biosystems) with a BamHI and EcoRI restriction digest sites at 5' and 3'end respectively of *Twist1* cDNA sequence using forward primer 5'-GGATCCCCACCATGATGCAGGACGTGTCC and reverse primer 5'-GAATTCCTAGTGGGACGCGGACATGG. The resulting PCR product was ligated into pGEMT-Easy (Promega, Madison, USA) and then subcloned into BamHI/EcoRI linearised pLEGO-iT2. The resulting pLEGO-mTWIST1-iT2 construct was sequence verified (Australian Genome Research Facility (AGRF) prior to use. Empty vector pLEGO-iT2 was used as a control.

Luciferase-tagged 5TGM1 cell line (as described previously[20, 24, 25])-overexpressing *Twist1* was generated by transfecting empty vector pLEGO-iT2 or pLEGO-mTWIST1-iT2 into HEK293T cells together with psPAX2 and pEQEco packaging plasmids. Lentiviral supernatant collected from transiently transfected HEK293T cells was used to transduce 5TGM1-luc cells as previously described.[26] Stable TWIST1-overexpressing 5TGM1 (5TGM1-TWIST1) or empty vector control (5TGM1-EV) cell lines were generated from the top 30% GFP-positive cells sorted using BD FACSFusion.

Library preparation

Total RNA was extracted from TWIST1-overexpressing 5TGM1 cells and empty vector control cells using RNA kit (Qiagen) according to the manufacturer's instructions. All samples used for cDNA library construction had a RIN score > 8, determined using a Bioanalyzer 2200 (Agilent). Library construction and RNA-sequencing with NextSeq500 (Illumina) was performed at David Gunn Genomics Facility (SAHMRI, Adelaide) as previously described (Chapter 3).

RNA-sequencing analysis

Raw fastq files containing single-end reads (1 × 75 bp) obtained from RNA-sequencing were analysed according to the workflow described previously in Chapter 3 (supplementary Figure S3.1). Briefly, read quality was assessed using FastQC and overrepresented adapter sequences were trimmed from the raw fastq files using Trimmomatic version 0.33 and repeated quality assessment. Single-end reads (1 ×75 bp)

were mapped to the reference genome mm9 using TopHat2. BAM files from individual lanes of the same sample were combined and the number of reads mapped were counted using featureCounts (part of Rsubread version 1.12.6). Transcripts expressed at levels below 1 count per million reads, in at least two libraries, were filtered out from downstream analysis. Differential expression was determined using exactTest from edgeR R package. Differentially regulated genes were determined using cut-off of 2-fold or greater changes in mean expression. Gene ontology (GO) annotation enrichment analysis was performed using DAVID Bioinformatics Resources 6.8[27] to classify genes by related functions and assess their level of enrichment.

Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

Total RNA was isolated from cell lines using TRIzol (Life Technologies, Mulgrave, Australia) and cDNA was synthesized using Superscript III First-Strand Synthesis System (Life Technologies). qPCR were performed using RT2 SYBR Green qPCR Mastermix (Qiagen, Chadstone, Ausrtalia) on the Bio-Rad CFX Connect (Bio-Rad) using the following primers murine transcripts: *Actb* (F: 5'-GATCATTGCTCCTCCTGAGC-3'; R: 5'-GTCATAGTCCGCCTAGAAGCAT-3'), *Twist1* (F: 5'-CAGCGGGTCATGGCTAAC-3'; R: 5'-TGAATCTTGCTCAGCTTGTCC-3'), *Ptprk* (F: 5'-GGGAACACACGTCAGGAGAT-3' and R: 5'-GTCGGCACAAGTGGATCATT-3'), *Tmem29* (F: 5'-GCAGGGTCTCAATGGCCTTA-3' and R: 5'-TACTTGAGCATCCTCTGGAACA-3'), *Mmp10* (F: 5'-TGACCCACTCACTTTCTCC-3'; R: 5'-GGGTGCAAGTGTCCATTTCT-3'). Gene expression was represented relative to *Actb* expression, calculated using the $2^{-\Delta CT}$ method.

Immunoblotting

Nuclear lysates were isolated using Nuclear Complex Co-IP kit (54001, Active Motif) according to the manufacturer's instruction. Briefly, nuclear proteins (30 μg) were prepared in non-reducing buffer, resolved on 12% SDS-polyacrylamide gels and transferred to PVDF membranes (GE Healthcare). Membranes were incubated in 2.5% enhanced chemiluminescence Blocking Agent (GE Healthcare) for 1 hour at room temperature and then probed overnight at 4°C with antibodies against TWIST1 (1:200 dilution; ab50887, Abcam) and histone H3 (1:5000 dilution; 9715S, Cell Signalling).

Proteins of interest were detected by incubating the membrane in alkaline phosphatase-conjugated anti-mouse IgG (1:5000 dilution; Millipore). Proteins were visualised using enhanced chemifluorescence substrate (ECF, GE Healthcare) on a ChemiDoc (BioRad).

Generation of conditioned medium (CM)

C57BL/KaLwRij-derived BM stromal cells were seeded until 80% confluence. Spent media was aspirated and replaced with serum free-DMEM supplemented with additives. For splenic cell CM, spleens were harvested from C57BL/KaLwRij mice and the cells recovered by mechanical disruption of the spleens prior to culture in RPMI-1640 medium supplemented with 10% FCS, 5 μM of 2-mercaptoethanol and additives. All conditioned media were collected after 48 hours and concentrated using Centriprep Centrifugal Filter Unit with Ultracel-3 membrane (Millipore) and stored at -20°C until use.

Cell proliferation assay

Bone marrow stromal cells (BMSC) were seeded at a density of 1×10^4 cells/well in 96-well clear-bottom black plates (Corning) and cultured overnight. Luciferase-expressing 5TGM1 cells were cultured with or without BMSC for 3 days. Firefly D-luciferin (150 ng/ml final concentration) was added to wells and incubated for 15 minutes. Bioluminescence signals (total flux, photon/sec) were measured using the Xenogen IVIS 100 bioluminescence imaging system (Caliper Life Sciences, Hopkinton, MA, USA). Absolute 5TGM1 cell numbers were calculated in relation to the bioluminescence signal obtained from a standard curve.

For WST-1 assays, 5TGM1 cells were plated at a density of 1×10^4 cells/well in splenic CM or RPMI-1640 media in 96-well plates in triplicates. Relative cell proliferation were measured at over 5 days by adding WST-1 reagent (Roche, Basel, Switzerland) according to the manufacturer's instruction.

Methylcellulose colony-forming unit assay

Tumour initiating potential of 5TGM1-EV and 5TGM1-TWIST1 were determined by culturing the cells in MethoCult^{TM} medium (H4230, Stem Cell Technology). Briefly, 5TGM1 cells were plated at 200 cells/mL in 1 mL of methylcellulose-based MethoCult culture medium mixture in 35 mm dishes in duplicate. The dishes were incubated at 37°C for 16 days and colonies enumerated using an Olympus SZX7 stereomicroscope. Images were photographed using a Olympus DP21 camera.

Adhesion assay

The adhesion of 5TGM1-TWIST and 5TGM1-EV cells to BM stromal cells and BM endothelial cells was performed as previously described.[24] BMSCs and BMECs were seeded at 1×10^4 cells per well in a 96-well plate and allowed to adhere overnight. Empty wells (plastic) were used as controls for adhesion. Non-adherent cells were gently aspirated followed by three washes with HBSS supplemented with 5% FCS. Adherent cells were quantitated using bioluminescence imaging as described above. The percentage of cell adhesion was determined in relation to the bioluminescence signal obtained from the total number of cells loaded.

For adhesion to fibronectin, 96-well plates were coated with 10 μg/ml recombinant fibronectin (source) or 1% BSA overnight at 37°C. 5TGM1 cells were added at 1×10^5 cells per well in serum-free culture media and incubated for 2 hours at 37˚C. Non-adherent cells were gently aspirated followed by three washes with Hank's buffer. Images for each well were taken at × 10 magnification. The number of adherent cells was counted using ImageJ software (http://fiji.sc). Cell adhesion to fibronectin was measured relative to cell adhesion to BSA.

Transendothelial migration assay

Migration assays were performed as previously described.[20] 5TGM1 cells were washed once in serum free IMDM media, seeded (1×10^5 cells) on 8 μm transwells, previously seeded with BMEC in serum-free IMDM media in quadruplicate. Cells were allowed to migrate towards BM stromal cell or spleenic cell conditioned-media and complete medium with CXCL12 at 37°C. Migrated cells in the bottom chamber were enumerated after 6 hours using microscopy as previously described.[24] Percentage cell migration is represented in relative to basal migration.

Gelatin zymography

5TGM1-EV and 5TGM1-TWIST1 cells were cultured in serum-free IMDM supplemented with additives for 48 hours. Conditioned media were collected and stored in -20°C until use. The conditioned media were electrophoresed on 10% SDS-PAGE gel containing 1% gelatin as previously described.[20] The gel was incubated in MMP activation buffer, stained with 0.25% Coomassie Blue and destained until clear zones of gelatin lysis against dark background of Coomassie Blue staining were visible. The gel was imaged on GelDoc (BioRad).

C57Bl/KaLwRij murine model of MM

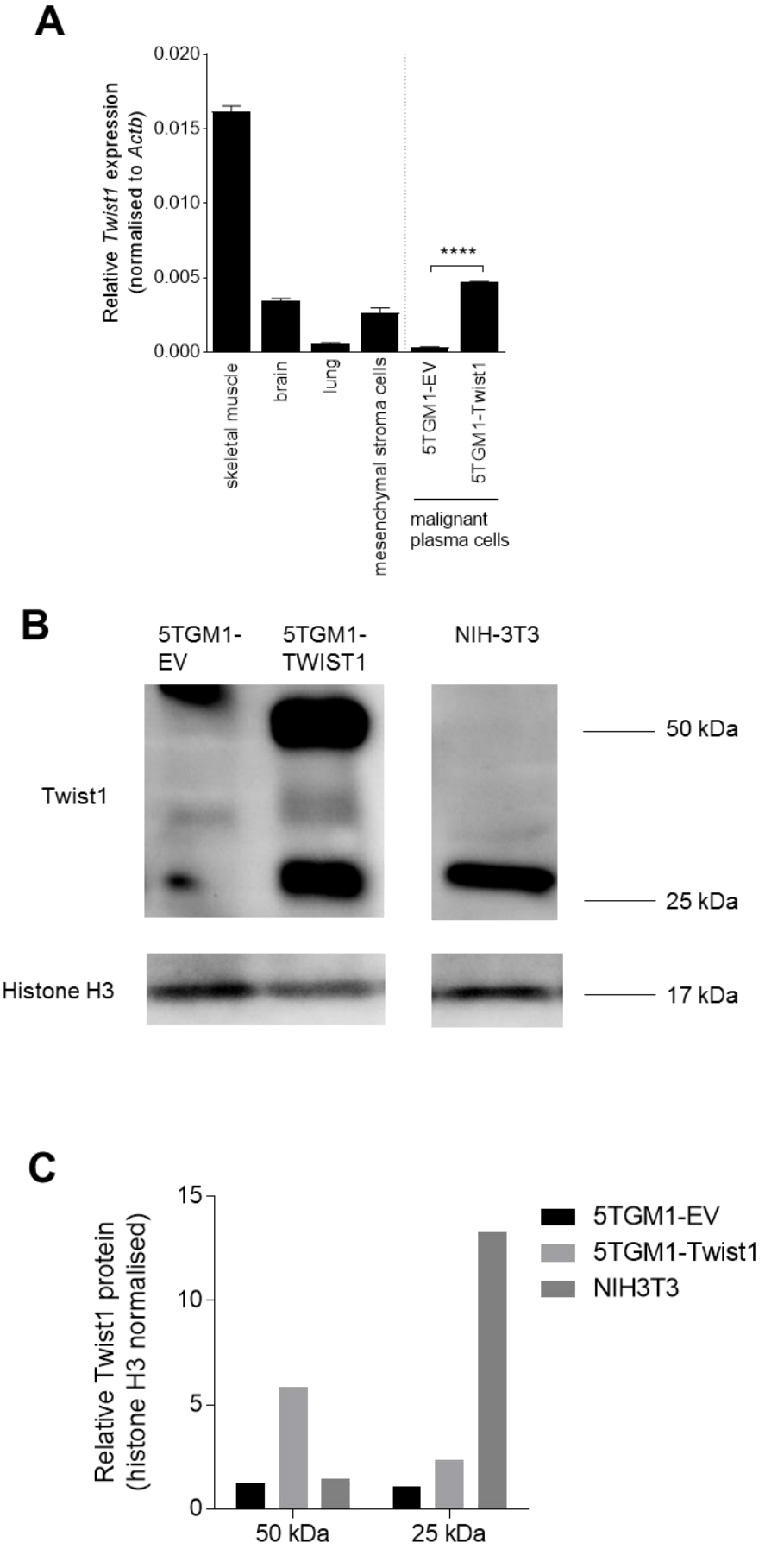
C57BL/KalwRij mice, kindly provided by Prof. Andrew Spencer (Monash University, Melbourne) were rederived, bred and housed at SAHMRI Bioresources facility (Adelaide). All procedures were performed with the approval of the SAHMRI and University of Adelaide Animal Ethics Committees. C57BL/KalwRij mice of 6- to 8-week of age were intravenously injected with 5×10^5 luciferase-expressing 5TGM1-EV or 5TGM1-TWIST1 cells (n=9 mice per group). Tumour development was monitored weekly by injecting firefly D-luciferin (150 mg/kg) intraperitoneally and imaged the Xenogen IVIS 100 system for bioluminescence signals as previously described.[20] Total tumour burden (total flux, photons/sec) was assessed using Living Image software. At the end of experiment, cardiac blood was collected for circulating tumour cells analysis and hind limbs were collected for histological analysis.

4.4 Results

4.4.1 TWIST1 expression in 5TGM1 cell line and C57BL/KaLwRij murine MM model.

The expression of murine *Twist1* was assessed in a range of C57BL/KaLwRij mouse-derived tissues using RT-qPCR. Consistent with previous findings[28, 29], *Twist1* was highly expressed in mesoderm-derived skeletal muscles and expressed at lower levels in endoderm-derived lung and ectoderm-derived brain tissue. To examine the role of *Twist1* expression in MM development, *Twist1* was overexpressed in the luciferase-tagged murine MM OC line, 5TGM1.[20, 24, 25] While *Twist1* expression was undetectable by RT-qPCR in the 5TGM1-EV cells, the levels of *Twist1* expression in the 5TGM1-TWIST1 cells was comparable with the *Twist1* levels observed mesenchymal stroma cells (Figure 4.1A). Western blotting was used to assess Twist1 protein expression in nuclear and cytoplasmic protein fractions of both 5TGM1-EV and 5TGM1-TWIST1. As seen in Figure 4.1B, elevated TWIST1 protein expression was observed in the nuclear fraction of 5TGM1-TWIST1 compared to the low endogenous level of TWIST1 in 5TGM1-EV cells (Figure 4.1B). In addition to the predicted 25 kDa monomeric TWIST1, additional 50kDa species was observed. (twice the predicted molecular mass of mTWIST1). As previously described[30], the larger molecular species most likely corresponds to the TWIST1/TWIST1 homodimer, and are due to ectopic expression of *Twist1* in 5TGM1- TWIST1 cells. As seen in Figure 4.1B, a singular 25kDa species was

Figure 4.1 Twist1 expression in C57BL/KaLwRijHsd mouse tissues and 5TGM1 cell lines. (A) TWIST1 expression is highly expressed in mesenchymal stromal cells and mesoderm-derived skeletal muscles and heart. TWIST1 was overexpressed in 5TGM1-TWIST1 compared with empty vector control 5TGM1-EV, to a comparable level with positive control C57BL/KaLwRij mouse-derived mesenchymal stromal cells, as shown by real-time polymerase chain reaction. Data was normalized to ACTB expression. Mean ± SD of triplicate. (B) Nuclear protein extracts (30 μg) of 5TGM1-EV, 5TGM1-TWIST1 and positive control NIH-3T3 cells were resolved on 12% SDS-PAGE gels and immunoblotted with the indicated antibodies. TWIST1 protein is overexpressed in 5TGM1-TWIST1 cells. (C) Protein levels of 25 kDa and 50 kDa bands were quantified by densitometry relative to the level of loading control histone H3.


A
Relative Twist1 expression (normalised to Actb)
0.020
0.015
0.010
0.005
0.000
****
skeletal muscle
brain
lung
mesenchymal stroma cells
5TGM1-EV
5TGM1-Twist1
malignant plasma cells
B
5TGM1-EV
5TGM1-TWIST1
NIH-3T3
Twist1
50 kDa
25 kDa
Histone H3
17 kDa
C
Relative Twist1 protein (histone H3 normalised)
15
10
5
0
50 kDa
25 kDa
5TGM1-EV
5TGM1-Twist1
NIH3T3

identified in the NIH-3T3 cell lysates which served as positive control for murine TWIST1 protein.

4.4.2 Transcriptome analysis of TWIST1-overexpressing 5TGM1 cells by RNA-sequencing

To investigate the transcriptome-wide effect of TWIST1 overexpression in MM cells, RNA-Seq was performed using polyA-isolated mRNA from 5TGM1-EV cells and 5TGM1-TWIST1 cells (n= 2 biological replicates per cell line). Using the Illumina NextSeq500, in excess of 29 million single-end reads were obtained per sample with average length of 75 bp. Approximately 93% of the reads were aligned to mm9 reference genome using Tophat2, and more than 60% were uniquely mapped to reference genes (supplementary Table S4.1). The percentage of mapped reads was similar between the two groups (5TGM1-EV, 93.7% ± 0.2; 5TGM1-TWIST1, 93.2% ± 0.1; mean ± range of duplicates), suggesting the absence of detectable biases in the sequenced data. The multidimensional scaling plots showed clear separation between 5TGM1-EV and 5TGM1-TWIST1 replicate libraries (supplementary Figure S4.1), indicating differences between transcriptomes due to TWIST1 overexpression.

As expected, overexpression of *TWIST1* in 5TGM1 cells resulted in a 12 log2 fold upregulation of *Twist1* in the 5TGM1-TWIST1 cells compared to the 5TGM1-EV cells. Compared to 5TGM1-EV cells, 149 genes were differentially regulated in the 5TGM1-TWIST1 cells, in which 88 (59.1%) genes were upregulated and 61 (40.9%) were downregulated (log2 fold change greater than 1 and p-value < 0.05, supplementary Table S4.1). Of these, 19 genes did not code for protein and 2 genes were of unknown function. The overall changes in gene expression associated with TWIST1-overexpression are presented in the heatmap (Figure 4.2A).

To determine the potential biological significance of *Twist1*, gene ontology (GO) analysis was performed on the significantly differentially expressed genes. A total of 17 and 4 GO annotations clusters were enriched (enrichment score > 1) in TWIST1-upregulated and -downregulated genes, respectively (supplementary Table S4.2, Table S4.3 and Figure 4.2B). As seen in Table 4.1, among the TWIST1-upregulated genes, genes involved in regulating cytokines secretion (*Clu, Tnfrsf8, Htr2b, Nlrp10,* and *Scamp5*), cell migration (*Ptprk, Ccr5, Sema4g,* and *Cxcl10*), proteolysis (*Mmp10* and

Table 4.1. List of genes differentially regulated in TWIST1-overexpressing 5TGM1 cells.

Symbol	Gene name	chromosome	Log2 Fold Change	P value
Cell migration				
Ptprk	Protein tyrosine phosphatase, receptor type, K	10	13.122	6.15E-100
Ccr5	Chemokine (C-C Motif) receptor 5		1.463	1.29E-02
Ccrl2	Chemokine (C-C motif) receptor-like 2	9	1.437	5.22E-03
Cxcl10	Chemokine (C-X-C motif) ligand 10	5	1.372	9.22E-03
Sema4g	Semaphorin 4G	19	1.193	1.62E-02
Proteolysis				
Mmp13	Matrix metallopeptidase 13	9	1.386	3.67E-02
Mmp10	Matrix metallopeptidase 10	9	1.658	2.79E-16
Actin cytoskeleton organisation				
Rac3	RAS-related C3 botulinum substrate 3	11	1.491	2.97E-03
Ccdc155	Coiled-coil domain containing 155	7	1.349	1.88E-02
Mybpc3	Myosin binding protein C, cardiac	2	1.624	2.24E-02
Xirp2	Xin actin-binding repeat containing 2	2	2.171	6.01E-03
Regulates cytokines secretion				
Clu	Clusterin	14	1.123	6.17E-03
Tnfrsf8	Tumor necrosis factor receptor superfamily, member 8	4	2.507	1.88E-03
Htr2b	5-hydroxytryptamine (serotonin) receptor 2B	1	2.175	1.52E-02
Nlrp10	NLR family, pyrin domain containing 10	7	1.307	2.68E-03
Scamp5	Secretory carrier membrane protein 5	9	1.015	9.78E-03
Positive regulation of metabolic processes				
Vdr	Vitamin D receptor	15	2.272	1.51E-04
Erbb3	Erb-B2 receptor tyrosine kinase 3	10	1.069	5.02E-03
Igf2	Insulin growth factor 2	7	2.081	1.87E-02
Cela1	Chymotrypsin Like Elastase Family Member 1	15	1.603	2.58E-02

Figure 4.2 Differentially expressed genes in TWIST1-overexpressing 5TGM1 cells identified from RNA-Seq and GO terms annotation clustering analysis. (A) Heatmap showed log2 cpm values and clustering for 97 genes differentially expressed in 5TGM1-EV and 5TGM1-TWIST1 cells (absolute fold change > 2, p-value less than 0.05) by RNA sequencing. Each column represents a biological replicate of 5TGM1-EV or 5TGM1-TWIST1 cells. Upregulated genes were coloured in red while downregulated genes were coloured in blue. (B) DAVID Functional Annotation Clustering analysis of upregulated genes upon *Twist1* overexpression in 5TGM1 cells. Annotation clusters with enrichment score greater than 1 was shown. (C) Venn diagram analysis showed that 3 genes (excluding TWIST1) were upregulated by TWIST1 in both 5TGM1-TWIST1 and WL2-TWIST1 cell lines. The details of overlapped genes were shown in Table 4.2.

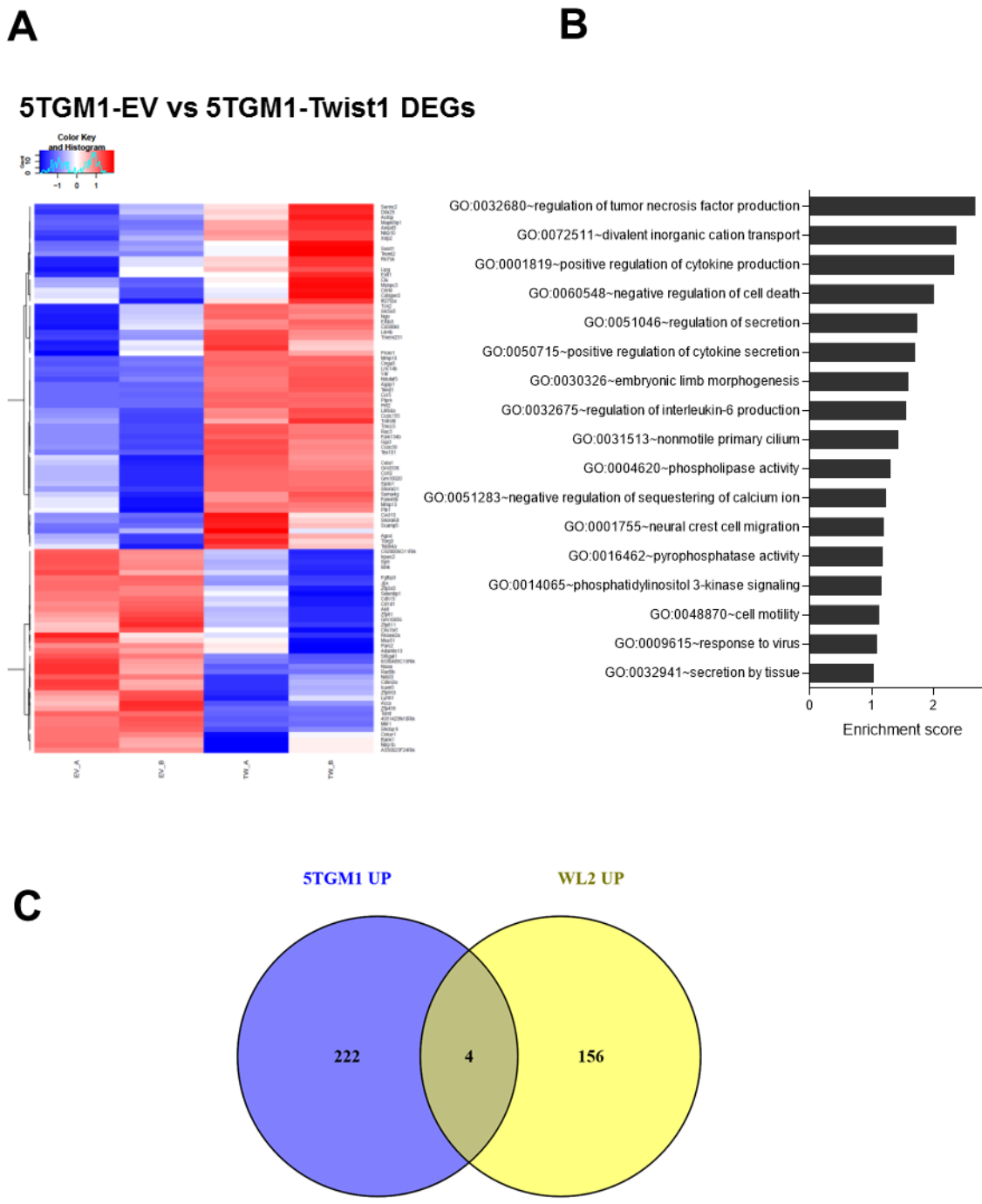


Table 4.2. List of 3 commonly upregulated genes in Twist1 overexpressing 5TGM1 and WL2 cells.

Gene symbol (human)	mean log2 FC WL2-EV vs WL2-TWIST1	FDR	mean log2 FC 5TGM1-EV vs 5TGM1-TWIST1	p-value
TWIST1	7.233	2.46E-16	12.179	7.04E-51
SERINC2	0.569	7.40E-03	1.073	4.37E-03
RRAGD	0.378	3.33E-02	0.798	2.47E-02
GAS7	0.390	1.73E-03	0.523	2.39E-02

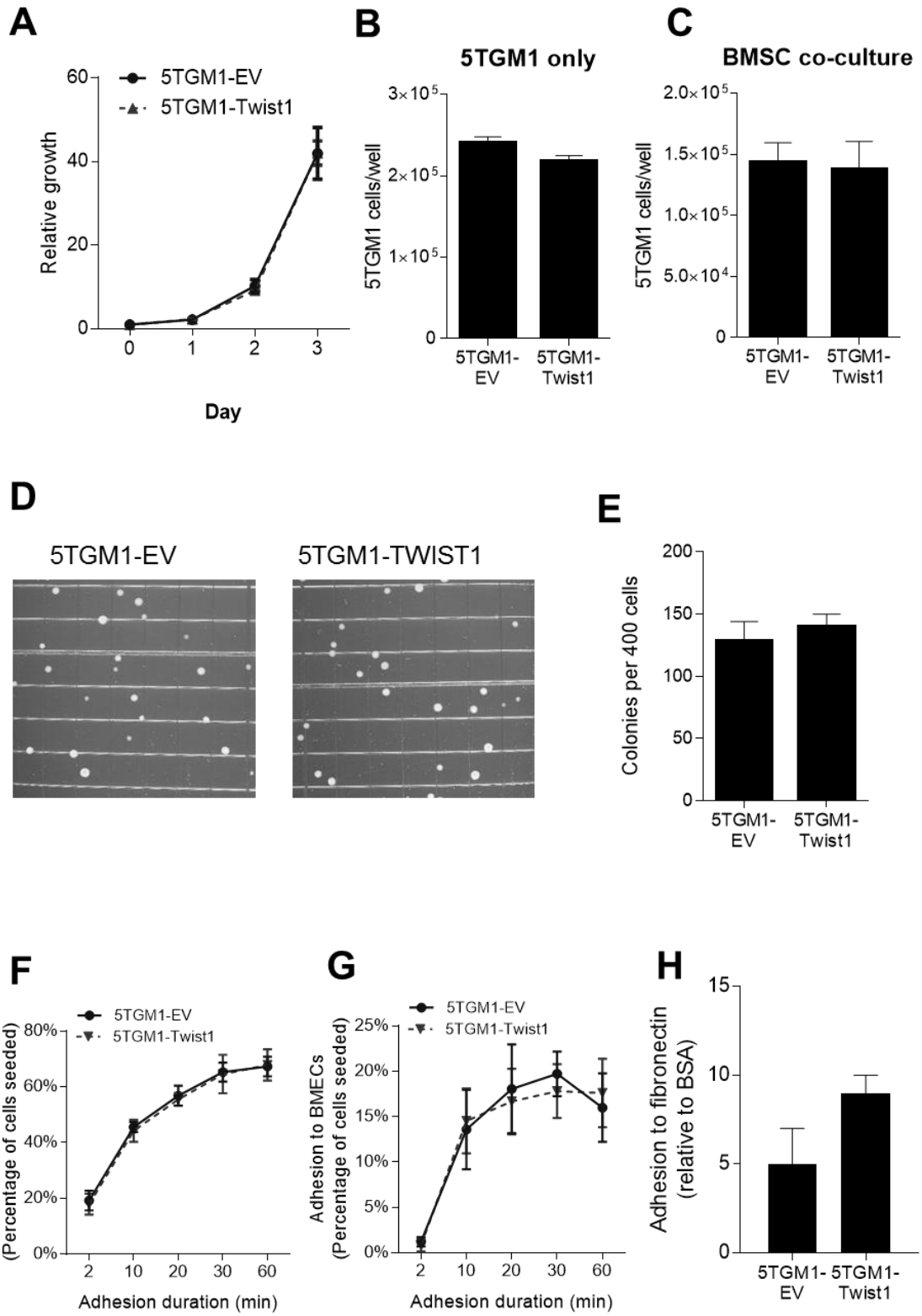
Mmp13) and actin cytoskeleton organisation (*Rac3, Ccdc155*, *Mybpc3* and *Xirp2*) were identified. Statistically significant genes of greatest fold change were validated using RT-qPCR. Consistent with RNA-seq analysis, RT-qPCR data showed that *PTPRK* was significantly upregulated ($p = 0.006$) while *TMEM29* was significantly downregulated ($p = 0.015$) in 5TGM1-Twist1 cells (supplementary Figure S4.2).

To determine whether the putative targets of TWIST1 in murine MM cells were similar to those targets of TWIST1 in the human WL2 cell line (see Chapter 3, Section 3.4.3, Figure 3.2 and supplementary Table S3.1), the list of DEGs were compared to the DEGs identified in TWIST1-overexpressing HMCL, WL2. Only 3 genes (*SERINC2*, *RRAGD*, and *GAS7*) were identified as being differentially upregulated in both TWIST1-overexpressing murine 5TGM1 and human WL2 MM cell lines (Figure 4.2C and Table 4.2) and no overlapping downregulated genes were identified (data not shown).

4.4.3 5TGM1 cell growth and adhesion to both BMSC and BMEC *in vitro* are unaffected by Twist1 overexpression.

To determine whether TWIST1 expression in 5TGM1 cells promotes autocrine signals/cytokines that stimulate mitogenesis, WST-1 proliferation assays were performed. As seen in Figure 4.3A, basal proliferation of 5TGM1-TWIST1 cells was not statistically different from that of 5TGM1-EV cells, over a three-day time course ($p = 0.919$, two-way ANOVA with Sidak's multiple comparisons test,). To assess the effects of TWIST1 expression on MM PC proliferation in the context of BM stromal cells, MM PCs were co-cultured with C57BL6/KaLwRij BMSC. As shown in Figure 4.3B and C, after three days of co-culture, 5TGM1-EV and 5TGM1-TWIST1 cell numbers were not significantly different (as assessed bioluminescence imaging), either in mono-culture or in co-culture with BMSC ($p = 0.89$ and 0.86, respectively; unpaired t-test). To determine the effect of *Twist1* on clonogenic capacity, 5TGM1-EV and 5TGM1-TWIST1 cells were cultured in methylcellulose. The number of colonies formed by 5TGM1-EV (128 ± 32 colonies) and 5TGM1-TWIST1 cells (139 ± 21 colonies) were not significantly different (Figure 4.3D) and of comparable size (Figure 4.3E). In addition, TWIST1 overexpression had no effect on 5TGM1 cell adhesion to BMSC and BMEC ($p = 0.86$ and 0.95 respectively, two-way ANOVA with Sidak's

Figure 4.3 5TGM1 cell growth, clonogenicity and adhesion to both BMSC and BMEC *in vitro* are unaffected by Twist1 overexpression. (A) Basal proliferation of 5TGM1 cells *in vitro* is not affected by TWIST1 expression after 3 days as assessed by WST-1 assay. Graph depict mean ± SEM of three independent experiments. TWIST1 overexpression has no effect on proliferation of 5TGM1 cells cultured with (B) or without (C) 57BL/KaLwRij mouse-derived bone marrow stromal cells as assessed by BLI after 3 days. Data represents the mean ± SEM of three independent experiments. (D) Representative photos of colonies formed by 5TGM1-EV and 5TGM1-TWIST1 cells. (E) TWIST1 overexpression in 5TGM1 does not affect the number of colonies formed in methylcellulose culture. Graph depicts mean ± range of a representative experiment in duplicate. 5TGM1-EV and 5TGM1-TWIST1 cells were adhered to monolayers of C57BL/KaLwRij BMSCs (F) and BMECs (G) for indicated duration. The degree of adhesion of 5TGM1-TWIST1 cells is no different from that of seen with 5TGM1-EV cells. Cell adhesion was quantitated using bioluminescence signal relative to cells seeded. Data represents the mean ± SEM of four and three independent experiments in triplicate, respectively. (H) 5TGM1 cells (1×10^5) were added to fibronectin-coated 96-wells and allowed to adhere for 2 hours at 37°C in serum free media. Data was presented in relative to BSA adhesion. Graph depict mean ± range of two independent experiments in triplicates.


A
B
C
5TGM1 only
BMSC co-culture
5TGM1-EV
5TGM1-Twist1
Relative growth
Day
5TGM1 cells/well
D
5TGM1-EV
5TGM1-TWIST1
E
Colonies per 400 cells
F
G
H
Adhesion to BMSCs (Percentage of cells seeded)
Adhesion to BMECs (Percentage of cells seeded)
Adhesion duration (min)
Adhesion to fibronectin (relative to BSA)

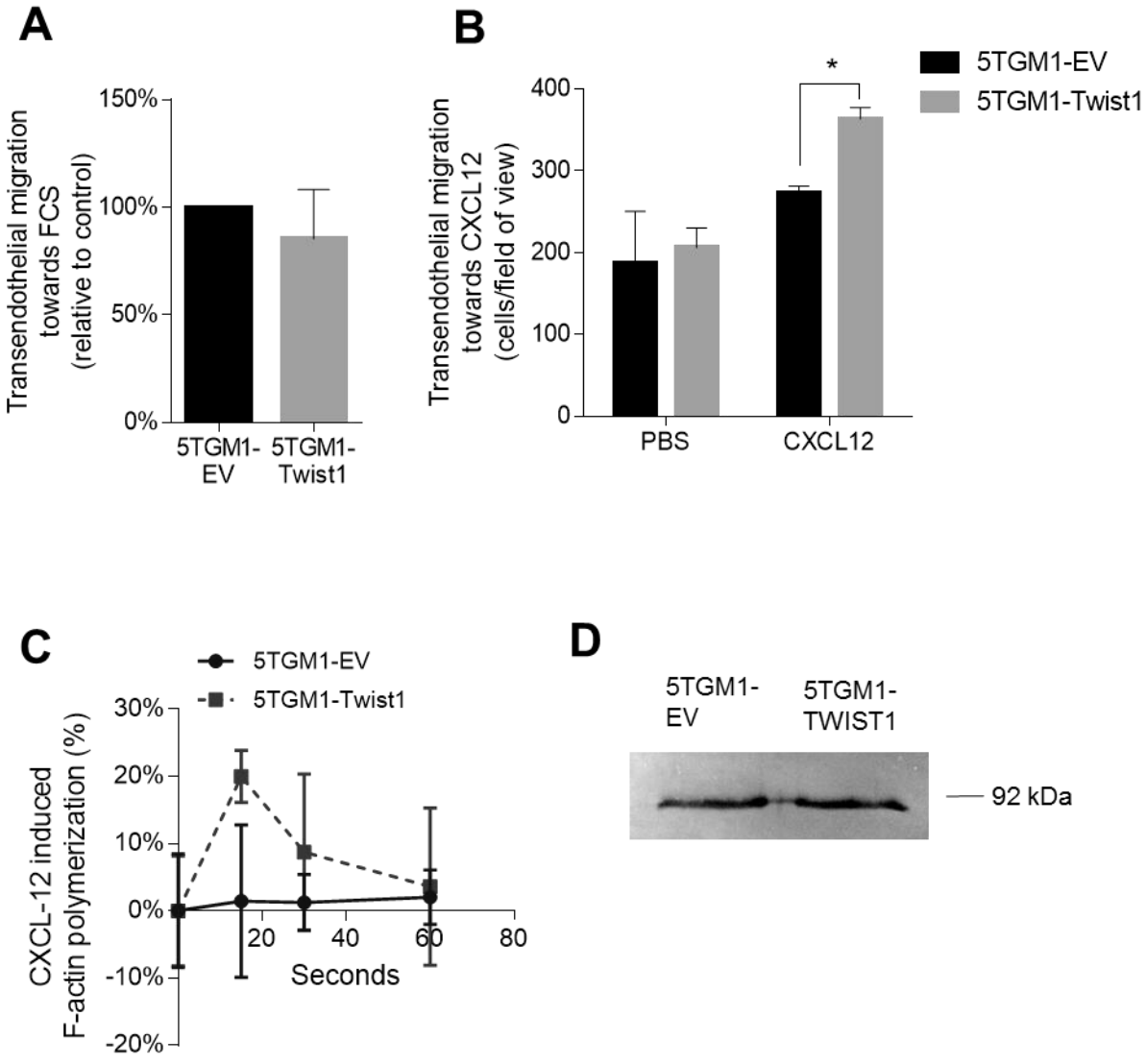
multiple comparisons test, Figure 4.3F and G), and adhesion to fibronectin (p = 0.215, unpaired t-test, Figure 4.3H).

4.4.4 TWIST1 expression increases 5TGM1 cell migration towards CXCL12 and actin polymerisation in 5TGM1 cells.

To determine whether TWIST1-regulated gene expressions affected 5TGM1 cell motility, transendothelial migration assays were performed. While no significant difference in migration of 5TGM1-EV cells and 5TGM1-TWIST1 cells was observed in response to a FCS concentration gradient (p = 0.3315, Figure 4.4A), TWIST1 expression induced a 33% increase in 5TGM1 cell migration towards CXCL12, a potent chemotactic cytokine, when compared to the vector control cells (p = 0.0047, two-tailed student's t-test). Furthermore, the effect of TWIST1 expression on F-actin remodeling in response to CXCL12 was examined. As shown in Figure 4.4C, while statistical significance was not reached (p = 0.0983, two-way ANOVA with Sidak's multiple comparisons test), 5TGM1-TWIST1 cells showed elevated actin polymerization compared with vector control cells. In the 5TGM1-TWIST1 cells, CXCL12-induced actin-polymerisation peaked at 15 seconds post-stimulation (20 ± 3.9%), while 5TGM1 cells showed limited changes in F-actin remodeling in response to CXCL12 stimulation (1.44 ± 11.3%).

Previous studies have highlighted a role for proetolytic enzymes, such as members of the matrix metalloproteinase family, in the trafficking of MM cells in and out of the BM.[31, 32] MM PC-derived proteolytic enzymes help degrade the subendothelial basement membrane proteins, thereby enabling invasion across basement membrane and interstitial tissues. As detailed above (section 4.4.2, Table 4.1), transcriptomic analysis of 5TGM1-TWIST1 cells revealed that *Twist1* expression in 5TGM1 cells leads to elevated transcript levels of MMP10 and MMP13. To determine differential secretion of MMPs regulated by TWIST1, conditioned media from 5TGM1-EV and 5TGM1-TWIST1 were analysed for the presence of MMP activity by gelatin zymography. As shown in Figure 4.4D, no differences were observed in gelatinolytic activity corresponding to 92 kDa MMP9 between 5TGM1-EV and 5TGM1-TWIST cells. The enzymatic activities of 72 kDa MMP2 and 60 kDa MMP13, as seen in previous study[33] were not detected in our study.

Figure 4.4 *Twist1* expression promotes 5TGM1 cell migration towards CXCL12. (A) Migration of 5TGM1 cells (1×10^5) through 8 μm transwells with bone marrow endothelial cells (BMECs) in response to FCS gradient were assessed. Cells migrated into lower chamber after 4 hours was photographed using inverted microscope and quantitated using ImageJ software. Data was presented as percentage of migrated cells in relative to non-stimulant control. Graphs depict mean ± SEM of three independent experiments in quadruplicate. (B) Transendothelial migration of 5TGM1 in response to CXLC12 (25 μM) was assessed as described in (A). Graphs depict mean ± SEM of three independent experiments in quadruplicate. (C) *Twist* expression enhanced CXCL12-induced F-actin polymerisation in 5TGM1 cells. Following CXCL12 stimulation (25 μM) for indicated duration, 5TGM1 cells were fixed in 2% paraformaldehyde for 15 min and permeabilised in 2% saponin. F-actin polymerisation was quantitated by flow cytometer after staining the cells with phalloidin-A680 for 30 min on ice in the dark. Data was represented as percentage change in MFI (mean ± SD of a representative experiment in triplicate). (D) MMP secretion regulated by TWIST1. Supernatants of 5TGM1-EV and 5TGM1-TWIST1 cells were analysed for gelatinase activity 48 h after incubation in serum-free medium. A 92 kDa band corresponding to MMP9 activity was detected. Gelatin zymogram from a representative of n = 3 experiments is shown.


A
Transendothelial migration towards FCS (relative to control)
150%
100%
50%
0%
5TGM1-EV
5TGM1-Twist1
B
Transendothelial migration towards CXCL12 (cells/field of view)
400
300
200
100
0
*
PBS
CXCL12
5TGM1-EV
5TGM1-Twist1
C
5TGM1-EV
5TGM1-Twist1
CXCL-12 induced F-actin polymerization (%)
30%
20%
10%
0%
-10%
-20%
20
40
60
80
Seconds
D
5TGM1-EV
5TGM1-TWIST1
92 kDa

4.4.5 TWIST1 promotes tumour development *in vivo*.

To investigate the role of TWIST1 expression in MM cell disease progression *in vivo*, the 5TGM1-EV or 5TGM1-TWIST1 cells were injected intravenously into C57BL/KaLwRij model of systemic MM ($n = 9$ mice/group). Tumour development was monitored over 4 weeks by *in vivo* bioluminescent imaging. 5TGM1-EV and 5TGM1-TWIST1 cells could be detected in tibiae, fermora and vertebral bodies 2 weeks after tumour inoculation (Figure 4.5A and Figure 4.5C). Ventral and dorsal imaging of 5TGM1-TWIST1- and 5TGM1-EV-bearing mice revealed no significant difference in tumour burden at week 2 and 3. However, by week 4, the tumour burden in 5TGM1-TWIST1-bearing mice was 34% and 47% higher, from ventral and dorsal imaging respectively ($p = 0.042$ and 0.00073, two-way ANOVA with Sidak's multiple comparisons test), when compared with mice inoculated with 5TGM1-EV cells, suggesting a role for TWIST1 in promoting tumour growth at the later stages of MM disease progression *in vivo*.

4.4.6 TWIST1 promote extramedullary tumour metastasis and growth

MM patients with advanced or relapsed disease often show evidence of increased extramedullary involvement and elevated numbers of circulating tumour cells.[34, 35] Spleenic infiltration and tumour establishment has been previously reported in the aggressive 5TGM1-C57BL/KaLwRij model 4 weeks after tumour inoculation.[12, 13, 17] To determine whether splenic infiltration also contributed to the difference in tumour burden observed between two groups of mice, the left and right flanks of the mice were imaged 4 weeks after tumour inoculation ($n = 5$ mice/group; Figure 4.6A). As shown in Figure 4.6B, while the bioluminescence signals from the right flank was not significantly different between two groups of mice, signals from the left flank of 5TGM1-TWIST1-bearing mice were 160% higher than that observed in 5TGM1-EV-bearing mice ($p = 0.002$, unpaired t-test), suggesting that TWIST1 promotes BM-independent growth of 5TGM1 cells.

Additionally, to investigate whether TWIST1 promotes 5TGM1 dissemination, the prevalence of circulating 5TGM1 cells was analysed using flow cytometry (Figure 4.6D). Cardiac blood was collected from 5TGM1-EV- and 5TGM1-TWIST1-bearing mice at the end of the experiment (week 4). While statistical significance was not reached, 5TGM1-TWIST1-bearing mice exhibited elevated numbers of circulating

Figure 4.5 *Twist1* overexpression in 5TGM1 cells promotes tumour development in C57BL/KaLwRij mice. Total growth was monitored by bioluminescence imaging at 2, 3 and 4 weeks following intravenous tumuor cell inoculation. Bioluminescence images of 5 representative mice from ventral (A) and dorsal (C) views are shown. For week 2, the colour scale was set at a radiance (photons/second/cm^2/steradian) of 10^3 to 10^4. For week 3 and 4, the scale was set at a radiance of 5×10^4 to 2.5×10^6. A significant increase of tumour burden was observed in the mice bearing 5TGM1-TWIST1 cells (n = 9) compared with empty vector control (n = 9) from both (B) ventral and (D) dorsal imaging. Graph depict mean ± SEM of 9 mice per group.

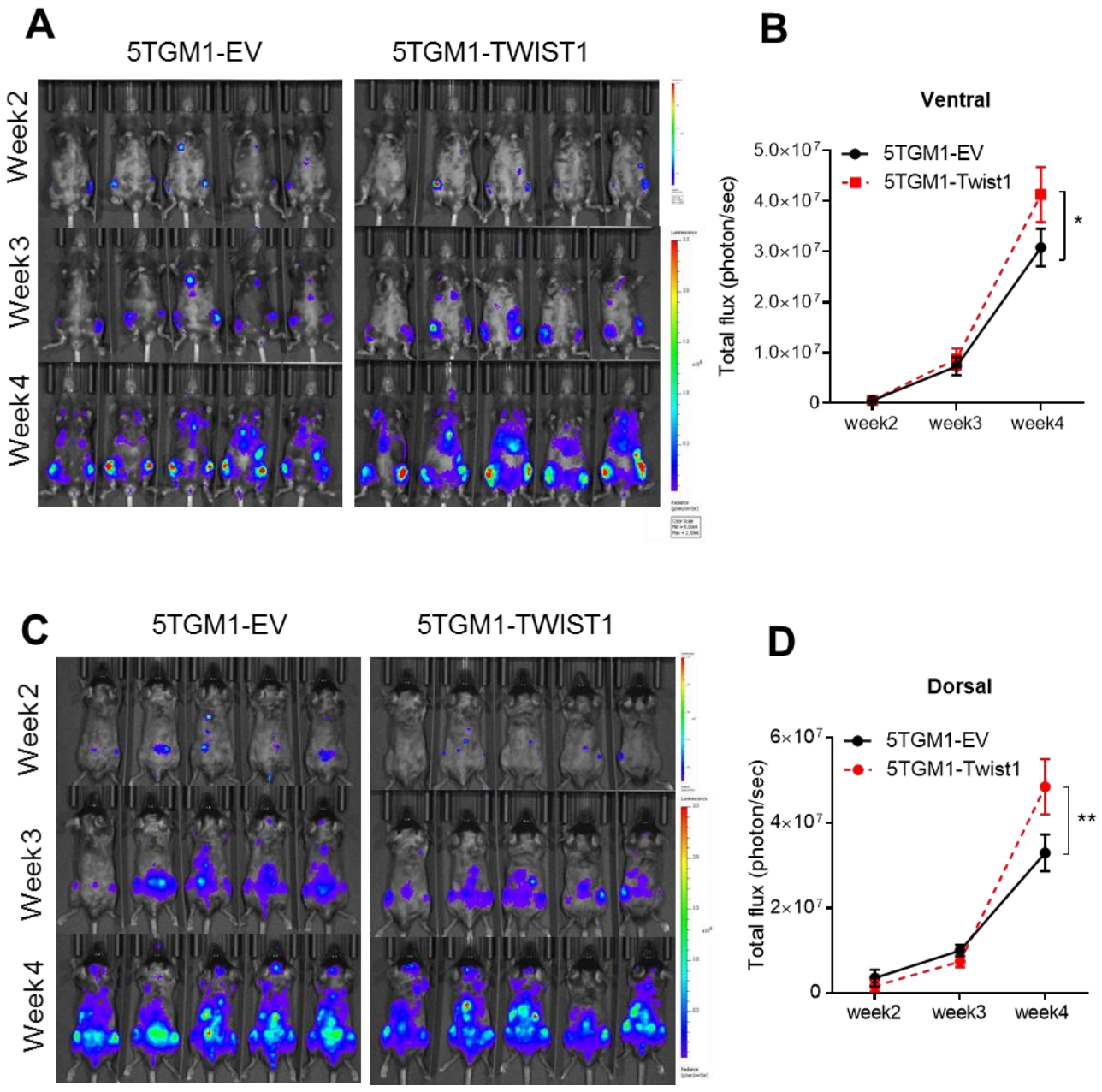
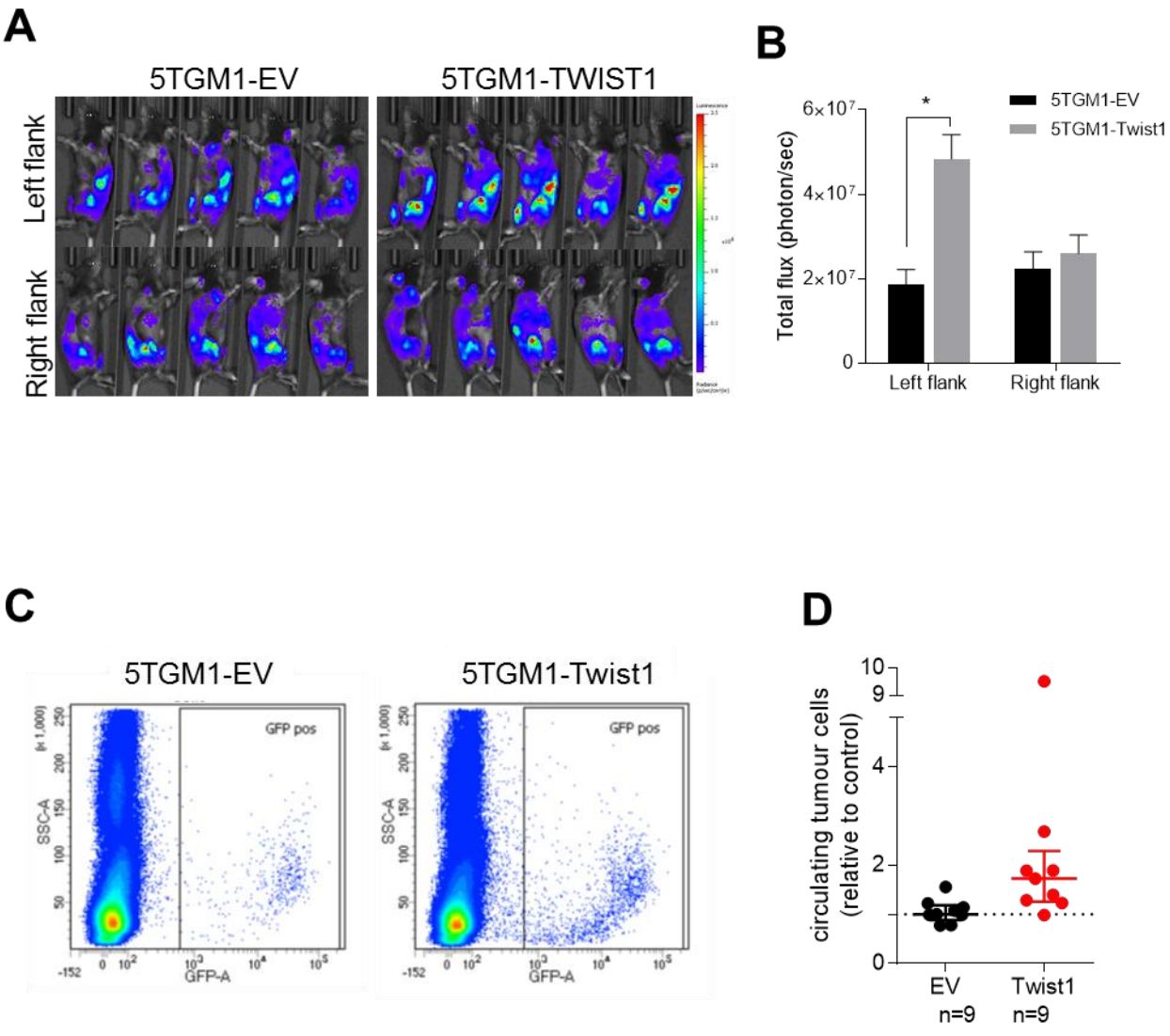

A
5TGM1-EV
5TGM1-TWIST1
Week2
Week3
Week4
B
Ventral
5TGM1-EV
5TGM1-Twist1
Total flux (photon/sec)
5.0×10^7
4.0×10^7
3.0×10^7
2.0×10^7
1.0×10^7
0
week2
week3
week4
*
C
5TGM1-EV
5TGM1-TWIST1
Week2
Week3
Week4
D
Dorsal
5TGM1-EV
5TGM1-Twist1
Total flux (photon/sec)
6×10^7
4×10^7
2×10^7
0
week2
week3
week4
**

Figure 4.6 *Twist1* expression promotes extramedullary growth via peripheral blood in vivo. (A) Bioluminescence images of 5 representative mice from left and right lateral views at 4 weeks after tumour inoculation are shown. Color scale was set at a radiance of 5×10^4 to 2.5×10^6. (B) A significant increase of tumour burden was observed on the left flank of mice bearing 5TGM1-TWIST1 cells (n = 5) compared with empty vector control (n = 5). Graph depict mean ± SEM of 5 mice per group. (C) Flow cytometry analysis of circulating GFP+ 5TGM1 cells detected in the cardiac blood from 5TGM1-EV- and 5TGM1-TWIST1-bearing mice. Representative plots of side scatter vs GFP with the gating on GFP+ cells are shown (D) Scatter plot depicts percentage of circulating tumour cells relative to vector control group. Median and interquartile ranges are shown.


A
5TGM1-EV
5TGM1-TWIST1
Left flank
Right flank
B
Total flux (photon/sec)
6×10^7
4×10^7
2×10^7
0
*
5TGM1-EV
5TGM1-Twist1
Left flank
Right flank
C
5TGM1-EV
5TGM1-Twist1
GFP pos
SSC-A
GFP-A
D
circulating tumour cells
(relative to control)
10
9
4
2
0
EV
n=9
Twist1
n=9

tumour cells, compared with that observed in 5TGM1-EV-bearing mice ($p = 0.1427$, unpaired t-test with Welch's correction; Figure 4.6E).

4.4.7 Splenic cell-derived soluble molecules stimulate proliferation of TWIST1-overexpressing 5TGM1 cells

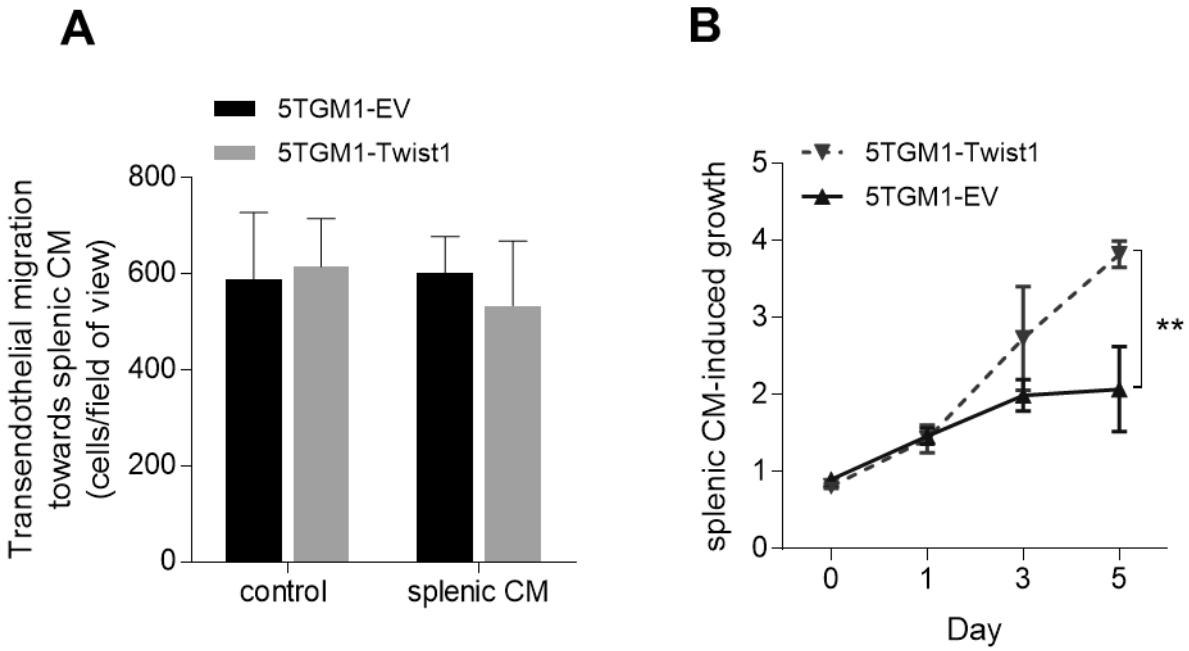
To determine whether the increased splenic tumour growth in the 5TGM1-Twist bearing mice was due to an increased migration capacity, transendothelial migration assays towards splenic cell explant culture-derived conditioned media (CM) were performed. As shown in Figure 4.7A, no significant difference in migration was observed between 5TGM1-EV cells and 5TGM1-TWIST1 cells in response to spleen cell CM ($p = 0.679$, unpaired t-test). Interestingly, the rate of 5TGM1-TWIST1 cell proliferation in the presence of splenic cells CM was 84% higher when compared with 5TGM1-EV cells after 5 days of culture, as assessed by WST-1 assay ($p = 0.0071$, two-way ANOVA with Sidak's multiple comparisons test Figure 4.6B). These data suggest that TWIST1 expression promotes 5TGM1 cell proliferation in response to soluble molecules secreted by spleen cells. While it remains to be confirmed, 5TGM1-TWIST cells exhibited elevated levels of CCR5 and ErbB3 (Table 4.1) suggesting that MIP1-α and IFNα, which are present in spleen cell condition media[41, 42], were likely to mediate enhanced proliferation observed in the 5TGM1-TWIST1 cells.[36, 37]

4.5 Discussion

Overexpression of Twist1 has been detected in various cancers and is involved in driving the invasion and metastasis of solid tumours.[21-23] We previously showed that overexpression of TWIST1 in human MM cells promotes cell motility *in vitro* (Chapter 3). In an effort to determine the role of TWIST1 in MM disease development *in vivo*, 5TGM1 cells that were genetically modified to overexpress TWIST1 and were injected intravenously into C57BL/KaLwRij mice. In addition, the biological effects of TWIST1 overexpression in 5TGM1 cells were also assessed *in vitro*.

Following intravenous inoculation, 5TGM1 cells were found to home to and establish within the BM leading to the formation of multiple tumours throughout the skeleton. Notably, TWIST1 overexpression in 5TGM1 cells also significantly increased overall tumour burden in C57BL/KaLwRij mice after 4 weeks. Previous studies have

Figure 4.7 *Twist1* expression increases cell proliferation in response to spleen cells CM. (A) Transendothelial migration of 5TGM1 cells (1×10^5) in response to C57BL/KaLwRij-derived spleen cells CM were assessed. Cells migrated into lower chamber after 4 hours was photographed using inverted microscope and quantitated using ImageJ software. Graphs depicts mean ± SEM of two independent experiments in quadruplicate. (B) *Twist1* expression promotes proliferation of 5TGM1 cells when cultured in C57BL/KaLwRij-derived spleen cells CM after 5 days, as assessed by WST-1 assay. Data were normalised to relative growth in RPMI-1640 media. Graph depict mean ± SEM of three independent experiments in triplicates.


A
5TGM1-EV
5TGM1-Twist1
Transendothelial migration towards splenic CM (cells/field of view)
800
600
400
200
0
control
splenic CM
B
5TGM1-Twist1
5TGM1-EV
splenic CM-induced growth
5
4
3
2
1
0
**
0
1
3
5
Day

shown that ectopic expression of TWIST1 in mammary carcinomas accelerated tumour establishment *in vivo*.[38, 39] The increased tumour burden observed at week 4 in mice receiving 5TGM1-TWIST1 cell, highlights a functional role for TWIST1 in more advanced stages of MM disease progression rather than affecting tumour establishment at the onset of MM disease. Furthermore, *in vitro* studies showed that TWIST1 promotes migration of 5TGM1 cells towards CXCL12, but did not affect the colony forming capability and proliferation of 5TGM1 cells under basal growth condition or in co-culture with BM stromal cells. Taken together, these results suggest that TWIST1 contributes to MM tumour development by promoting dissemination of 5TGM1 cells to distant, CXCL12-rich sites throughout the BM, during MM disease progression.

More advanced cases of human MM disease are often associated with both an increase in the number of circulating tumour cells and an increased incidence of extramedullary disease and is associated with adverse prognostic features.[34, 35, 40, 41] In the present study, we showed that TWIST1 expression in 5TGM1 cells was associated with an increase in splenic tumour burden and an increase in the frequency of circulating tumour cells. Splenic infiltration is frequently observed in the 5T33/5TGM1-C57BL/KaLwRij model[12, 13, 17], and is associated with the murine spleens capacity to support hematopoiesis and thus its capacity to support MM growth. In contrast, hematopoiesis is restricted to the BM in adult human[42], and extramedullary myeloma development within the spleen is seen in aggressive, relapsed MM.[43] Notably, our data showed that TWIST1 expression in 5TGM1 enhances their capacity to respond to mitogenic factors found within C57BL/KaLwRij derived spleen cell-conditioned media, potentially by increasing the expression of cytokine receptors (Table 4.1). Furthermore, TWIST1-mediated extramedullary infiltration of MM was recently reported in a study by Yang and colleagues[44] who showed, using immunohistochemistry, that nuclear-localised TWIST1 expression was higher in extramedullary tumours compared with BM from MM patients with extramedullary disease.[44] Collectively, these observations highlight the need to perform further studies to identify the cytokines and their cognate receptors that promote Twist1-mediated extramedullary dissemination, survival and proliferation of MM plasma cells.

Transcriptomic analysis of 5TGM1-TWIST1 cells revealed a significant upregulation of genes involved in cell migration, proteolysis and actin cytoskeleton organisation;, processes which are crucial for MM cell trafficking during dissemination to distant sites and homing to appropriate niches.[45] Although there was little overlap

between TWIST1-regulated candidate genes in human and murine MM cells, there was evidence of an enrichment of genes that play roles in regulating cell migration. Previous studies have identified the chemokine CXCL12 and its cognate receptor CXCR4 as a key factors in mediating homing and dissemination of MM cells.[46] Inhibition of CXCR4[46] or CXCL12[47] function using AMD3100[46] or neutralizing antibodies were shown to decrease MM cell migration. In this present study, we showed that TWIST1 overexpression in 5TGM1 cells increased transendothelial migration towards CXCL12 and was associated with increased actin cytoskeleton polymerization. Previous studies have shown a positive correlation between TWIST1 and CXCR4 protein expression in head and neck squamous cell[48] and papillary thyroid carcinomas.[49] Interestingly, the CXCL12/CXCR4 axis has been shown to upregulate TWIST1 protein in MM[50] and glioblastoma.[51] Although our RNA-seq data indicated that Twist1 overexpression did not upregulate transcript level of CXCR4, the elevation in CCR5 may promote surface relocation of CXCR4 by forming heterodimer and subsequently activate signaling in response to CXCL12 stimulation, as seen in T cells.[52, 53]

In addition to the dynamic cytoskeletal remodeling, MM cell migration also relies on the degradation of basement membrane proteins, mediated by matrix metalloproteinases (MMP). MMPs are a family of zinc-dependent endopeptidases that are involved in the turnover of extracellular matrix components during normal physiological processes such as embryogenesis, wound healing and angiogenesis as well as during pathological process, such as osteoarthritis and tumour invasion.[54] Excessive expression of several MMPs, such as the gelatinases (MMP-2 and MMP-9) and the collagenases (MMP-1, MMP-8 and MMP-13) have been previously reported in MM cells from patients and in human and murine MM cell lines (reviewed in [32, 55]), and have been implicated in mediating angiogenesis[56, 57], homing[58, 59] and osteolysis[60] during MM development. In this present study, the expression of MMP-10 and MMP-13 were elevated in TWIST1-overexpressing 5TGM1 cells, suggesting that Twist1 may upregulate MMP expression in MM PCs. This finding is consistent with studies which show that MMP-1 and MMP-2 are upregulated by Twist1 in giant-cell tumour of the bone[61] and melanoma[62], respectively. Although differential proteolytic activity of MMP-9 was not detected between 5TGM1-EV and 5TGM1-TWIST1, TWIST1-mediated upregulation of MMPs may contribute to progression of MM disease by promoting bone resorption activity and angiogenesis.[63, 64]

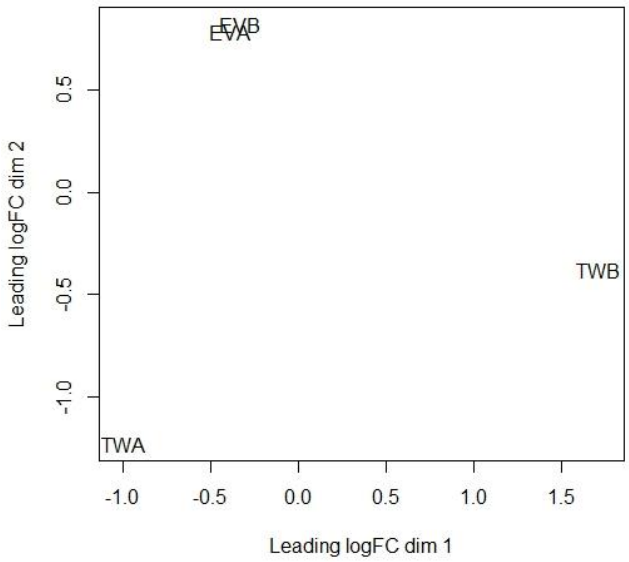
TWIST1 is known to promote epithelial-to-mesenchymal transition in carcinomas during tumour metastasis, by mediating the loss E-cadherin expression and the gain of N-cadherin expression.[23, 65] Although we (Chapter 1) and others[66] have identified a role for EMT in the dissemination of MM cells, overexpression of TWIST1 in 5TGM1 cells was not associated with upregulation of EMT-related genes. In contrast to the cadherin switch often seen in carcinomas expressing high level of TWIST1[67], overexpression of TWIST1 in 5TGM1 was found to decrease the expression of both E- and N- cadherin (supplementary Figure S4.3B). These findings suggest that TWIST1-mediated migration in 5TGM1 cell is independent of the cadherin switch, most likely owing to the hematopoietic origin of 5TGM1 MM cells. This finding is consistent with our previous study on human MM cell lines, which showed an absence of the cadherin switch upon TWIST1 overexpression or knockdown (Chapter 3, supplementary Figure S3.6B).

In summary, overexpression of TWIST1 enhances tumour development of MM *in vivo*, by promoting tumour cell dissemination. Moreover, elevated extramedullary proliferation of 5TGM1 cells suggest that TWIST1 contributes to aggressive MM disease progression by favoring tumour growth independent of BM microenvironment. Further investigations are warranted to determine the mechanism underlying BM-independence of MM cells growth associated with TWIST1 expression.

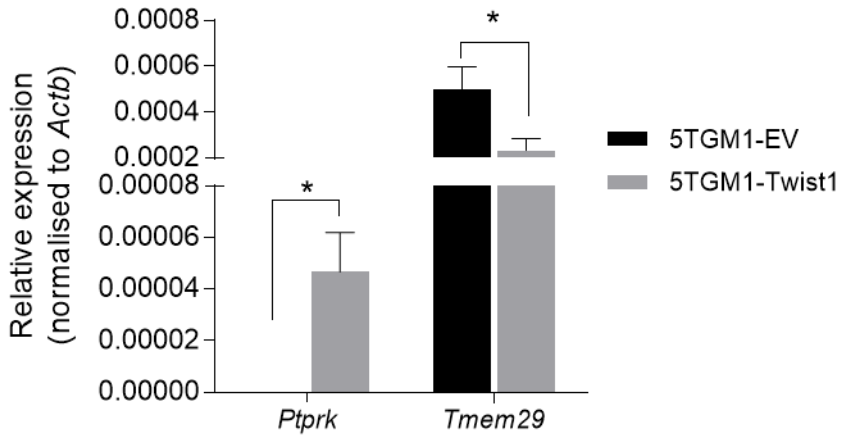
4.6 Supplementary Figures and Tables

Supplementary Table S4.1. Raw and mapped reads of each 5TGM1-EV or 5TGM1-TWIST1 sample from RNA-seq analysis.

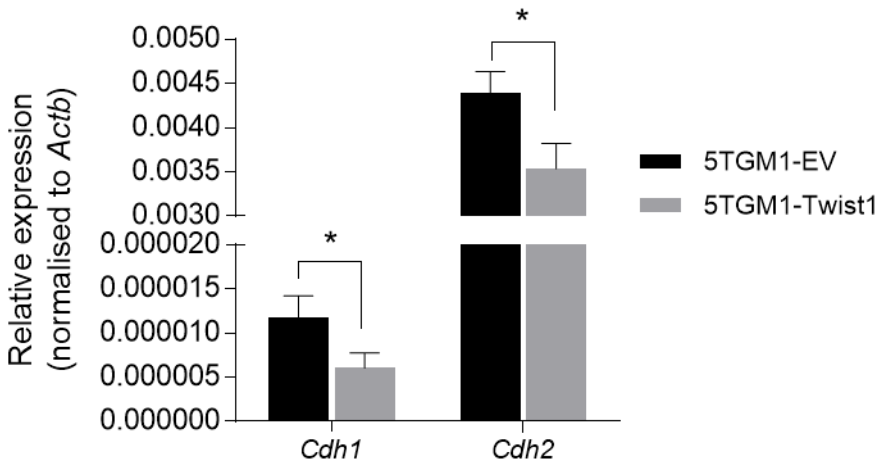
	Total reads	Mapped to genome		Unique gene mapping	
		Reads	Frequency	Reads	Frequency
5TGM1-EV_A	36975040	34597545	93.6%	22390850	60.6%
5TGM1-EV_B	38188847	35813501	93.8%	23373556	61.2%
5TGM1-TWIST1_A	31453201	29279785	93.1%	18772743	59.7%
5TGM1-TWIST1_B	42803908	39906083	93.2%	26380378	61.6%



supplementary Figure S4.1 Multidimensional scaling plot to evaluate libraries similarities. Graph was plotted using edgeR package in R to visualised similarity of 5TGM1-EV and 5TGM1-TWIST1 replicate samples. The root-mean-square of top 500 genes with largest absolute log2 fold change between two sample libraries (leading logFC) was calculated. Strong separation is observed amongst libraries from two groups in the second dimension.



supplementary Figure S4.2 Validation of RNA-seq data by RT-qPCR. Relative mRNA levels of representative DEG with greatest fold change (*Ptprk* and *Tmem39*) were upregulated and downregulated respectively in 5TGM1-TWIST1 cells compared with vector control cells. Data were normalised to reference gene *GAPDH* (mean ± SD of triplicates).



supplementary Figure S4.3 Modulation of *Twist1* expression does not generate cadherin switch in 5TGM1 cells. *CDH1* and *CDH2* mRNA levels in 5TGM1-EV and 5TGM1-TWIST1 cells were assessed by RT-qPCR and normalised to housekeeping gene *Actb*. Graph depict mean ± SD of triplicates from a representative experiment (* $p < 0.05$ unpaired t-test).

supplemetary Table S4.1. List of genes differentially regulated in TWIST1 overexpression in 5TGM1 cells.

GeneID	Symbol	chromosome	type_of_gene	log2 Fold Change	PValue
19272	PTPRK	10	protein-coding	13.122	6.15E-100
22160	TWIST1	12	protein-coding	12.179	7.04E-51
18630	PET2	X	protein-coding	9.648	2.28E-12
66270	FAM134B	15	protein-coding	5.638	5.49E-10
54646	PPP1R3F	X	protein-coding	3.831	2.56E-04
97908	HIST1H3G	13	protein-coding	3.495	1.98E-04
83671	SYTL2	7	protein-coding	3.403	2.38E-03
30927	SNAI3	8	protein-coding	3.354	1.19E-03
140795	P2RY14	3	protein-coding	3.349	6.27E-04
11981	ATP9A	2	protein-coding	3.331	3.80E-04
242700	IFNLR1	4	protein-coding	3.237	4.36E-03
56707	ZFP111	7	protein-coding	3.060	5.18E-04
1.01E+08	GM3336	8	protein-coding	2.926	4.77E-05
66528	SMIM5	11	protein-coding	2.920	8.85E-03
217305	CD300LD	11	protein-coding	2.664	2.64E-03
12829	COL4A4	1	protein-coding	2.636	1.14E-03
1.01E+08	GM19605	12	ncRNA	2.613	8.53E-03
56746	TEX101	7	protein-coding	2.570	2.63E-08
21941	TNFRSF8	4	protein-coding	2.507	1.88E-03
69206	2010016I18RIK	3	ncRNA	2.340	8.42E-03
232790	OSCAR	7	protein-coding	2.317	3.82E-03
74646	SPSB1	4	protein-coding	2.315	2.94E-03
319880	TMCC3	10	protein-coding	2.313	1.15E-06
22337	VDR	15	protein-coding	2.272	1.51E-04
331474	RGAG4	X	protein-coding	2.226	9.57E-03
15559	HTR2B	1	protein-coding	2.175	1.52E-02
241431	XIRP2	2	protein-coding	2.171	6.01E-03
30959	DDX25	9	protein-coding	2.161	3.48E-03
171530	UCN2	9	protein-coding	2.142	1.22E-02
16002	IGF2	7	protein-coding	2.081	1.87E-02
328573	4930407I10RIK	15	protein-coding	2.041	1.73E-02
347722	AGAP1	1	protein-coding	2.027	1.61E-03
51938	CCDC39	3	protein-coding	2.023	7.78E-03
19099	MAPK8IP1	2	protein-coding	1.999	3.19E-04
114479	SLC5A5	8	protein-coding	1.975	1.05E-02
75641	1700029I15RIK	2	protein-coding	1.973	3.35E-02
58182	PROKR1	6	protein-coding	1.960	3.08E-02
214922	SLC39A2	14	protein-coding	1.951	2.13E-02
229499	FCRL1	3	protein-coding	1.946	2.07E-02
64242	NGB	12	protein-coding	1.868	1.21E-02
269615	PLCH2	4	protein-coding	1.809	3.41E-02
212670	CATSPER2	2	protein-coding	1.804	4.25E-02
14727	LILR4B	10	protein-coding	1.790	5.58E-07
100342	FAM46B	4	protein-coding	1.774	2.74E-02
22153	TUBB4A	17	protein-coding	1.773	2.34E-02

GeneID	Symbol	chromosome	type_of_gene	log2 Fold Change	PValue
21378	TBRG3	15	ncRNA	1.761	2.63E-02
665270	PLB1	5	protein-coding	1.752	3.57E-02
634731	SUSD1	4	protein-coding	1.720	3.72E-02
54137	ACRBP	6	protein-coding	1.697	1.84E-03
272322	ARNTL2	6	protein-coding	1.679	3.52E-02
17384	MMP10	9	protein-coding	1.658	2.79E-16
1E+08	SNORA31	14	snoRNA	1.631	1.93E-02
12790	CNGA3	1	protein-coding	1.625	1.51E-02
17868	MYBPC3	2	protein-coding	1.624	2.24E-02
70957	STAMOS	2	ncRNA	1.616	4.04E-02
104370	SNORA68	8	snoRNA	1.611	4.69E-03
109901	CELA1	15	protein-coding	1.603	2.58E-02
353211	PRUNE2	19	protein-coding	1.557	4.82E-02
14728	LILRB4A	10	protein-coding	1.531	2.55E-06
170758	RAC3	11	protein-coding	1.491	2.97E-03
18514	PBX1	1	protein-coding	1.475	4.92E-02
76850	AGO4	4	protein-coding	1.466	2.53E-02
12774	CCR5	9	protein-coding	1.463	1.29E-02
54199	CCRL2	9	protein-coding	1.437	5.22E-03
17386	MMP13	9	protein-coding	1.386	3.67E-02
15945	CXCL10	5	protein-coding	1.372	9.22E-03
384619	CCDC155	7	protein-coding	1.349	1.88E-02
11717	AMPD3	7	protein-coding	1.324	3.71E-02
244202	NLRP10	7	protein-coding	1.307	2.68E-03
228094	CERKL	2	protein-coding	1.292	3.56E-02
432779	LRRC14B	13	protein-coding	1.283	2.82E-02
1E+08	GM10020	15	pseudo	1.280	2.32E-02
328833	TREML2	17	protein-coding	1.265	3.85E-02
19782	RMRP	4	ncRNA	1.229	4.84E-02
19817	RN7SK	9	snRNA	1.194	4.50E-03
26456	SEMA4G	19	protein-coding	1.193	1.62E-02
16891	LIPG	18	protein-coding	1.193	3.74E-02
76933	IFI27L2A	12	protein-coding	1.125	3.14E-02
56219	EXTL1	4	protein-coding	1.125	1.70E-02
12759	CLU	14	protein-coding	1.123	6.17E-03
66706	NDUFAF3	9	protein-coding	1.082	8.67E-03
230779	SERINC2	4	protein-coding	1.073	4.37E-03
13867	ERBB3	10	protein-coding	1.069	5.02E-03
110175	GGCT	6	protein-coding	1.058	6.35E-03
234740	TMEM231	8	protein-coding	1.051	2.66E-02
21452	TCN2	11	protein-coding	1.024	5.58E-03
66566	NTPCR	8	protein-coding	1.022	4.29E-02
56807	SCAMP5	9	protein-coding	1.015	9.78E-03
93726	RNASE2A	14	protein-coding	-1.021	4.18E-02
629016	ZFP953	13	protein-coding	-1.022	1.70E-03
83398	NDST3	3	protein-coding	-1.043	2.61E-02

GeneID	Symbol	chromosome	type_of_gene	log2 Fold Change	PValue
320295	C920006O11RIK	9	ncRNA	-1.048	1.40E-02
545471	ZFP345	2	protein-coding	-1.056	8.16E-03
194231	CNKSR1	4	protein-coding	-1.071	3.53E-02
245666	IQSEC2	X	protein-coding	-1.102	1.64E-02
73919	LYRM1	7	protein-coding	-1.106	4.08E-02
20440	ST6GAL1	16	protein-coding	-1.122	8.37E-10
75731	IDNK	13	protein-coding	-1.142	3.00E-02
67111	NAAA	5	protein-coding	-1.145	1.49E-04
224694	ZFP81	17	protein-coding	-1.204	1.15E-03
242248	BANK1	3	protein-coding	-1.207	3.87E-02
329470	ACCS	2	protein-coding	-1.221	3.33E-02
70252	JPX	X	ncRNA	-1.243	8.11E-03
381373	SP9	2	protein-coding	-1.261	1.09E-03
218311	ZFP455	13	protein-coding	-1.312	2.02E-02
72514	FGFBP3	19	protein-coding	-1.340	1.08E-03
110786	IGLC2	16	other	-1.348	4.76E-02
76224	6530409C15RIK	6	protein-coding	-1.389	2.68E-02
77810	A930015D03RIK	17	ncRNA	-1.484	4.92E-02
230577	PARS2	4	protein-coding	-1.504	3.27E-02
71836	SHCBP1L	1	protein-coding	-1.516	3.41E-02
11864	ARNT2	7	protein-coding	-1.547	4.30E-02
624855	CKS1BRT	8	protein-coding	-1.606	4.11E-02
240063	ZFP811	17	protein-coding	-1.613	3.28E-02
320642	A630066F11RIK	10	ncRNA	-1.665	3.46E-02
1E+08	GM10856	15	unknown	-1.680	4.13E-03
12578	CDKN2A	4	protein-coding	-1.688	1.81E-02
12554	CDH13	8	protein-coding	-1.694	5.53E-03
434863	GM15128	X	protein-coding	-1.694	2.50E-02
15898	ICAM5	9	protein-coding	-1.729	2.37E-02
1.01E+08	GM20300	10	ncRNA	-1.773	3.66E-02
791260	TOMT	7	protein-coding	-1.786	9.77E-03
231724	RAD9B	5	protein-coding	-1.833	3.49E-03
11423	ACHE	5	protein-coding	-1.835	3.40E-02
637515	NLRP1B	11	protein-coding	-1.932	4.17E-02
70981	4931423N10RIK	2	protein-coding	-1.936	8.63E-03
380732	MILR1	11	protein-coding	-1.942	8.58E-03
433791	ZFP992	4	protein-coding	-2.082	1.09E-02
68870	AK8	2	protein-coding	-2.096	6.74E-03
74843	MSS51	14	protein-coding	-2.156	4.01E-02
319166	HIST1H2AE	13	protein-coding	-2.307	4.16E-02
12479	CD1D1	3	protein-coding	-2.341	5.43E-03
71004	4931440P22RIK	3	ncRNA	-2.512	2.44E-03
72303	CYP2C65	19	protein-coding	-2.563	6.48E-03
14403	GABRD	4	protein-coding	-2.606	2.85E-02
330790	HAPLN4	8	protein-coding	-2.769	2.68E-02
110877	SLC18A1	8	protein-coding	-2.780	1.88E-02

GeneID	Symbol	chromosome	type_of_gene	log2 Fold Change	PValue
320977	A330023F24RIK	1	ncRNA	-2.874	7.68E-03
77705	9230104L09RIK	2	protein-coding	-2.977	6.19E-04
1.01E+08	GM17238	18	ncRNA	-3.037	7.85E-03
50524	SALL2	14	protein-coding	-3.101	2.07E-03
279028	ADAMTS13	2	protein-coding	-3.310	6.74E-03
73844	ANKRD45	1	protein-coding	-3.764	9.82E-04
19866	RNU7	6	snRNA	-3.969	2.55E-04
20341	SELENBP1	3	protein-coding	-4.475	7.85E-06
16582	KIFC3	8	protein-coding	-5.463	1.17E-05
22177	TYROBP	7	protein-coding	-7.967	1.07E-06
1E+08	GM12758	7	unknown	-7.994	6.60E-07
382245	TMEM29	X	protein-coding	-7.997	6.42E-07

supplementary Table S4.2 Function annotation clustering GO terms enriched in Twist1-upregulated genes.

Annotation Cluster 1	Enrichment Score: 2.6675464418670964												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0032680~regulation of tumor necrosis factor production	5	5.681818182	0.001121508	CCR5, CLU, TNFRSF8, CD300LD, TWIST1	67	124	17958	10.80765527	0.903639348	0.541541625	1.920019039	
GOTERM_BP_FAT	GO:1903555~regulation of tumor necrosis factor superfamily cytokine production	5	5.681818182	0.001190027	CCR5, CLU, TNFRSF8, CD300LD, TWIST1	67	126	17958	10.63610519	0.916481148	0.462416254	2.036188778	
GOTERM_BP_FAT	GO:0071706~tumor necrosis factor superfamily cytokine production	5	5.681818182	0.001261429	CCR5, CLU, TNFRSF8, CD300LD, TWIST1	67	128	17958	10.46991604	0.928046767	0.409241387	2.157110217	
GOTERM_BP_FAT	GO:0032760~positive regulation of tumor necrosis factor production	4	4.545454545	0.002454503	CCR5, CLU, TNFRSF8, TWIST1	67	73	17958	14.68656716	0.994047417	0.434093021	4.15692768	
GOTERM_BP_FAT	GO:1903557~positive regulation of tumor necrosis factor superfamily cytokine production	4	4.545454545	0.002551547	CCR5, CLU, TNFRSF8, TWIST1	67	74	17958	14.48810004	0.995140275	0.412968815	4.317885738	
GOTERM_BP_FAT	GO:0032640~tumor necrosis factor production	4	4.545454545	0.009369844	CCR5, CLU, TNFRSF8, TWIST1	67	118	17958	9.085757652	0.999999997	0.458483666	15.01060759	
Annotation Cluster 2	Enrichment Score: 2.376591840117395												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0072511~divalent inorganic cation transport	8	9.090909091	8.40E-04	TEX101, VDR, CATSPER2, CCR5, TCN2, HTR2B, SLC39A2, CXCL10	67	418	17958	5.129757909	0.826702493	0.583709828	1.441866981	
GOTERM_BP_FAT	GO:0070838~divalent metal ion transport	7	7.954545455	0.004148816	TEX101, VDR, CATSPER2, CCR5, HTR2B, SLC39A2, CXCL10	67	415	17958	4.520985434	0.999828045	0.461604899	6.930831086	
GOTERM_BP_FAT	GO:0030001~metal ion transport	9	10.22727273	0.008159205	TEX101, VDR, SLC5A5, CATSPER2, CCR5, TCN2, HTR2B, SLC39A2, CXCL10	67	790	17958	3.053504629	0.999999962	0.423640336	13.1982045	
GOTERM_BP_FAT	GO:0006816~calcium ion transport	6	6.818181818	0.010955451	TEX101, VDR, CATSPER2, CCR5, HTR2B, CXCL10	67	365	17958	4.405970149	1	0.481198098	17.33047551	
Annotation Cluster 3	Enrichment Score: 2.341228054803135												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0001819~positive regulation of cytokine production	7	7.954545455	0.002554621	CCR5, CLU, TNFRSF8, HTR2B, NLRP10, SCAMP5, TWIST1	67	376	17958	4.989917434	0.995171408	0.384202064	4.322981207	
GOTERM_BP_FAT	GO:0001817~regulation of cytokine production	8	9.090909091	0.004825128	CCR5, CLU, TNFRSF8, HTR2B, NLRP10, SCAMP5, CD300LD, TWIST1	67	570	17958	3.761822467	0.999958289	0.411851538	8.016811167	
GOTERM_BP_FAT	GO:0001816~cytokine production	8	9.090909091	0.007682109	CCR5, CLU, TNFRSF8, HTR2B, NLRP10, SCAMP5, CD300LD, TWIST1	67	622	17958	3.44732927	0.999999896	0.448724121	12.47400847	
Annotation Cluster 4	Enrichment Score: 2.001910763861947												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0060548~negative regulation of cell death	11	12.5	0.004795973	FAM134B, CERKL, CCR5, ERBB3, CLU, PROKR1, AGO4, NGB, MAPK8IP1, HTR2B, TWIST1	67	1051	17958	2.805260093	0.999955662	0.445465905	7.97024247	
GOTERM_BP_FAT	GO:0006915~apoptotic process	15	17.04545455	0.004995135	GGCT, ERBB3, CLU, TNFRSF8, PROKR1, FAM134B, VDR, PRUNE2, CERKL, CCR5, AGO4, MAPK8IP1, NC	67	1794	17958	2.241052264	0.999970789	0.406695641	8.287914204	
GOTERM_BP_FAT	GO:0043066~negative regulation of apoptotic process	10	11.36363636	0.00738295	FAM134B, CERKL, CCR5, ERBB3, CLU, PROKR1, AGO4, MAPK8IP1, HTR2B, TWIST1	67	944	17958	2.839299266	0.999999805	0.448024823	12.0170084	
GOTERM_BP_FAT	GO:0012501~programmed cell death	15	17.04545455	0.007699294	GGCT, ERBB3, CLU, TNFRSF8, PROKR1, FAM134B, VDR, PRUNE2, CERKL, CCR5, AGO4, MAPK8IP1, NC	67	1885	17958	2.132863534	0.9999999	0.43759941	12.50019178	
GOTERM_BP_FAT	GO:0043069~negative regulation of programmed cell death	10	11.36363636	0.008098679	FAM134B, CERKL, CCR5, ERBB3, CLU, PROKR1, AGO4, MAPK8IP1, HTR2B, TWIST1	67	958	17958	2.797806375	0.999999957	0.442692262	13.10664355	
GOTERM_BP_FAT	GO:0008219~cell death	15	17.04545455	0.013654986	GGCT, ERBB3, CLU, TNFRSF8, PROKR1, FAM134B, VDR, PRUNE2, CERKL, CCR5, AGO4, MAPK8IP1, NC	67	2018	17958	1.992293241	1	0.529702555	21.14342277	
GOTERM_BP_FAT	GO:0010941~regulation of cell death	13	14.77272727	0.015200707	FAM134B, VDR, CERKL, CCR5, ERBB3, CLU, TNFRSF8, PROKR1, AGO4, NGB, MAPK8IP1, HTR2B, TWIS	67	1637	17958	2.128520501	1	0.541111017	23.25143223	
GOTERM_BP_FAT	GO:0042981~regulation of apoptotic process	12	13.63636364	0.019778018	FAM134B, VDR, CERKL, CCR5, ERBB3, CLU, TNFRSF8, PROKR1, AGO4, MAPK8IP1, HTR2B, TWIST1	67	1497	17958	2.148535878	1	0.580090929	29.18677527	
GOTERM_BP_FAT	GO:0043067~regulation of programmed cell death	12	13.63636364	0.021229822	FAM134B, VDR, CERKL, CCR5, ERBB3, CLU, TNFRSF8, PROKR1, AGO4, MAPK8IP1, HTR2B, TWIST1	67	1513	17958	2.125815075	1	0.598712636	30.97710637	
Annotation Cluster 5	Enrichment Score: 1.7398873518485685												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0051046~regulation of secretion	9	10.22727273	0.006319863	UCN2, CCR5, ERBB3, SYTL2, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	756	17958	3.190831557	0.999998184	0.42350027	10.37476465	
GOTERM_BP_FAT	GO:0046903~secretion	10	11.36363636	0.016989376	VDR, UCN2, CCR5, ERBB3, SYTL2, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	1082	17958	2.477170524	1	0.540068816	25.62448358	
GOTERM_BP_FAT	GO:0032940~secretion by cell	8	9.090909091	0.056163327	UCN2, CCR5, SYTL2, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	939	17958	2.283534405	1	0.790945845	63.16176498	
Annotation Cluster 6	Enrichment Score: 1.7001658498852519												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0050715~positive regulation of cytokine secretion	5	5.681818182	7.93E-04	CCR5, HTR2B, NLRP10, SCAMP5, TWIST1	67	113	17958	11.85972791	0.808854666	0.808854666	1.361780609	
GOTERM_BP_FAT	GO:0050714~positive regulation of protein secretion	6	6.818181818	0.002060408	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	245	17958	6.563996345	0.986436634	0.415821008	3.50064912	
GOTERM_BP_FAT	GO:0050707~regulation of cytokine secretion	5	5.681818182	0.00311281	CCR5, HTR2B, NLRP10, SCAMP5, TWIST1	67	164	17958	8.171641791	0.998497058	0.418238223	5.243822889	
GOTERM_BP_FAT	GO:0050663~cytokine secretion	5	5.681818182	0.004338344	CCR5, HTR2B, NLRP10, SCAMP5, TWIST1	67	180	17958	7.445273632	0.999884371	0.453565934	7.236374178	
GOTERM_BP_FAT	GO:0051047~positive regulation of secretion	7	7.954545455	0.004816408	CCR5, SYTL2, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	428	17958	4.383665783	0.99995752	0.428361478	8.002885083	
GOTERM_BP_FAT	GO:1903532~positive regulation of secretion by cell	6	6.818181818	0.015592359	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	399	17958	4.030524071	1	0.533270014	23.77706157	
GOTERM_BP_FAT	GO:0050708~regulation of protein secretion	6	6.818181818	0.022158019	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	437	17958	3.680043717	1	0.607170682	32.09929653	
GOTERM_BP_FAT	GO:0051222~positive regulation of protein transport	6	6.818181818	0.034504157	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	492	17958	3.268656716	1	0.716988904	45.48243765	
GOTERM_BP_FAT	GO:0009306~protein secretion	6	6.818181818	0.04129743	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	517	17958	3.110597881	1	0.736137613	51.74337385	
GOTERM_BP_FAT	GO:1903530~regulation of secretion by cell	7	7.954545455	0.04405601	UCN2, CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	703	17958	2.668860534	1	0.748797305	54.08696069	
GOTERM_BP_FAT	GO:1904951~positive regulation of establishment of protein localization	6	6.818181818	0.044227966	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	527	17958	3.051573253	1	0.740080209	54.22943869	
GOTERM_BP_FAT	GO:0032940~secretion by cell	8	9.090909091	0.056163327	UCN2, CCR5, SYTL2, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	939	17958	2.283534405	1	0.790945845	63.16176498	
GOTERM_BP_FAT	GO:0032880~regulation of protein localization	7	7.954545455	0.165287965	TMEM231, CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	1009	17958	1.859473692	1	0.905043477	95.59031252	
GOTERM_BP_FAT	GO:0051223~regulation of protein transport	6	6.818181818	0.165565118	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	792	17958	2.030529172	1	0.904058208	95.61554036	
GOTERM_BP_FAT	GO:0070201~regulation of establishment of protein localization	6	6.818181818	0.200590571	CCR5, HTR2B, MMP13, NLRP10, SCAMP5, TWIST1	67	848	17958	1.896437623	1	0.919798341	97.90989671	
Annotation Cluster 7	Enrichment Score: 1.5941159020149094												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0030326~embryonic limb morphogenesis	4	4.545454545	0.016589208	TMEM231, PBX1, MMP13, TWIST1	67	146	17958	7.343283582	1	0.539333169	25.09965719	
GOTERM_BP_FAT	GO:0035113~embryonic appendage morphogenesis	4	4.545454545	0.016589208	TMEM231, PBX1, MMP13, TWIST1	67	146	17958	7.343283582	1	0.539333169	25.09965719	
GOTERM_BP_FAT	GO:0035107~appendage morphogenesis	4	4.545454545	0.026994897	TMEM231, PBX1, MMP13, TWIST1	67	176	17958	6.091587517	1	0.66621875	37.67422045	
GOTERM_BP_FAT	GO:0035108~limb morphogenesis	4	4.545454545	0.026994897	TMEM231, PBX1, MMP13, TWIST1	67	176	17958	6.091587517	1	0.66621875	37.67422045	
GOTERM_BP_FAT	GO:0048736~appendage development	4	4.545454545	0.03685912	TMEM231, PBX1, MMP13, TWIST1	67	199	17958	5.387534688	1	0.728842811	47.73477302	
GOTERM_BP_FAT	GO:0060173~limb development	4	4.545454545	0.03685912	TMEM231, PBX1, MMP13, TWIST1	67	199	17958	5.387534688	1	0.728842811	47.73477302	
Annotation Cluster 8	Enrichment Score: 1.5510003290957888												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0032675~regulation of interleukin-6 production	4	4.545454545	0.01003096	CCR5, NLRP10, CD300LD, TWIST1	67	121	17958	8.860490934	0.999999999	0.461107877	15.98522981	
GOTERM_BP_FAT	GO:0032755~positive regulation of interleukin-6 production	3	3.409090909	0.030324207	CCR5, NLRP10, TWIST1	67	74	17958	10.86607503	1	0.695465575	41.25780687	
GOTERM_BP_FAT	GO:0032635~interleukin-6 production	3	3.409090909	0.073091529	CCR5, NLRP10, TWIST1	67	121	17958	6.6453682	1	0.837811116	73.05342829	
Annotation Cluster 9	Enrichment Score: 1.4317284654406535												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_CC_FAT	GO:0031513~nonmotile primary cilium	4	4.545454545	0.027047764	PTPRK, TMEM231, CERKL, CNGA3	62	154	14491	6.070800168	0.998791011	0.965229474	29.88850447	
GOTERM_CC_FAT	GO:0001750~photoreceptor outer segment	3	3.409090909	0.03690414	PTPRK, CERKL, CNGA3	62	72	14491	9.738575269	0.999900223	0.953618603	38.54944969	
GOTERM_CC_FAT	GO:0072372~primary cilium	4	4.545454545	0.050770116	PTPRK, TMEM231, CERKL, CNGA3	62	198	14491	4.721733464	0.999997142	0.958885178	49.07074275	

supplementary Table S4.2 Function annotation clustering GO terms enriched in Twist1-upregulated genes.

Annotation Cluster 10	Enrichment Score: 1.305885848837165												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0004620~phospholipase activity	3	3.409090909	0.032248644	PLB1, PLCH2, LIPG	59	82	16904	10.48201736	0.999990253	0.999990253	36.15463136	
GOTERM_MF_FAT	GO:0016298~lipase activity	3	3.409090909	0.045403921	PLB1, PLCH2, LIPG	59	99	16904	8.682074987	0.999999921	0.995712819	47.06246073	
GOTERM_BP_FAT	GO:0016042~lipid catabolic process	4	4.545454545	0.082553907	PLB1, PLCH2, LIPG, TWIST1	67	278	17958	3.856544615	1	0.86111914	77.43095616	
Annotation Cluster 11	Enrichment Score: 1.2370482612486071												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0051283~negative regulation of sequestering of calcium ion	4	4.545454545	0.005623818	TEX101, CCR5, HTR2B, CXCL10	67	98	17958	10.93999391	0.999992179	0.414031309	9.28392424	
GOTERM_BP_FAT	GO:0051209~release of sequestered calcium ion into cytosol	4	4.545454545	0.005623818	TEX101, CCR5, HTR2B, CXCL10	67	98	17958	10.93999391	0.999992179	0.414031309	9.28392424	
GOTERM_BP_FAT	GO:0097553~calcium ion transmembrane import into cytosol	4	4.545454545	0.006114081	TEX101, CCR5, HTR2B, CXCL10	67	101	17958	10.61504359	0.999997203	0.426476307	10.05355611	
GOTERM_BP_FAT	GO:1902656~calcium ion import into cytosol	4	4.545454545	0.006114081	TEX101, CCR5, HTR2B, CXCL10	67	101	17958	10.61504359	0.999997203	0.426476307	10.05355611	
GOTERM_BP_FAT	GO:0051282~regulation of sequestering of calcium ion	4	4.545454545	0.008126431	TEX101, CCR5, HTR2B, CXCL10	67	112	17958	9.57249467	0.999999959	0.432828942	13.14863681	
GOTERM_BP_FAT	GO:0051208~sequestering of calcium ion	4	4.545454545	0.008126431	TEX101, CCR5, HTR2B, CXCL10	67	112	17958	9.57249467	0.999999959	0.432828942	13.14863681	
GOTERM_BP_FAT	GO:0006816~calcium ion transport	6	6.818181818	0.010955451	TEX101, VDR, CATSPER2, CCR5, HTR2B, CXCL10	67	365	17958	4.405970149	1	0.481198098	17.33047551	
GOTERM_BP_FAT	GO:0051238~sequestering of metal ion	4	4.545454545	0.011191938	TEX101, CCR5, HTR2B, CXCL10	67	126	17958	8.508884151	1	0.47892138	17.67131789	
GOTERM_BP_FAT	GO:0070588~calcium ion transmembrane transport	5	5.681818182	0.011579834	TEX101, CATSPER2, CCR5, HTR2B, CXCL10	67	239	17958	5.607319053	1	0.481257086	18.22751762	
GOTERM_BP_FAT	GO:0060402~calcium ion transport into cytosol	4	4.545454545	0.013738571	TEX101, CCR5, HTR2B, CXCL10	67	136	17958	7.883230904	1	0.522684326	21.25879441	
GOTERM_BP_FAT	GO:0060401~cytosolic calcium ion transport	4	4.545454545	0.014284232	TEX101, CCR5, HTR2B, CXCL10	67	138	17958	7.768981181	1	0.527602544	22.00806958	
GOTERM_BP_FAT	GO:0070509~calcium ion import	4	4.545454545	0.026603038	TEX101, CCR5, HTR2B, CXCL10	67	175	17958	6.126396588	1	0.667901365	37.23914066	
GOTERM_BP_FAT	GO:0032845~negative regulation of homeostatic process	4	4.545454545	0.038739843	TEX101, CCR5, HTR2B, CXCL10	67	203	17958	5.281376369	1	0.723948536	49.47026352	
GOTERM_BP_FAT	GO:2000021~regulation of ion homeostasis	4	4.545454545	0.03921759	TEX101, CCR5, HTR2B, CXCL10	67	204	17958	5.25548727	1	0.722881926	49.90238989	
GOTERM_BP_FAT	GO:0006874~cellular calcium ion homeostasis	5	5.681818182	0.066271081	TEX101, VDR, CCR5, HTR2B, CXCL10	67	415	17958	3.22927531	1	0.828816623	69.41475106	
GOTERM_BP_FAT	GO:0055074~calcium ion homeostasis	5	5.681818182	0.072941974	TEX101, VDR, CCR5, HTR2B, CXCL10	67	429	17958	3.123891034	1	0.840583229	72.97821378	
GOTERM_BP_FAT	GO:0072503~cellular divalent inorganic cation homeostasis	5	5.681818182	0.075903083	TEX101, VDR, CCR5, HTR2B, CXCL10	67	435	17958	3.080802882	1	0.842650441	74.43123367	
GOTERM_BP_FAT	GO:0072507~divalent inorganic cation homeostasis	5	5.681818182	0.08834798	TEX101, VDR, CCR5, HTR2B, CXCL10	67	459	17958	2.91971515	1	0.863059256	79.77088456	
GOTERM_BP_FAT	GO:0007204~positive regulation of cytosolic calcium ion concentration	4	4.545454545	0.088824345	TEX101, CCR5, HTR2B, CXCL10	67	287	17958	3.735607676	1	0.856218189	79.95273046	
GOTERM_BP_FAT	GO:0051235~maintenance of location	4	4.545454545	0.097488848	TEX101, CCR5, HTR2B, CXCL10	67	299	17958	3.585683622	1	0.851849885	83.00329352	
GOTERM_BP_FAT	GO:0034220~ion transmembrane transport	6	6.818181818	0.110339676	TEX101, CATSPER2, CCR5, CNGA3, HTR2B, CXCL10	67	692	17958	2.323958243	1	0.875505659	86.73342008	
GOTERM_BP_FAT	GO:0098662~inorganic cation transmembrane transport	5	5.681818182	0.111740059	TEX101, CATSPER2, CCR5, HTR2B, CXCL10	67	500	17958	2.680298507	1	0.870200439	87.08961631	
GOTERM_BP_FAT	GO:0051480~regulation of cytosolic calcium ion concentration	4	4.545454545	0.112660477	TEX101, CCR5, HTR2B, CXCL10	67	319	17958	3.360875871	1	0.865984344	87.31880237	
GOTERM_BP_FAT	GO:0006875~cellular metal ion homeostasis	5	5.681818182	0.123439007	TEX101, VDR, CCR5, HTR2B, CXCL10	67	519	17958	2.582175826	1	0.87913138	89.73259576	
GOTERM_BP_FAT	GO:0098655~cation transmembrane transport	5	5.681818182	0.132385753	TEX101, CATSPER2, CCR5, HTR2B, CXCL10	67	533	17958	2.51435132	1	0.888444816	91.40027354	
GOTERM_BP_FAT	GO:0098660~inorganic ion transmembrane transport	5	5.681818182	0.144275568	TEX101, CATSPER2, CCR5, HTR2B, CXCL10	67	551	17958	2.432212802	1	0.896866363	93.22437381	
GOTERM_BP_FAT	GO:0030003~cellular cation homeostasis	5	5.681818182	0.168550064	TEX101, VDR, CCR5, HTR2B, CXCL10	67	586	17958	2.286944119	1	0.904315155	95.87876433	
GOTERM_BP_FAT	GO:0055065~metal ion homeostasis	5	5.681818182	0.172853659	TEX101, VDR, CCR5, HTR2B, CXCL10	67	592	17958	2.263765631	1	0.907782256	96.23218194	
GOTERM_BP_FAT	GO:0006873~cellular ion homeostasis	5	5.681818182	0.179381874	TEX101, VDR, CCR5, HTR2B, CXCL10	67	601	17958	2.229865647	1	0.911495387	96.71423646	
GOTERM_BP_FAT	GO:0032846~positive regulation of homeostatic process	3	3.409090909	0.201730866	TEX101, CCR5, CXCL10	67	226	17958	3.557918373	1	0.919993434	97.96081147	
GOTERM_BP_FAT	GO:0055080~cation homeostasis	5	5.681818182	0.230394175	TEX101, VDR, CCR5, HTR2B, CXCL10	67	668	17958	2.006211458	1	0.932100048	98.91586023	
GOTERM_BP_FAT	GO:0098771~inorganic ion homeostasis	5	5.681818182	0.24310711	TEX101, VDR, CCR5, HTR2B, CXCL10	67	684	17958	1.959282535	1	0.938703394	99.18697002	
GOTERM_BP_FAT	GO:0050801~ion homeostasis	5	5.681818182	0.273920959	TEX101, VDR, CCR5, HTR2B, CXCL10	67	722	17958	1.856162401	1	0.942981639	99.60349203	
GOTERM_BP_FAT	GO:0032844~regulation of homeostatic process	4	4.545454545	0.274731545	TEX101, CCR5, HTR2B, CXCL10	67	496	17958	2.161531054	1	0.942851026	99.61107061	
GOTERM_BP_FAT	GO:0055082~cellular chemical homeostasis	5	5.681818182	0.308741496	TEX101, VDR, CCR5, HTR2B, CXCL10	67	764	17958	1.75412206	1	0.951731583	99.83036836	
GOTERM_BP_FAT	GO:0048878~chemical homeostasis	6	6.818181818	0.385654087	TEX101, VDR, CCR5, LIPG, HTR2B, CXCL10	67	1108	17958	1.451425185	1	0.965003077	99.97789649	
GOTERM_BP_FAT	GO:0019725~cellular homeostasis	5	5.681818182	0.396356384	TEX101, VDR, CCR5, HTR2B, CXCL10	67	868	17958	1.543950753	1	0.965349882	99.98368456	
Annotation Cluster 12	Enrichment Score: 1.2042169204877133												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0001755~neural crest cell migration	3	3.409090909	0.018645395	SEMA4G, HTR2B, TWIST1	67	57	17958	14.10683425	1	0.566103771	27.75976749	
GOTERM_BP_FAT	GO:0014032~neural crest cell development	3	3.409090909	0.032621146	SEMA4G, HTR2B, TWIST1	67	77	17958	10.44272146	1	0.715567103	43.61602182	
GOTERM_BP_FAT	GO:0014033~neural crest cell differentiation	3	3.409090909	0.038235676	SEMA4G, HTR2B, TWIST1	67	84	17958	9.57249467	1	0.730464781	49.01043358	
GOTERM_BP_FAT	GO:0060485~mesenchyme development	4	4.545454545	0.061343913	SEMA4G, ERBB3, HTR2B, TWIST1	67	245	17958	4.375997563	1	0.807933918	66.50335145	
GOTERM_BP_FAT	GO:0014031~mesenchymal cell development	3	3.409090909	0.13175976	SEMA4G, HTR2B, TWIST1	67	172	17958	4.674939257	1	0.890836667	91.29244299	
GOTERM_BP_FAT	GO:0001667~ameboidal-type cell migration	4	4.545454545	0.133636732	CCR5, SEMA4G, HTR2B, TWIST1	67	345	17958	3.107592472	1	0.889111503	91.61200327	
GOTERM_BP_FAT	GO:0048762~mesenchymal cell differentiation	3	3.409090909	0.14806822	SEMA4G, HTR2B, TWIST1	67	185	17958	4.346430012	1	0.896988745	93.72490045	
Annotation Cluster 13	Enrichment Score: 1.18216280431716												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_MF_FAT	GO:0016462~pyrophosphatase activity	7	7.954545455	0.054179546	PRUNE2, RAC3, ATP9A, DDX25, NTPCR, AGAP1, TUBB4A	59	795	16904	2.522716128	0.999999997	0.992566972	53.34905552	
GOTERM_MF_FAT	GO:0016818~hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	7	7.954545455	0.054720179	PRUNE2, RAC3, ATP9A, DDX25, NTPCR, AGAP1, TUBB4A	59	797	16904	2.516385599	0.999999998	0.980969455	53.71274647	
GOTERM_MF_FAT	GO:0016817~hydrolase activity, acting on acid anhydrides	7	7.954545455	0.055537353	PRUNE2, RAC3, ATP9A, DDX25, NTPCR, AGAP1, TUBB4A	59	800	16904	2.506949153	0.999999998	0.964991448	54.25748341	
GOTERM_MF_FAT	GO:0017111~nucleoside-triphosphatase activity	6	6.818181818	0.113443792	RAC3, ATP9A, DDX25, NTPCR, AGAP1, TUBB4A	59	749	16904	2.295127967	1	0.985570181	80.76093823	
Annotation Cluster 14	Enrichment Score: 1.1624012532564059												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0014065~phosphatidylinositol 3-kinase signaling	3	3.409090909	0.051385456	ERBB3, HTR2B, TWIST1	67	99	17958	8.122116689	1	0.773801856	59.80374803	
GOTERM_BP_FAT	GO:0060485~mesenchyme development	4	4.545454545	0.061343913	SEMA4G, ERBB3, HTR2B, TWIST1	67	245	17958	4.375997563	1	0.807933918	66.50335145	
GOTERM_BP_FAT	GO:0048015~phosphatidylinositol-mediated signaling	3	3.409090909	0.082687804	ERBB3, HTR2B, TWIST1	67	130	17958	6.185304248	1	0.858575407	77.48779566	
GOTERM_BP_FAT	GO:0048017~inositol lipid-mediated signaling	3	3.409090909	0.085968992	ERBB3, HTR2B, TWIST1	67	133	17958	6.045786107	1	0.863830393	78.83924001	
Annotation Cluster 15	Enrichment Score: 1.1190876854095297												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0040011~locomotion	11	12.5	0.053311171	CCRL2, PTPRK, MMP10, CATSPER2, CCR5, SEMA4G, CCDC39, HTR2B, NLRP10, CXCL10, TWIST1	67	1538	17958	1.916988529	1	0.781947222	61.19047157	
GOTERM_BP_FAT	GO:0048870~cell motility	10	11.36363636	0.058858548	PTPRK, MMP10, CATSPER2, CCR5, SEMA4G, CCDC39, HTR2B, NLRP10, CXCL10, TWIST1	67	1355	17958	1.978080079	1	0.802405312	64.93757003	
GOTERM_BP_FAT	GO:0051674~localization of cell	10	11.36363636	0.058858548	PTPRK, MMP10, CATSPER2, CCR5, SEMA4G, CCDC39, HTR2B, NLRP10, CXCL10, TWIST1	67	1355	17958	1.978080079	1	0.802405312	64.93757003	
GOTERM_BP_FAT	GO:0006928~movement of cell or subcellular component	10	11.36363636	0.18080563	PTPRK, MMP10, CATSPER2, CCR5, SEMA4G, CCDC39, HTR2B, NLRP10, CXCL10, TWIST1	67	1728	17958	1.551098673	1	0.910860816	96.81134822	
Annotation Cluster 16	Enrichment Score: 1.093490641506508												

supplementary Table S4.2 Function annotation clustering GO terms enriched in Twist1-upregulated genes.

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0009615~response to virus	5	5.681818182	0.04675476	IFI27L2A, CLU, HIST1H3G, IFNLR1, CXCL10	67	369	17958	3.631840796	1	0.750079897	56.27562523
GOTERM_BP_FAT	GO:0016032~viral process	6	6.818181818	0.068982387	CCR5, IFI27L2A, CLU, HIST1H3G, IFNLR1, CXCL10	67	599	17958	2.684773129	1	0.83037469	70.91339919
GOTERM_BP_FAT	GO:0044764~multi-organism cellular process	6	6.818181818	0.070546189	CCR5, IFI27L2A, CLU, HIST1H3G, IFNLR1, CXCL10	67	603	17958	2.666963689	1	0.833793912	71.74603029
GOTERM_BP_FAT	GO:0044403~symbiosis, encompassing mutualism through parasitism	6	6.818181818	0.08859359	CCR5, IFI27L2A, CLU, HIST1H3G, IFNLR1, CXCL10	67	646	17958	2.489441338	1	0.858253627	79.86483622
GOTERM_BP_FAT	GO:0044419~interspecies interaction between organisms	6	6.818181818	0.08859359	CCR5, IFI27L2A, CLU, HIST1H3G, IFNLR1, CXCL10	67	646	17958	2.489441338	1	0.858253627	79.86483622
GOTERM_BP_FAT	GO:0051707~response to other organism	8	9.090909091	0.093347482	CCR5, IFI27L2A, CLU, TNFRSF8, HIST1H3G, NLRP10, IFNLR1, CXCL10	67	1060	17958	2.022866798	1	0.849210053	81.60431042
GOTERM_BP_FAT	GO:0043207~response to external biotic stimulus	8	9.090909091	0.093702171	CCR5, IFI27L2A, CLU, TNFRSF8, HIST1H3G, NLRP10, IFNLR1, CXCL10	67	1061	17958	2.020960232	1	0.847714092	81.72824802
GOTERM_BP_FAT	GO:0009607~response to biotic stimulus	8	9.090909091	0.114387981	CCR5, IFI27L2A, CLU, TNFRSF8, HIST1H3G, NLRP10, IFNLR1, CXCL10	67	1116	17958	1.921360937	1	0.868167765	87.73864299

Annotation Cluster 17 Enrichment Score: 1.0333545045718506

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0032941~secretion by tissue	3	3.409090909	0.044190161	VDR, SYTL2, MMP13	67	91	17958	8.836148926	1	0.744801662	54.19814969
GOTERM_BP_FAT	GO:0007589~body fluid secretion	3	3.409090909	0.048640181	VDR, SYTL2, MMP13	67	96	17958	8.375932836	1	0.759291168	57.74594973
GOTERM_BP_FAT	GO:0050878~regulation of body fluid levels	3	3.409090909	0.369501087	VDR, SYTL2, MMP13	67	350	17958	2.297398721	1	0.963293022	99.9653906

supplementary Table S4.3 Function annotation clustering GO terms enriched in Twist1-downregulated genes.

Annotation Cluster 1	Enrichment Score: 1.2625908356625735												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0002694~regulation of leukocyte activation	5	8.196721311	0.027186186	CDKN2A, MILR1, BANK1, CD1D1, IGLC2	44	479	17958	4.260296071	1	1	35.77638956	
GOTERM_BP_FAT	GO:0050865~regulation of cell activation	5	8.196721311	0.034608932	CDKN2A, MILR1, BANK1, CD1D1, IGLC2	44	517	17958	3.94716019	1	0.999999998	43.21227758	
GOTERM_BP_FAT	GO:0045321~leukocyte activation	6	9.836065574	0.041621855	CDKN2A, MILR1, BANK1, CD1D1, IGLC2, TYROBP	44	802	17958	3.053389254	1	0.999994587	49.48901014	
GOTERM_BP_FAT	GO:0001775~cell activation	6	9.836065574	0.06919152	CDKN2A, MILR1, BANK1, CD1D1, IGLC2, TYROBP	44	927	17958	2.641659312	1	0.995721326	68.39660629	
GOTERM_BP_FAT	GO:0002682~regulation of immune system process	6	9.836065574	0.179529178	CDKN2A, MILR1, BANK1, CD1D1, IGLC2, TYROBP	44	1256	17958	1.949696005	1	0.995372121	95.83702622	
Annotation Cluster 2	Enrichment Score: 1.158456256716628												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0002694~regulation of leukocyte activation	5	8.196721311	0.027186186	CDKN2A, MILR1, BANK1, CD1D1, IGLC2	44	479	17958	4.260296071	1	1	35.77638956	
GOTERM_BP_FAT	GO:0050865~regulation of cell activation	5	8.196721311	0.034608932	CDKN2A, MILR1, BANK1, CD1D1, IGLC2	44	517	17958	3.94716019	1	0.999999998	43.21227758	
GOTERM_BP_FAT	GO:0050864~regulation of B cell activation	3	4.918032787	0.047135346	CDKN2A, BANK1, IGLC2	44	145	17958	8.444200627	1	0.999983597	53.96040851	
GOTERM_BP_FAT	GO:0051249~regulation of lymphocyte activation	4	6.557377049	0.075891864	CDKN2A, BANK1, CD1D1, IGLC2	44	413	17958	3.952894563	1	0.983317565	71.85971871	
GOTERM_BP_FAT	GO:0042113~B cell activation	3	4.918032787	0.147861924	CDKN2A, BANK1, IGLC2	44	284	17958	4.311299616	1	0.99372589	92.35062488	
GOTERM_BP_FAT	GO:0046649~lymphocyte activation	4	6.557377049	0.225075236	CDKN2A, BANK1, CD1D1, IGLC2	44	685	17958	2.383278036	1	0.997950733	98.3368637	
Annotation Cluster 3	Enrichment Score: 1.114406200024837												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0010556~regulation of macromolecule biosynthetic process	16	26.2295082	0.038989157	JPX, ZFP345, CKS1BRT, ZFP992, HIST1H2AE, ARNT2, CD1D1, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDI	44	3984	17958	1.639101862	1	0.99999973	47.21314569	
GOTERM_BP_FAT	GO:0034645~cellular macromolecule biosynthetic process	18	29.50819672	0.048929183	JPX, ZFP345, ST6GAL1, NDST3, CKS1BRT, ZFP992, HIST1H2AE, ARNT2, ZFP455, ZFP953, CDH13, SALL2, Z	44	4836	17958	1.519117979	1	0.999928086	55.33325751	
GOTERM_BP_FAT	GO:0006355~regulation of transcription, DNA-templated	14	22.95081967	0.057425935	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3465	17958	1.649035813	1	0.999782857	61.33018459	
GOTERM_BP_FAT	GO:1903506~regulation of nucleic acid-templated transcription	14	22.95081967	0.057766816	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3468	17958	1.647609311	1	0.999470509	61.55424413	
GOTERM_BP_FAT	GO:2001141~regulation of RNA biosynthetic process	14	22.95081967	0.058452864	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3474	17958	1.644763699	1	0.998963952	62.00149541	
GOTERM_BP_FAT	GO:2000112~regulation of cellular macromolecule biosynthetic process	15	24.59016393	0.065324985	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3892	17958	1.572981874	1	0.997339848	66.22028224	
GOTERM_BP_FAT	GO:0097659~nucleic acid-templated transcription	14	22.95081967	0.069571411	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3565	17958	1.602779549	1	0.994156006	68.60318464	
GOTERM_BP_FAT	GO:0032774~RNA biosynthetic process	14	22.95081967	0.071799727	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3582	17958	1.595172834	1	0.988602793	69.78963391	
GOTERM_BP_FAT	GO:0051252~regulation of RNA metabolic process	14	22.95081967	0.073536444	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3595	17958	1.589404476	1	0.987189773	70.68504015	
GOTERM_BP_FAT	GO:0034654~nucleobase-containing compound biosynthetic process	15	24.59016393	0.07495252	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, AK8, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN:	44	3969	17958	1.542465471	1	0.985492965	71.39664968	
GOTERM_BP_FAT	GO:0016070~RNA metabolic process	16	26.2295082	0.077426648	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, RNU7, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDK	44	4357	17958	1.498779394	1	0.981645582	72.6012257	
GOTERM_BP_FAT	GO:0018130~heterocycle biosynthetic process	15	24.59016393	0.082680888	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, AK8, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN:	44	4026	17958	1.520627286	1	0.980526995	75.00334613	
GOTERM_BP_FAT	GO:0019438~aromatic compound biosynthetic process	15	24.59016393	0.084943986	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, AK8, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN:	44	4042	17958	1.51460798	1	0.976514084	75.9758687	
GOTERM_BP_FAT	GO:0051171~regulation of nitrogen compound metabolic process	15	24.59016393	0.116607848	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	4241	17958	1.443538188	1	0.989574359	86.35594062	
GOTERM_BP_FAT	GO:0010468~regulation of gene expression	15	24.59016393	0.125417047	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	4290	17958	1.427050223	1	0.991589105	88.38500076	
GOTERM_BP_FAT	GO:0019219~regulation of nucleobase-containing compound metabolic process	14	22.95081967	0.131144279	JPX, ZFP345, ZFP992, CKS1BRT, HIST1H2AE, ARNT2, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDKN2A, ZI	44	3944	17958	1.448759911	1	0.992254147	89.5484779	
GOTERM_BP_FAT	GO:0010467~gene expression	17	27.86885246	0.163608842	JPX, ZFP345, CKS1BRT, ZFP992, HIST1H2AE, ARNT2, RNU7, ZFP455, ZFP953, SALL2, CDH13, ZFP81, CDK	44	5256	17958	1.320075758	1	0.993880315	94.33122971	
Annotation Cluster 4	Enrichment Score: 1.0315874184290663												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR	
GOTERM_BP_FAT	GO:0002695~negative regulation of leukocyte activation	3	4.918032787	0.05067627	CDKN2A, MILR1, BANK1	44	151	17958	8.108669476	1	0.999791773	56.63334915	
GOTERM_BP_FAT	GO:0050866~negative regulation of cell activation	3	4.918032787	0.061846682	CDKN2A, MILR1, BANK1	44	169	17958	7.245024207	1	0.99768952	64.1431423	
GOTERM_BP_FAT	GO:0002683~negative regulation of immune system process	3	4.918032787	0.256517322	CDKN2A, MILR1, BANK1	44	409	17958	2.993665259	1	0.998994343	99.14506068	

4.7 References

1. Bergsagel PL, Mateos MV, Gutierrez NC, Rajkumar SV, San Miguel JF. Improving overall survival and overcoming adverse prognosis in the treatment of cytogenetically high-risk multiple myeloma. *Blood* 2013 Feb 07; **121**(6)**:** 884-892.

2. Lwin ST, Edwards CM, Silbermann R. Preclinical animal models of multiple myeloma. *Bonekey Rep* 2016; **5:** 772.

3. Mitsiades CS, Anderson KC, Carrasco DR. Mouse models of human myeloma. *Hematol Oncol Clin North Am* 2007 Dec; **21**(6)**:** 1051-1069, viii.

4. Schuler J, Ewerth D, Waldschmidt J, Wasch R, Engelhardt M. Preclinical models of multiple myeloma: a critical appraisal. *Expert Opin Biol Ther* 2013 Jun; **13 Suppl 1:** S111-123.

5. Mitsiades CS, Mitsiades NS, Bronson RT, Chauhan D, Munshi N, Treon SP, *et al.* Fluorescence imaging of multiple myeloma cells in a clinically relevant SCID/NOD in vivo model: biologic and clinical implications. *Cancer research* 2003 Oct 15; **63**(20)**:** 6689-6696.

6. Chauhan D, Hideshima T, Anderson KC. A novel proteasome inhibitor NPI-0052 as an anticancer therapy. *British journal of cancer* 2006 Oct 23; **95**(8)**:** 961-965.

7. Neri P, Tagliaferri P, Di Martino MT, Calimeri T, Amodio N, Bulotta A, *et al.* In vivo anti-myeloma activity and modulation of gene expression profile induced by valproic acid, a histone deacetylase inhibitor. *Br J Haematol* 2008 Nov; **143**(4)**:** 520-531.

8. Urashima M, Chen BP, Chen S, Pinkus GS, Bronson RT, Dedera DA, *et al.* The development of a model for the homing of multiple myeloma cells to human bone marrow. *Blood* 1997 Jul 15; **90**(2)**:** 754-765.

9. Yaccoby S, Barlogie B, Epstein J. Primary myeloma cells growing in SCID-hu mice: a model for studying the biology and treatment of myeloma and its manifestations. *Blood* 1998 Oct 15; **92**(8)**:** 2908-2913.

10. Yaccoby S. The phenotypic plasticity of myeloma plasma cells as expressed by dedifferentiation into an immature, resilient, and apoptosis-resistant phenotype. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2005 Nov 1; **11**(21)**:** 7599-7606.

11. Asosingh K, Radl J, Van Riet I, Van Camp B, Vanderkerken K. The 5TMM series: a useful in vivo mouse model of human multiple myeloma. *Hematol J* 2000; **1**(5)**:** 351-356.

12. Radl J, Croese JW, Zurcher C, Van den Enden-Vieveen MH, de Leeuw AM. Animal model of human disease. Multiple myeloma. *Am J Pathol* 1988 Sep; **132**(3)**:** 593-597.

13. Vanderkerken K, De Raeve H, Goes E, Van Meirvenne S, Radl J, Van Riet I*, et al.* Organ involvement and phenotypic adhesion profile of 5T2 and 5T33 myeloma cells in the C57BL/KaLwRij mouse. *British journal of cancer* 1997; **76**(4)**:** 451-460.

14. Garrett IR, Dallas S, Radl J, Mundy GR. A murine model of human myeloma bone disease. *Bone* 1997 Jun; **20**(6)**:** 515-520.

15. Manning LS, Berger JD, O'Donoghue HL, Sheridan GN, Claringbold PG, Turner JH. A model of multiple myeloma: culture of 5T33 murine myeloma cells and evaluation of tumorigenicity in the C57BL/KaLwRij mouse. *British journal of cancer* 1992 Dec; **66**(6)**:** 1088-1093.

16. Dallas SL, Garrett IR, Oyajobi BO, Dallas MR, Boyce BF, Bauss F*, et al.* Ibandronate reduces osteolytic lesions but not tumor burden in a murine model of myeloma bone disease. *Blood* 1999 Mar 01; **93**(5)**:** 1697-1706.

17. Oyajobi BO, Munoz S, Kakonen R, Williams PJ, Gupta A, Wideman CL*, et al.* Detection of myeloma in skeleton of mice by whole-body optical fluorescence imaging. *Mol Cancer Ther* 2007 Jun; **6**(6)**:** 1701-1708.

18. Edwards CM, Lwin ST, Fowler JA, Oyajobi BO, Zhuang J, Bates AL*, et al.* Myeloma cells exhibit an increase in proteasome activity and an enhanced response to proteasome inhibition in the bone marrow microenvironment in vivo. *Am J Hematol* 2009 May; **84**(5)**:** 268-272.

19. Oyajobi BO, Franchin G, Williams PJ, Pulkrabek D, Gupta A, Munoz S*, et al.* Dual effects of macrophage inflammatory protein-1alpha on osteolysis and tumor burden in the murine 5TGM1 model of myeloma bone disease. *Blood* 2003 Jul 01; **102**(1)**:** 311-319.

20. Mrozik KM, Cheong CM, Hewett D, Chow AW, Blaschuk OW, Zannettino AC*, et al.* Therapeutic targeting of N-cadherin is an effective treatment for multiple myeloma. *Br J Haematol* 2015 Nov; **171**(3)**:** 387-399.

21. Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nature reviews Cancer* 2005 Sep; **5**(9)**:** 744-749.

22. Kwok WK, Ling MT, Lee TW, Lau TC, Zhou C, Zhang X*, et al.* Up-regulation of TWIST in prostate cancer and its implication as a therapeutic target. *Cancer research* 2005 Jun 15; **65**(12)**:** 5153-5162.

23. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C*, et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 2004 Jun 25; **117**(7)**:** 927-939.

24. Cheong CM, Chow AW, Fitter S, Hewett DR, Martin SK, Williams SA, *et al.* Tetraspanin 7 (TSPAN7) expression is upregulated in multiple myeloma patients and inhibits myeloma tumour development in vivo. *Exp Cell Res* 2015 Mar 1; **332**(1)**:** 24-38.

25. Noll JE, Hewett DR, Williams SA, Vandyke K, Kok C, To LB, *et al.* SAMSN1 is a tumor suppressor gene in multiple myeloma. *Neoplasia* 2014 Jul; **16**(7)**:** 572-585.

26. Isenmann S, Arthur A, Zannettino AC, Turner JL, Shi S, Glackin CA, *et al.* TWIST family of basic helix-loop-helix transcription factors mediate human mesenchymal stem cell growth and commitment. *Stem Cells* 2009 Oct; **27**(10)**:** 2457-2468.

27. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocols* 2008 12//print; **4**(1)**:** 44-57.

28. Figeac N, Daczewska M, Marcelle C, Jagla K. Muscle stem cells and model systems for their investigation. *Dev Dyn* 2007 Dec; **236**(12)**:** 3332-3342.

29. Zhao P, Hoffman EP. Embryonic myogenesis pathways in muscle regeneration. *Dev Dyn* 2004 Feb; **229**(2)**:** 380-392.

30. Danciu TE, Whitman M. Oxidative stress drives disulfide bond formation between basic helix-loop-helix transcription factors. *J Cell Biochem* 2010 Feb 01; **109**(2)**:** 417-424.

31. Parmo-Cabanas M, Molina-Ortiz I, Matias-Roman S, Garcia-Bernal D, Carvajal-Vergara X, Valle I, *et al.* Role of metalloproteinases MMP-9 and MT1-MMP in CXCL12-promoted myeloma cell invasion across basement membranes. *J Pathol* 2006 Jan; **208**(1)**:** 108-118.

32. Van Valckenborgh E, Croucher PI, De Raeve H, Carron C, De Leenheer E, Blacher S, *et al.* Multifunctional role of matrix metalloproteinases in multiple myeloma: a study in the 5T2MM mouse model. *Am J Pathol* 2004 Sep; **165**(3)**:** 869-878.

33. Leeman MF, McKay JA, Murray GI. Matrix metalloproteinase 13 activity is associated with poor prognosis in colorectal cancer. *Journal of Clinical Pathology* 2002 05/01/accepted; **55**(10)**:** 758-762.

34. Varettoni M, Corso A, Pica G, Mangiacavalli S, Pascutto C, Lazzarino M. Incidence, presenting features and outcome of extramedullary disease in multiple myeloma: a longitudinal study on 1003 consecutive patients. *Ann Oncol* 2010 Feb; **21**(2)**:** 325-330.

35. Gonsalves WI, Morice WG, Rajkumar V, Gupta V, Timm MM, Dispenzieri A, *et al.* Quantification of clonal circulating plasma cells in relapsed multiple myeloma. *Br J Haematol* 2014 Nov; **167**(4)**:** 500-505.

36. Lentzsch S, Gries M, Janz M, Bargou R, Dorken B, Mapara MY. Macrophage inflammatory protein 1-alpha (MIP-1 alpha) triggers migration and signaling cascades mediating survival and proliferation in multiple myeloma (MM) cells. *Blood* 2003 May 01; **101**(9)**:** 3568-3573.

37. Walters DK, French JD, Arendt BK, Jelinek DF. Atypical expression of ErbB3 in myeloma cells: cross-talk between ErbB3 and the interferon-alpha signaling complex. *Oncogene* 2003 Jun 05; **22**(23)**:** 3598-3607.

38. Croset M, Goehrig D, Frackowiak A, Bonnelye E, Ansieau S, Puisieux A, *et al.* TWIST1 expression in breast cancer cells facilitates bone metastasis formation. *J Bone Miner Res* 2014 Aug; **29**(8)**:** 1886-1899.

39. Mironchik Y, Winnard PT, Jr., Vesuna F, Kato Y, Wildes F, Pathak AP, *et al.* Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer. *Cancer research* 2005 Dec 1; **65**(23)**:** 10801-10809.

40. Usmani SZ, Rodriguez-Otero P, Bhutani M, Mateos MV, Miguel JS. Defining and treating high-risk multiple myeloma. *Leukemia* 2015 Nov; **29**(11)**:** 2119-2125.

41. Gonsalves WI, Rajkumar SV, Gupta V, Morice WG, Timm MM, Singh PP, *et al.* Quantification of clonal circulating plasma cells in newly diagnosed multiple myeloma: implications for redefining high-risk myeloma. *Leukemia* 2014 Oct; **28**(10)**:** 2060-2065.

42. Noll JE, Williams SA, Purton LE, Zannettino AC. Tug of war in the haematopoietic stem cell niche: do myeloma plasma cells compete for the HSC niche? *Blood cancer journal* 2012 Sep 14; **2:** e91.

43. Saboo SS, Krajewski KM, O'Regan KN, Giardino A, Brown JR, Ramaiya N, *et al.* Spleen in haematological malignancies: spectrum of imaging findings. *Br J Radiol* 2012 Jan; **85**(1009)**:** 81-92.

44. Yang JZ, Lian WG, Sun LX, Qi DW, Ding Y, Zhang XH. High nuclear expression of Twist1 in the skeletal extramedullary disease of myeloma patients predicts inferior survival. *Pathol Res Pract* 2016 Mar; **212**(3)**:** 210-216.

45. Vande Broek I, Vanderkerken K, Van Camp B, Van Riet I. Extravasation and homing mechanisms in multiple myeloma. *Clin Exp Metastasis* 2008; **25**(4)**:** 325-334.

46. Alsayed Y, Ngo H, Runnels J, Leleu X, Singha UK, Pitsillides CM, *et al.* Mechanisms of regulation of CXCR4/SDF-1 (CXCL12)-dependent migration and homing in multiple myeloma. *Blood* 2007 Apr 1; **109**(7)**:** 2708-2717.

47. Roccaro AM, Sacco A, Purschke WG, Moschetta M, Buchner K, Maasch C, *et al.* SDF-1 inhibition targets the bone marrow niche for cancer therapy. *Cell Rep* 2014 Oct 09; **9**(1)**:** 118-128.

48. Ou DL, Chien HF, Chen CL, Lin TC, Lin LI. Role of Twist in head and neck carcinoma with lymph node metastasis. *Anticancer Res* 2008 Mar-Apr; **28**(2B)**:** 1355-1359.

49. Wang N, Luo HJ, Yin GB, Dong CR, Xu M, Chen GG, *et al.* Overexpression of HIF-2alpha, TWIST, and CXCR4 is associated with lymph node metastasis in papillary thyroid carcinoma. *Clin Dev Immunol* 2013; **2013:** 589423.

50. Roccaro AM, Mishima Y, Sacco A, Moschetta M, Tai YT, Shi J, *et al.* CXCR4 Regulates Extra-Medullary Myeloma through Epithelial-Mesenchymal-Transition-like Transcriptional Activation. *Cell Rep* 2015 Jul 28; **12**(4)**:** 622-635.

51. Yao C, Li P, Song H, Song F, Qu Y, Ma X, *et al.* CXCL12/CXCR4 Axis Upregulates Twist to Induce EMT in Human Glioblastoma. *Mol Neurobiol* 2016 Aug; **53**(6)**:** 3948-3953.

52. Aggarwal R, Ghobrial IM, Roodman GD. Chemokines in multiple myeloma. *Exp Hematol* 2006 Oct; **34**(10)**:** 1289-1295.

53. Contento RL, Molon B, Boularan C, Pozzan T, Manes S, Marullo S, *et al.* CXCR4-CCR5: a couple modulating T cell functions. *Proceedings of the National Academy of Sciences of the United States of America* 2008 Jul 22; **105**(29)**:** 10101-10106.

54. Barille S, Akhoundi C, Collette M, Mellerin MP, Rapp MJ, Harousseau JL, *et al.* Metalloproteinases in multiple myeloma: production of matrix metalloproteinase-9 (MMP-9), activation of proMMP-2, and induction of MMP-1 by myeloma cells. *Blood* 1997 Aug 15; **90**(4)**:** 1649-1655.

55. Kelly T, Borset M, Abe E, Gaddy-Kurten D, Sanderson RD. Matrix metalloproteinases in multiple myeloma. *Leukemia & lymphoma* 2000 Apr; **37**(3-4)**:** 273-281.

56. Vacca A, Ria R, Semeraro F, Merchionne F, Coluccia M, Boccarelli A, *et al.* Endothelial cells in the bone marrow of patients with multiple myeloma. *Blood* 2003 Nov 01; **102**(9)**:** 3340-3348.

57. Vacca A, Ribatti D, Presta M, Minischetti M, Iurlaro M, Ria R, *et al.* Bone marrow neovascularization, plasma cell angiogenic potential, and matrix metalloproteinase-2 secretion parallel progression of human multiple myeloma. *Blood* 1999 May 01; **93**(9)**:** 3064-3073.

58. Vande Broek I, Asosingh K, Allegaert V, Leleu X, Facon T, Vanderkerken K, *et al.* Bone marrow endothelial cells increase the invasiveness of human multiple

myeloma cells through upregulation of MMP-9: evidence for a role of hepatocyte growth factor. *Leukemia* 2004 May; **18**(5)**:** 976-982.

59. Van Valckenborgh E, Bakkus M, Munaut C, Noel A, St Pierre Y, Asosingh K*, et al.* Upregulation of matrix metalloproteinase-9 in murine 5T33 multiple myeloma cells by interaction with bone marrow endothelial cells. *Int J Cancer* 2002 Oct 20; **101**(6)**:** 512-518.

60. Fowler JA, Mundy GR, Lwin ST, Lynch CC, Edwards CM. A murine model of myeloma that allows genetic manipulation of the host microenvironment. *Dis Model Mech* 2009 Nov-Dec; **2**(11-12)**:** 604-611.

61. Singh S, Mak IWY, Handa D, Ghert M. The Role of TWIST in Angiogenesis and Cell Migration in Giant Cell Tumor of Bone. *Advances in Biology* 2014; **2014:** 8.

62. Weiss MB, Abel EV, Mayberry MM, Basile KJ, Berger AC, Aplin AE. TWIST1 is an ERK1/2 effector that promotes invasion and regulates MMP-1 expression in human melanoma cells. *Cancer research* 2012 Dec 15; **72**(24)**:** 6382-6392.

63. Fu J, Li S, Feng R, Ma H, Sabeh F, Roodman GD*, et al.* Multiple myeloma-derived MMP-13 mediates osteoclast fusogenesis and osteolytic disease. *J Clin Invest* 2016 May 02; **126**(5)**:** 1759-1772.

64. Kudo Y, Iizuka S, Yoshida M, Tsunematsu T, Kondo T, Subarnbhesaj A*, et al.* Matrix metalloproteinase-13 (MMP-13) directly and indirectly promotes tumor angiogenesis. *The Journal of biological chemistry* 2012 Nov 09; **287**(46)**:** 38716-38728.

65. Eckert MA, Lwin TM, Chang AT, Kim J, Danis E, Ohno-Machado L*, et al.* Twist1-induced invadopodia formation promotes tumor metastasis. *Cancer Cell* 2011 Mar 08; **19**(3)**:** 372-386.

66. Azab AK, Hu J, Quang P, Azab F, Pitsillides C, Awwad R*, et al.* Hypoxia promotes dissemination of multiple myeloma through acquisition of epithelial to mesenchymal transition-like features. *Blood* 2012 June 14, 2012; **119**(24)**:** 5782-5794.

67. Wheelock MJ, Shintani Y, Maeda M, Fukumoto Y, Johnson KR. Cadherin switching. *J Cell Sci* 2008 Mar 15; **121**(Pt 6)**:** 727-735.

Chapter 5:

Discussion

5.1 General discussion

Multiple myeloma (MM) is the second most common haematological malignancy and is characterised by clonal proliferation of malignant plasma cells (PC) within the bone marrow (BM). While the development of proteasome inhibitors such as Bortezomib and immunomodulatory drugs, such as Lenalidomide, have significantly improved the clinical outcomes of myeloma patients, a subset of clinically and genetically high-risk patients are yet to significantly benefit from these therapeutic advancements.[1, 2, 3, 4] This highlights the need to determine the molecular mechanisms that underlie the aggressive progression of MM disease in these high-risk subtypes.

The t(4;14)(p16;q32) chromosomal translocation, characterised by overexpression of MMSET, is associated with adverse prognosis resulting from poor response to conventional chemotherapy, short period of remission and aggressive relapse.[5-8] Recent data has demonstrated that the incorporation of bortezomib into the treatment regime for t(4;14) patients improves survival outcome[9, 10] leading to this subtype of MM being reclassified as intermediate-risk cytogenetic disease.[11] Despite this, the presence of the t(4;14) translocation remains an adverse prognostic factor as these patients display shorter event-free survival and overall survival when compared with MM patients lacking the t(4;14) translocation[9], leading to the International Myeloma Working Group recommending that t(4;14) MM be considered a high risk subgroup, requiring aggressive treatment.[2, 12] As such, it is important to identify putative targets that are regulated by MMSET and to determine the downstream implications of MMSET overexpression and how it contributes to aggressive progression in t(4;14) MM patients.

In most cases of MM, multiple focal tumours are evident throughout the skeleton at diagnosis, suggesting that myeloma cells continuously traffic to distant BM niches, leading to PC dissemination and MM disease progression. Although the mechanisms underlying MM PC homing and egress to new BM niches remain to be fully elucidated, emerging evidence suggests that a process resembling the metastasis of epithelial cancers and involving the activation of EMT-related gene expression occurs in MM PC.[13] In the present study, we hypothesised that overexpression of MMSET activates the expression of EMT-related genes, which, in turn, promote MM PC metastasis, leading to aggressive disease progression in t(4;14) MM. To test this hypothesis, in

Chapter 2, GEPs of MM PC from 1012 MM patients, derived from 4 microarray studies were examined. In addition, the expression levels of EMT-related genes were examined by evaluation of RNA-seq data of 64 HMCL.

To date, GEP data has been used to molecularly classify and risk-stratify MM patients.[12] While gene expression data has not been widely used clinically for prognostic stratification of MM patients[14], their use in evaluating the transcriptional patterns associated with cytogenetic aberrations has provided much-needed insight into the pathogenesis of MM PCs.[15-19] In Chapter 2, meta-analysis was performed by combining patients GEPs from four independent microarray datasets using Fisher's inverse chi-square method.[20] Importantly, meta-analysis of these data served to increase the statistical power of the studies and thus enable more accurate detection of differentially expressed genes. Additionally, this approach enabled the averaging of conflicting results to be estimated. In contrast to previous studies that revealed an association of EMT-like process in MM based on protein expression of a relatively small number of EMT-related proteins[13, 21, 22], the present study has identified a larger set of MMSET-regulated mesenchymal genetic markers, which are independent of laboratory- or cohort-specific biases. To this end, we identified 17 mesenchymal markers that were significantly upregulated in MMSET-high patients and HMCLs compared with their MMSET-low counterparts, including the key EMT transcription activator TWIST1. Notably, no epithelial markers were found to be downregulated using this approach. While GSEA demonstrated a lack of enrichment of EMT-related genes in MMSET-high MM patients' GEPs, the transcriptomes of t(4;14) HMCL were significantly enriched for EMT-related genes. These differences could be accounted for by the relative homogeneity in gene expression between the different HMCL[23, 24] compared to subclonally heterogeneous primary MM cells.[4] Furthermore, the majority of HMCL have been obtained from patients at relapse or at advanced disease stages from extramedullary sites[23], and as such the transcriptomes of the t(4;14) HMCL may reflect gene expression changes that accompany extramedullary disease progression and/or clonal selection following *in vitro* culture.

Importantly, among the mesenchymal genes upregulated, the key EMT transcription activator TWIST1 was found to be overexpressed in approximately 50% of t(4;14) MM patients. Although TWIST1 was not shortlisted as a MMSET-regulated

gene in previous GEP studies[18, 24, 25], the use of meta-analysis in this study enabled the subtle elevation of TWIST1 expression in t(4;14) to be detected when compared with non-t(4;14) MM cases. Notably, the more pronounced fold-change of TWIST1 expression in t(4;14) HMCL further suggests that TWIST1 expression exerts clonal advantage to MM PCs allowing them to survive and proliferate in the absence of BM microenvironment cues. Notably, the results presented herein are consistent with a previous study, showing that overexpression of MMSET activates TWIST1 expression in prostate cancer cells promotes cell invasion and migration.[26] As such, the association between TWIST1 and MMSET highlights a potential role for TWIST1 in mediating MMSET-regulated, aggressive progression in t(4;14) MM.

Accumulating data suggests that the expression of TWIST1 represents a novel risk factor in multiple haematological malignancies and acts by inhibiting apoptosis[27, 28] and enhancing therapeutic resistance.[29, 30] Furthermore, the expression of TWIST1 initiates the metastasis of epithelial cancers by modulating transcription of a range of genes involve in cell migration, invasion and overcoming apoptosis. However, the functional roles of TWIST1 in MM development and progression have not been investigated. Therefore, understanding the transcriptomic changes that accompany TWIST1 expression in MM PC is essential to decipher functional changes driven by TWIST1 in MM development. This question was addressed in Chapter 3 and 4 using RNA-sequencing technology to identify putative targets regulated by TWIST1 in human and murine MM cells. RNA-sequencing is a quantitative sequence-based approach that allows the nucleotide sequence of cDNA species, reverse transcribed from RNA, to be evaluated. Compared with the hybridisation-based microarray technology, this sequence-based approach has a higher dynamic range, enabling the accurate detection of genes that are expressed at either very low or very high levels.[31] Although the number of genes commonly upregulated by TWIST1 in both human and murine cell lines were limited, functional analyses showed that the genes upregulated in both WL2-TWIST1 and 5TGM1-TWIST1 cells were enriched for cellular migration. However, whether TWIST1 mediates cell migration directly by binding to the E-box of target genes, or indirectly, by regulating other transcription factors requires further investigation.

The studies presented in this thesis are the first to examine the functional role of TWIST1 overexpression in MM PCs. Consistent with functional analyses, the *in vitro*

data presented herein, showed that TWIST1 overexpression increased MM cell migration towards CXCL12, a potent chemoattractant expressed by cells of the BM stroma, by stimulating F-actin polymerisation. Notably, complementary studies in which TWIST1 was transiently knocked down in the KMS11 HMCL by siRNA, were found to result in reduced MM cell migration. Although the molecular mechanisms and signalling pathways activated by TWIST1-mediated directional cell movement of MM PCs remains to be fully elucidated, pathway analysis from transcriptomic data and GTPase activation assays performed in HMCL suggest that Rac1 GTPase signalling may be activated, as previously reported in head and neck squamous cell carcinoma.[32]

The migration promoting effects of TWIST1, observed *in vitro*, were in accord with *in vivo* data presented in Chapter 4. To this end, inoculation of TWIST1-overexpressing 5TGM1 cells into mice increased overall tumour burden after 4 weeks, and was accompanied by an increased in the number of circulating tumour cells. These results are consistent with previous studies performed in other tumour models, which showed that overexpression of TWIST1 promotes tumour establishment and metastasis in breast cancer[33, 34], gastric carcinoma[35], glioma[36], oral squamous cell carcinoma[37] and pancreatic cancer.[38] Notably, *in vitro* data from Chapter 4 demonstrated that TWIST1 did not affect the colony forming capability and proliferation of 5TGM1 cells, suggesting that TWIST1 contributes to MM tumour development by promoting dissemination of MM PCs to distant BM niches enriched in CXCL12 as MM disease progresses.

Interestingly, a previous study by Xu et al. has showed that TWIST1 expression in AML cells promotes extramedullary infiltration leading to aggressive phenotype *in vivo*.[39] In agreement with these findings, the studies presented in Chapter 4 showed that TWIST1 increased tumour burden in the spleen *in vivo*, suggesting that TWIST1 confers survival advantage to MM PCs at extramedullary sites. The increased splenic tumour burden could be attributed to higher overall tumour burden as splenic infiltration is observed frequently in 5TGM1-C57Bl/KaLwRij model due to the murine spleens' capacity to support haematopoiesis and thus support MM growth.[40-42] However, our *in vitro* data showed that mitogenic responsiveness of TWIST1-overexpressing MM PC was enhanced when cultured with spleen cell conditioned-media but was unaffected when co-cultured with BM stromal cells. These results further support the role of

TWIST1 in mediating extramedullary growth of MM cells. Future *in vivo* studies utilising intratibial injection of MM cells directly into the BM could serve as a strategy to assess the metastatic capacity of TWIST1-expressing MM cells to intra- and extramedullary sites as the disease progresses.

Extramedullary progression in MM is often observed during relapsed and is associated with poor prognosis.[43, 44] While the precise mechanisms leading to extramedullary dissemination are yet to be fully understood, high risk cytogenetic cases of MM, including del 13q, del 17p and t(4;14) are frequently observed in patients with extramedullary progression.[45-48] A recent study by Yang et al.[49] showed that TWIST1 expression was significantly elevated in MM PCs derived from extramedullary tumour sites when compared with PCs from the BM of MM patient with soft tissues involvement. In view of this finding, we postulate that MMSET acts as an upstream activator of TWIST1 in t(4;14) MM, which subsequently confers survival advantage to MM PCs at extramedullary sites, leading to aggressive disease progression. Further longitudinal studies, examining the association of TWIST1 with MMSET levels and its implication in survival in bigger cohorts are required to determine the potential of TWIST1 to serve as prognostic marker for aggressive, extramedullary progression in t(4;14) MM.

5.2 Therapeutic implications of TWIST1 in MM

The established role of TWIST1 in mediating numerous cancer metastasis-promoting pathways provided a rationale to develop therapeutic strategies to target TWIST1. To this end, experimental silencing of TWIST1, using small interfering RNA (siRNA), has proven effective at overcoming tumour cell invasion, metastasis and drug resistance in epithelial cancers.[50] Although germ line deletion of TWIST1 in mice is lethal[51], inducible TWIST1 knockout in adult mice has no major side effects, with the exception that it accelerates hair growth.[52] As TWIST1 expression in adult humans is restricted to a subpopulation of mesenchymal derived tissues[53], off-target effects of TWIST1 inhibition are likely to be minimal. However, RNAi-based therapeutic targeting of transcription factors is technically challenging with respect to target site delivery and siRNA instability.[54] Furthermore, screening of pharmacological compounds that targets TWIST1 are still in their early stages and there are only a limited number of studies that have shown success with siRNA-based therapeutic targeting of TWIST1 in inhibiting xenograft tumour *in vivo*.[55, 56]

Notably, TWIST1 has been implicated in maintaining dormancy and self-renewal of hematopoietic stem cells (HSC)[57] with TWIST1 deletion in HSC significantly reduces their long term repopulating ability following transplantation into recipient mice.[57] Given that autologous stem-cell transplantation is a standard treatment for transplant-eligible MM patients[58], targeting TWIST1 may compromise the capacity for HSC to re-establish the hematopoietic system after transplantation and thus complicates the practicality of using TWIST1 inhibition as a treatment for MM.

While therapeutic targeting of TWIST1 in MM may be limited, development of inhibitors against its upstream activator, MMSET, has made significant progress. To this end, compounds that can molecular inhibit H3K36 catalytic activity, specific to MMSET, have been developed.[59-61] Among these compound, the N-alkyl sinefungin derivatives reported by Tisi and colleagues[60] demonstrated remarkable efficacy to inhibit the level of H3K36Me2 catalysed by MMSET on recombinant nucleosomes at low concentration and highlights the need for further investigation of the efficacy of these compound as anti-tumour therapeutics. Given the dependency of t(4;14) MM on MMSET expression for proliferation, survival and tumorigenicity, MMSET inhibitors are anticipated to potentiate the clinically efficacy of bortezomib as observed *in vivo*[62] to overcome poor prognosis of t(4;14) cytogenetic in MM.

5.3 Conclusion

Combining GEP analyses from newly diagnosed MM patients, *in vitro* migration assays and *in vivo* animal studies, the data presented in this thesis demonstrate a role for TWIST1 in stimulating F-actin polymerisation, thereby promoting MM PCs migration, and promoting metastasis of t(4;14) MM PCs, downstream of MMSET overexpression (Figure 5.1). While further studies are required, data presented herein suggest that TWIST1 confers an advantage to MM PCs to colonise extramedullary sites. Taken together, these data show that TWIST1-mediated MM PCs dissemination and extramedullary growth may, in part, explain the aggressive disease progression of t(4;14) MM and adds to our current understanding of the role of TWIST1 in haematological malignancies.

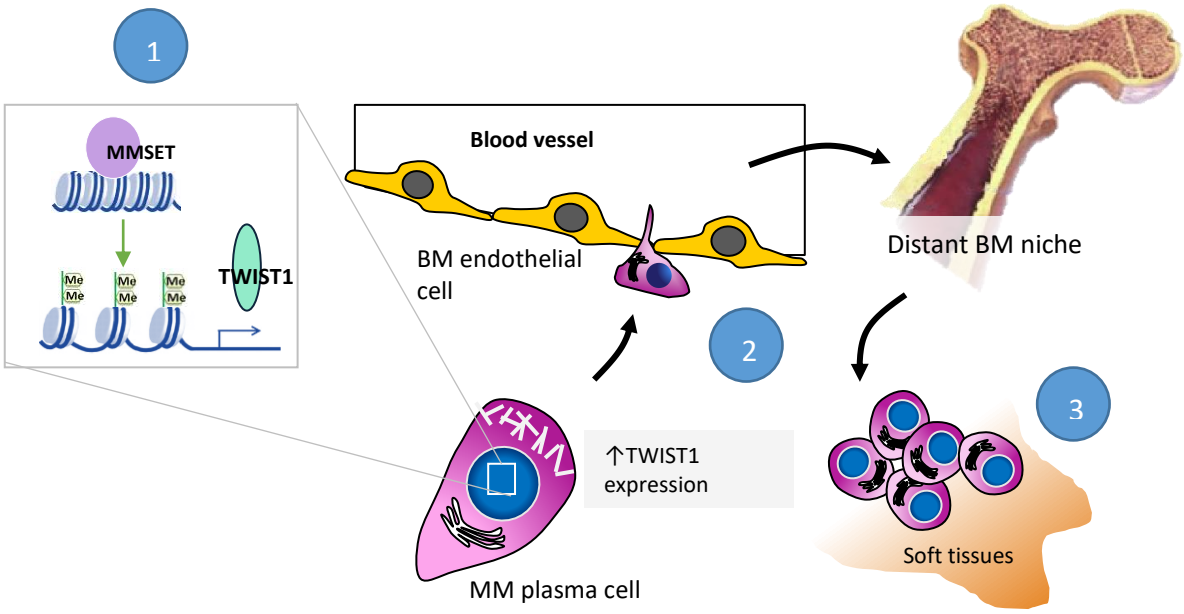


Figure 5.1. Actions of TWIST1 overexpression in MM PCs. (1) MMSET expression increases TWIST1 level in MM plasma cells. (2) TWIST1 regulates cytoskeleton organization to increase MM cell migration to distant BM niche. (3) TWIST1 stimulates MM growth in response to splenic-derived soluble factors.

5.4 References

1. Lonial S, Boise LH, Kaufman J. How I treat high-risk myeloma. *Blood* 2015 Sep 24; **126**(13)**:** 1536-1543.

2. Sonneveld P, Avet-Loiseau H, Lonial S, Usmani S, Siegel D, Anderson KC*, et al.* Treatment of multiple myeloma with high-risk cytogenetics: a consensus of the International Myeloma Working Group. *Blood* 2016 Jun 16; **127**(24)**:** 2955-2962.

3. Corre J, Munshi N, Avet-Loiseau H. Genetics of multiple myeloma: another heterogeneity level? *Blood* 2015 Mar 19; **125**(12)**:** 1870-1876.

4. Manier S, Salem KZ, Park J, Landau DA, Getz G, Ghobrial IM. Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol* 2016 Aug 17.

5. Jaksic W, Trudel S, Chang H, Trieu Y, Qi X, Mikhael J*, et al.* Clinical outcomes in t(4;14) multiple myeloma: a chemotherapy-sensitive disease characterized by rapid relapse and alkylating agent resistance. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 Oct 01; **23**(28)**:** 7069-7073.

6. Karlin L, Soulier J, Chandesris O, Choquet S, Belhadj K, Macro M*, et al.* Clinical and biological features of t(4;14) multiple myeloma: a prospective study. *Leukemia & lymphoma* 2011 Feb; **52**(2)**:** 238-246.

7. Chang H, Sloan S, Li D, Zhuang L, Yi QL, Chen CI*, et al.* The t(4;14) is associated with poor prognosis in myeloma patients undergoing autologous stem cell transplant. *Br J Haematol* 2004 Apr; **125**(1)**:** 64-68.

8. Stewart AK, Bergsagel PL, Greipp PR, Dispenzieri A, Gertz MA, Hayman SR*, et al.* A practical guide to defining high-risk myeloma for clinical trials, patient counseling and choice of therapy. *Leukemia* 2007 Mar; **21**(3)**:** 529-534.

9. Avet-Loiseau H, Leleu X, Roussel M, Moreau P, Guerin-Charbonnel C, Caillot D*, et al.* Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010 Oct 20; **28**(30)**:** 4630-4634.

10. Cavo M, Tacchetti P, Patriarca F, Petrucci MT, Pantani L, Galli M*, et al.* Bortezomib with thalidomide plus dexamethasone compared with thalidomide plus dexamethasone as induction therapy before, and consolidation therapy after, double autologous stem-cell transplantation in newly diagnosed multiple myeloma: a randomised phase 3 study. *Lancet* 2010 Dec 18; **376**(9758)**:** 2075-2085.

11. Kumar SK, Mikhael JR, Buadi FK, Dingli D, Dispenzieri A, Fonseca R, *et al.* Management of newly diagnosed symptomatic multiple myeloma: updated Mayo Stratification of Myeloma and Risk-Adapted Therapy (mSMART) consensus guidelines. *Mayo Clin Proc* 2009 Dec; **84**(12)**:** 1095-1110.

12. Chng WJ, Dispenzieri A, Chim CS, Fonseca R, Goldschmidt H, Lentzsch S, *et al.* IMWG consensus on risk stratification in multiple myeloma. *Leukemia* 2014 Feb; **28**(2)**:** 269-277.

13. Azab AK, Hu J, Quang P, Azab F, Pitsillides C, Awwad R, *et al.* Hypoxia promotes dissemination of multiple myeloma through acquisition of epithelial to mesenchymal transition-like features. *Blood* 2012 June 14, 2012; **119**(24)**:** 5782-5794.

14. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. *Nature reviews Cancer* 2012 Apr 12; **12**(5)**:** 335-348.

15. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 Sep 10; **23**(26)**:** 6333-6338.

16. Chng WJ, Glebov O, Bergsagel PL, Kuehl WM. Genetic events in the pathogenesis of multiple myeloma. *Best practice & research Clinical haematology* 2007 Dec; **20**(4)**:** 571-596.

17. Shaughnessy JD, Jr., Zhan F, Burington BE, Huang Y, Colla S, Hanamura I, *et al.* A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* 2007 Mar 15; **109**(6)**:** 2276-2284.

18. Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S, *et al.* The molecular classification of multiple myeloma. *Blood* 2006 Sep 15; **108**(6)**:** 2020-2028.

19. Mattioli M, Agnelli L, Fabris S, Baldini L, Morabito F, Bicciato S, *et al.* Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma. *Oncogene* 2005 Apr 07; **24**(15)**:** 2461-2473.

20. Campain A, Yang YH. Comparison study of microarray meta-analysis methods. *BMC Bioinformatics* 2010 Aug 03; **11:** 408.

21. Li J, Pan Q, Rowan PD, Trotter TN, Peker D, Regal KM, *et al.* Heparanase promotes myeloma progression by inducing mesenchymal features and motility of myeloma cells. *Oncotarget* 2016 Mar 8; **7**(10)**:** 11299-11309.

22. Sun Y, Pan J, Mao S, Jin J. IL-17/miR-192/IL-17Rs regulatory feedback loop facilitates multiple myeloma progression. *PloS one* 2014; **9**(12)**:** e114647.

23. Drexler HG, Matsuo Y. Malignant hematopoietic cell lines: in vitro models for the study of multiple myeloma and plasma cell leukemia. *Leuk Res* 2000 Aug; **24**(8)**:** 681-703.

24. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A*, et al.* Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* 2010 Oct 7; **116**(14)**:** 2543-2553.

25. Wu SP, Pfeiffer RM, Ahn IE, Mailankody S, Sonneveld P, van Duin M*, et al.* Impact of Genes Highly Correlated with MMSET Myeloma on the Survival of Non-MMSET Myeloma Patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2016 Aug 15; **22**(16)**:** 4039-4044.

26. Ezponda T, Popovic R, Shah MY, Martinez-Garcia E, Zheng Y, Min DJ*, et al.* The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. *Oncogene* 2012 Jul 16.

27. Li X, Marcondes AM, Gooley TA, Deeg HJ. The helix-loop-helix transcription factor TWIST is dysregulated in myelodysplastic syndromes. *Blood* 2010 Sep 30; **116**(13)**:** 2304-2314.

28. Kim MS, Kim GM, Choi YJ, Kim HJ, Kim YJ, Jin W. TrkC promotes survival and growth of leukemia cells through Akt-mTOR-dependent up-regulation of PLK-1 and Twist-1. *Mol Cells* 2013 Aug; **36**(2)**:** 177-184.

29. Cosset E, Hamdan G, Jeanpierre S, Voeltzel T, Sagorny K, Hayette S*, et al.* Deregulation of TWIST-1 in the CD34+ compartment represents a novel prognostic factor in chronic myeloid leukemia. *Blood* 2011 Feb 3; **117**(5)**:** 1673-1676.

30. Zhang J, Wang P, Wu F, Li M, Sharon D, Ingham RJ*, et al.* Aberrant expression of the transcriptional factor Twist1 promotes invasiveness in ALK-positive anaplastic large cell lymphoma. *Cell Signal* 2012 Apr; **24**(4)**:** 852-858.

31. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009 Jan; **10**(1)**:** 57-63.

32. Yang WH, Lan HY, Huang CH, Tai SK, Tzeng CH, Kao SY*, et al.* RAC1 activation mediates Twist1-induced cancer cell migration. *Nature cell biology* 2012 Mar 11; **14**(4)**:** 366-374.

33. Croset M, Goehrig D, Frackowiak A, Bonnelye E, Ansieau S, Puisieux A*, et al.* TWIST1 expression in breast cancer cells facilitates bone metastasis formation. *J Bone Miner Res* 2014 Aug; **29**(8)**:** 1886-1899.

34. Mironchik Y, Winnard PT, Jr., Vesuna F, Kato Y, Wildes F, Pathak AP*, et al.* Twist overexpression induces in vivo angiogenesis and correlates with

chromosomal instability in breast cancer. *Cancer research* 2005 Dec 1; **65**(23)**:** 10801-10809.

35. Luo GQ, Li JH, Wen JF, Zhou YH, Hu YB, Zhou JH. Effect and mechanism of the Twist gene on invasion and metastasis of gastric carcinoma cells. *World J Gastroenterol* 2008 Apr 28; **14**(16)**:** 2487-2493.

36. Mikheeva SA, Mikheev AM, Petit A, Beyer R, Oxford RG, Khorasani L*, et al.* TWIST1 promotes invasion through mesenchymal change in human glioblastoma. *Mol Cancer* 2010 Jul 20; **9:** 194.

37. da Silva SD, Alaoui-Jamali MA, Soares FA, Carraro DM, Brentani HP, Hier M*, et al.* TWIST1 is a molecular marker for a poor prognosis in oral cancer and represents a potential therapeutic target. *Cancer* 2014 Feb 01; **120**(3)**:** 352-362.

38. Chen S, Chen JZ, Zhang JQ, Chen HX, Yan ML, Huang L*, et al.* Hypoxia induces TWIST-activated epithelial-mesenchymal transition and proliferation of pancreatic cancer cells in vitro and in nude mice. *Cancer Lett* 2016 Dec 01; **383**(1)**:** 73-84.

39. Xu J, Zhang W, Yan XJ, Lin XQ, Li W, Mi JQ*, et al.* DNMT3A mutation leads to leukemic extramedullary infiltration mediated by TWIST1. *J Hematol Oncol* 2016 Oct 10; **9**(1)**:** 106.

40. Oyajobi BO, Munoz S, Kakonen R, Williams PJ, Gupta A, Wideman CL*, et al.* Detection of myeloma in skeleton of mice by whole-body optical fluorescence imaging. *Mol Cancer Ther* 2007 Jun; **6**(6)**:** 1701-1708.

41. Radl J, Croese JW, Zurcher C, Van den Enden-Vieveen MH, de Leeuw AM. Animal model of human disease. Multiple myeloma. *Am J Pathol* 1988 Sep; **132**(3)**:** 593-597.

42. Vanderkerken K, De Raeve H, Goes E, Van Meirvenne S, Radl J, Van Riet I*, et al.* Organ involvement and phenotypic adhesion profile of 5T2 and 5T33 myeloma cells in the C57BL/KaLwRij mouse. *British journal of cancer* 1997; **76**(4)**:** 451-460.

43. Usmani SZ, Rodriguez-Otero P, Bhutani M, Mateos MV, Miguel JS. Defining and treating high-risk multiple myeloma. *Leukemia* 2015 Nov; **29**(11)**:** 2119-2125.

44. Varettoni M, Corso A, Pica G, Mangiacavalli S, Pascutto C, Lazzarino M. Incidence, presenting features and outcome of extramedullary disease in multiple myeloma: a longitudinal study on 1003 consecutive patients. *Ann Oncol* 2010 Feb; **21**(2)**:** 325-330.

45. Rasche L, Bernard C, Topp MS, Kapp M, Duell J, Wesemeier C*, et al.* Features of extramedullary myeloma relapse: high proliferation, minimal marrow involvement, adverse cytogenetics: a retrospective single-center study of 24 cases. *Ann Hematol* 2012 Jul; **91**(7)**:** 1031-1037.

46. Besse L, Sedlarikova L, Greslikova H, Kupska R, Almasi M, Penka M, *et al.* Cytogenetics in multiple myeloma patients progressing into extramedullary disease. *Eur J Haematol* 2016 Jul; **97**(1)**:** 93-100.

47. Billecke L, Murga Penas EM, May AM, Engelhardt M, Nagler A, Leiba M, *et al.* Cytogenetics of extramedullary manifestations in multiple myeloma. *Br J Haematol* 2013 Apr; **161**(1)**:** 87-94.

48. Qu X, Chen L, Qiu H, Lu H, Wu H, Qiu H, *et al.* Extramedullary manifestation in multiple myeloma bears high incidence of poor cytogenetic aberration and novel agents resistance. *Biomed Res Int* 2015; **2015:** 787809.

49. Yang JZ, Lian WG, Sun LX, Qi DW, Ding Y, Zhang XH. High nuclear expression of Twist1 in the skeletal extramedullary disease of myeloma patients predicts inferior survival. *Pathol Res Pract* 2016 Mar; **212**(3)**:** 210-216.

50. Qin Q, Xu Y, He T, Qin C, Xu J. Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. *Cell research* 2012 Jan; **22**(1)**:** 90-106.

51. Chen ZF, Behringer RR. twist is required in head mesenchyme for cranial neural tube morphogenesis. *Genes & development* 1995 Mar 15; **9**(6)**:** 686-699.

52. Xu Y, Xu Y, Liao L, Zhou N, Theissen SM, Liao XH, *et al.* Inducible knockout of Twist1 in young and adult mice prolongs hair growth cycle and has mild effects on general health, supporting Twist1 as a preferential cancer target. *Am J Pathol* 2013 Oct; **183**(4)**:** 1281-1292.

53. Wang SM, Coljee VW, Pignolo RJ, Rotenberg MO, Cristofalo VJ, Sierra F. Cloning of the human twist gene: its expression is retained in adult mesodermally-derived tissues. *Gene* 1997 Mar 10; **187**(1)**:** 83-92.

54. Redell MS, Tweardy DJ. Targeting transcription factors in cancer: Challenges and evolving strategies. *Drug Discov Today Technol* 2006 Autumn; **3**(3)**:** 261-267.

55. Finlay J, Roberts CM, Dong J, Zink JI, Tamanoi F, Glackin CA. Mesoporous silica nanoparticle delivery of chemically modified siRNA against TWIST1 leads to reduced tumor burden. *Nanomedicine* 2015 Oct; **11**(7)**:** 1657-1666.

56. Finlay J, Roberts CM, Lowe G, Loeza J, Rossi JJ, Glackin CA. RNA-based TWIST1 inhibition via dendrimer complex to reduce breast cancer cell metastasis. *Biomed Res Int* 2015; **2015:** 382745.

57. Dong CY, Liu XY, Wang N, Wang LN, Yang BX, Ren Q, *et al.* Twist-1, a novel regulator of hematopoietic stem cell self-renewal and myeloid lineage development. *Stem Cells* 2014 Dec; **32**(12)**:** 3173-3182.

58. Kumar S, Giralt S, Stadtmauer EA, Harousseau JL, Palumbo A, Bensinger W, *et al.* Mobilization in myeloma revisited: IMWG consensus perspectives on stem cell collection following initial therapy with thalidomide-, lenalidomide-, or bortezomib-containing regimens. *Blood* 2009 Aug 27; **114**(9)**:** 1729-1735.

59. di Luccio E. Inhibition of Nuclear Receptor Binding SET Domain 2/Multiple Myeloma SET Domain by LEM-06 Implication for Epigenetic Cancer Therapies. *J Cancer Prev* 2015 Jun; **20**(2)**:** 113-120.

60. Tisi D, Chiarparin E, Tamanini E, Pathuri P, Coyle JE, Hold A, *et al.* Structure of the Epigenetic Oncogene MMSET and Inhibition by N-Alkyl Sinefungin Derivatives. *ACS Chem Biol* 2016 Nov 18; **11**(11)**:** 3093-3105.

61. Will CM, Scholle M, St. Pierre R, Schiltz G, Bradner JE, Mrksich M, *et al.* High Throughput Screening Identifies Potential Inhibitors of WHSC1/MMSET, a Histone Methyltransferase Oncoprotein in Multiple Myeloma and Acute Lymphocytic Leukemia. *Blood* 2015; **126**(23)**:** 3251-3251.

62. Xie Z, Bi C, Chooi JY, Chan ZL, Mustafa N, Chng WJ. MMSET regulates expression of IRF4 in t(4;14) myeloma and its silencing potentiates the effect of bortezomib. *Leukemia* 2015 Dec; **29**(12)**:** 2347-2354.