

Investigating the role of WWP2 Ubiquitin  
Ligase isoforms in TGF $\beta$ -dependent  
oncogenic signalling

---

HIU TUNG TIFFANY YIM

A thesis submitted for the degree of Doctor of Philosophy at the  
School of Biological Sciences, University of East Anglia.

**September 2017**

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

## Abstract

Ubiquitination is a post-translational modification involving the attachment of ubiquitin molecules onto target substrates. The most common function of ubiquitination is proteasomal degradation, which is commonly facilitated by the ubiquitination of target proteins at Lys-48. E3 ligases are crucial to the process of ubiquitination, and include WWP2 ubiquitin Ligase, an E3 enzyme that has been shown to interact with Smad mediators of the TGF $\beta$  family and might subsequently have a role in TGF $\beta$ -dependent oncogenic processes such as EMT. To elucidate the function of WWP2 in TGF $\beta$  signalling and increase our understanding of this protein, is therefore the forefront of our research. Here, we found that the overexpression of WWP2 WW3-4 recognition domain led to an increase in Smad3-dependent TGF $\beta$  gene expression, providing strong evidence that in the absence of TGF $\beta$ , WW3-4 binds to the Smad3 mediator. We also predicted and validated the expression of two novel WWP2 isoforms, WWP2- $\Delta$ HECT and WWP2-N- $\Delta$ C2. Results suggest that the incomplete HECT domain within WWP2- $\Delta$ HECT is catalytically non-functional. Furthermore, the transcription of WWP2- $\Delta$ HECT has been linked to a 2kb region within intron 9/10 with its expression positively regulated by EGF stimulation. On the contrary, when the transcriptional mechanism of the existing WWP2-C isoform was investigated, a 0.5kb region at the end of intron 10/11 was identified to enhance promoter activity responsible for expression, which was negatively regulated by the SOX9 transcription factor. It was also found that ESRP splicing factor is involved in the negative regulation of not only the existing isoform WWP2-N, but also the novel WWP2- $\Delta$ HECT isoform. Using expression studies, we found that WWP2-N and  $\Delta$ HECT were expressed at higher levels in epithelial cells, therefore suggesting their role as guardians of the epithelial phenotype during EMT. Overall, the data provided here can help clarify the role of WWP2 in TGF $\beta$ -dependent oncogenic signalling.

## Table of Contents

ABSTRACT .....	II
TABLE OF CONTENTS .....	III
ACKNOWLEDGMENT.....	VII
LIST OF ABBREVIATIONS .....	VIII
LIST OF FIGURES .....	XII
LIST OF TABLES .....	XVIII
<b>CHAPTER 1 INTRODUCTION .....</b>	<b>1</b>
1.1 INTRODUCTION .....	2
1.2 UBIQUITINATION .....	4
1.2.1 <i>The mechanism of Ubiquitination</i> .....	4
1.2.2 <i>Deubiquitination</i> .....	8
1.2.3 <i>Functions of ubiquitination</i> .....	9
1.3 NEDD4 UBIQUITIN LIGASES .....	12
1.3.1 <i>Structure of WWP2 and its isoforms</i> .....	15
1.3.2 <i>WWP2 and its role in signalling pathways</i> .....	18
1.4 TGF $\beta$ PATHWAY .....	21
1.4.1 <i>The canonical pathway of TGF<math>\beta</math> signalling</i> .....	21
1.4.2 <i>Regulation of the TGF<math>\beta</math> canonical pathway</i> .....	25
1.4.3 <i>TGF<math>\beta</math>, BMP and Cancer</i> .....	27
1.5 CANCER AND MALIGNANT MELANOMA .....	30
1.5.1 <i>Cancer progression and EMT</i> .....	31
1.5.2 <i>Role of Ubiquitin ligases in Cancer</i> .....	33
1.6 AIMS .....	36
<b>CHAPTER 2 METHODS AND MATERIALS .....</b>	<b>37</b>
2.1 CLONING .....	38
2.1.1 <i>Genomic DNA extraction</i> .....	39
2.1.2 <i>High fidelity PCR</i> .....	39
2.1.3 <i>Gel extraction</i> .....	41
2.1.4 <i>Restriction enzyme digestion</i> .....	42
2.1.5 <i>TA Cloning</i> .....	42

2.1.6	Standard T4 ligation .....	43
2.1.7	In-Fusion HD Cloning .....	44
2.1.8	Heat-Shock transformation into DH5 $\alpha$ and other competent cells.....	44
2.1.9	Plasmid isolation .....	44
2.1.10	DNA quantification and sequencing .....	45
2.1.11	Preparation of DH5 $\alpha$ competent cell stocks .....	45
2.1.12	Site-directed mutagenesis .....	46
2.2	CELL CULTURE .....	47
2.2.1	Cell passage .....	47
2.2.2	Preparation of cell stocks to freeze .....	48
2.2.3	TGF $\beta$ , BMP and EGF stimulation of cells.....	48
2.3	IMMUNOLABELLING AND FLUORESCENCE MICROSCOPY .....	49
2.3.1	PEI Transfection .....	49
2.3.2	Immunolabelling.....	49
2.4	PROTEIN ANALYSIS.....	51
2.4.1	PEI Transfection and co-transfections with p-CAGAC12-luc Luciferase reporter .....	51
2.4.2	Cell Harvest of total cell lysates.....	51
2.4.3	Streptavidin pull-downs.....	52
2.4.4	SDS Polyacrylamide gel electrophoresis .....	52
2.4.5	Protein preparation for mass spectrometry .....	53
2.4.6	Western blotting and dot blotting.....	54
2.4.7	Luciferase assays, protein harvest and statistical analysis.....	56
2.5	RNA ANALYSIS .....	57
2.5.1	RNA Isolation .....	57
2.5.2	cDNA Synthesis .....	57
2.5.3	Semi-quantitative PCR.....	58
2.5.4	Agarose gel electrophoresis .....	59
2.5.5	Taqman Quantitative PCR .....	60

### **CHAPTER 3 EXPLORING THE BIOLOGICAL ROLES OF WWP2 WW**

<b>DOMAINS.....</b>	<b>62</b>
3.1 INTRODUCTION .....	63

3.2	RESULTS .....	65
3.2.1	<i>Studying the effects of WWP2 WW domains on Smad3-dependent gene expression using the p-CAGAC12-luc reporter vector in luciferase assays.....</i>	65
3.2.2	<i>The study of WWP2 WW3-4 tandem domain and associated proteins using Quantitative Mass spectrometry .....</i>	76
3.3	DISCUSSION.....	79
3.3.1	<i>Investigating WWP2 WW domain and Smad3-dependent gene expression.....</i>	79
3.3.2	<i>Effects of TGF<math>\beta</math> on WWP2 WW3-4 protein associations .....</i>	87
 <b>CHAPTER 4 PREDICTION AND VALIDATION OF NOVEL AND EXISTING WWP2 ISOFORMS .....</b>		
<b>93</b>		
3.4	INTRODUCTION .....	94
3.5	RESULTS .....	98
3.5.1	<i>Identification of novel isoforms using bioinformatics.....</i>	98
3.5.2	<i>Characterizing the expression pattern of existing and novel WWP2 isoforms.....</i>	109
4.3	DISCUSSION.....	125
4.3.1	<i>Identification and validation of novel WWP2 isoforms .....</i>	125
4.3.2	<i>Characterizing expression of WWP2 isoforms.....</i>	126
 <b>CHAPTER 5 DECIPHERING THE REGULATION AND FUNCTIONAL ROLES OF WWP2 ISOFORMS .....</b>		
<b>135</b>		
5.1	INTRODUCTION .....	136
5.2	RESULTS .....	139
5.2.1	<i>Transcriptional regulation of WWP2-C and WWP2-<math>\Delta</math>HECT... ..</i>	139
5.2.2	<i>Studying the effects of WWP2-<math>\Delta</math>HECT on Smad-3 dependent gene expression using p-CAGAC12-luc reporter in luciferase assays. ....</i>	155
5.2.3	<i>Investigating ESRPs as a post-transcriptional regulator of WWP2-N expression.....</i>	158
5.2.4	<i>Cellular localization of WWP2 isoforms.....</i>	162
5.3	DISCUSSION.....	171

5.3.1	<i>Regulation of WWP2-C promoter region and the relevance of SOX9 transcription factor</i> .....	171
5.3.2	<i>Investigating the regulation of WWP2-ΔHECT promoter activity and the effects of growth factor stimulation</i> .....	178
5.3.3	<i>The effects of WWP2-ΔHECT overexpression on Smad3-dependent gene expression</i> .....	181
5.3.4	<i>Studying the role of ESRP1 and 2 as a post-transcriptional mechanism in WWP2-N production</i> .....	184
5.3.5	<i>Cellular localization of WWP2 isoforms and the hypothesized Smad interactions</i> .....	185
<b>CHAPTER 6</b>	<b>DISCUSSION</b> .....	<b>192</b>
6.1	DISCUSSION.....	193
6.2	FUNCTIONAL IMPLICATIONS OF WWP2 WW DOMAINS .....	194
6.3	CHARACTERIZATION OF NOVEL AND ESTABLISHED WWP2 ISOFORMS .....	202
6.4	THE ROLE OF WWP2 ISOFORMS IN EMT .....	210
6.5	OVERALL EFFECTS OF TGFβ AND BMP ON ONCOGENESIS IN MELANOMA ...	212
6.6	FUTURE WORK .....	214
<b>REFERENCES</b>	.....	<b>216</b>
<b>APPENDIX</b>	.....	<b>237</b>

## Acknowledgment

First and foremost, I would like to thank my primary supervisor Dr Andrew Chantry for providing me with guidance and his scientific knowledge, not only through my 3 years in the lab, but especially through the write up period, where he has so kindly corrected drafts upon drafts of my thesis. Secondly, I would like to thank my secondary supervisor Dr Samuel Fountain for evaluating my progress through the duration of my PhD and providing constructive criticism where needed. I would also like to thank Dr Mette Mogensen, her lab and Dr Paul Thomas for sharing their wealth of knowledge in immunolabelling and immunofluorescence microscopy. Special thanks also to Dr Rosemary Davidson for helping me with all my qPCR needs, Dr Gerhard Salbaach for answering all my mass spec queries, Dr Simon Moxon for helping me with bioinformatic problems, and the Nieto lab for our collaboration in WWP2 expression studies.

Throughout my PhD life, there are people who I have worked with on a daily basis, that I now consider more a friend rather than just colleagues. Dr Lloyd Wahl for one, has been there since the very beginning and has answered, if not mocked all my stupid questions. I thank you for making lab life so entertaining. I would also like to thank Jess Watt and Greg Hughes (Drs to be), for helping me around the lab and being good lab mates. And to Dr Janice Layhadi, thank you for all the help you've given me and for being one of the best friends I have made during these four years.

Finally, my family and my partner Bradley has played a big part in my emotional and physical wellbeing for the past 4 years. I would like to thank my family for continually supporting me through the difficult times and providing me with motivation when I most needed it. I would also like to thank Bradley, who still loves me and cooks me nice food even when thesis writing has driven me to a state of complete laziness. My PhD life has been a long journey of which would not have been possible without all the people I have mentioned above, and for that, I thank everyone of you wholeheartedly.

## List of abbreviations

aa - Amino acid

AKT - Protein Kinase B

AMP - Adenosine monophosphate

ATP - Adenosine triphosphate

BMP - Bone Morphogenic protein

C2 domain- Calcium binding domain

CCR5 - C-C motif chemokine receptor 5

CDK - Cyclin-dependent kinase

CYLD - Ubiquitin Carboxyl-Terminal hydrolase

CTT - Chaperonin-containing t-complex polypeptide 1

CXCR3 - C-X-C motif chemokine receptor 3

DAPI - 4',6-diamidino-2-phenylindole

DEP domain - Dishevelled, Egl-10 and Pleckstrin domain

DMSO - Dimethyl sulfoxide

DNA - Deoxyribose nucleic acid

DPBS - Dulbecco's Phosphate-buffered Saline

DUB - Deubiquitinating enzymes

Dsh - Dishevelled

E-cadherin - Epithelial cadherin

ECM - Extracellular matrix

EDTA - Ethylene diamine tetra-acetic acid

EGF - Epidermal growth factor

EGFR - Epidermal growth factor receptor

EMT - Epithelial to mesenchymal transition

ENaC - Epithelial Sodium Channel

ERK - Extracellular signal-regulated kinase

ESRP - Epithelial splicing regulatory protein

EST- Expressed sequence tag

FCS - Foetal calf serum

FL - Full length

GC-SBE - GC-rich Smad Binding Element

GF - Growth Factor  
Grb2 - Growth factor receptor-bound protein 2  
H2A - Histone 2A  
HDAC - Histone Deacetylase  
HECT E3 - Homologous to E6AP Carboxy terminus E3  
HERC E3 - HECT and RLD domains containing E3  
I-Smad - Inhibitory-Smad  
IL - Interleukin  
ITCH - Itchy E3 ubiquitin ligase  
JAMM - JAB1/MPN/Mov34 metalloenzyme  
LAP - Latency-associated protein  
LB - Lysogeny Broth  
Lod score - Log odd score  
LTBP - Latent TGF $\beta$ -binding protein  
MAPK - Mitogen-activated Protein Kinase  
MCC - Colorectal mutant cancer protein  
Mdm2 - Mouse double minute 2 homolog  
MEK - Mitogen-activated Protein Kinase Kinase  
MET - Mesenchymal to epithelial transition  
MH1 - Mad-homology 1  
MH2 - Mad-homology 2  
MINDY- Motif-interacting with ubiquitin-containing novel DUB family  
MITF - Microphthalmia-associated transcription factor  
MJD - Machado-Joseph Disease Protease  
mTOR - Mechanistic Target of Rapamycin  
N-Cad - Neural cadherin  
NCC - Neural crest cells  
NEAA - Non-essential amino acid  
NEDD4 - Neural precursor cell-expressed developmental  
downregulated gene 4  
NES - Nuclear export signal  
NHLF Cells - Normal Human Lung Fibroblast cells  
NICD3 - Notch 3

NK cells - Natural killer cells  
NLS - Nuclear localization signal  
NMR - Nuclear magnetic resonance  
NMuMG Cells- Non-transformed Mouse Mammary Gland epithelial cells  
OD - Optical density  
OTU - Otubain Protease  
PAGE - Polyacrylamide Gel Electrophoresis  
PAI-1 - Plasminogen activator inhibitor-1  
PBS - Phosphate buffered saline  
PCAF - P300/CBP-associated factor  
PCNA - Proliferating-cell nuclear antigen  
PCR - Polymerase chain reaction  
PDGFR - Platelet-derived growth factor receptor  
PDK1 - Phosphoinositide-dependent kinase-1  
PEI - Polyethylenimine  
Pen/Strep - Penicillin/Streptomycin  
PI3K - Phosphoinositide 3-kinase  
PIP2 - Phosphatidylinositol 4,5-bisphosphate  
PIP3 - Phosphatidylinositol 3,4,5-trisphosphate  
PTEN - Phosphatase and Tensin Homolog  
R-Smad - Regulatory-Smad  
Rb - Retinoblastoma  
RING E3 - RING finger-containing E3  
ROS - Reactive oxygen species  
RTK - Receptor Tyrosine Kinase  
SARA - Smad anchor for receptor activation protein  
SBE - Smad binding element  
SCC - Squamous Cell Carcinoma  
SDS - Sodium Dodecyl Sulphate  
Shc - Src homology domain 2 containing  
SLUG - Zinc finger protein SNAI2  
SMURF - Smad ubiquitin regulatory factor  
SNAIL - Zinc finger protein SNAI1

Sos - Son of sevenless homolog  
Src - Proto-oncogene tyrosine protein kinase  
SRN - Splicing Regulatory Network  
SS-B - Lupus La protein  
TAE Buffer - Tris acetate EDTA Buffer  
TEMED - Tetramethylethylenediamine  
TF - transcription factor  
TGF $\beta$  - Transforming growth factor  $\beta$   
TGF $\beta$ R - Transforming growth factor  $\beta$  receptor  
U-BOX - UFD2 homology  
Ub - Ubiquitin  
UBA1 - Ubiquitin-activating enzyme E1  
UBC - ubiquitin-conjugating  
UCH - Ubiquitin C-terminal Hydrolase  
uH2A - Ubiquitinated Histone 2A  
USP - Ubiquitin-specific Protease  
VEGF - Vascular endothelial growth factor  
WW domain - tryptophan-tryptophan recognition domain  
WWP1 - WW domain containing E3 ubiquitin protein ligase 1  
WWP2 - WW domain containing E3 ubiquitin protein ligase 2

## List of Figures

FIGURE 1.2.1- THE E1 UBIQUITIN-ACTIVATING ENZYME AND ITS ISOFORMS.....	5
FIGURE 1.2.2- A DIAGRAM TO SUMMARIZE THE PROCESS OF UBIQUITINATION. ....	6
FIGURE 1.2.3- A DIAGRAM SHOWING THE SUB-FAMILIES OF E3 UBIQUITIN LIGASES .....	7
FIGURE 1.2.4- THE SIMPLIFIED REPRESENTATION OF THE UBIQUITIN-DEPENDENT 26S PROTEASOMAL DEGRADATION MECHANISM .....	10
FIGURE 1.2.5- A SUMMARY DEPICTING DIFFERENT TYPES OF UBIQUITIN POST- TRANSLATIONAL MODIFICATION .....	11
FIGURE 1.3.1- THE NEDD4 FAMILY OF E3 UBIQUITIN LIGASES .....	13
FIGURE 1.3.2- A DIAGRAM OF THE WWP2 GENE LOCUS AND EXPRESSED ISOFORMS....	17
FIGURE 1.3.3 - THE AMINO ACID SEQUENCE OF ESTABLISHED WWP2 ISOFORMS .....	18
FIGURE 1.3.4- EFFECTS OF WWP2-N/FL BINDING AND TGF $\beta$ STIMULATION ON SMAD3 LEVELS.....	20
FIGURE 1.4.1- A SIMPLIFIED SCHEMATIC SHOWING THE DOMAINS AND KEY MOTIFS PRESENT IN THE SMAD FAMILY OF TGF $\beta$ SIGNALLING MEDIATORS .....	22
FIGURE 1.4.2- A DIAGRAM DEPICTING THE CANONICAL TGF $\beta$ PATHWAY .....	24
FIGURE 1.5.1- THE PROCESS OF EMT AND ANGIOGENESIS IN ONCOGENESIS .....	32
FIGURE 2.1.1- THE PRK5 PLASMID MAP .....	38
FIGURE 2.1.2- THE PGL4.27 PLASMID MAP .....	39
FIGURE 2.1.3- A PLASMID MAP SHOWING THE PGEM-T EASY VECTOR .....	43
FIGURE 3.2.1- PROTEIN EXPRESSION OF WWP2 WW TANDEM DOMAINS AND SMAD7 IN HEK293A CELLS .....	66
FIGURE 3.2.2- THE EFFECTS OF TGF $\beta$ ON SMAD3-DEPENDENT GENE EXPRESSION OF WWP2 WW TANDEM DOMAIN TRANSFECTED HEK293A.....	67
FIGURE 3.2.3- THE EFFECTS OF TGF $\beta$ ON SMAD3-DEPENDENT GENE EXPRESSION OF WWP2 WW TANDEM DOMAIN TRANSFECTED A375 CELLS .....	68
FIGURE 3.2.4- PROTEIN EXPRESSION OF WWP2 WW INDIVIDUAL DOMAINS IN HEK293A.....	69
FIGURE 3.2.5- THE EFFECTS OF TGF $\beta$ ON SMAD3-DEPENDENT GENE EXPRESSION OF WWP2 WW INDIVIDUAL DOMAIN TRANSFECTED HEK293A.....	70
FIGURE 3.2.6- THE EFFECTS OF TGF $\beta$ ON SMAD3-DEPENDENT GENE EXPRESSION OF WWP2 WW INDIVIDUAL DOMAIN TRANSFECTED A375 .....	70

<b>FIGURE 3.2.7- THE EFFECTS OF TGF<math>\beta</math> STIMULATION ON WW3 AND WW4 DOMAIN TRANSFECTED A375 CELLS .....</b>	<b>72</b>
<b>FIGURE 3.2.8- CELLULAR SUB-LOCALIZATION OF WWP2 WW3-4 TANDEM DOMAIN IN ARPE-19.....</b>	<b>73</b>
<b>FIGURE 3.2.9- PROTEIN EXPRESSION OF SMAD 2, 3, AND 7 IN WW TANDEM DOMAIN TRANSFECTED HEK293A CELLS .....</b>	<b>75</b>
<b>FIGURE 3.2.10- THE PROTEIN EXPRESSION OF WWP2 WW3-4 IN HEK293A FROM A STREPTAVIDIN PULL-DOWN EXPERIMENT) .....</b>	<b>77</b>
<b>FIGURE 3.3.1- SUGGESTED MECHANISM OF WW3-4/SMAD BINDING .....</b>	<b>81</b>
<b>FIGURE 3.3.2 A SEQUENCE ALIGNMENT OF THE WW2 DOMAIN WITHIN WWP2 AND SMURF2.....</b>	<b>84</b>
<b>FIGURE 3.3.3- PREDICTION OF NES SEQUENCES IN WWP2 HECT DOMAIN.....</b>	<b>87</b>
<b>FIGURE 4.1.1- A SIMPLIFIED REPRESENTATION OF THE EFFECTS OF TGF<math>\beta</math> AND BMP ON LEVELS OF ONCOGENIC PROCESSES. ....</b>	<b>97</b>
<b>FIGURE 4.2.1- THE PROTEIN EXPRESSION OF WWP2 ISOFORMS PREVIOUSLY IDENTIFIED IN WESTERN BLOTS OF HEK293A CELLS USING PRIMARY WWP2-N AND WWP2-C TERMINAL ANTIBODIES .....</b>	<b>99</b>
<b>FIGURE 4.2.2- WHOLE GENOME ANALYSIS RESULTING IN PREDICTED REGIONS OF WWP2- <math>\Delta</math>HECT. ....</b>	<b>100</b>
<b>FIGURE 4.2.3- A FLOW CHART SUMMARIZING THE DETECTION, PREDICTION AND VALIDATION OF WWP2 NOVEL ISOFORMS.....</b>	<b>101</b>
<b>FIGURE 4.2.4- WHOLE GENOME ANALYSIS RESULTING IN PREDICTED REGIONS OF WWP2- N-<math>\Delta</math>C2 .....</b>	<b>102</b>
<b>FIGURE 4.2.5- MRNA EXPRESSION OF NOVEL WWP2 ISOFORMS .....</b>	<b>104</b>
<b>FIGURE 4.2.6- NOVEL AND ESTABLISHED WWP2 ISOFORMS FL, N, C, <math>\Delta</math>HECT, N-<math>\Delta</math>C2 AND FL-<math>\Delta</math>C2 ISOFORMS HIGHLIGHTING THE DOMAINS INCLUDED AND THE TRANSCRIPTION REGION. ....</b>	<b>105</b>
<b>FIGURE 4.2.7- WWP2 NOVEL ISOFORMS HIGHLIGHTED IN THE WWP2-FL AMINO ACID SEQUENCE.....</b>	<b>106</b>
<b>FIGURE 4.2.8- A SUMMARY OF THE NOVEL WWP2-FL-<math>\Delta</math>C2 ISOFORM DOMAIN COMPOSITION AND AMINO ACID SEQUENCE.....</b>	<b>107</b>
<b>FIGURE 4.2.9- MRNA EXPRESSION OF WWP2-FL-<math>\Delta</math>C2 IN MELANOMA CELL LINES ...</b>	<b>108</b>

FIGURE 4.2.10- RT-PCR ANALYSIS OF <b>WWP2</b> ISOFORM EXPRESSION IN <b>SKMEL28</b> AND <b>M229</b> STIMULATED WITH A <b>TGF<math>\beta</math></b> AND <b>BMP</b> TIME COURSE OF <b>1, 6, AND 18 HR</b> .....	110
FIGURE 4.2.11- THE EFFECTS OF <b>TGF<math>\beta</math></b> AND <b>BMP</b> STIMULATION ON mRNA EXPRESSION OF <b>WWP2</b> ISOFORMS IN MELANOMA CELL LINES .....	112
FIGURE 4.2.12- A BIOINFORMATIC APPROACH TO COMPARE <b>WWP2-N</b> AND <b><math>\Delta</math>HECT</b> . .....	114
FIGURE 4.2.13- THE mRNA EXPRESSION OF <b>WWP2</b> ISOFORMS IN <b>A375</b> CELLS STIMULATED WITH <b>TGF<math>\beta</math></b> FOR <b>1, 6 AND 18 HOURS</b> .....	116
FIGURE 4.2.14- THE mRNA EXPRESSION OF <b>WWP2</b> ISOFORMS IN <b>A375</b> CELLS STIMULATED WITH <b>BMP</b> FOR <b>1, 6 AND 18 HOURS</b> .....	116
FIGURE 4.2.15- THE EFFECTS OF <b>TGF<math>\beta</math></b> AND <b>BMP</b> ON <b>WWP2</b> mRNA EXPRESSION IN <b>MCF-7</b> CELLS .....	117
FIGURE 4.2.16- THE mRNA EXPRESSION OF <b>WWP2</b> ISOFORMS IN MELANOMA TISSUE SAMPLES USING <b>TISSUESCAN</b> CDNA ARRAYS PURCHASED FROM <b>ORIGENE</b> .....	120
FIGURE 4.2.17- THE mRNA EXPRESSION OF <b>EMT</b> MARKERS IN MELANOMA CELL LINES .....	123
FIGURE 4.2.18- QPCR ANALYSIS OF <b>WWP2</b> ISOFORM EXPRESSION IN CANCEROUS AND NON-CANCEROUS CELL LINES TO INVESTIGATE THE EXPRESSION IN EPITHELIAL AND MESENCHYMAL CELLS .....	124
FIGURE 4.3.1- A SIMPLE DIAGRAM TO ILLUSTRATE EFFECTS OF <b>BMP4</b> ON CELL INVASION AND MIGRATION .....	129
FIGURE 5.2.1- A SCHEMATIC REPRESENTATION OF THE <b>WWP2-C</b> AND <b><math>\Delta</math>HECT</b> ISOFORMS ON THE <b>WWP2</b> GENE LOCUS. ....	139
FIGURE 5.2.2- A DIAGRAMMATIC REPRESENTATION OF THE <b>WWP2</b> GENE SEQUENCE BETWEEN INTRON <b>10/11</b> AND <b>13/14</b> AND THE <b>WWP2-C</b> PROMOTER CONSTRUCTS IN <b>PGL4.27</b> .....	141
FIGURE 5.2.3- REGIONS OF PREDICTED CONSERVATION WITHIN THE <b>WWP2-C</b> PREDICTED PROMOTER REGION USING <b>PHASTCONS</b> AND <b>PHYLOP</b> IN THE <b>UCSC</b> GENOME BROWSER OF <b>20</b> MAMMALIAN SPECIES.....	142
FIGURE 5.2.4- <b>WWP2-C</b> PREDICTED PROMOTER REGION NUCLEOTIDE SEQUENCE .....	143
FIGURE 5.2.5- CONSERVATION OF <b>SOX9</b> BINDING SITE WITHIN PREDICTED <b>WWP2-C</b> PROMOTER REGION USING <b>MULTIZ</b> ALIGNMENT OF <b>20</b> MAMMALS .....	144

<b>FIGURE 5.2.6- A WESTERN BLOT USING PRIMARY ANTI-FLAG ANTIBODIES TO DETECT SOX9 PROTEIN EXPRESSION IN HEK293A CELLS .....</b>	<b>145</b>
<b>FIGURE 5.2.7- THE EFFECTS OF SOX9 AND TGF<math>\beta</math> ON WWP2-C PREDICTED PROMOTER ACTIVITY.....</b>	<b>146</b>
<b>FIGURE 5.2.8- THE EFFECT OF TGF<math>\beta</math> STIMULATION ON MRNA EXPRESSION OF SOX9 IN MELANOMA CELL LINES.....</b>	<b>147</b>
<b>FIGURE 5.2.9- WWP2-C WILD-TYPE AND MUTANT PROMOTER ACTIVITY IN HEK293A, A375 AND MCF-7 .....</b>	<b>150</b>
<b>FIGURE 5.2.10- THE PREDICTED PROMOTER REGION OF WWP2-<math>\Delta</math>HECT .....</b>	<b>152</b>
<b>FIGURE 5.2.11- THE EFFECTS OF TGF<math>\beta</math> ON WWP2-<math>\Delta</math>HECT PREDICTED PROMOTER ACTIVITY.....</b>	<b>153</b>
<b>FIGURE 5.2.12- THE EFFECTS OF TGF<math>\beta</math>, BMP AND EGF ON WWP2-<math>\Delta</math>HECT 2 KB PREDICTED PROMOTER ACTIVITY.....</b>	<b>154</b>
<b>FIGURE 5.2.13- PROTEIN EXPRESSION OF CLONED WWP2-<math>\Delta</math>HECT CONSTRUCT IN HEK293A.....</b>	<b>156</b>
<b>FIGURE 5.2.14- THE EFFECTS OF WWP2-<math>\Delta</math>HECT OVEREXPRESSION AND TGF<math>\beta</math>.....</b>	<b>157</b>
<b>FIGURE 5.2.15- THE WWP2 GENE LOCUS AND THE WWP2-N ISOFORM.....</b>	<b>158</b>
<b>FIGURE 5.2.16- PROTEIN EXPRESSION OF ESRP1 AND ESRP2 IN HEK293A.....</b>	<b>159</b>
<b>FIGURE 5.2.17- THE EFFECTS OF ESRP ON THE MRNA EXPRESSION OF WWP2 ISOFORMS IN A375 .....</b>	<b>160</b>
<b>FIGURE 5.2.18- THE EFFECTS OF ESRP ON THE MRNA EXPRESSION OF WWP2 ISOFORMS IN MCF-7 CELLS.....</b>	<b>161</b>
<b>FIGURE 5.2.19 - WESTERN BLOTS USING PRIMARY ANTI-HA ANTIBODIES TO DETECT THE PROTEIN EXPRESSION OF WWP2- FL-<math>\Delta</math>C2, N-<math>\Delta</math>C2, <math>\Delta</math>HECT, WWP2-N AND WWP2-C .....</b>	<b>163</b>
<b>FIGURE 5.2.20- FLUORESCENCE MICROSCOPY OF ARPE-19 CELLS EXPRESSING TRANSFECTED WWP2-FL .....</b>	<b>164</b>
<b>FIGURE 5.2.21- FLUORESCENCE MICROSCOPY OF ARPE-19 CELLS EXPRESSING TRANSFECTED WWP2-N .....</b>	<b>165</b>
<b>FIGURE 5.2.22- FLUORESCENCE MICROSCOPY OF ARPE-19 CELLS EXPRESSING TRANSFECTED WWP2-C.....</b>	<b>166</b>
<b>FIGURE 5.2.23- SUB-CELLULAR LOCALIZATION OF NOVEL WWP2-<math>\Delta</math>HECT ISOFORM IN ARPE-19 .....</b>	<b>167</b>

FIGURE 5.2.24- SUB-CELLULAR LOCALIZATION OF NOVEL <b>WWP2-N-ΔC2</b> ISOFORM IN <b>ARPE-19</b> .....	168
FIGURE 5.2.25- SUB-CELLULAR LOCALIZATION OF NOVEL <b>WWP2-FL-ΔC2</b> ISOFORM IN <b>ARPE-19</b> .....	169
FIGURE 5.2.26- A CELLULAR LOCALIZATION STUDY THAT EXPLORES THE RELATIONSHIP BETWEEN <b>WWP2-N</b> AND <b>TUBULIN</b> .....	170
FIGURE 5.3.1- A SCHEMATIC OF THE TRANSCRIPTION MECHANISM INVOLVED IN <b>WWP2- C</b> AND <b>ΔHECT</b> PRODUCTION. ....	172
FIGURE 5.3.2- THE DEPICTION OF PREDICTED <b>WWP2C</b> -PROMOTER SEQUENCES WITHIN INTRON 10/11 OF THE <b>WWP2</b> GENE LOCUS AND <b>H3K4ME3</b> MARKS USING THE <b>UCSC</b> GENOME BROWSER .....	173
FIGURE 5.3.3-PUTATIVE <b>SOX9</b> BINDING SITES IN THE PREDICTED <b>WWP2-C</b> PROMOTER REGION .....	177
FIGURE 5.3.4- THE PREDICTED <b>WWP2-ΔHECT</b> PROMOTER REGION AND HIGHLIGHTED <b>H3K4ME3</b> MARKS .....	180
FIGURE 5.3.5- THE SUGGESTED PREFERENTIAL BINDING PARTNERS OF <b>ΔHECT</b> IN PRESENCE AND ABSENCE OF <b>TGFβ</b> AND ITS EFFECTS ON <b>SMAD3</b> DEPENDENT GENE TRANSCRIPTION. ....	183
FIGURE 5.3.6- A SIMPLIFIED SCHEMATIC REPRESENTATION SHOWING THE CELLULAR LOCALIZATION OF <b>WWP2</b> ISOFORMS AND THE RESPECTIVE <b>SMAD</b> BINDING IN NORMAL AND DISEASE STATES. ....	186
FIGURE 5.3.7- PREDICTION OF <b>NES</b> SIGNAL WITHIN THE <b>WWP2 WW3</b> DOMAIN REGION .....	190
FIGURE 5.3.8- THE PREDICTION OF <b>NUCLEAR EXPORT SIGNALS (NES)</b> IN RESIDUE <b>383- 440</b> OF <b>WWP2</b> USING THE <b>NETNES</b> PROGRAM .....	190
FIGURE 6.2.1- A DIAGRAM DEPICTING THE PROPOSED <b>WW3-4</b> BINDING FROM LUCIFERASE DATA .....	195
FIGURE 6.2.2- A SCHEMATIC REPRESENTATION OF THE SUGGESTED MECHANISM RESPONSIBLE FOR THE <b>SMAD3</b> -DEPENDENT GENE EXPRESSION IN <b>WWP2-FL</b> TRANSFECTED <b>A375</b> CELLS .....	196
FIGURE 6.2.3- AN AMINO ACID SEQUENCE ALIGNMENT SHOWING THE <b>WW3-4</b> LINKER REGIONS OF <b>WWP2</b> , <b>WWP1</b> AND <b>ITCH</b> . ....	198

<b>FIGURE 6.2.4- A SEQUENCE ALIGNMENT OF SMURF2 AND WWP2 WW2 DOMAINS..</b>	
.....	200
<b>FIGURE 6.2.5- A SIMPLIFIED DIAGRAM DEPICTING THE PROPOSED MECHANISMS</b>	
<b>RESULTING IN THE WW3-4-DEPENDENT INCREASE IN H2A BINDING IN MASS</b>	
<b>SPECTROMETRY.....</b>	201
<b>FIGURE 6.3.1- A SIMPLIFIED DIAGRAM TO SHOW THE PREDICTED DUAL ROLE OF WWP2-</b>	
<b>FL IN ONCOGENESIS .....</b>	203
<b>FIGURE 6.3.2- A DIAGRAM SHOWING THE PROPOSED BINDING FROM ΔHECT LUCIFERASE</b>	
<b>ASSAYS.....</b>	209
<b>FIGURE 6.4.1- DIAGRAM EXPLAINING WWP2-C, N AND ΔHECT AS GUARDIANS OF</b>	
<b>EMT .....</b>	211

List of tables

TABLE 2.1-1- CLONING OF <b>WWP2</b> CONSTRUCTS AND THE CLONING METHODS USED ...	38
TABLE 2.1-2-PCR REACTION CONDITIONS USED FOR PLATINUM HIGH FIDELITY DNA TAQ PCR .....	40
TABLE 2.1-3- PCR REACTION CONDITIONS USED FOR HIGH FIDELITY PHUSION PCR .....	40
TABLE 2.1-4- CLONING PRIMERS OBTAINED FROM EUROFINS GENOMIC USED FOR VARIOUS CLONING PROCESSES .....	41
TABLE 2.1-5- COMPONENTS IN A 40 $\mu$ L XHOI HINDIII DOUBLE RESTRICTION DIGEST ....	42
TABLE 2.1-6- A 10 $\mu$ L T4 LIGATION REACTION TO LIGATE PCR INSERT INTO PGEM-T EASY VECTOR .....	42
TABLE 2.1-7- COMPONENTS IN A STANDARD 10 $\mu$ L T4 REACTION .....	43
TABLE 2.1-8- COMPONENTS INCLUDED IN A 5 $\mu$ L IN-FUSION REACTION .....	44
TABLE 2.1-9- RECIPE FOR TBI USED IN DH5A PREPARATION .....	45
TABLE 2.1-10- RECIPE FOR TBII USED IN DH5A PREPARATION .....	46
TABLE 2.1-11- PRIMERS USED FOR QUIKCHANGE DELETION OF SOX9 BINDING SITE IN PREDICTED <b>WWP2-C</b> PROMOTER SEQUENCE .....	46
TABLE 2.2-1- A TABLE SHOWING THE DIFFERENT MEDIUMS USED FOR VARIOUS CELL TYPES .....	47
TABLE 2.3-1- STANDARD 1X PHOSPHATE BUFFERED SALINE ( <b>PBS</b> ) RECIPE .....	50
TABLE 2.3-2- PRIMARY ANTIBODIES USED IN IMMUNOLABELLING .....	50
TABLE 2.3-3- SECONDARY ANTIBODIES USED IN IMMUNOLABELLING .....	50
TABLE 2.4-1- RECIPE FOR NP40 CELL LYSIS BUFFER .....	52
TABLE 2.4-2- RECIPE FOR 2X LAMMLI BUFFER.....	52
TABLE 2.4-3- 10% SDS PAGE GEL RECIPE .....	53
TABLE 2.4-4- 15% SDS PAGE GEL RECIPE .....	53
TABLE 2.4-5- RECIPE FOR 10X SDS PAGE RUNNING BUFFER .....	53
TABLE 2.4-6- RECIPE FOR 1X TRANSFER BUFFER.....	55
TABLE 2.4-7- RECIPE FOR PONCEAU S STAINING SOLUTION (0.1%w/v PONCEAU S IN 5%v/v ACETIC ACID) .....	55
TABLE 2.4-8- RECIPE FOR 1X TBST PH7.6 .....	55
TABLE 2.4-9- PRIMARY ANTIBODIES USED IN WESTERN BLOTTING .....	55
TABLE 2.4-10- SECONDARY ANTIBODIES USED IN WESTERN BLOTTING .....	56

<b>TABLE 2.5-1-MIXTURE ONE REQUIRED IN A 40<math>\mu</math>L GOSCRIPT CDNA SYNTHESIS REACTION</b>	57
<b>TABLE 2.5-3- MIXTURE TWO REQUIRED IN A 40<math>\mu</math>L GOSCRIPT CDNA SYNTHESIS REACTION</b>	57
<b>TABLE 2.5-4 -PCR SETTINGS USED FOR GOSCRIPT CDNA SYNTHESIS</b>	58
<b>TABLE 2.5-5- COMPONENTS IN A 20<math>\mu</math>L GOTAQ RT-PCR REACTION</b>	58
<b>TABLE 2.5-6- PCR SETTINGS USED FOR GOTAQ RT-PCR</b>	58
<b>TABLE 2.5-7- ANNEALING TEMPERATURES USED IN RT-PCR</b>	59
<b>TABLE 2.5-8- RT-PCR PRIMER SEQUENCE</b>	59
<b>TABLE 2.5-9- RECIPE FOR 20X TAE BUFFER</b>	60
<b>TABLE 2.5-10- COMPONENTS USED IN A 20 <math>\mu</math>L TAQMAN REACTION</b>	61
<b>TABLE 2.5-11- TAQMAN QPCR PRIMER SEQUENCES PROVIDED BY PRIMERDESIGN</b>	61
<b>TABLE 2.5-12- CYCLE INFORMATION FOR TAQMAN QPCR</b>	61
<b>TABLE 3.2-1- QUANTITATIVE MASS SPECTROMETRY RESULTS SHOWING EFFECTS OF TGF<math>\beta</math> ON PROTEINS ASSOCIATED WITH WWP2 WW3-4</b>	77
<b>TABLE 4.3-1- BRAF STATUS OF MELANOMA CELL LINES USED IN WWP2 EXPRESSION STUDY</b>	128
<b>TABLE 5.3-1- A TABLE SHOWING THE NRAS STATUSES OF MELANOMA CELL LINES USED IN SOX9 RT-PCR</b>	175

## Chapter 1 Introduction

## 1.1 Introduction

The human body is a delicately tuned biological machine made of countless proteins and many different types of cells. The precursors of protein are amino acids coded by nucleotide codons and are essentially the building blocks of life. For normal function, there must be a tight regulation on all aspects of the body. The term homeostasis is used to describe the upkeep of optimum conditions needed for a healthy body, which includes the fine balance of protein levels facilitated by protein turnover. Protein regulation is controlled by an array of different mechanisms, an example being ubiquitin-mediated protein degradation which can involve an enzyme known as WWP2 (WW domain-containing protein 2) ubiquitin ligase, and are most likely triggered by signalling molecules such as cytokines. The complex crosstalk between signalling pathways, proteins and genetic material is responsible for appropriate actions taken to keep proteins at homeostatic levels. Any mis-regulation, caused by mutation or otherwise will likely lead to abnormal functions and ultimately disease. One of the most common modern day diseases caused by mutation-mediated protein mis-regulation is cancer, which has been linked to the Transforming growth factor  $\beta$  (TGF $\beta$ ) signalling pathway. This signalling pathway employs Smad signalling molecules to carry out signal transduction suggesting that mis-regulated levels of such molecules could lead to a change in the TGF $\beta$  signalling pathway making the development of cancer more favourable. Thus, the crucial idea of TGF $\beta$ -mediated oncogenesis is a key theme that is investigated in this thesis.

In addition to TGF $\beta$  and cancer development, elements of the post-translational modification known as ubiquitination will also be explored. This process is one of the key mechanisms used in protein turnover by sending target molecules, such as Smad mediators of the TGF $\beta$  pathway, for proteasomal degradation using three different enzymes known as the E1 ubiquitin-activation enzyme, E2 ubiquitin-conjugating enzyme and E3 ubiquitin ligase. Moreover, the crosstalk between the signalling molecule

TGF $\beta$  and E3 ubiquitin ligases via interactions between Smad mediators, presents a fascinating opportunity to investigate the relationship between TGF $\beta$  signalling and cancer through mechanisms such as TGF $\beta$ -mediated epithelial to mesenchymal transition (EMT), and therefore uncover the role ubiquitination might have on TGF $\beta$  and subsequently cancer. WWP2 is an E3 ubiquitin ligase shown to interact with Smad mediators of the TGF $\beta$  pathway. Studying the relationship between WWP2 isoforms and TGF $\beta$  pathway may therefore prove useful in understanding the effects on TGF $\beta$  dependent oncogenic processes such as EMT due to the effects WWP2 may have on Smad protein levels and subsequently TGF $\beta$  signalling. The characterization of existing and novel WWP2 isoforms is therefore the focus in this thesis where we look at both artificial tandem domains within the isoforms and the biological function of expressed isoforms in relation to TGF $\beta$ -dependent oncogenic activities.

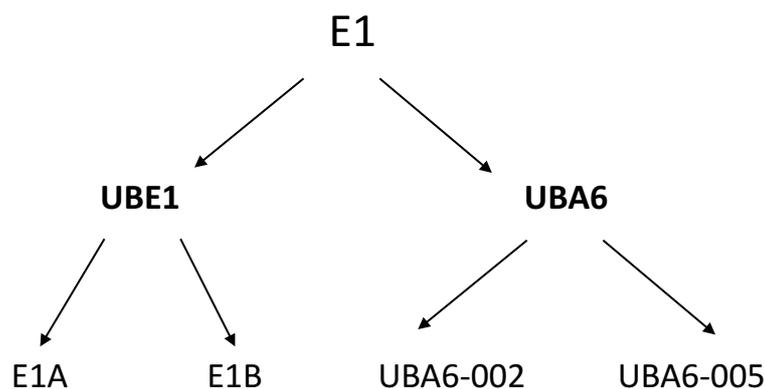
## 1.2 Ubiquitination

Ubiquitination is a type of post-translational modification on proteins akin to phosphorylation. These two mechanisms involve the reversible attachment of ubiquitin or phosphate molecules respectively, onto target substrate by covalent interactions (Sun and Chen, 2004). Ubiquitin (Ub) is a 76-amino acid globular protein containing a mixture of  $\alpha$ -helices and  $\beta$ -sheets packed tightly by numerous hydrogen bonds forming a hydrophobic core, and is found ubiquitously in eukaryotic cells (Vijay-kumar *et al*, 1987). The process of ubiquitination requires a cascade of three enzymes, E1 ubiquitin-activating enzyme, E2 ubiquitin-conjugating enzyme, and E3 ubiquitin ligase, to mediate its array of functions. Through the attachment of mono-Ub and poly-Ub onto specific lysine residues of target proteins, a wide range of functions can be triggered including DNA repair and signalling (Johnson, 2002). In fact, the implications of this process are so vast that research has shown 0.1-5 % of all eukaryotic cellular proteins at any one time, is ubiquitin (Kimura and Tanaka, 2010). The most well-known role of ubiquitination is arguably proteasomal degradation mediated by the formation of an iso-peptide link between a poly-Ub chain and the target substrate at Lys-48 (Tenno *et al*, 2004). Therefore, due to the high specificity of this process, it is important to further explore the mechanism and associated functions of ubiquitination.

### 1.2.1 The mechanism of Ubiquitination

As mentioned above, there are three key mediators of ubiquitination including the E1 Ub-activating enzyme, E2 Ub-conjugating enzyme, and E3 Ub-ligase. To date, two E1 enzymes have been found in eukaryotes, however, the high specificity of this dynamic process to the large amount of target substrate is the result of more than 40 E2s and ~617 E3s (Li *et al*, 2008). The two E1 enzymes, UBE1 and UBA6 both have truncated versions of their full-length protein, known as isoforms. The established isoforms of UBE1 ubiquitin activating enzyme, initially referred to as E1<sub>117kDa</sub> and E1<sub>110 kDa</sub> are now known as UBE1-E1a and UBE1-E1B (Cook and Chock, 1992).

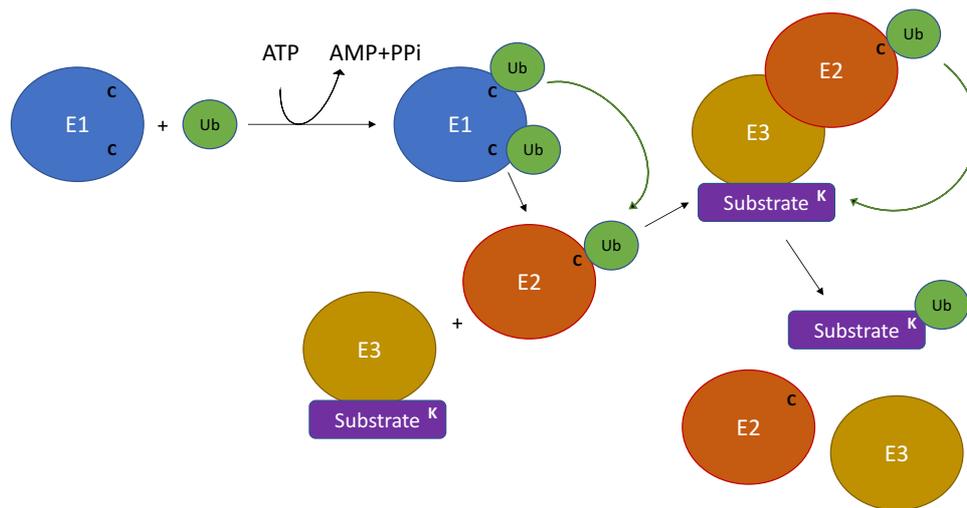
These isoforms have been suggested to be a product of putative start codons within the UBE1 gene locus (Handley-gearharts *et al*, 1994), and interestingly have been shown to differentially localize in the cell. E1a was found localized predominantly in the nucleus whilst E1B was detected primarily in the cytosol (Stephen *et al*, 1996). The second E1 ubiquitin-activating enzyme UBA6, also has two truncated isoforms expressed in the human body; UBA6-002 and UBA6-005. The difference in combination of functional domains in these isoforms suggest that they may have different properties in regards to their roles in ubiquitination (Soond and Chantry, 2011). Together with the increased specificity created by the E1 isoforms (summarized in fig 1.2.1) and their ability to bind preferentially to different E2 Ub-conjugating enzyme (Stephen *et al*, 1996), the range of E3 isoforms in the human genome which allows the binding of vast amounts of different proteins creates even more specificity and diversity in the ubiquitination mechanism. Similar to the E1 ubiquitin-activating enzymes, E2 ubiquitin-conjugating enzymes also have multiple expressed protein isoforms. All in all, the larger group of E2 enzymes and their relative isoforms mean that once again, the level of specificity and diversity increases due to the ability of different E1, E2 and E3 isoforms to bind preferentially to a diverse range of specific targets, ultimately mediating an array of ubiquitin-dependent cellular processes (Pickart and Eddins, 2004; Johnson, 2002; Soond and Chantry, 2011).



*Figure 1.2.1- The E1 ubiquitin-activating enzyme and its isoforms. A simple schematic showing the E1 ubiquitin-activating enzymes and protein isoforms expressed in the human body.*

To begin the process of ubiquitination, a 3-step ATP-dependent mechanism is required to attach ubiquitin onto the monomeric E1 Ub-activating

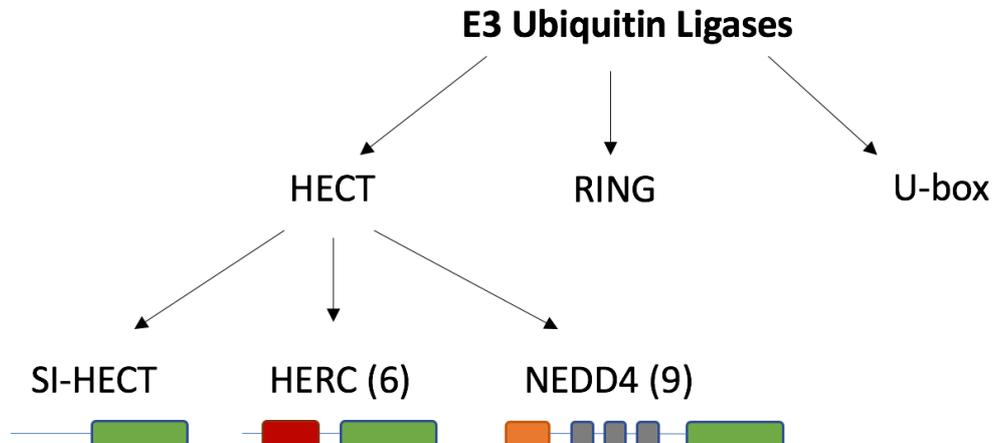
enzyme. This involves the conversion of E1, Ub and ATP first into an E1-Ub-adenylate intermediate complex ( $E1+Ub+ATP \rightleftharpoons E1-Ub-AMP+PPi$ ) which leads to the covalent thioester bonding between Ub and E1 ( $E1-Ub-AMP \rightleftharpoons E1-Ub+AMP$ ). Interestingly, an additional Ub is then added onto E1-Ub using ATP to create a complex containing one E1 enzyme and two Ub molecules ( $E1-Ub+Ub+ATP \rightleftharpoons E1-Ub-Ub-AMP+PPi \rightleftharpoons E1-Ub-Ub+AMP$ ) (Haas and Rose, 1982). The Ub is then transferred onto the catalytic cysteine residue within the Ub-conjugating (UBC) domain of E2 Ub-conjugating enzyme and covalently linked using a thioester bond. The function of E2 is to transfer the Ub molecule onto lysine residues of target substrate with the help of E3 Ub-ligases (shown in fig 1.2.2). Notably, E2 is also responsible for the formation of poly-Ub chains by the transfer of Ub onto another Ub at one of its seven lysine residues (Valimberti *et al*, 2015).



**Figure 1.2.2-** A diagram to summarize the process of ubiquitination. A Ubiquitin molecule is bound to the Cysteine (C) residue of E1 ubiquitin-activating enzyme in an ATP dependent reaction. The Ub is then transferred onto the catalytic Cysteine (C) of E2 Ub-conjugating enzyme shown by the green arrow. The E2-Ub complex then binds to E3-Substrate complex allowing the transfer of Ub onto Lysine (K) residue of target substrate.

The final step of ubiquitination utilizes a class of enzymes known as the E3 Ub-ligases. The main function of this vast collection of enzymes is to increase specificity within the process allowing specific ubiquitin conjugation onto a wide range of target substrates. There are three different families of E3 Ub-ligases including the RING finger-containing (RING) E3s, the UFD2 Homology

(U-box) E3s and the Homologous to E6AP Carboxy Terminus (HECT) E3s. The RING, HECT and U-box within the naming of these E3s all represent the domains responsible for E2 binding (Pickart and Eddins, 2004). In this thesis, we focus on a Ub-ligase from the HECT family of E3s. In contrast to RING and U-box E3s where the enzyme act simply as a bridging protein for the binding of Ub onto target substrate, the Ub molecule forms a temporary thiol-ester bond with the catalytic cysteine residue in the HECT domain of HECT E3s before the final conjugation with target substrate (Maspero *et al*, 2011). HECT E3 ligases are further categorized into three groups; the HECT and RLD domains containing (HERC) E3s, Neural precursor cell-expressed developmental downregulated gene 4 (Nedd4) E3s which contains a C2 lipid binding domain, tryptophan-tryptophan (WW) domains and a HECT domain, and finally the SI-HECT E3s which contain neither WW or RLD domains responsible for substrate recognition. A summary of E3 ubiquitin ligases is shown in fig 1.2.3. To date, six members of HERC E3s and nine members of NEDD4 E3s have been found (Scheffner and Staub, 2007).



**Figure 1.2.3-** A diagram showing the sub-families of E3 ubiquitin ligases. E3 Ub ligases are separated into three main families; HECT E3s, RING E3s and U-box E3s. The HECT E3 family is further categorized into three subfamilies; HERC E3s containing an RLD domain (red) along with a HECT catalytic domain (green), NEDD4 E3s contain a N-terminal C2 domain (orange), 2-4 WW recognition domains (grey) and the C-terminal HECT domain, and finally the SI-HECT E3s which do not contain WW or RLD Domains.

### 1.2.2 Deubiquitination

Just as phosphatases remove phosphate molecules from its target substrate in phosphorylation, ubiquitination also has its own class of enzymes that specifically remove ubiquitin to reverse the effects of this process. These are known as deubiquitinating enzymes (DUBs). There are around 100 DUBs found in the human genome separated into six classes, five of which are cysteine proteases whilst the other is a zinc-dependent metalloprotease. The ubiquitin-specific protease (USP), ubiquitin C-terminal hydrolase (UCH), Otubain protease (OTU), Machado-Joseph disease protease (MJD) and motif-interacting with ubiquitin-containing novel DUB family (MINDY) are all subclasses of cysteine protease DUBs named after their Ub-protease domains. The Ub-protease domain in metalloprotease DUBs are known as JAB1/MPN/Mov34 metalloenzyme (JAMM), and accounts for 14 of the 95 DUB genes overall (Harrigan *et al.*, 2017).

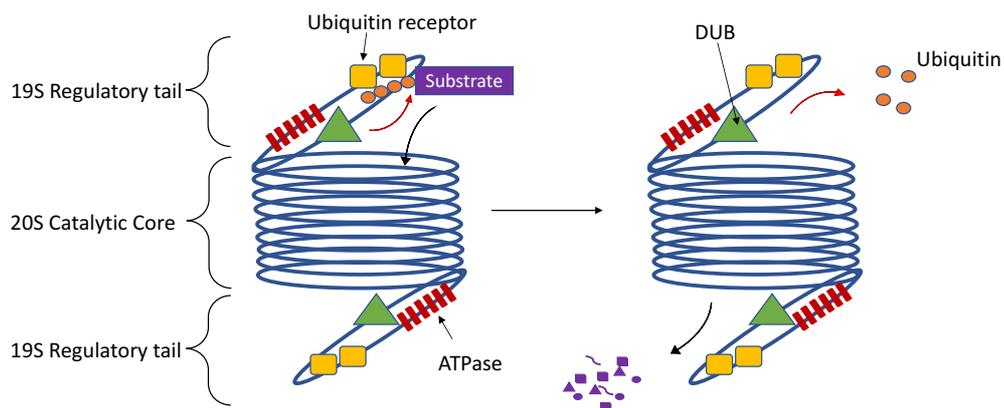
DUBs function by the cleavage of peptide bonds next to the c-terminal Gly-76 which is the final residue in active Ub. Since all Ub molecules are produced as inactive precursors containing a C-terminal extension past Gly-76, it is assumed that DUBs play an essential role in activation of Ub precursors by the cleavage of c-terminal extension via a hydrolysis reaction (Grou *et al.*, 2015). Another important function of DUBs is to stabilize homeostatic levels of ubiquitin within the cell. This includes the cleaving of Ub from target substrate committed for proteasomal degradation to avoid the inappropriate degradation of Ub (Amerik and Hochstrasser, 2004). Unanchored ubiquitin in cells not only exist in its monomeric form but substantial levels of poly-ubiquitin chains are also found. Studies have suggested that DUBs are responsible for the dissemination of such polypeptides, and that a fault in this mechanism leads to abnormalities in ubiquitin-dependent proteasomal degradation (Amerik *et al.*, 1997). Finally, DUBs are also involved in the removal of Ub signals from proteins incorrectly tagged, as means of a proof-reading function (Komander *et al.*, 2009).

### 1.2.3 Functions of ubiquitination

Ubiquitination as a form of post-translational modification can lead to many outcomes depending on the number of Ub molecules attached and the position of the iso-peptide bond. Ubiquitination can occur in many forms to increase specificity and range of functions. The addition of a poly-ubiquitin chain onto target substrate is known as poly-ubiquitination, whilst the addition of one Ub molecule is known as mono-ubiquitination. There are two types of poly-ubiquitin branches, homotypic chains contain Ub bound together at the same position on each Ub whilst heterotypic chains contain Ub that bind at different lysine positions (Swatek and Komander, 2016). It is also possible for target substrates to be mono-ubiquitinated at multiple sites (multi-ubiquitination), and with the help of certain RING E3s, branched poly-ubiquitination has also been observed. There are a total of seven lysine residues on ubiquitin of which poly-peptide bonds can be formed, Lys-6, 11, 27, 29, 33, 48 and 63 (Sadowski *et al*, 2012). In addition, branched poly-ubiquitination can also occur at Lys-6, 27 and 48 (Ben-Saadon *et al*, 2006). To add a further layer of complexity, ubiquitin molecules can be modified by phosphorylation and acetylation to facilitate a wider range of functions. Furthermore, the addition of a phosphate or an acetyl group onto Ub molecules may also act as an extra level of regulation to the post-translational modification of ubiquitination (Swatek and Komander, 2016).

The most extensively studied ubiquitination modification is the Lys-48 poly-ubiquitination of target substrates which results in degradation by the 26S proteasome. The attachment of a poly-ubiquitin chain at Lys-48 containing a minimum of four Ub molecules and subsequent increase of binding affinity between substrate and 26S proteasome is needed for target recognition and degradation (Thrower *et al*, 2000). The 26S proteasome contains the 20S proteolytic core sandwiched by two 19S regulatory tails. The substrate recognition of ubiquitinated targets occur through the binding of substrate-bound poly-ubiquitin chains and ubiquitin receptors Rpn10 and Rpn13 on the 19S tail (Grice and Nathan, 2016). In addition, the 19S

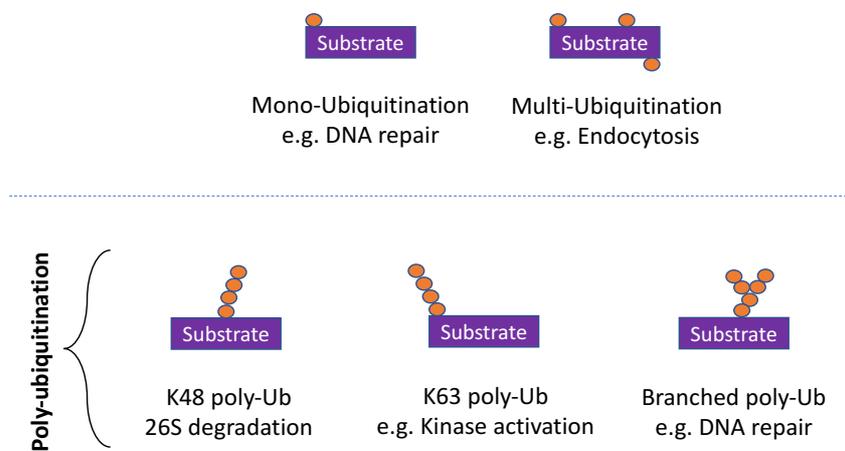
regulatory domains also has six ATPases (Rpt1-6) responsible for the partial unfolding of target proteins and the subsequent opening of the 20S catalytic core. The 19S complex also contains DUBs tasked with the detachment of Ub from target protein which has been suggested as a signal for proteolytic degradation of substrates (Peth *et al*, 2010). A simplified diagram in fig 1.2.4 shows the mechanism of ubiquitin-dependent 26S proteasomal degradation.



**Figure 1.2.4- The simplified representation of the ubiquitin-dependent 26S proteasomal degradation mechanism.** The 26S proteasome is made up of a 20S proteolytic core flanked by two 19S regulatory tails. The substrate tagged by a poly-Ub chain (shown in orange) is recognized by the ubiquitin receptors shown in yellow. The DUB shown in green, is responsible for cleaving the isopeptide bond between the ubiquitin chain and target substrate. The presence of six ATPases in the 19S tail (in red) collectively mediate the unfolding of target protein and the opening of the 20S catalytic core.

Contrasting with the above, poly-ubiquitination at Lys-63 does not result in proteasomal degradation of target proteins, and has shown to be implicated in roles such as DNA repair and kinase activation. The Lys-63 poly-ubiquitination of proliferating-cell nuclear antigen (PCNA) has been shown to mediate an “error-free” DNA repair mechanism in the RAD6 E2 pathway. Interestingly, the mono-ubiquitination of PCNA displayed an “error-prone” mechanism within the same RAD6 DNA repair pathway (Sun and Chen, 2004). Multi-ubiquitination is the binding of a single Ub molecule to multiple sites on a target substrate. This mechanism has been shown to be pivotal in the endocytosis and subsequent degradation of Receptor Tyrosine Kinases (RTKs). For example, Haglund *et al* demonstrated that the multi-

ubiquitination of epidermal growth factor receptor (EGFR) and platelet-derived growth factor receptor (PDGFR) leads to its proteasome-independent degradation by the lysosome (Haglund *et al*, 2003). Finally, branched poly-ubiquitination of Ring1B RING E3 is a critical intermediate for the mono-ubiquitination of histone H2A, allowing the transcriptional regulation of target genes. uH2A is most often associated with the silencing of specific target genes suggesting that branched poly-ubiquitination has an important role to play in gene regulation (Cao and Yan, 2012). A summary of the different types of ubiquitination processes is shown in fig 1.2.5.



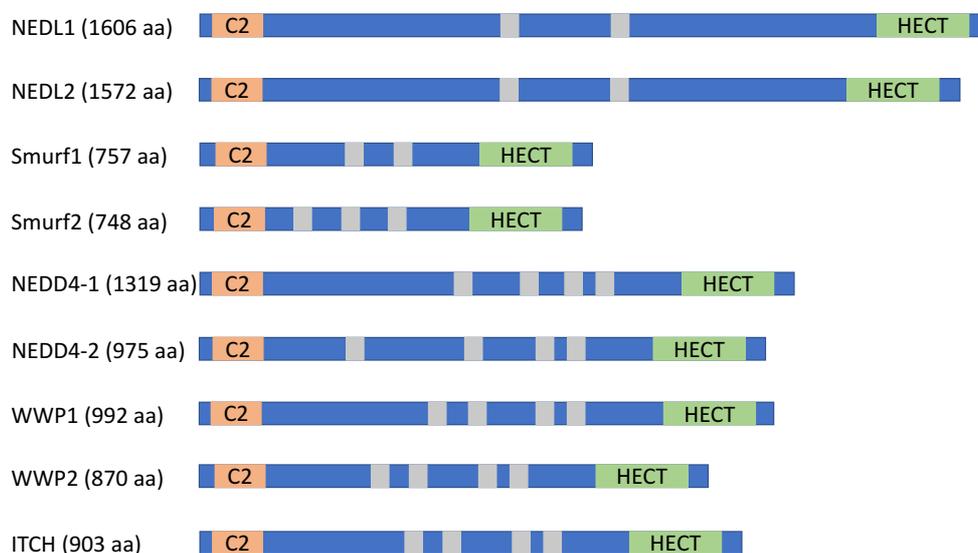
**Figure 1.2.5- A summary depicting different types of ubiquitin post-translational modification.** Ubiquitin molecules are shown in orange and are bound to substrates in various manners as shown above. Poly-ubiquitination involves the addition of a ubiquitin chain onto target substrate whilst mono-ubiquitination is the covalent binding of one Ub. Multi-ubiquitination is the conjugation of one Ub molecule onto multiple sites of the target protein.

### 1.3 NEDD4 Ubiquitin Ligases

Due to the diversity in the functions of ubiquitination including those facilitated by NEDD4 E3 ubiquitin ligases, the regulation of this E3 becomes a crucial part of normal cellular function. The auto-inhibition of these ligases is one way of negatively regulating their catalytic activity. The auto-inhibition of E3 ligases results in a conformation that inhibits substrate binding and subsequent ubiquitination. However, proteins such as NEDD4 Family-interacting proteins (NDFIPs) containing multiple PY motifs bind to the WW recognition domains of NEDD4 ligases allowing the release of the auto-inhibitory conformation and therefore positively regulating NEDD4 function (Mund and Pelham, 2009). Furthermore, research has also shown that the binding of NDFIPs to NEDD4 not only functions to release its catalytic activity, but also to recruit the NEDD4 ligase to the target substrates, therefore facilitating the NEDD4-mediated ubiquitination of the NDFIP-bound target substrates. In a study by Foot *et al*, the ability of NDFIP to bind to and mediate the WWP2 NEDD4 ligase-mediated degradation of Divalent metal transporter 1 (DMT1) was highlighted (Foot *et al*, 2008). Thus, the auto-inhibition of NEDD4 along with the binding of NDFIPs and other NEDD4 regulatory proteins result in the tight regulation of NEDD4 function.

Currently, there are nine NEDD4 E3 ubiquitin ligases identified including NEDD4-1, NEDD4-2, WWP1, WWP2, ITCH, NEDL1, NEDL2, Smurf1 and Smurf2. These are grouped together under the same umbrella as they all contain an N-terminal C2 lipid binding domain, WW (tryptophan-tryptophan) substrate recognition domains and a C-terminal HECT catalytic domain. However, these NEDD4 ligases differ in size, ranging from 748 aa (Smurf2) to 1606 aa (NEDL1), and number of WW domains as shown in fig 1.3.1. The difference in WW domain composition dictates the specific function of the ligase. For example, Smurf1 contains 2 WW domains which bind to and facilitate the degradation of Smad1 and 5 (Zhu *et al*, 1999), whilst Smurf2, containing 3 WW domains mono-ubiquitinates Smad3 to inhibit the TGF $\beta$  pathway (Tang *et al*, 2011). The transcription of isoforms is

not an uncommon occurrence in E3 ubiquitin ligases. Most of the nine NEDD4 family of E3 enzymes have a set of expressed truncated isoforms or transcripts of novel isoforms that have unconfirmed protein expression. For example, WWP1 is a 992 aa E3 ligase that contains C2 lipid binding domain, four WW domains and a HECT catalytic domain. There are currently six WWP1 isoforms all containing different combinations of functional domains which suggests that they may all have slightly different binding partners and functions. Furthermore, the expression of these isoforms has been shown to differ depending on the cell type. The alternate transcription regulation responsible for the generation of these isoforms are still unclear, however, most speculate the involvement of putative promoters or alternate splicing events resulting in retained introns (Flasza *et al*, 2002). The subsequent increased amount of functional truncated isoforms with different roles amplify the overall complexity of the mechanism by introducing extra levels of regulation to ultimately multiply specificity of the ubiquitination system.



**Figure 1.3.1- The NEDD4 family of E3 ubiquitin ligases.** A diagrammatical representation of NEDD4 ubiquitin ligases containing the C2 lipid binding domain in orange, WW substrate recognition domains in grey, HECT catalytic domain in green and the linker regions in blue. Note that drawings are not to scale and all gene information was obtained from UniProt.

Many studies have explored the different roles of NEDD4 isoforms caused by the difference in domain composition of truncated isoforms. For example, the overexpression of WWP1 has been linked to prostate cancer progression by inhibiting the tumour-suppressant properties of TGF $\beta$ . However, in the same study, it was found that the expression of a truncated WWP1 isoform in both normal and prostate cancer cell lines indicates that the shortened WWP1 isoform is unlikely to be involved in prostate cancer development (Chen *et al*, 2007). Additionally, the NEDD4-2 ubiquitin ligase and its isoforms has also been studied with results suggesting that the presence of different domains alter their ability to carry out specific functions. In a study by Raikwar and Thomas investigating the NEDD4-2-mediated degradation of epithelial Na<sup>+</sup> channel (ENaC), it was found that the NEDD4-2- $\Delta$ C2 isoform (an isoform lacking the C2 domain), is a better inhibitor of ENaC in *Xenopus* oocytes than the NEDD4-2- $\Delta$ WW2,3 isoform (one that lacks the WW2,3 domains). Furthermore, cell surface expression of ENaC in HEK293A cells were also inhibited by NEDD4-2 with the  $\Delta$ C2 isoform being the more potent inhibitor compared to the full-length NEDD4-2 ligase (Raikwar and Thomas, 2008). Collectively, the above evidence supports the importance of domain composition in the function of protein isoforms.

The WW-domain containing E3 ubiquitin-protein ligase 2 (WWP2) is a member of the NEDD4 family of E3s that we have focused our research on. As mentioned previously, E3s works closely with E1 ubiquitin-activating enzymes and E2 ubiquitin-conjugating enzymes to ubiquitinate target substrates via an iso-peptide bond which signals for a multitude of different cellular responses. The most prominent role of WWP2 is the facilitation of target substrates for proteasomal degradation. WWP2 has also been shown to control proliferation amongst many other cellular functions through a multitude of signalling pathways (Choi *et al*, 2015). Undeniably, one of the most studied WWP2 substrates is the Phosphatase and Tensin homologue (PTEN) tumour suppressor, a lipid phosphatase which carries out its function by reversing the effects of Phosphatidylinositol 3-kinase (PI3K), namely by

converting Phosphatidylinositol 3,4,5, triphosphate (PIP<sub>3</sub>) to Phosphatidylinositol 4,5, bisphosphate (PIP<sub>2</sub>) (An *et al*, 2014). The role of WWP2 has been shown using WWP2 knockdowns, suggesting that WWP2 increases levels of PIP<sub>3</sub> by signalling for PTEN degradation, ultimately leading to the increase in activated Protein Kinase B (Akt) by Phosphoinositide-dependent Kinase 1 (PDK1) phosphorylation. The increase in levels of phosphorylated Akt leads to an increase in levels of cell growth and proliferation (Maddika *et al*, 2011).

### 1.3.1 Structure of WWP2 and its isoforms

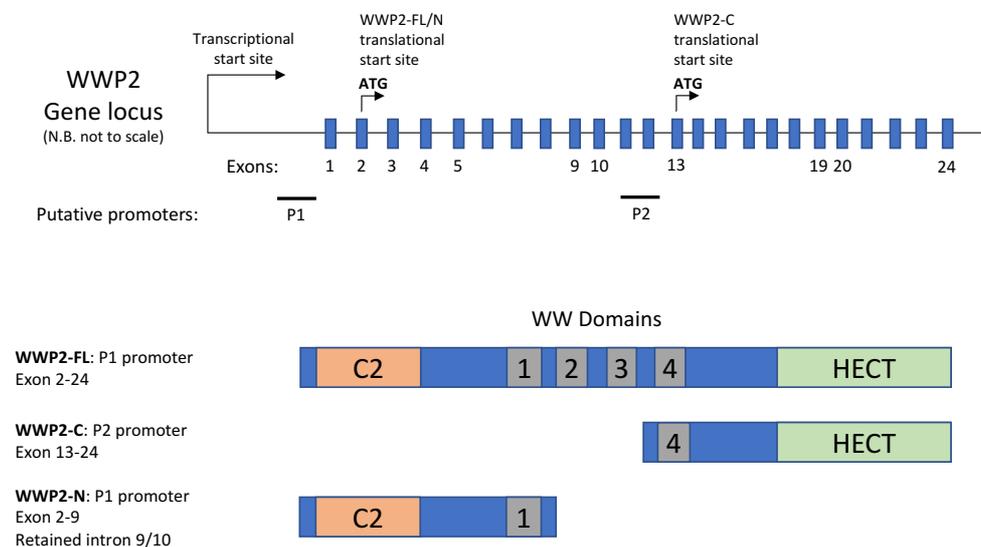
WWP2 is an 870 amino acid protein encoded by the WWP2 gene located on chromosome 16q22.1 (Chen *et al*, 2014). Translation of the full length WWP2 isoform (FL) begins at exon 2 and concludes at exon 24. This produces an isoform that contains an N-terminal C2 lipid binding domain, four tryptophan-tryptophan (WW1-4) substrate recognition domains and a C-terminal HECT catalytic domain. The N-terminal C2 domain is responsible for the binding of the protein to the membrane lipid bilayer. This domain contains a pair of 4-stranded  $\beta$ -sheets connected by three loops on the top and four on the bottom. Calcium ions (Ca<sup>2+</sup>) bind exclusively to the top three loops of the C2 domain allowing the electrostatic binding of C2 to negatively charged phospholipids on the membrane (Rizo and Südhof, 1998). At the c-terminal of the C2 domain are four WW substrate recognition domains. Each domain comprise of  $\beta$  strands arranged to give a globular structure (Sudol *et al*, 1995). Its ability in recognize PPXY motifs on target substrates is related to its proline-rich nature and the appearance of double-tryptophan in its sequence. The presence of hydrophobic residues within the binding pocket of WW domains allow the binding of hydrophobic prolines within the PPXY motif. WW domains most commonly refer to two tryptophan residues spaced 20-22 amino acids apart that are responsible for target recognition. However, contrary to what the name suggests, WW domains do not necessarily contain the two conserved tryptophan residues as either the first or the second tryptophan may be replaced by an aromatic phenylalanine or tyrosine (Sudol *et al*, 2005). Finally, the HECT ligase domain

lies at the c-terminus of WWP2 and is arranged as the N and C lobes. Unlike other E3s, HECT ligases contain a catalytic cysteine residue in its C-lobe that permits transthioesterification meaning that Ub binds to HECT E3 before conjugation with target substrate (Huibregtse *et al*, 1995; Berndsen and Wolberger, 2014).

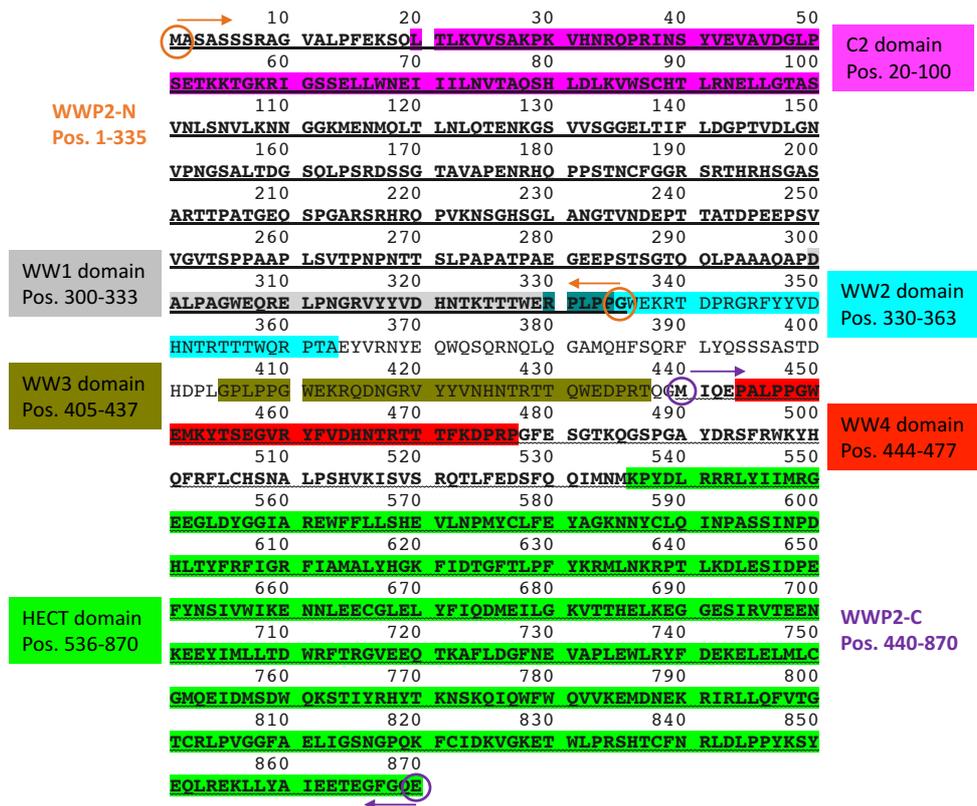
Interestingly, E3 ubiquitin ligases including WWP2, have the ability to auto-inhibit as a form of negative regulation. This form of inactive WWP2 has proven to be relatively stable as demonstrated by its unchanging cellular levels in response to Heclin, an inhibitor of HECT ligases (Mund *et al*, 2015). The mechanism of auto-inhibition involves the binding of the C2 domain to the HECT domain near its catalytic cysteine, effectively inhibiting the formation of thioester bond to Ub (Wiesner *et al*, 2007). The release and consequent activation of E3s can be mediated by binding of different substrates. For example, it has been shown that calcium binding to C2 domain of NEDD4s releases the protein from auto-inhibition (Wang *et al*, 2010), whilst in WWP2, Dishevelled (DVL2) has been suggested to bind with WW domains via its PPXY motif to relieve auto-inhibition (Mund *et al*, 2015).

Currently there are three established WWP2 isoforms inclusive of the FL protein mentioned above (ENST00000359154.6) and are shown in fig 1.3.2. WWP2-N (ENST00000569174.5) is a shortened isoform whereby transcription of the gene ends prematurely at exon 9 due to the presence of a stop codon at the retained intron 9/10. This subsequently produces a truncated version of WWP2 that spans aa 1-335 containing only the C2 and WW1 domains (shown in fig 1.3.3). WWP2-C (ENST00000568684.1) on the other hand is a 431 aa protein that runs from residue 440-871 (fig 1.3.3) and contains the WW4 and HECT domains and is the result of the presence of a putative promoter leading to the translation of a start codon at exon 13. The alternate splicing events resulting in the retained intron in WWP2-N has been suggested to be a result of splicing factors such as Epithelial splicing regulatory protein 1 (ESRP1) and ESRP2 (Chantry, 2011). As suggested by their names, ESRPs are highly expressed in epithelial cells whilst

mesenchymal expression is substantially down-regulated. The ability of these proteins to potentially regulate the transcription of different WWP2 isoforms therefore suggests that the down-regulation of ESRPs in mesenchymal cells will induce differential expression of different protein isoforms compared to epithelial cells (Warzecha *et al*, 2009a). Many properties of these isoforms such as transcriptional regulation and functional binding are yet unclear, however, it has been shown that WWP2-N can relieve auto-inhibition of WWP2-FL in the absence of TGF $\beta$  suggesting an activating role in the WWP2 cascade (Soond and Chantry, 2011).



**Figure 1.3.2- A diagram of the WWP2 gene locus and expressed isoforms.** A summary of WWP2 isoforms showing the C2 domain highlighted in orange, the WW domains in grey and the HECT domain in green. The transcription of different isoforms is a result of a combination of retained introns and putative promoters leading to two different established truncations of WWP2-FL. The simplified WWP2 gene locus located in chromosome 16q22.1 is also depicted above, showing positions of various start codons and predicted promoters.



**Figure 1.3.3 - The amino acid sequence of established WWP2 isoforms.** The full length WWP2 protein has the C2 domain highlighted in pink, WW1 domain in grey, WW2 in blue, WW3 in dark green, WW4 in red and the HECT domain in green. The WWP2-N isoform (underlined and in bold) spans position 1-335 (first and last amino acids circled in orange) whilst the C-isoform spans amino acid 440-870 (underlined and in bold, with first and last amino acids circled in purple).

### 1.3.2 WWP2 and its role in signalling pathways

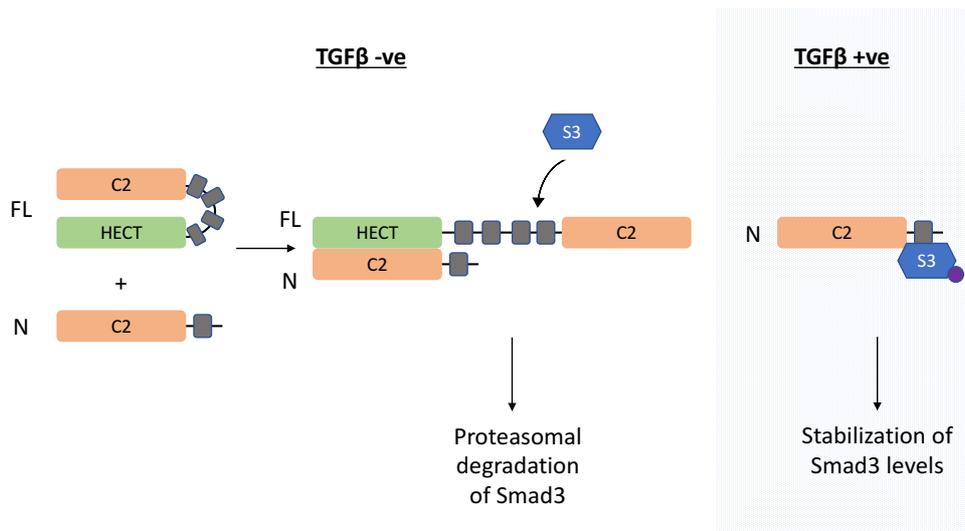
The WWP2 ubiquitin ligase has been highly implicated in many signalling pathways by the binding and subsequent inhibition of pathway modulators. Recently, it was shown that WWP2 acts as a mediator allowing the crosstalk between the WNT signalling pathway and the Notch pathway. Dishevelled homolog (DVL2) is a key protein in the WNT pathway that binds to Axin leading to the stabilization of  $\beta$ -catenin consequently activating transcription of target genes. DVL2 was shown to bind to WWP2 primarily by the PPXY motif, but also the YXY motif and the DEP domain on DVL2 which binds to the C2 domain of WWP2. The binding of DEP to C2 allows the activation of WWP2 by relieving the auto-inhibition of WWP2 through the C2 and HECT domains. More importantly, the disinhibition of WWP2 is reliant on the polymerization of DVL2. The DVL2-activated WWP2 then binds to and ubiquitinates the intracellular domains of the Notch3 receptor

(NICD3) in the signalling pathway subsequently causing the inhibition of Notch signal transduction (Mund *et al*, 2015).

Another signalling pathway that has been shown to interact with WWP2 is the TGF $\beta$  pathway (details of this pathway will be discussed in section 1.4). TGF $\beta$  stimulation affects levels of WWP2 isoforms differently. It was shown that whilst WWP2-FL levels increased with TGF $\beta$  stimulation, WWP2-N expression ultimately disappeared 4 days post-TGF $\beta$  treatment. Studies have also shown that not only does TGF $\beta$  affect levels of WWP2, but the reverse is also true as WWP2 can also alter levels of TGF $\beta$  signalling. This is carried out by the binding of TGF $\beta$  ligand-stimulated Smad mediators (Smad2, 3 and 7) to WWP2 ubiquitin ligase resulting in the proteasomal degradation of Smads. Although WWP2-FL binds to Smad 2, 3 and 7, preference lies with Smad2/3 binding. Additionally, this binding can be enhanced by TGF $\beta$  stimulation, causing a shift in binding preferences resulting in higher levels of Smad7 degradation by WWP2-mediated ubiquitination in comparison to Smad2/3. Notably, the presence of TGF $\beta$  increases rate of ubiquitin-mediated degradation of all Smads by WWP2 (Soond and Chantry, 2011).

WWP2-N, a truncated isoform containing only the C2 and WW1 domain was found to bind preferentially to Smad3 in a TGF $\beta$ -dose dependent manner. Further investigations showed that the WW1 domain was responsible for the binding to residue 180-240 of Smad3. In the absence of TGF $\beta$ , WWP2-N binds to WWP2-FL to relieve auto-inhibition leading to the binding and subsequent degradation of Smad3 (as shown in fig 1.3.4). However, in the presence of TGF $\beta$ , Smad3 phosphorylation by ligand binding releases its auto-inhibition mechanism therefore allowing the binding of WWP2-N and Smad3 ultimately stabilizing levels of Smad3 due to the lack of HECT domain in the N-isoform. On the other hand, WWP2-C has the ability to bind to Smad7 via the WW4 domain, also in a TGF $\beta$ -dependent manner which presumably signals for degradation due to the presence of an intact HECT domain (Soond and Chantry, 2011). Recent evidence has suggested the

binding of WW4 domain to not only Smad7 but also Smad3. NMR studies have shown that whilst the WW4 domain has the greatest binding affinity to phosphorylated Smad7, in fact the affinity to native Smad3 is higher than to native Smad7 (Wahl, 2016). To study this complex relationship between WWP2 and TGF $\beta$ , we must first obtain a full understanding of the TGF $\beta$  pathway and the role of Smad mediators.



**Figure 1.3.4- Effects of WWP2-N/FL binding and TGF $\beta$  stimulation on Smad3 levels.** A diagram showing the binding of WWP2-N to FL which releases auto-inhibition allowing the attachment of and consequent degradation of Smad3 in the absence of TGF $\beta$ . Under TGF $\beta$  conditions, Smad3 is phosphorylated to relieve its auto-inhibition allowing WWP2-N to bind directly to Smad3, however, due to the lack of HECT catalytic domain, Smad3 is not sent to proteasome for degradation.

## 1.4 TGF $\beta$ pathway

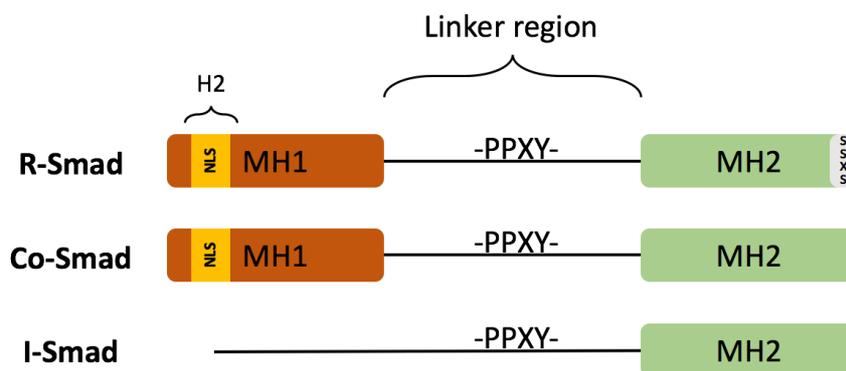
The TGF $\beta$  pathway is one of many signalling pathways in cells responsible for activating different genes which initiates an array of different responses such as cell differentiation and migration (Saitoh and Miyazawa, 2012). It also has a complex relationship with WWP2 E3 ligase as discussed above. This pathway is triggered by ligands including TGF $\beta$ , Bone Morphogenic protein (BMP) and Activin (Guo and Wang, 2009), binding to transmembrane Serine/ Threonine kinase receptors (Feng and Derynck, 2005). There are three TGF $\beta$  isoforms, TGF $\beta$ 1, TGF $\beta$ 2 and TGF $\beta$ 3. These are all produced in a 75 kDa latent form, bound to a homodimeric latency-associated protein (LAP). This complex is bound by disulfide bridges via cysteine residues to LAP and a latent TGF $\beta$ -binding protein (LTBP). Transglutaminase facilitates the covalent binding between N-terminal cysteine residue on the LTBP and the extracellular matrix (ECM) therefore anchoring all three components of the latent TGF $\beta$  to the ECM. There has been evidence suggesting that activation of TGF $\beta$  can be triggered by low pH levels, reactive oxygen species (ROS), and other proteases which would in turn elicit furin-type enzyme facilitated cleavage of mature 24 kDa TGF $\beta$  from LAP (Annes *et al*, 2003).

TGF $\beta$  signal transduction most often uses the canonical pathway which involves binding of ligand to TGF $\beta$ RI and TGF $\beta$ RII and uses traditional downstream mediators of Smad signalling molecules. However, although this is the most common way of TGF $\beta$  signal transduction, it is not the only pathway stimulated by TGF $\beta$  ligand binding. The non-canonical TGF $\beta$  pathway relies on other signal transducers to confer a signal in response to TGF $\beta$  binding that does not use Smad molecules as mediators of the signal.

### 1.4.1 The canonical pathway of TGF $\beta$ signalling

TGF $\beta$  signalling can be mediated through the canonical and the non-canonical pathway. The most widely used and studied canonical pathway involves the use of Smad signal mediators. These molecules are categorized into Regulatory-Smads (R-Smad), Common-mediator Smads (Co-Smad) and

Inhibitory Smads (I-Smad). Whilst stimulation of BMP induces signal transduction mediated by R-Smad1, 5, 8, Co-Smad4 and I-Smad6, TGF $\beta$  binding to membrane receptors trigger a signalling cascade that involves R-Smad2/3, Co-Smad4 and I-Smad7. All three categories of Smads contain the highly conserved C-terminal Mad-homology 2 (MH2) domain, however, the SSXS motif phosphorylation site is exclusive to the C-terminus of MH2 domains within R-Smads. The Mad-homology 1 (MH1) domain is present in both R-Smads and Co-Smads and contains elements such as the H2 lysine-rich helix and Nuclear location signal (NLS) that may facilitate the nuclear translocation of Smad mediators (Xiao *et al*, 2000). It has also been suggested that an 11-residue  $\beta$ -hairpin motif present in MH1 is responsible for their DNA binding abilities (Chai *et al*, 2003). The MH1 and MH2 domains are joined by a linker region that contains a PPXY motif containing specificity for E3 ubiquitin ligases such as WWP2 (Soond and Chantry, 2011). Although I-Smads lack the N-terminal MH1 domain, the linker region and PPXY motif remain intact. The structure of Smad proteins is summarized in fig 1.4.1. Prior to ligand activation, R-Smad2, and to a lesser degree R-Smad3 has been shown to be anchored to the intracellular side of the membrane by Smad anchor for receptor activation protein (SARA). The proximity between the R-Smads and TGF $\beta$  receptors increases the efficiency of TGF $\beta$  signal transduction (Goto *et al*, 2001).

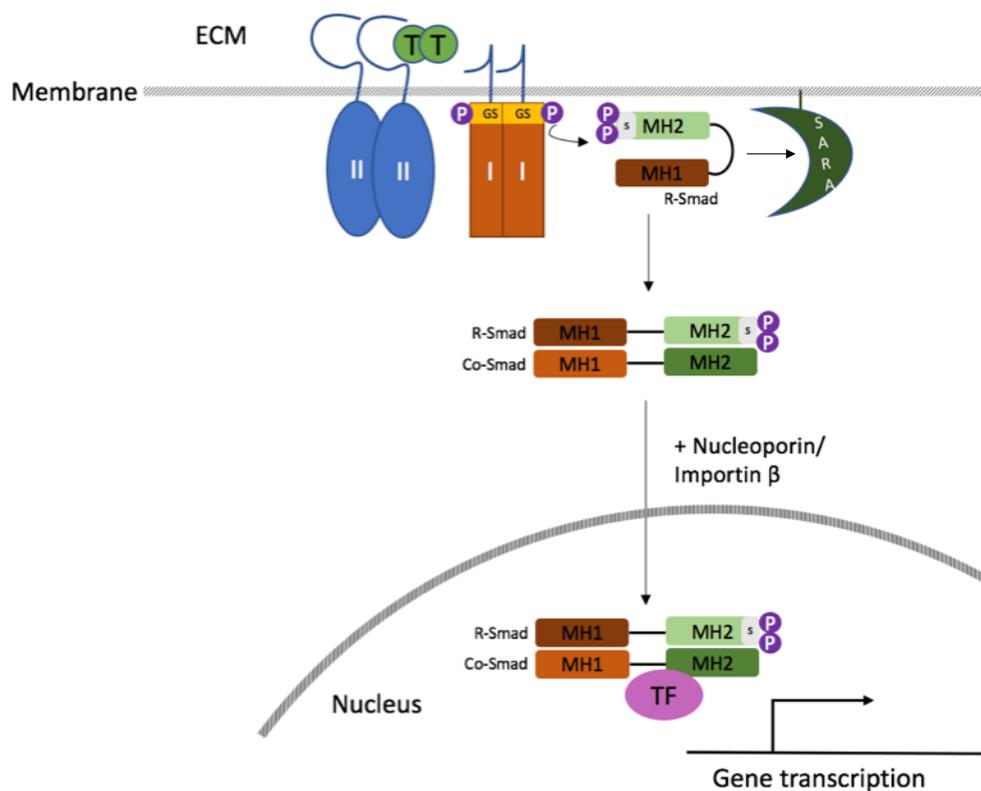


**Figure 1.4.1- A simplified schematic showing the domains and key motifs present in the Smad family of TGF $\beta$  signalling mediators.** The MH1 domain (in brown) contains the H2 lysine-rich helix and the NLS (in yellow) whilst the C-terminal SSXS motif is only expressed in the MH2 domain (in green) within R-Smads. I-Smads contain the linker region, PPXY recognition motif and MH2 domain but lack the N-terminal MH1 domain.

Type I TGF $\beta$  receptor (TGFR1) and Type II TGF $\beta$  receptor (TGFR2) are Serine/Threonine kinases responsible for conveying the extracellular signal into the inside of the cell. Both receptors are homodimeric transmembrane proteins that have an extracellular ligand binding domain, and an intracellular kinase domain (Shi and Massagué, 2003). As shown in fig 1.4.2, the first step of the cascade involves the binding of a mature TGF $\beta$  dimer to the extracellular binding site on TGFR2. This induces a conformational change in TGFR2 allowing the binding of TGFR1 to both the TGF $\beta$  ligand, and TGFR2. The proximity between the two receptors allows the transphosphorylation of Ser/Thr residues in the GS motif at the N-terminal TGFR1 kinase domain by the constitutively active TGFR2 (Massagué, 1998). The signal is then mediated by a combination of Smad molecules. R-Smad2/3 is recruited through recognition of the L3 loop next to the SSXS motif in MH2 domain by the L45 loop in TGFR1 adjacent to the GS motif allowing the phosphorylation of the SSXS motif in R-Smad2/3 (Chen *et al*, 1998). A positively charged pocket at the surface of Co-Smad4 then interacts with the phosphorylated SSXS motif of R-Smads to form a heterodimer ready for nuclear translocation (Wu *et al*, 2001).

Two major theories explain the mechanism employed for the nuclear import of Co-Smad-R-Smad complexes. The first one suggests that the N-terminal KKLKK motif within the H2 helix in MH1 of Smad3 is responsible for the binding of Importin- $\beta$  which mediates nuclear membrane translocation (Xiao *et al*, 2000). The second theory dictates that binding of R-Smad2 via a hydrophobic corridor in the MH2 domain to FG repeats in Nucleoporins such as CAN/Nup214 and Nup153 allow direct import into the nucleus (Xu *et al*, 2002). In the nucleus, an 11 residue  $\beta$ -hairpin present in the MH1 domain of R-Smad3 and Co-Smad4 within the heterodimeric Smad complex binds to the Smad binding element (SBE) sequence -CAGAC- in the DNA to allow regulation of target genes (Chai *et al*, 2003). To increase specificity of Smad-DNA binding, the Smad complex binds to both transcription coactivators and corepressors in the nucleus to regulate target gene transcription. An example of a transcription coactivator is GLI1. This transcription factor (TF)

forms a complex with Smad2/Smad4 in the nucleus and ultimately regulates the transcription of BCL2 gene via binding to its promoter. The GLI1 mediated transcription of BCL2 has been implicated in cancer by its function in cell growth and apoptosis. However, the regulation of BCL2 transcription by Smad2/Smad4/GLI1 is also dependent on the presence of p300/CREB-binding protein-associated factor (PCAF), a histone acetyltransferase that adds another level of specificity in this transcription regulation mechanism (Nye *et al*, 2014). On the other hand, the binding of Smad complexes with transcription corepressors such as TGIF in the nucleus leads to the recruitment of Histone deacetylases (HDACs) which changes the conformation of DNA in the nucleosome ultimately negatively regulating the transcription of target genes (Wotton *et al*, 1999).



**Figure 1.4.2- A diagram depicting the Canonical TGFβ pathway.** Upon TGFβ binding shown in green, the GS domain of TGFβR1 shown in orange is transphosphorylated by TGFβR2 which is shown in blue. This in turn phosphorylates the SSXS motif of MH2 domain within R-Smad allowing the release from SARA. The phosphorylated R-Smad then associates with Co-Smad4 and is subsequently translocated into the nucleus with the help of Nucleoporins and Importin β. This Smad complex then binds to various transcription factors (TF) to regulate target gene transcription.

#### 1.4.2 Regulation of the TGF $\beta$ canonical pathway

Due to the important role the TGF $\beta$  signalling cascade has on genes controlling growth and differentiation, it is vital that there are strict regulations to prevent overexpression and mis-expression of target genes. Smad mediators themselves are regulated by auto-inhibition where the N-terminus of Smad2 and Smad4 has been shown to bind to the C-terminal MH2 domain. This effectively shields binding sites within Smads to prevent nuclear translocation and DNA binding. Ligand activation of the cascade leads to the phosphorylation of Smads subsequently releasing it from auto-inhibition (Hata *et al*, 1997).

Another mode of negative regulation within the TGF $\beta$  signalling pathway is the Inhibitory I-Smads. These inhibitory Smads employ several different mechanisms to successfully inhibit TGF $\beta$  signalling. Following TGF $\beta$  stimulation, I-Smad7 binds to activated TGF $\beta$ 1 via its MH2 domain subsequently inhibiting the binding of R-Smads therefore meaning that phosphorylation of R-Smads does not occur (Hayashi *et al*, 1997). Additionally, I-Smad7 also recruits E3 ubiquitin ligases including Smurfs and WWP1 leading to the proteasomal degradation of both TGF $\beta$ 1 and I-Smad7 (Xu *et al*, 2012). Although I-Smad7 is degraded in the proteasome along with the TGF $\beta$ 1, continual TGF $\beta$  stimulation leads to the transcription of I-Smad7 through the binding of phospho-Smad2/3/4 to the SBE sequences in the Smad7 promoter region, thus stabilizing the levels of Smad7 (Stopa *et al*, 2000).

The direct binding of I-Smad7 through the MH2 domain in the nucleus to SBE sequences in target gene promoter regions competitively inhibits the binding of functional Smad complexes therefore reducing TGF $\beta$ -mediated transcription of targeted genes (Zhang *et al*, 2007). Furthermore, I-Smad7 not only binds to activated R-Smads in a competitive manner via the MH2 domain, hindering the R-Smad-Co-Smad complex formation, but it also acts as an adaptor protein supporting the proteasomal degradation of Smad2/3 mediated by NEDD4L (Yan *et al*, 2016; Gao *et al*, 2009).

As mentioned previously, a mechanism involved in the control of TGF $\beta$  signalling levels is ubiquitin-dependent proteasomal degradation. Ubiquitinated Smad mediators can be signalled for 26S proteasomal degradation subsequently affecting levels of TGF $\beta$  signalling. A positive or negative effect on signal transduction depends on the Smad mediator signalled for degradation. For example, if the I-Smad was ubiquitinated and subsequently sent to the 26S proteasome, the TGF $\beta$  signal would increase, however, if R-Smads or Co-Smads were signalled for degradation, the levels of TGF $\beta$  transduction would decrease. As its name suggests, the Smad Ubiquitin Regulatory Factor (SMURF), is one of the E3 ubiquitin ligases involved in the ubiquitin-dependent regulation of Smad mediators. It has been found that Smurf regulators have the ability to ubiquitinate and subsequently send both R-Smads and I-Smads for proteasomal degradation, suggesting that Smurfs can positively and negatively affect TGF $\beta$  signal transduction. Furthermore, Smurfs can also inhibit TGF $\beta$  signalling by using Smad7 as a carrier to translocate from the nucleus to TGF $\beta$ R<sub>s</sub> at the membrane to tag the receptor for proteasomal degradation (Kavsak *et al*, 2000) (Ebisawa *et al*, 2001). Although Smurfs are one of the main ubiquitin ligases to regulate TGF $\beta$  signalling, many other ligases in the E3 family are also involved in this regulation, including but not exclusive to ITCH and WWP1 (Inoue and Imamura, 2008).

Deubiquitinating enzymes (DUBs) counteract the effects of E3 ligases by removing ubiquitin molecules from ubiquitinated proteins or by inhibiting the ubiquitination of target proteins (Herhaus and Sapkota, 2014). The ability of DUBs to reverse ubiquitination of Smad mediators means that they also play an important part in the regulation of TGF $\beta$  signalling. USP4, 11 and 15 are three DUBs of the USP family all identified as DUBs that facilitate the deubiquitination of TGF $\beta$ R<sub>1</sub> therefore enhancing the levels of TGF $\beta$  signalling (Zhang *et al*, 2012; Al-Salihi *et al*, 2012; Eichhorn *et al*, 2012). On the contrary, CYLD, another member of the USP family has been shown to bind to and deubiquitinate Smad7 mediators which increases the natural inhibition of TGF $\beta$  signalling (Herhaus and Sapkota, 2014). Together, the

multitude of mechanisms and proteins involved in the regulation of this signalling pathway maintains the normal functions of TGF $\beta$  and consequently TGF $\beta$ -dependent processes.

#### 1.4.3 TGF $\beta$ , BMP and Cancer

TGF $\beta$  plays an important role in normal cell functions such as cell survival, proliferation and differentiation. However, it is also highly implicated in many diseases, most notably cancer. The effects of TGF $\beta$  on cancer development is not as straightforward as having strictly a positive or a negative influence, rather, it has been well established that TGF $\beta$  has a dual-role in oncogenesis. TGF $\beta$  has been shown to exert an inhibitory force and acting as a tumour suppressor by being a key regulator of cell proliferation, inducing the cyclin-dependent kinase inhibitor (p15) and inhibiting other cell cycle mediators such as cyclin. This lowers activation of E2F via inactivation of retinoblastoma (Rb), effectively halting the cell cycle by a decrease of mediator genes such as c-myc (Epstein *et al*, 2000). Nevertheless, increased levels of TGF $\beta$  stimulation and the increase in aggressiveness of cancer has been suggested as a switch that shifts the tumour-suppressing properties of TGF $\beta$  to having oncogenic effects instead. This could be due to the development of resistance by oncogenic cells to TGF $\beta$ -mediated growth inhibition, but not the effect of TGF $\beta$  on oncogenic mechanisms. Moreover, it has been shown that at this increased level of TGF $\beta$ , this cytokine is capable of stimulating its own production therefore creating a positive feedback loop that further favours cancer development (Lasfar and Cohen-Solal, 2010).

TGF $\beta$  increases the aggressiveness and aids development of cancer in several ways. Firstly, and perhaps most established is its ability to induce Epithelial to mesenchymal transition (EMT). This form of cell trans-differentiation is activated by many EMT markers including SNAIL, Twist and Vimentin, whilst a decrease in levels of E-cadherin is advantageous to this process. TGF $\beta$  induces EMT by increasing Smad-dependent transcription of Twist, fibronectin, Vimentin and SNAIL which also represses

levels of E-cadherin transcription ultimately changing aspects of cell adhesion and motility (Lamouille *et al*, 2014).

Another mechanism crucial to the growth of cancer is angiogenesis which is linked to the growth factor VEGF. TGF $\beta$  increases levels of VEGF which in turn promotes TGF $\beta$ -mediated apoptosis, an important step in angiogenesis (Ferrari *et al*, 2009). TGF $\beta$  further supports tumour development by the suppression of Natural killer (NK) cells in the body that are normally tasked with the detection and obliteration of foreign pathogens and tumour cells. It was found that in response to Interleukin-15 (IL-15), TGF $\beta$  stimulation inhibits the mammalian Target of Rapamycin kinase (mTOR) which is an important component in NK cell signal transduction (Viel *et al*, 2016). Combined with the other effects of TGF $\beta$ , this growth factor proves to be a potent oncogenic factor through altering the functions in many aspects of cellular homeostasis.

Specifically in metastatic and invasive melanoma, a significant increase in TGF $\beta$  levels have been observed in comparison to healthy melanocytes. Intriguingly, the expression of TGF $\beta$ 2 was not found in melanocytes but exclusively in melanoma cells (Lasfar and Cohen-Solal, 2010). As mentioned above, the dual role of TGF $\beta$  is not lost in melanoma as evidence show that in early stages of cutaneous melanoma, TGF $\beta$  increases PAI-1, a plasminogen activator inhibitor which inhibits cells invasion and migration by lowering plasmin levels (Humbert and Lebrun, 2013). With the progression of melanoma, it has been suggested that the dual role of TGF $\beta$  switch-over is a result of the development of resistance to TGF $\beta$ -mediated tumour suppressive effects. Oncogenic effects of TGF $\beta$ , for example linked to the phosphorylation of Smad3 linker region in malignant melanoma, is on the other hand unaffected which subsequently leads to the decrease in p15 and p21 tumour suppressor gene transcription (Cohen-Solal *et al*, 2011).

Not only has the TGF $\beta$  cytokine within the TGF $\beta$  superfamily shown oncogenic effects in melanoma, but many members of the BMP family have

also shown the ability to assist cancer development by inducing angiogenesis, migration and invasion alike (Rothhammer *et al*, 2007). In a study by Rothhammer, the expression profiles of BMP in nine melanoma cell lines was investigated and results suggested that whilst BMP2 only showed an increased expression in two cell lines, BMP4 and 7 showed high expression in all nine cell lines. In the same study, Rothhammer *et al* were also able to pinpoint BMP4 as the main promoter of cell migration and invasion (Rothhammer *et al*, 2005). Conversely, there have also been many studies that have shown BMP to have a negative effect on oncogenesis. In one study, it was shown that not only was the growth of melanoma cells stunted by BMP7, but these cells also displayed high levels of morphological change from the mesenchymal phenotype to the epithelial phenotype suggesting that BMP7 has a role to play in mesenchymal to epithelial transition (MET) (Na *et al*, 2009).

## 1.5 Cancer and malignant melanoma

Cancer is a disease that can be described as the abnormal growth of cells deviating from homeostatic levels. This abnormal increase is the result of multiple cellular mechanisms combined including, but not exclusive to, atopic cell proliferation to begin with, then angiogenesis and EMT which increases the aggressiveness of the cancer leading to metastasis (Perrot *et al*, 2013). Signals that initiate atopic proliferation usually arise from mutations within fundamental genes guarding the normal cell cycle and are known as tumour suppressors and proto-oncogenes. Due to the complex nature of cell homeostasis and signalling pathway cross-talk, cancer can result from any one or several mutations in a signalling cascade component, or regulatory signalling molecules (Mclaughlin *et al*, 2003).

Cutaneous melanoma is a type of skin cancer originating from melanocytes, cells responsible for the production of pigments known as melanin. The development of melanoma can present as abnormal moles or lesions (Dummer *et al*, 2009). In normal development, Microphthalmia-associated transcription factor (MITF) regulates the differentiation of neural crest cells (NCC) into melanoblasts, a progenitor cell of melanocytes. This happens in a section of the neural crest known as the trunk, where the production of glia cells also occurs. Cells originating from NCCs have high EMT potential to allow ease of migration during development, and can also self-renew to ensure steady levels of NCCs. The aggressiveness of melanoma has therefore been suggested to be a result of the properties melanoma cells attained from its cell origin (Mort *et al*, 2015).

Melanoma can be categorised into different stages dependent on its appearance and aggressiveness. Stage 0 is the most benign form and describes a tumour that is localised to the epidermis. Stage IA and IB include tumours that are up to 2 mm in thickness with or without ulcerations which has not metastasized. Stage IIA, IIB and IIC are used to describe tumours that are between 1-4 mm of thickness, with or without ulcerations that have not

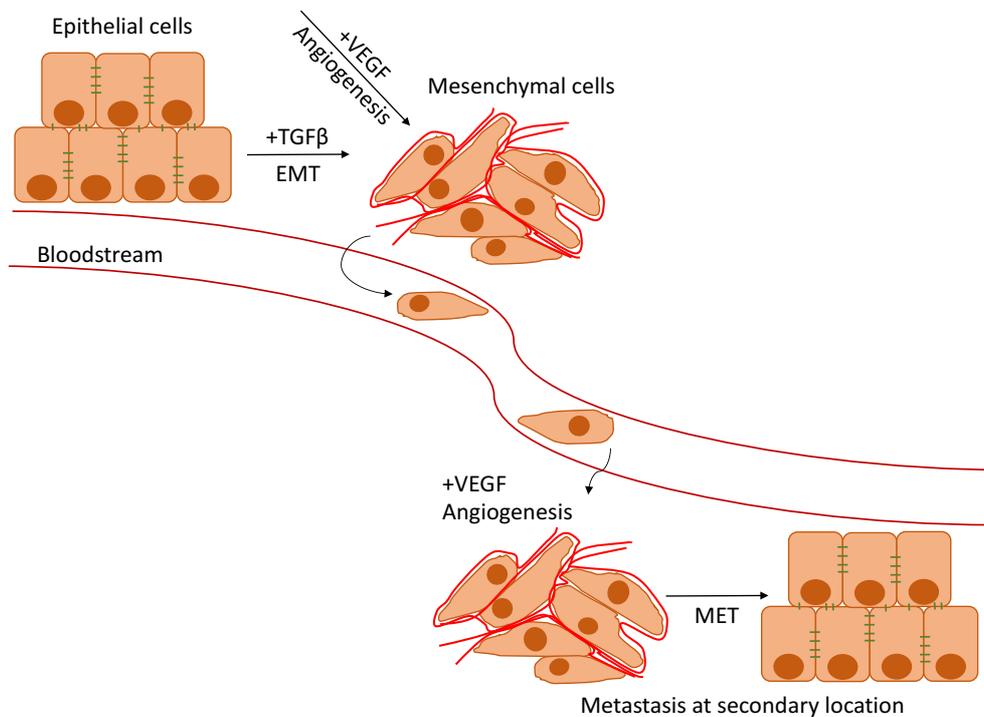
metastasized whilst stage IIIA, IIIB and IIIC are used to describe tumours of any thickness that has metastasized to nearby lymph nodes and tissues. Finally, the most aggressive form of melanoma is at stage IV, where the tumour has metastasized to lymph nodes and other organs away from the primary tumour and will require a combination of treatments such as surgical removal of tumour, radiotherapy and chemotherapy (Melanoma Research Foundation, 2017). Since the 1990s, incidences of melanoma in the UK has increased by 119% with the likelihood of melanoma development enhanced by factors such as UV-exposure, hereditary mutations, skin colour and increased number of moles or the presence of benign moles (Cancer Research UK, 2014; Jilaveanu *et al*, 2009). Oncogenes such as CDK4 and BRAF are under strict regulation by many signalling pathways in the body including the TGF $\beta$  pathway, however, mutations in these genes found in melanoma patients can cause abnormal signalling in regulation of cell growth and differentiation subsequently favouring oncogenesis (Perrot *et al*, 2013).

#### 1.5.1 Cancer progression and EMT

The aggressiveness of a cancer can depend on the tumour size, and metastasis. Benign cancers are tumours that have stayed in their primary location without having developed the ability to metastasize to a secondary location within the body. The increase in size of the tumour relies highly on a mechanism known as angiogenesis. This is the development of blood vessels within the tumour to supply cancerous cells with adequate oxygen to survive. Without this mechanism, the tumour is restricted in its size which results in a less aggressive cancer and a decreased chance of metastasis. The VEGF signalling pathway is predominantly responsible for angiogenesis, therefore making VEGF a key player in tumorigenesis (Hoeben *et al*, 2004).

For the cancer to metastasize, a crucial event that must occur is EMT. Together with other occurrences, EMT contributes to oncogenesis by altering cell adhesion, migration and proteolysis (Bogenrieder and Herlyn, 2003). The trans-differentiation of epithelial cells to mesenchymal cells is characterized by three major features including the change in morphology

from globular cells arranged in a mono-layer associated with the basement membrane via apical-basal polarity, to spindle shaped cells containing protrusions for migration purposes; a shift in cell-cell interaction proteins to vimentin and fibronectin; and the acquisition of ability to invade the extracellular matrix (ECM) due to the motile trait of mesenchymal cells. These changes collectively allow cancerous cells to break away from primary tumour into the ECM or blood vessels and migrate to secondary sites thus negatively affecting prognosis of the cancer (Yang and Weinberg, 2008). At the secondary location, the cells that had undergone EMT must be reversed by the process of MET to allow the cancer cells to anchor and form a secondary tumour (Kalluri and Weinberg, 2009). The process of metastasis is demonstrated in fig 1.5.1.



**Figure 1.5.1- The process of EMT and angiogenesis in oncogenesis.** Epithelial melanoma cells are uniform cells held together by cell-cell interactions which undergo EMT as a result of  $TGF\beta$  stimulation. Mesenchymal cells have properties allowing them to escape the primary location of the tumour and metastasize via the bloodstream to a secondary location. VEGF is an essential growth factor that induces angiogenesis allowing blood vessels to develop within tumours. At the secondary location, melanoma cells undergo MET, the reverse of EMT, allowing metastatic cells to anchor and form tumour at a secondary location.

Although EMT is a crucial step in oncogenesis, it must not be forgotten that it is also a mechanism that is needed for normal embryogenesis and organ development. There are three types of EMT in total. Type 1 EMT generates mesenchymal cells that lack invasive and fibrotic properties and acts strictly as a tool for organ development and embryo formation. Type 2 EMT occurs in response to inflammation and results in fibrotic cells that act as part of a wound healing process. Finally, Type 3 EMT is the result of mutations caused by the malfunction of oncogenes and tumour suppressors consequently leading to the metastasis of cancer cells to a secondary location (Kalluri and Weinberg, 2009).

There are a wide range of markers that can be used to detect EMT. The levels of these EMT markers provide a good indication to the levels of EMT occurrence in cells due to their respective biological functions. For example, N-Cadherin and E-cadherin are adhesion proteins that decrease during EMT whilst transcription factors SNAIL and Twist, which function to represses E-cadherin, increase during EMT (Kalluri and Weinberg, 2009). Another EMT marker that increase during EMT is Vimentin, an intermediate filament protein found in mesenchymal cells (Zavadil and Böttinger, 2005).

Overall, a multitude of signalling pathways and mechanisms contribute to the many different aspects of cancer development, such as EMT. A mechanism shown to be highly involved in oncogenesis is ubiquitination, mediated by E3 ubiquitin ligases.

#### 1.5.2 Role of Ubiquitin ligases in Cancer

Many studies have investigated the role E3 ubiquitin ligases play in oncogenesis. Some identify the overexpression of E3 ubiquitin ligases in cancer suggesting their role in oncogenesis, whilst others show a decrease in E3 ligase levels in various cancers. Mdm2 is a RING E3 ubiquitin ligase that interacts with p53 and ultimately facilitates the proteasomal degradation of this tumour suppressor protein (Honda *et al*, 1997). More than half of all human cancers show the inhibition of p53 suggesting that amongst other

regulatory processes, E3 ligases such as Mdm2 might be involved in the process of oncogenesis (Sun, 2006). A HECT E3 ubiquitin ligase that has been suggested as a regulator of cancer development is EDD1. The role of EDD1 includes mediating the key DNA damage response and cell cycle checkpoints in normal cell proliferation (Munoz *et al*, 2007). The overexpression of this tumour suppressor gene has been observed in ovarian, breast, gastric and colon cancers (Subbaiah *et al*, 2016). Together, evidences suggest that E3 ubiquitin ligases can have both positive and negative impacts on oncogenesis by marking specific molecules for proteasomal degradation.

Studies on NEDD4 E3 ubiquitin ligases have demonstrated mostly oncogenic properties in melanoma. In one study, it was found that the NEDD4-L E3 ligase, a homologue of NEDD4-2, was expressed in many metastatic melanomas *in vitro*, but not in benign melanocytes. In the same study, down-regulation of NEDD4-L was shown to decrease cell growth in melanoma cell lines. *In vivo* studies on mice were performed and highlighted that increased expression of NEDD4-L led to melanoma cell growth. All in all, NEDD4-L has shown strong oncogenic properties that contributes to the progression and aggressiveness of melanoma (Kito *et al*, 2014). In another study, evidence was provided showing the oncogenic properties Smurf2 has in melanoma. Mitogen-activated protein kinase kinase (MEK) inhibitors represent one therapeutic approach for treating melanoma, however it was shown that these inhibitors were not as efficient a treatment as originally thought. It was later found that Smurf2 increases levels of MITF in melanoma cells which increased resistance to the cytotoxicity induced by MEK inhibitors. Consequently, follow-up studies showed that the downregulation of Smurf2 lowered levels of MITF and significantly increased the cytotoxicity of MEK inhibitors in melanoma therefore improving the efficiency of this therapeutic measure (Smith *et al*, 2013).

As mentioned previously, WWP2 was identified as the E3 ubiquitin ligase responsible for the ubiquitination of PTEN tumour suppressor. The inhibition of this suppressor leads to the increased activity of Akt and subsequent

growth and proliferation, making WWP2 a potential onco-protein (Maddika *et al*, 2011). It was also found that in liver cancer cells, WWP2 upregulated levels of chemokine receptors CXCR3 and CCR5 which are responsible for cell proliferation and motility (Qin *et al*, 2016), whilst in lung adenocarcinoma, the down-regulation of WWP2 directly correlates to the decrease in cell migration and invasion (Yang *et al*, 2016). Together, these studies provide evidence that suggest WWP2 is an oncogene with negative effects on cancer prognosis. On the other hand, conflicting evidence has suggested WWP2 as a tumour suppressor in ovarian cancer. It was identified that WWP2 is a negative regulator of the Notch3 signalling pathway, one that is highly implicated in ovarian cancer. This interaction leads to the decrease in Notch3-mediated cell morphological change and platinum resistance (Jung *et al*, 2014). Although studies have shown implications of Smurf2 and NEDD4-2 in melanoma, no links have yet been found between WWP2 and melanoma. The importance of elucidating the role of WWP2 in melanoma is therefore highlighted by the lack of previous research into their relationship and the conflicting evidence suggesting that WWP2 may have alternate roles in oncogenesis dependent on the different cell types and cancers.

## 1.6 Aims

WWP2 E3 ubiquitin ligase and the TGF $\beta$  pathway is highly implicated in the disease of cancer. Currently, little is known about the WWP2 protein oncogenic potential, especially the truncated WWP2 isoforms which may have vastly different functions in oncogenesis compared to the full-length protein. In the above section, the complex interplay between TGF $\beta$ , WWP2 and cancer was explained, however, many aspects of their relationship is still unclear. We postulate that TGF $\beta$ -dependent EMT in oncogenesis is partly mediated by WWP2 and that the relationship between WWP2 and TGF $\beta$  is a complex feedback loop, namely due to the ability of TGF $\beta$  in affecting expression of WWP2 but also the ability of WWP2 in regulating the levels of TGF $\beta$  signalling. The overall aim of this thesis is to explore different aspects of the WWP2 gene to create a clearer picture of its relevance in oncogenesis and its relationship with TGF $\beta$ . To achieve this, a few key areas were explored, which are highlighted below.

- I. Investigating the role of WWP2 WW recognition domains in Smad3-dependent TGF $\beta$  gene expression in the hope of discovering a WW peptide which could inhibit Smad function
- II. Identification and validation of novel truncated WWP2 isoforms
- III. Studying the expression profiles of novel and established WWP2 isoforms in cancerous and non-cancerous cells
- IV. Examining the role of WWP2 isoforms in TGF $\beta$ -dependent EMT
- V. Exploring the transcriptional regulation responsible for the generation of truncated WWP2 isoforms

The collective data from experiments performed during this thesis should help us achieve a more comprehensive idea of WWP2 and its role in TGF $\beta$ -mediated oncogenic signalling.

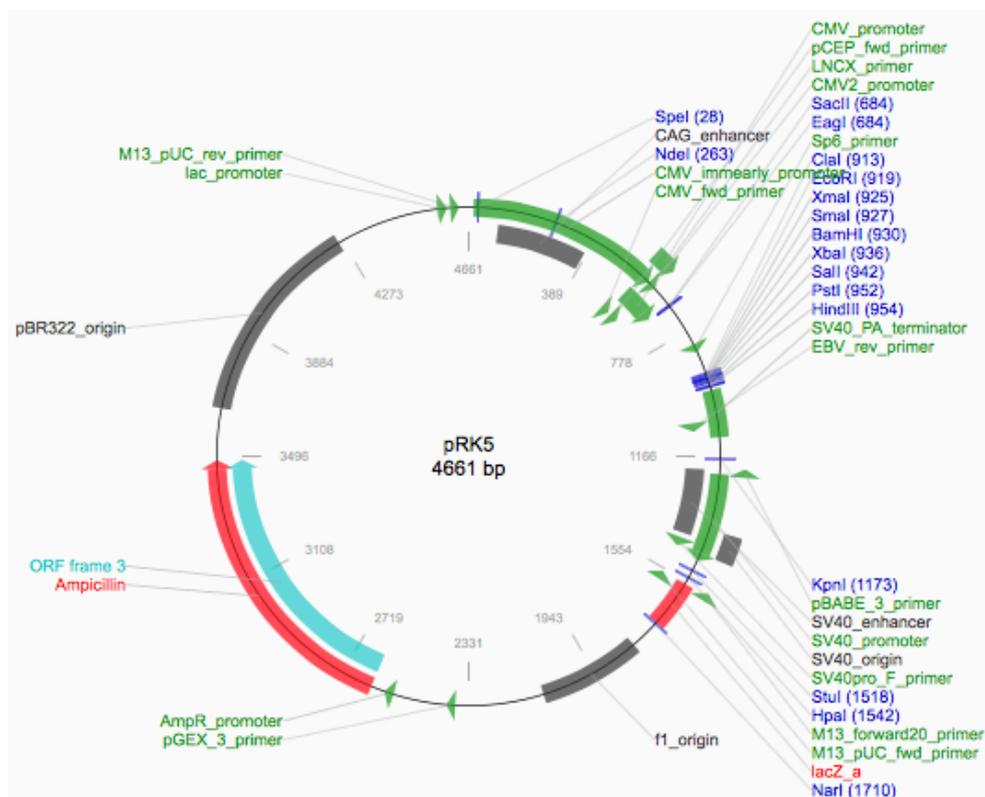
## Chapter 2 Methods and Materials

## 2.1 Cloning

Cloning was a substantial part of the research as most experiments required WWP2 constructs. A few sets of constructs were cloned during this research period using a range of different cloning methods. TA cloning was used for Promoter region cloning into pGL4.27 due to the size of the constructs, whilst T4 and In-fusion was used for all other constructs cloned. These are shown below in table 2.1-1.

*Table 2.1-1- Cloning of WWP2 constructs and the cloning methods used*

Constructs	Vector	Cloning methods
WWP2-C promoter 0.5, 1.2, 2 kb	pGL4.27	TA, T4
WWP2-ΔHECT promoter 0.5, 0.7, 1, 2 kb	pGL4.27	TA, T4, In-fusion
WWP2-FL, WW1-2, 2-3, 3-4, 1-4, ΔHECT	pRK5-N-HA	T4
WWP2 WW1, 2, 3, 4	pRK5-N-HA	T4
WWP2 WW3-4, WW 1-4, ΔHECT, ΔC2	pRK5-N-HA-C-STREP	T4



*Figure 2.1.1- The pRK5 plasmid map (www.addgene.org accessed on 22/12/2015)*

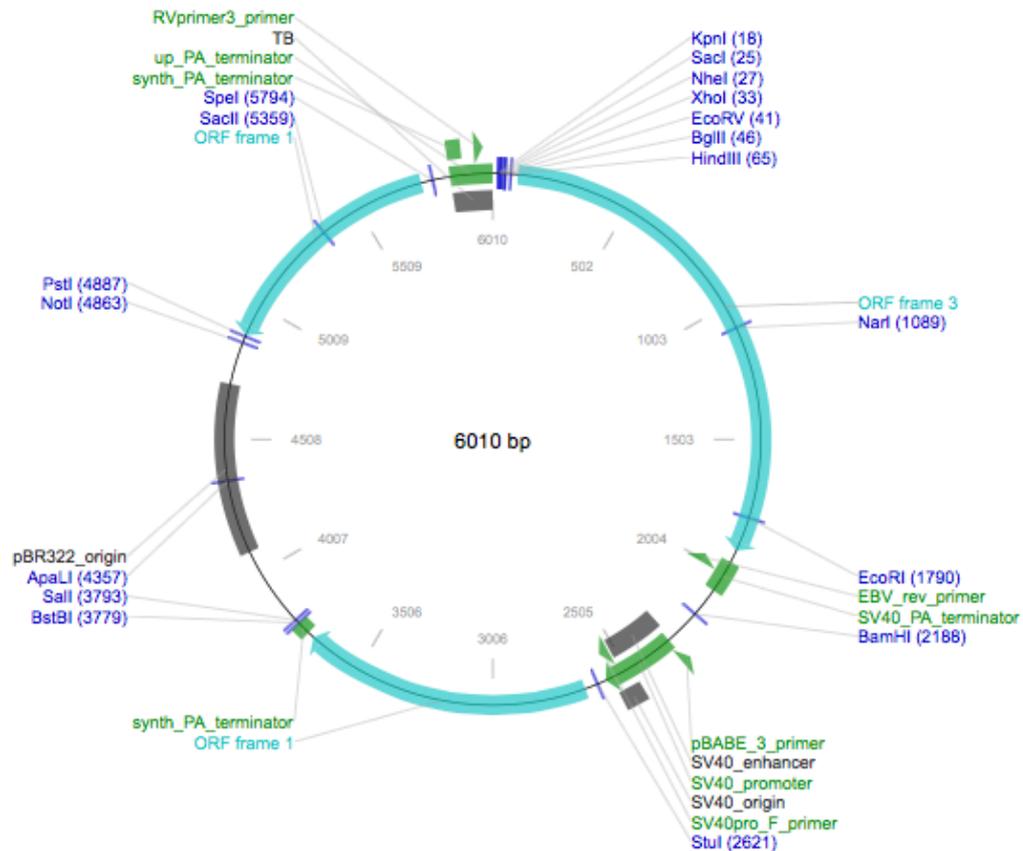


Figure 2.1.2- The pGL4.27 plasmid map ([www.addgene.org](http://www.addgene.org) accessed on 22/12/2015)

### 2.1.1 Genomic DNA extraction

Genomic DNA from HEK293A/T cells were used as templates for cloning. This was extracted using the Promega Wizard Genomic DNA purification kit. Manufacturers protocol was followed and DNA was rehydrated in 100 µl of DNA rehydration solution provided with the kit.

### 2.1.2 High fidelity PCR

#### 2.1.2a- Taq DNA PCR

This method was used prior to TA cloning as it produces poly-A-tails. The ThermoFisher Platinum High fidelity DNA polymerase guide was followed using the PCR reaction conditions described in table 2.1-2 and primers listed in table 2.1-4. The products were ran on a 1% agarose gel electrophoresis for 1 hr at 100 V.

### 2.1.2b Phusion PCR

Thermo Scientific High fidelity Phusion Polymerase was used to amplify inserts that were cloned using T4 ligation. The 3-step protocol provided by supplier was followed, using PCR reaction conditions provided in table 2.1-3 and cloning primers shown in table 2.1-4. Products were then ran on a 1% agarose gel electrophoresis for 1 hr at 100 V.

*Table 2.1-2-PCR reaction conditions used for Platinum High fidelity DNA Taq PCR*

<b>Step</b>	<b>Temperature</b>	<b>Time</b>	<b>Cycles</b>
Initial Denaturation	94°C	30	1
Denaturation	94°C	15	30
Annealing	70°C	30	
Extension	68°C	1 min per kb	
Hold	4°C	indefinite	1

*Table 2.1-3- PCR reaction conditions used for High fidelity Phusion PCR*

<b>Step</b>	<b>Temperature</b>	<b>Time</b>	<b>Cycles</b>
Initial Denaturation	98°C	30	1
Denaturation	98°C	10	30
Annealing	*	10	
Extension	72°C	10	
Final extension	72°C	5 min	1
Hold	4°C	indefinite	1

*Table 2.1-4- Cloning primers obtained from Eurofins Genomic used for various cloning processes*

<b>Construct</b>	<b>Vector</b>	<b>FW Primer sequence 5'-3'</b>	<b>RV Primer sequence 5'-3'</b>
WWP2-C promoter 0.5kb	PGL4.27	CTCGAGCTCTGGGCGTGGG CA GGGCTT	AAGCTTCTAAACACAAGAAACA CAGCAGAACACAG
WWP2-C promoter 1.2kb	PGL4.27	CTCGAGGAGGGCTGTGGAA GG GGGCTT	AAGCTTCTAAACACAAGAAACA CAGCAGAACACAG
WWP2-C promoter 2kb	PGL4.27	CTCGAGAAAGGCAAGGATT CAGAAAAGGGC	AAGCTTCTAAACACAAGAAACA CAGCAGACACAG
WWP2-ΔHECT promoter 0.5kb	PGL4.27	CTCGAGTACTCAACAATATC ACCAAGGAA	AAGCTTCTGGAAATAGGAACCT CAGAATG
WWP2-ΔHECT promoter 1kb	PGL4.27	CTCGCTAGCCTCGAGGAAG AAAGGCCTGGGTGT	CCTCTAGTGTCTAAGCTTCTGGA AATAGGAACCTCAGA
WWP2-ΔHECT promoter 2kb	PGL4.27	CTCGCTAGCCTCGAGGCAG TCCTGTTTATGTTT	CCTCTAGTGTCTAAGCTTCTGGA AATAGGAACCTCAGA
WWP2-FL	PRK5-N- HA	CGCTAGCTTGCTCGAGATG GCATCTGCCAGCTCTAGCC	CGGTATCGATAAGCTTTTACTCC TGTCCAAAGCCCT
WWP2 WW1-2	PRK5-N- HA	CGCTAGCTTGCTCGAGCAG GCCCCGACGCTCT	CGGTATCGATAAGCTTTTACTCG CACGTA CTCCGCG
WWP2 WW2-3	PRK5-N- HA	CGCTAGCTTGCTCGAGGAG CGGCCCTTCCTC	CGGTATCGATAAGCTTTTACTG GGTCCGGGGATCC
WWP2 WW3-4	PRK5-N- HA	CGCTAGCTTGCTCGAGGAT CCCCTGGGCCCCCT	CGGTATCGATAAGCTTTTACTCA AACCCCGGGCGA
WWP2 WW1-4	PRK5-N- HA	CGCTAGCTTGCTCGAGCAG GCCCCGACGCTCT	CGGTATCGATAAGCTTTTACTCA AACCCCGGGCGA
WWP2-ΔHECT	PRK5-N- HA	CGCTAGCTTGCTCGAGATG CAGCACTTCAGCCAAAGAT	CGGTATCGATAAGCTTTTATCAC ATGATGTACTCTTCTTGTTCTC C
WWP2-FL-ΔC2	PRK5-N- HA	CGCTAGCTTGCTCGAGATG CAGCTGACCCTGAAC	CGGTATCGATAAGCTTTTACTCC TGTCCAAAGCCCT
WWP2 WW1	PRK5-N- HA	CGCTAGCTTGCTCGAGCAG GCCCCGACGCTCT	CGGTATCGATAAGCTTTTAAAG AAGGGGCCGCTCC
WWP2 WW2	PRK5-N- HA	CGCTAGCTTGCTCGAGGAG CGGCCCTTCCTC	CGGTATCGATAAGCTTTTACTCG CACGTA CTCCGCG
WWP2 WW3	PRK5-N- HA	CGCTAGCTTGCTCGAGGAT CCCCTGGGCCCCCT	GGTATCGATAAGCTTTTACTGG GTCCGGGGATCC
WWP2 WW4	PRK5-N- HA	CGCTAGCTTGCTCGAGCAG GGGATGATCCAGGA	CGGTATCGATAAGCTTTTACTCA AACCCCGGGCGA
WWP2 WW3-4 Strep	PRK5-N- HA	CGCTAGCTTGCTCGAGGAT CCCCTGGGCCCCCT	GGTATCGATAAGCTTTTACTTTT CGAACTGCGGGTGGCTCCACTC AAACCCCGGGCGA
WWP2 WW1-4 Strep	PRK5-N- HA	CGCTAGCTTGCTCGAGCAG GCCCCGACGCTCT	GGTATCGATAAGCTTTTACTTTT CGAACTGCGGGTGGCTCCACTC AAACCCCGGGCGA
WWP2-ΔHECT Strep	PRK5-N- HA	CGCTAGCTTGCTCGAGATG CAGCACTTCAGCCAAAGAT	GATAAGCTTTTACTTTTCAACT GCGGGTGGCTCCACATGATGTA CTTCTCTGTT
WWP2-N-ΔC2 Strep	PRK5-N- HA	CGCTAGCTTGCTCGAGATG CAGCTGACCCTGAAC	CGATAAGCTTTTACTTTTCAAC TGCGGGTGGCTCCACCAGCCTG GAGGAAG

### 2.1.3 Gel extraction

The Qiagen QIAquick gel extraction kit was used to purify PCR insert from agarose gels. Standard manufacturers protocol followed and inserts were

eluted with 30 µl of elution buffer. For quantification of PCR product, please refer to section 2.1.10.

#### 2.1.4 Restriction enzyme digestion

PCR inserts and plasmid vectors were digested with *XhoI* and *HindIII* restriction enzymes provided by Promega before T4 ligation. The digestion reactions were incubated at 37 °C for 3-4 hrs using components listed in table 2.1-5. These were subsequently ran on a 1% agarose gel for 1 hr at 100 V, then gel extracted as described in section 2.1.3.

*Table 2.1-5- Components in a 40 µl XhoI HindIII double restriction digest*

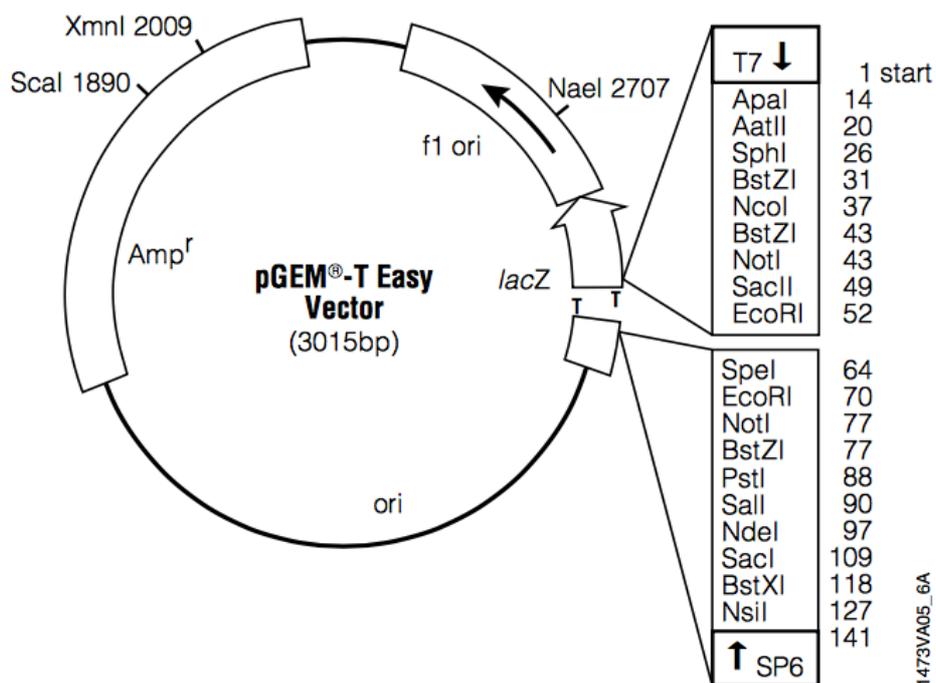
<b>Components</b>	<b>Amount</b>
10x <i>HindIII</i> Buffer	4 µl
BSA	0.4 µl
<i>HindIII</i>	2 µl
<i>XhoI</i>	2 µl
Insert / vector DNA	1-2 µg
dH <sub>2</sub> O	Up to 40 µl

#### 2.1.5 TA Cloning

Promega pGEM-T easy vector systems were used for TA cloning with a 1:1 ratio of insert and linearized pGEM-T easy vector. Ligation reactions were incubated at room temperature for 1 hr using components listed in table 2.1-6 and the pGEM-T easy vector is shown in figure 2.1.3. Cloned DNA was then added to a standard heat shock transformation described in section 2.1.8, using XL10-gold ultra-competent cells (Agilent Technologies).

*Table 2.1-6- A 10 µl T4 ligation reaction to ligate PCR insert into pGEM-T easy vector*

<b>Components</b>	<b>Amount</b>
2x rapid ligation buffer	5 µl
pGEM-T easy vector (50ng)	1 µl
Insert DNA	50 ng
T4 DNA ligase	1 µl
Nuclease free water	Up to 10 µl



**Figure 2.1.3- A plasmid map showing the pGEM-T easy vector** (Promega pGEM-T easy vector system protocol 22/12/2015)

### 2.1.6 Standard T4 ligation

Digested PCR inserts were cloned into linearized vectors using Promega T4 ligase. A vector:insert ratio of 3:1 or 1:1 was used as suggested by Promega and reactions were incubated at room temperature for 3 hrs. Cloned constructs were then transformed into Stellar high efficiency competent cells (Clontech) using the standard heat shock transformation protocol described in section 2.1.8. Details of the ligation reaction is shown below in table 2.1-7.

**Table 2.1-7- Components in a standard 10 µl T4 reaction**

Components	Amount
Ligase 10x Buffer	1 µl
Linearized vector	100 ng
Insert DNA	*
T4 Ligase	0.5 µl
dH <sub>2</sub> O	Up to 10 µl

### 2.1.7 In-Fusion HD Cloning

The Clontech In-Fusion cloning kit is one that allows you to bypass the need of digesting PCR inserts. It ligates the vector complementary sequence of the insert to the linearized vector during a 15 mins 50 °C incubation step. The In-Fusion reaction uses components listed in table 2.1-8. The cloned product is then transformed into high efficiency competent Stellar cells using the standard heat shock protocol.

*Table 2.1-8- Components included in a 5 µl In-Fusion reaction*

<b>Components</b>	<b>Amount</b>
5x In-Fusion HD enzyme premix	0.5 µl
Linearized vector	50-100 ng
PCR insert	50-100 ng
dH <sub>2</sub> O	Up to 5 µl

### 2.1.8 Heat-Shock transformation into DH5α and other competent cells

In a 1.5 ml microcentrifuge tube, 1-3 µl of DNA was added to 200 µl of DH5α cells and incubated for 30 mins on ice. The sample was then heat shocked at 42 °C for 45 s and immediately placed back on ice for a further 10 mins. It was then incubated in 37 °C shaker for 45-60 mins after the addition of 300 µl of LB. Agar plates containing ampicillin (100 µg/ml) were used to spread around 100 µl of transformation sample onto and incubated overnight at 37 °C.

### 2.1.9 Plasmid isolation

The Qiagen plasmid plus midiprep kit and the Qiagen QIAprep spin miniprep kit were used to isolate plasmid from bacterial inoculations. A 10 ml LB inoculation containing the appropriate antibiotic was used for a miniprep isolation whilst a 50 ml inoculation was needed for a midiprep. These inoculations were incubated in the 37 °C shaker overnight in preparation of plasmid isolation. All standard protocols provided by Qiagen followed.

#### 2.1.10 DNA quantification and sequencing

Concentrations of DNA and RNA samples were measured by the detection of the absorbance at 260 nm ( $A_{260}$ ) using the Thermo Scientific NanoDrop spectrophotometer. The quality of samples was measured using the  $A_{260/280}$  ratio. A ratio of around 1.8 for DNA was deemed pure whilst a ratio of around 2 was used as the quality reference point for RNA. All cloned products were sequenced by Eurofins Genomics using appropriate sequencing primers.

#### 2.1.11 Preparation of DH5 $\alpha$ competent cell stocks

On an agar plate containing no antibiotics, 30  $\mu$ l of a DH5 $\alpha$  glycerol stock was added and incubated at 37 °C overnight. A colony was then picked from the agar plate and inoculated into 20 ml of plain LB with no antibiotic and placed in the 37 °C shaker at 275 rpm overnight. 5 ml of this culture was placed into 200 ml of fresh LB and incubated in 37 °C shaker until  $OD_{600}$  of 0.3-0.4 was reached. This was then split into four 50 ml falcon tubes and iced for 15 mins. After this, the tubes were spun at 4 °C for 5 mins at 700-1500x g relative centrifugal force (rcf). The supernatant was discarded and the pellet in each tube was then resuspended in 16 ml of ice cold TBI and incubated on ice for a further 15 mins. The tubes were then spun again at 700-1500 rcf at 4 °C for 5 mins. Finally, the supernatant was discarded and each pellet resuspended in 4 ml of ice cold TBII. These were aliquoted into 200  $\mu$ l tubes and stored in the -80°C freezer. Recipe for TBI and TBII are shown in table 2.1-9 and 2.1-10.

*Table 2.1-9- Recipe for TBI used in DH5 $\alpha$  preparation*

<b>Component</b>	<b>Amount</b>
RbCl <sub>2</sub>	3 g
MnCl <sub>2</sub> .4H <sub>2</sub> O	2.97 g
CaCl <sub>2</sub> .2H <sub>2</sub> O	0.38 g
KAc	10.73 g
H <sub>2</sub> O	Up to 250 ml

*Table 2.1-10- Recipe for TBII used in DH5 $\alpha$  preparation*

<b>Component</b>	<b>Amount</b>
RbCl <sub>2</sub>	0.3 g
CaCl <sub>2</sub> .2H <sub>2</sub> O	2.75 g
0.5M MOPS pH6.8	5 ml
Glycerol	37.5 ml
H <sub>2</sub> O	Up to 250 ml

#### 2.1.12 Site-directed mutagenesis

The QuikChange XL Site-Directed Mutagenesis kit from Agilent Technologies was used to delete the SOX9 binding site (-CCTTGAG-) within the predicted WWP2-C promoter sequence. All standard manufacturer's instructions were followed using the primers listed in table 2.1-11.

*Table 2.1-11- Primers used for QuikChange deletion of SOX9 binding site in predicted WWP2-C promoter sequence*

<b>Primer</b>	<b>Primer sequence</b>
QuikChange SOX9 1	cgccttgccgccgggagccggagctga
QuikChange SOX9 2	tcagctccggctcccggcgccaaggcg

## 2.2 Cell Culture

The wide range of mammalian cells used in the experiments were all incubated at the standard conditions of 37 °C and 5% CO<sub>2</sub>. However, they required slightly different media for optimum growth which are listed below in table 2.2-1.

*Table 2.2-1- A table showing the different mediums used for various cell types*

Cell line	Cell type		Media used
HEK293A	Human Kidney	Embryonic	Dulbecco's Modified Eagle's Medium (DMEM)- 10% Foetal calf serum (FCS), 1% Penicillin/Streptomycin (pen/strep), 1% None essential amino acid (NEAA), 1% Pyruvate
ARPE-19	Human Pigment	Retinal	DMEM F12 1:1- 10% FCS, 0.1% gentamicin, 1% L-Glu
MCF-7	Human Cancer	Breast	DMEM- 10% FCS, 1% pen/strep,
A2058, A375, HTB63, M202, M229, M285, M296, MEL501, SKMEL28	Human cell lines	melanoma	Roswell Park memorial Institute 1640 (RPMI1640)- 10% FCS, 1% pen/strep

### 2.2.1 Cell passage

Cell passage was normally performed at 80-90 % cell confluency. In addition to this, growth medium was replaced with 10ml fresh media preheated to 37 °C per T75 flask, every 2-3 days. To begin the passage, media was removed and 3 ml ice cold Dulbecco's Phosphate buffer saline (DPBS) (Gibco) was used to wash the cells in a T75 flask. Then, the cells were incubated at 37 °C with 5 ml of preheated TryPLE Express enzyme (ThermoFisher Scientific) until cells had dissociated from the flask. This was then centrifuged at 24000x g rcf for 5 mins at room temperature, and the supernatant was discarded. Cell pellets were then resuspended in 1 ml fresh media and a proportion of this was added to a T75 flask containing 10ml of preheated media depending on the ratio at which cells needed to be split.

### 2.2.2 Preparation of cell stocks to freeze

Cells were trypsinized and centrifuged as described in section 2.2.1. After the removal of supernatant, cell pellets were resuspended in 5 ml FCS with 10% DMSO. This was then aliquoted into five 1 ml tubes and stored in the -80 °C freezer.

### 2.2.3 TGF $\beta$ , BMP and EGF stimulation of cells

Cells to be stimulated by recombinant human TGF $\beta$ -1, recombinant human BMP-4 and EGF (all supplied from R and D systems) were starved in media containing 0.5% FCS 24 hrs before stimulation. In addition, cell medium was replaced with fresh preheated media containing 0.5% FCS an hour prior to stimulation. A standard stimulation concentration of 5 ng ml<sup>-1</sup> was used for TGF $\beta$ -1 and 10 ng ml<sup>-1</sup> for BMP-4 and EGF. Cells were then incubated at 37 °C and 5% CO<sub>2</sub> overnight unless otherwise stated.

## 2.3 Immunolabelling and fluorescence microscopy

### 2.3.1 PEI Transfection

Two days prior to transfection,  $1 \times 10^5$  ARPE-19 cells were seeded in a 3cm tissue culture dish containing three 1 cm coverslips and 2 ml of media. This was then incubated at 37 °C and 5% CO<sub>2</sub> until cells reach 60-70% confluence. Polyethylenimine transfection agent (PEI) obtained from Polysciences was used to transfect WWP2 construct into ARPE-19 cells. First, 7 µl of 1mg/ml PEI and 3 µg of DNA was added to 300 µl of DMEM 0% FCS. This was vortexed for 10 s and incubated at room temperature for 15 mins. The mixture was then added drop-wise to the cells, and placed back in the 37 °C incubator overnight. Note that 1hr prior to transfection, media in culture dishes was replaced with 2 ml of fresh preheated DMEM with 10% FCS.

### 2.3.2 Immunolabelling

Media was removed from the culture dish, and cells were fixed using 2 ml of methanol at -20 °C and placed in the freezer for 5 mins. The methanol was then discarded and cells washed with 1% goat serum in PBS three times (Please refer to table 2.3-1 for PBS recipe). The cells were then blocked with 10% goat serum in PBS for 30 mins at room temperature. Primary antibodies were then incubated with cells at room temperature for 60 mins by placing cells on the coverslip face down onto parafilm containing 50 µl of antibody in 1% goat serum PBS (note that the concentrations used are highlighted in table 2.3-2). Coverslips containing the cells were then placed in a culture dish and washed with 1% goat serum in PBS six times for 5 mins each. The cells were then incubated in darkness with secondary alexafluor antibody in 1% goat serum PBS on parafilm as described previously, for 30 mins (concentrations shown in table 2.3-3). The cells were then washed twice in 1% goat serum PBS for 10 mins then DAPI stained using NucBlue provided by Life Technologies following manufacturers protocol and finally mounted onto cover glass using Hydromount (National Diagnostics). After refrigeration at 4 °C overnight, samples were ready for examination under the Zeiss LSM510 META confocal microscope using the 63x Plan Apochromat

(N.A. 1.4) oil immersion objective. DAPI staining was excited at 364 nm using a UV laser and emission collected between 385 nm and 470 nm. The alexafluor 488 antibodies were excited at 488 nm using an Argon laser and emitted wavelengths of between 505 nm and 50 nm. The alexafluor 647 antibodies were excited at 633 nm using a He-Ne laser and wavelengths of between 650 nm and 750 nm were collected. Antibodies used are listed in table 2.3-2 and 2.3-3.

*Table 2.3-1- Standard 1x Phosphate Buffered saline (PBS) recipe*

<b>Component</b>	<b>Amount</b>
NaCl	8 g
KCL	0.2 g
Na <sub>2</sub> HPO <sub>4</sub>	1.44 g
KH <sub>2</sub> PO <sub>4</sub>	0.24 g
H <sub>2</sub> O	Up to 1 L

*Table 2.3-2- Primary antibodies used in immunolabelling*

<b>Primary antibody</b>	<b>Manufacturer</b>	<b>Product code</b>	<b>Concentration used</b>
Rat anti-HA	Sigma Aldrich	Clone3F10	1:1000
Rabbit anti- $\alpha$ Tubulin	Abcam	Ab15246	1:100

*Table 2.3-3- Secondary antibodies used in immunolabelling*

<b>Secondary antibody</b>	<b>Manufacturer</b>	<b>Product code</b>	<b>Concentration used</b>
Anti-Rat Alexafluor 488	ThermoFisher Scientific	A11006	1:1000
Anti-Rabbit Alexafluor 647	Invitrogen	21245	1:1000

## 2.4 Protein Analysis

### 2.4.1 PEI Transfection and co-transfections with p-CAGAC12-luc Luciferase reporter

HEK293A cells were transfected when cell confluence reached 50-60%. For transfection in 6-well plates, 3000 ng of DNA, 7  $\mu$ l PEI and 300  $\mu$ l 0% FCS DMEM were used according to the standard protocol described in section 2.3.1.

For co-transfections, cells were seeded onto 24-well plates and allowed to reach 50% confluence. Cells were then co-transfected with 600 ng of WWP2, SOX9 construct DNA (hSOX9/pcDNA-5'UT-Flag obtained from the Lefebvre lab at The Cleveland Clinic Foundation) or ESRP expression vector (courtesy of the Carsten lab at the University of Pennsylvania School of Medicine), 200 ng p-CAGAC12-luc reporter (courtesy of Caroline Hill from CRUK) and 50 ng pRSV- $\beta$ galactosidase ( $\beta$ gal) control vector (provided by Professor Simak Ali at the School of Medicine, Imperial College London) using 2  $\mu$ l PEI and 100  $\mu$ l 0% FCS DMEM per 1 ml well.

### 2.4.2 Cell Harvest of total cell lysates

To harvest cells from a 6-well plate, cells were first washed with 500  $\mu$ l of ice cold DPBS, then 300  $\mu$ l of NP40 cell lysis buffer (table 2.4-1) containing cOmplete mini protease inhibitor (Roche) and PhosStop phosphatase inhibitors (Roche) was added to each well. This was then incubated at 4  $^{\circ}$ C on a rotator plate for 30 mins before cell lysates were transferred into microcentrifuge tubes and spun at 11700x g rcf at 4  $^{\circ}$ C for 20 mins. The supernatant containing the cellular proteins was then added to equal amounts of 2x Lamlli buffer (table 2.4-2) with 20% DTT and boiled for 5 mins. Samples were then stored at -20  $^{\circ}$ C until analysis using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

*Table 2.4-1- Recipe for NP40 Cell lysis buffer*

<b>Component</b>	<b>Amount</b>
NaCl	0.88 g
Tris-HCl pH8	0.79 g
IGEPAL CA-630	1%
dH <sub>2</sub> O	Up to 100 ml

*Table 2.4-2- Recipe for 2x Lamml buffer*

<b>Component</b>	<b>Amount</b>
SDS 10%	4 ml
1M Tris pH6.8	1.2 ml
Bromophenol Blue 1%	200 µl
Glycerol	2 ml
dH <sub>2</sub> O	2.6 ml

### 2.4.3 Streptavidin pull-downs

Constructs cloned containing a Strep-tag (WSHPQFEK), were purified using the IBA Lifesciences Strep-Tactin Sepharose 50% suspension, allowing the identification of molecules bound to strep constructs when western blots were performed. First, 800 µl of cleared strep tagged construct transfected cell lysates were incubated with 40 µl of Strep beads at 4 °C with rotation overnight. The samples were then centrifuged at 4 °C and the supernatant discarded. The pellets were then washed three times with 400 µl of 0.1% NP40, then resuspended in 150 µl of 2xLamml with 20% DTT and boiled for 5 mins. Samples were stored at -20 °C until SDS PAGE was performed.

### 2.4.4 SDS Polyacrylamide gel electrophoresis

The percentage of SDS PAGE gel used is dependent upon the expected size of proteins to be detected. The higher the percentage, the smaller the protein that can be detected on the gel. The 10% and 15% SDS PAGE gels were the most commonly used during this research and are shown in table 2.4-3 and 2.4-4 respectively. The amount of protein sample loaded into each well depended on the level of protein within each sample which was determined by normalisation using β-actin antibodies. Gels were ran in SDS PAGE running buffer (Table 2.4-5) at 180 V for 90 mins or until dye front had reached the bottom of the gel.

*Table 2.4-3- 10% SDS PAGE gel recipe*

<b>Component</b>	<b>Volume in separating gel</b>	<b>Volume in stacking gel</b>
ddH <sub>2</sub> O	3.2 ml	2.6 ml
30% Acrylamide	2.67 ml	1 ml
1.5M Tris pH 8.8	2 ml	1.25 ml
10% SDS	80 µl	50 µl
10% APS	80 µl	50 µl
TEMED	8 µl	5 µl

*Table 2.4-4- 15% SDS PAGE gel recipe*

<b>Component</b>	<b>Volume in separating gel</b>	<b>Volume in stacking gel</b>
ddH <sub>2</sub> O	1.8 ml	2.6 ml
30% Acrylamide	4 ml	1 ml
1.5M Tris pH 8.8	2 ml	1.25 ml
10% SDS	80 µl	50 µl
10% APS	80 µl	50 µl
TEMED	8 µl	5 µl

*Table 2.4-5- Recipe for 10x SDS PAGE running buffer*

<b>Component</b>	<b>Amount</b>
Tris Base	242 g
Glacial acetic acid	57.1 ml
0.5M EDTA pH 8	100 ml
dH <sub>2</sub> O	Up to 1 L

#### 2.4.5 Protein preparation for mass spectrometry

Mass spectrometry was performed on protein samples that were pulled down using streptavidin beads as described in section 2.4.3. After incubation of the proteins with the beads overnight at 4°C, samples were spun at 11700x g rcf for 1 min and supernatants were discarded. The leftover beads were then washed twice with 400ul of 0.1% NP40 in dH<sub>2</sub>O, then three times with dH<sub>2</sub>O. For the quantitative iTRAQ approach, samples were handed over to the proteomics facility at the John Innes Centre to carry out trypsin digestion and quantitative mass spectrometry.

For the initial qualitative mass spectrometry analysis, proteins were eluted from beads by boiling in 50 µl of 2x lamml buffer with 20% DTT for 5 mins. These samples were then ran on a 10% separating only gel until samples were a few millimetres into the top of the gel. Gel cassettes were then rinsed

with water and samples were extracted using clean blades. The gel slices were then placed into low bind tubes and destained with 1 ml of 30% ethanol at 65 °C for 30 mins. After this, the ethanol was removed and samples were washed for 15 mins with 1 ml Triethylammonium bicarbonate (TEAB) buffer (Sigma) containing 50% acetonitrile (ACN). Samples were then incubated with 1 ml of 10 mM DTT at 55 °C for 30 mins before DTT was removed completely. A solution of 30 mM Iodoacetamide (IAA) in 50 mM TEAB buffer was then incubated with the samples at room temperature in the dark for 30 mins before another 15 mins wash with TEAB buffer containing 50% ACN. Following another 15 mins wash with 50 mM TEAB buffer, gel slices were cut into approximately 1x1 mm pieces and washed again with 50 mM TEAB buffer containing 50% ACN. After the final 15 mins wash step using 1 ml of 100% acetonitrile, samples were placed into a SpeedVac concentrator for 30 mins, then handed over to the Proteomics facility in the John Innes Centre for trypsin digestion and MS analysis. All solutions and buffers used in this process were provided by the John Innes proteomics facility.

#### 2.4.6 Western blotting and dot blotting

Standard wet transfers were performed to transfer proteins from SDS PAGE gels onto 0.45 µm Nitrocellulose membrane (GE Healthcare) and were ran at 100 V for 60 mins in transfer buffer (Table 2.4-6). Ponceau S staining solution (Table 2.4-7) was then used to reversibly stain the membrane to show the protein markers and to confirm a successful transfer. After washing the membranes three times with TBST (Table 2.4-8), they were then blocked with 5% milk powder (Marvel) in TBST at 4 °C for 1 hr. Membranes were then washed with TBST three times for 10 mins before the addition of primary antibody (Ab). Incubation of primary antibody occurs at 4 °C overnight on a rotator plate. After this, membranes were washed again with TBST at room temperature three times for 10 mins before secondary Ab incubation, which takes place on a rotator plate at room temperature for 1 hr. Table 2.4-9 and 2.4-10 show the antibodies and the dilutions used for western blotting. After another three 5 mins TBST washes, membranes were

placed onto cling film and Amersham ECL prime western blot reagent mix was applied according to manufacturer's guide. The blots were then ready to be developed onto Amersham Hyperfilm ECL in the dark room.

For proteins that were too small to be detected using PAGE, dot blot was performed. To do this, 30 µl of cell lysates were loaded onto a gridded 0.45 µm Nitrocellulose membrane. This was then blocked, incubated with antibodies and developed in the same way as described above.

*Table 2.4-6- Recipe for 1x Transfer buffer*

<b>Component</b>	<b>Amount</b>
Tris Base	3 g
Glycine	15 g
Methanol	200 ml
dH <sub>2</sub> O	Up to 1 L

*Table 2.4-7- Recipe for Ponceau S staining solution (0.1%w/v Ponceau S in 5%v/v acetic acid)*

<b>Component</b>	<b>Amount</b>
Ponceau S (Sigma-Aldrich)	1 g
Acetic acid	50 ml
dH <sub>2</sub> O	Up to 1 L

*Table 2.4-8- Recipe for 1x TBST pH7.6*

<b>Component</b>	<b>Amount</b>
Tris HCL	2.423 g
NaCl	8 g
Tween20	1 ml
dH <sub>2</sub> O	Up to 1 L

*Table 2.4-9- Primary antibodies used in western blotting*

<b>Primary antibody</b>	<b>Manufacturer</b>	<b>Product code</b>	<b>Concentration used</b>
Rat anti-HA	Sigma Aldrich	Clone3F10	1:1000
Rabbit anti-WWP2C	Abcam	Ab60130	1 µg/ml
Rabbit anti-WWP2N	Abcam	Ab103527	1:10000
Rabbit anti-Smad2	Cell signalling technology	86F7	1:1000
Rabbit anti-Smad3	Cell signalling technology	C67H9	1:1000
Mouse anti-Smad7	R&D Systems	MAB2029	1 µg/ml
Mouse anti-Flag	Sigma Aldrich	F3165	1:10000
Rabbit anti-βactin	Abcam	Ab75186	1 µg/ml

*Table 2.4-10- Secondary antibodies used in western blotting*

<b>HRP-conjugated Secondary antibody</b>	<b>Manufacturer</b>	<b>Product code</b>	<b>Concentration used</b>
Anti-Rat	Sigma Aldrich	A5795	1:100000
Anti-Rabbit	Sigma Aldrich	A0545	1:100000
Anti-mouse	Sigma Aldrich	A4416	1:4000

#### 2.4.7 Luciferase assays, protein harvest and statistical analysis

To harvest from a 24 well plate, 5x Cell culture lysis buffer provided by the Promega Luciferase Assay system was made into a 1x solution, and 100  $\mu$ l of this was added into each well. The plates were then placed onto a plate rotator and agitated at 4 °C for 30 mins. The wells containing the samples were then scraped to remove any remaining cells which were then transferred into microcentrifuge tubes to be spun at 4 °C for 20 mins at 11700x g rcf. To a 96 well reading plate, 20  $\mu$ l of each sample were added to its corresponding well followed by 100  $\mu$ l of Luciferase or  $\beta$ -galactosidase reporter substrates (Promega) which were added according to manufacturer's instructions. Results were read using the FlexStation fluorometric plate reader and normalised with  $\beta$ -glo readings. Fold changes were calculated using empty vector negative controls.

The standard error was used to work out error bars shown and the two-tailed, paired T-test was used to work out the significance of results with \*P= <0.05, \*\*P=<0.005 and \*\*\*P=<0.001. Normal distribution of results was assumed after inputting data into goodness-of-fit tests including the Kolmogorov-Smirnov test and the Shapiro-Wilk test.

## 2.5 RNA Analysis

### 2.5.1 RNA Isolation

At 80-90% cell confluence, total RNA isolation was performed using the Promega Total RNA Isolation System. Manufacture's protocol was followed to obtain RNA samples which were then measured using the NanoDrop Spectrophotometer (as described in section 2.1.10). These RNA samples were used to synthesis cDNA on the same day and the remaining RNA were stored in the -80 °C freezer.

### 2.5.2 cDNA Synthesis

The Promega GoScript Reverse Transcription System was used to synthesis cDNA from the RNA extracted previously. The reaction consists of two main step. First, mixture one (table 2.5-1) was heated at 70 °C for 5 mins then placed on ice until the addition of mixture two (table 2.5-2). The samples were then incubated in the PCR machine using the settings shown in table 2.5-3. Note that the volume of mixture one is larger than proposed in the manufacture's protocol. This was to maximize the amount of RNA in the reaction when RNA samples were diluted.

*Table 2.5-1-Mixture one required in a 40µl GoScript cDNA Synthesis reaction*

<b>Component</b>	<b>Amount</b>
RNA	Up to 5 µg
Random primers	2.5 µl
dH <sub>2</sub> O	Up to 25.1 µl

*Table 2.5-2- Mixture two required in a 40µl GoScript cDNA Synthesis reaction*

<b>Component</b>	<b>Amount</b>
GoScript 5x Reaction Buffer	8 µl
MgCl <sub>2</sub>	2.4 µl
10mM dNTP	2 µl
Recombinant RNasin Ribonuclease Inhibitor	0.5 µl
GoScript Reverse Transcriptase	2 µl

*Table 2.5-3 -PCR settings used for GoScript cDNA Synthesis*

	<b>Temperature</b>	<b>Time</b>
<b>Anneal</b>	25 °C	5 mins
<b>Extend</b>	42 °C	60 mins
<b>Inactivate Reverse Transcriptase</b>	70 °C	15 mins

### 2.5.3 Semi-quantitative PCR

Semi-quantitative RT-PCR was performed using the Promega GoTaq Flexi DNA Polymerase. All standard protocols were followed using the components listed in table 2.5-4 and PCR settings in table 2.5-5 and 2.5-6. GAPDH was used as the housekeeping gene of choice to normalise amount of cDNA added into each reaction. A list of RT-PCR primers is shown in table 2.5-7.

*Table 2.5-4- Components in a 20µl GoTaq RT-PCR reaction*

<b>Component</b>	<b>Amount</b>
5x Green GoTaq Flexi buffer	4 µl
10 mM dNTP	0.4 µl
MgCl <sub>2</sub>	1.2 µl
10 µM FW Primer	1 µl
10 µM RV Primer	1 µl
Template cDNA	dependent on optimization
GoTaq DNA Polymerase	0.1 µl
dH <sub>2</sub> O	Up to 20 µl

*Table 2.5-5- PCR settings used for GoTaq RT-PCR*

<b>Step</b>	<b>Temperature</b>	<b>Time</b>	<b>Cycles</b>
Initial Denaturation	95 °C	2 mins	1
Denaturation	95 °C	30 s	35
Annealing	*	30 s	
Extension	72 °C	45 s	
Final Extension	72 °C	5 mins	1
Hold	4 °C	indefinite	1

**Table 2.5-6- Annealing temperatures used in RT-PCR**

<b>Target Isoform</b>	<b>Annealing temperature used</b>
WWP2-FL	49 °C
WWP2-C	49 °C
WWP2-C-ΔHECT	49 °C
N-Cadherin	49 °C
SLUG	49 °C
SNAIL	49 °C
Twist	49 °C
WWP2-N	53 °C
GAPDH	53 °C
SOX9	56 °C
Vimentin	56 °C

**Table 2.5-7- RT-PCR primer sequence**

<b>Target Isoform</b>	<b>Forward primer 5'-3'</b>	<b>Reverse primer 5'-3'</b>
WWP2-FL	ACCTTGAGAAATGAACTGCT	TCCACATAGTAAAACCTGCC
WWP2-N	ATCTGTCAACCTCTCCAACG	ACTCATGGTGTGAGTTCCAG
WWP2-C	CTGCTGTGTTTCTTGTGTTTA	ATGGCGATGAATCTGCCTATAA
WWP2-ΔHECT	GCCGAAAGAACAATTACTG	TTCTCTGCAACCCACTAGAT
WWP2-N-ΔC2	AGGAATGATCTGCTTGGATG	GAAAAGCAATCAAGTGACCC
GAPDH	ACCACAGTCCATGCCATCAC	TCCACCACCTGTTGCTGTA
N-Cadherin	GATGTTGAGGTACAGAATCGT	GGTCGGTCTGGATGGCGA
SLUG	GGTCAAGAAGCATTTC AAC	GGTAATGTGTGGGTCCGA
SNAIL	CAGACCCACTCAGATGTCAA	CATAGTTAGTCACACCTCGT
Twist	GGGAGTCCGCAGTCTTAC	CCTGTCTCGCTTTCTCTTT
SOX9	GGGAAGGCCGCCAGGGCGA	TGCCTTGCCCGACTGCAGTTCT
Vimentin	GAACGCCAGATGCGTGAAATG	CCAGAGGGAGTGAATCCAGATTA

#### 2.5.4 Agarose gel electrophoresis

A 1x agarose gel was made using 0.6 g of agarose in a total volume of 60 ml 1x TAE buffer with 6 µl of SYBR Safe gel dye for a small electrophoresis tank or 2 g in 200 ml of 1x TAE buffer with 20 µl of SYBR Safe gel dye in a larger tank. The recipe for TAE buffer is shown in table 2.5-8. Once the gel had set and was placed in the electrophoresis tank filled with TAE buffer, samples were loaded into wells and the gel was then ran at 100 V for 60 mins. To mark DNA sizes, 10 µl of the 1 Kb Plus DNA ladder from ThermoFisher Scientific was used. DNA was visualised under ultra-violet (UV) light and the density of bands was measured using Image J.

*Table 2.5-8- Recipe for 20x TAE Buffer*

<b>Component</b>	<b>Amount</b>
Tris base	242 g
Glacial acetic acid	57.1 ml
0.5 M EDTA pH 8	100 ml
dH <sub>2</sub> O	Up to 1 L

### 2.5.5 Taqman Quantitative PCR

The geNorm Reference gene selection kit provided by Primer Designs was used to determine the best reference genes to use in future experiments. After testing against all 12 of the available reference genes, it was found that 18S and UBC were the most suitable for the range of cDNA samples that were to be analysed. Standard Taqman protocols provided by Primer Designs were followed using components shown in table 2.5-9 and cycle information in table 2.5-11. All Taqman primers used are shown in table 2.5-10. Taqman qPCR reactions were carried out in 96 well plates provided by Applied Biosystems, in the ABI PRISM 7500 PCR system. In addition to performing qPCR analysis on cDNA samples obtained from self-cultured cells, the Origene TissueScan Melanoma cDNA array II was also purchased and used to probe for different WWP2 isoforms. Negative controls without cDNA were used in all qPCR reactions to ensure there were no false positive results.

To work out the fold changes ( $2^{-\Delta\Delta CT}$ ) used in the comparison of qPCR results, the following calculations were used:

$$CT\ value(Gene\ Of\ Interest) - CT(Housekeeping\ gene) = \Delta CT$$

$$\Delta CT(Sample\ of\ interest) - \Delta CT(baseline\ sample) = \Delta\Delta CT$$

$$Fold\ change = 2^{-\Delta\Delta CT}$$

Note that CT values of over 35 were disregarded and the statistical significance of results were worked out using methods described in section 2.4.7.

*Table 2.5-9- Components used in a 20 µl Taqman reaction*

<b>Component</b>	<b>Amount</b>
Gene probe mix	1 µl
Primer Designs Precision Plus 2x Mastermix	10 µl
cDNA (4 ng/µl)	1 µl
dH <sub>2</sub> O	8 µl

*Table 2.5-10- Taqman qPCR primer sequences provided by Primerdesign*

<b>Target gene of interest</b>	<b>Sense</b>	<b>Anti-sense</b>
WWP2-FL	GCAGTCGCAGCGGAATCA	AGGGGATCATGGTCAGTCGA
WWP2-C	CTGCCTTTTCCAAAACACTT	AGGGAATGCAGTGGGACTT
WWP2-N	GACCTGAGACTCTGGAAGTGA	CGTCACTGTGCTATGCTACC
WWP2-ΔHECT intron 9/10	GTGCTAGGCATCCTGAGACTT	CTGGAAATAGGAACCTCAGAATGAA
WWP2-ΔHECT intron 19/20	CCATTGTCTGGATCAAGTGAGTT	AGTACCTGCTTCCCACCCA
WWP2-N-ΔC2	TATATGTGGGTGTCTGCCTGTT	ATACCTGATTCATTACATCCAAGCA

*Table 2.5-11- Cycle information for Taqman qPCR*

<b>Step</b>	<b>Temperature</b>	<b>Time</b>	<b>Cycles</b>
Activation	95 °C	2 mins	1
Denaturation	95 °C	15 s	40
Extension	60 °C	1 mins	

Chapter 3 Exploring the biological roles of WWP2  
WW domains

### 3.1 Introduction

WWP2 is an E3 ubiquitin ligase containing a C2 lipid binding domain, four WW recognition domains and a HECT catalytic domain. The central function of WWP2 WW domains is to create binding specificity to target substrates. These domains recognize a plethora of PPXY motifs which are present in naturally occurring proteins including many signalling mediators. The high specificity in target binding allows the ubiquitination of proteins and subsequently controls the fate of the protein. WWP2 has most commonly been linked to facilitating the proteasomal degradation of target proteins, therefore uncovering WWP2 WW domain binding partners will reveal protein substrates that are involved in Ub-dependent proteasomal degradation. One key pathway that WWP2 has been linked to is the TGF $\beta$  signalling pathway, and because this signalling pathway is known to have a critical role in oncogenesis, we surmise that WWP2 could be a key player in oncogenesis through its effects on TGF $\beta$  signal transduction. The focus in this chapter is therefore the effects of WWP2 WW recognition domains on TGF $\beta$ -dependent gene expression.

First, we focus on investigating the effects of WW recognition domains on Smad-3 dependent TGF $\beta$  gene expression using recombinant WW constructs that were cloned into mammalian expression plasmids. These constructs were then co-transfected into mammalian cells along with the p-CAGAC12-luc Smad3 gene reporter vector. Lysates harvested from these cells were then used in Luciferase reporter assays. To further explore the relationship between WWP2 WW domains and TGF $\beta$ , western blots were also performed on WW tandem domain transfected cells using anti-Smad antibodies to detect levels of Smad expression in response to increases in WW tandem domain expression. Earlier studies from this lab suggested that WW1 recognition domain binds preferentially to Smad2/3 whilst WW4 domain binds to Smad7 (Soond and Chantry, 2011). With the new information that is obtained from this study, we will be able to build on existing knowledge and further clarify the interaction between WW domains

and Smad mediators, consequently unveiling the effects of WW domains on TGF $\beta$  activity.

Originally, the overall aim of WWP2 WW domain interaction studies was to create peptide inhibitors containing the minimal WW domain sequence required to competitively bind to Smads and lower functional WWP2-Smad binding and subsequent degradation of Smad mediators. It was hypothesized that a WW domain peptide with the ability to bind to R-Smad2/3 might stabilize its levels and could have been used to aid the tumour suppressive role of TGF $\beta$  in early stages of cancer to prevent cancer spread. On the other hand, a WW domain peptide that stabilizes inhibitor Smad7 levels could have negatively regulated the oncogenic TGF $\beta$  pathway in advanced stages of oncogenesis and therefore work as a tumour suppressant. However, emerging NMR studies at the time suggested that due to the complex structural and binding properties of WWP2 WW domains, the development of peptides would not be feasible. It is for this reason that we progressed onto other aspects of our studies.

Indeed, the effects of WWP2 on Smad-dependent TGF $\beta$  gene expression is a crucial element of our studies due to the relationship between WWP2 and the TGF $\beta$  pathway in oncogenesis, however, it is also useful to explore levels of other proteins present in cell lysates transfected with the WWP2 WW domains and the possible effects of TGF $\beta$  on these interactions. The results from these experiments may highlight specific oncogenes or tumour-suppressors previously unrelated to WWP2, that may be key to the role of WWP2 in TGF $\beta$ -dependent oncogenesis. For this reason, quantitative mass spectrometry was used to identify proteins associated with transfected WWP2 WW tandem domains. To do this, a streptavidin tag was first cloned onto the WW tandem construct, then used in Strep-Tactin purification. Cells were then harvested using the standard mass spectrometry protein preparation protocol (please refer to section 2.4.5), then digested with trypsin and processed using quantitative iTRAQ mass spectrometry.

## 3.2 Results

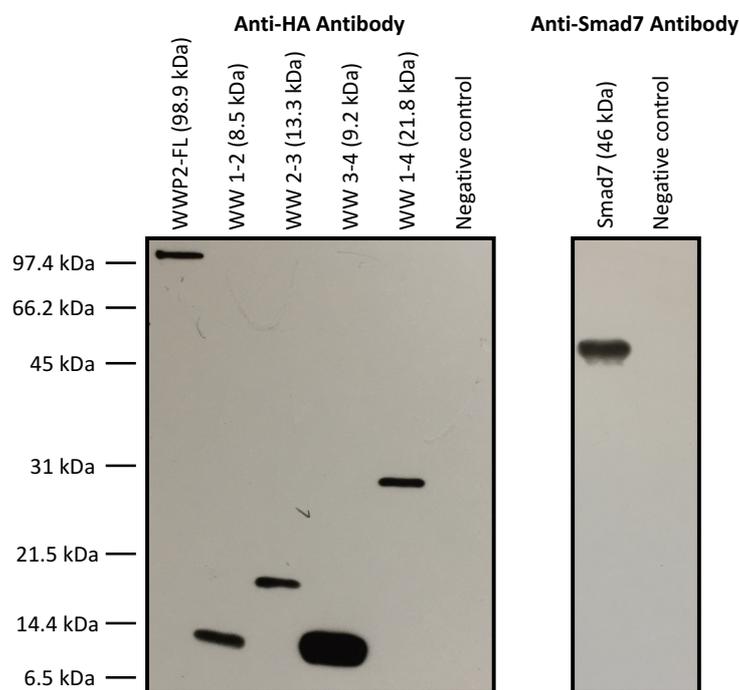
### 3.2.1 Studying the effects of WWP2 WW domains on Smad3-dependent gene expression using the p-CAGAC12-luc reporter vector in luciferase assays

Four different WW recognition domains, WW1, 2, 3 and 4 are present in the full length WWP2 with truncated isoforms containing different combinations of these domains. To test the Smad3-dependent TGF $\beta$  gene expression of different WW domains, tandem constructs and individual WW constructs were generated by cloning the corresponding domains into pRK5-HA vectors. These constructs were then co-transfected into mammalian cancerous and non-cancerous cells at 600 ng/ml along with 200 ng of p-CAGAC12-luc Smad3 reporter vector and 50 ng of pRSV- $\beta$ -galactosidase construct for normalization. Cell lysates were harvested 3 days post transfection following standard harvest protocols provided in section 2.4.2 and finally, the luciferase assays were performed according to standard protocol described in section 2.4.7. Fold changes were calculated using results from the empty pRK5-HA vector. The reading corresponds to the levels of Luciferin which is the result of Smad3 binding to the CAGAC12 reporter sequence therefore inferring that the greater the reading, the higher the levels of Smad3-dependent TGF $\beta$  gene expression. Smad7, a natural inhibitor of the TGF $\beta$  pathway was used to demonstrate the negative effects it has on the Smad3-dependent gene expression levels which results in a decrease in luciferase activity.

#### 3.2.1.1 Luciferase assays investigating the effects of WWP2 WW tandem domains on Smad3-dependent gene expression

To begin the investigation, four WW tandem constructs were cloned into the pRK5-HA plasmid including WW1-2, WW2-3, WW3-4, and WW1-4. The expression of these peptides along with a Smad7 construct, an inhibitor of Smad3-dependent gene expression, were verified by performing western blots using anti-HA and anti-Smad7 antibodies on HEK293A cell lysates transfected with the WW tandem domains and the Smad7 constructs. These

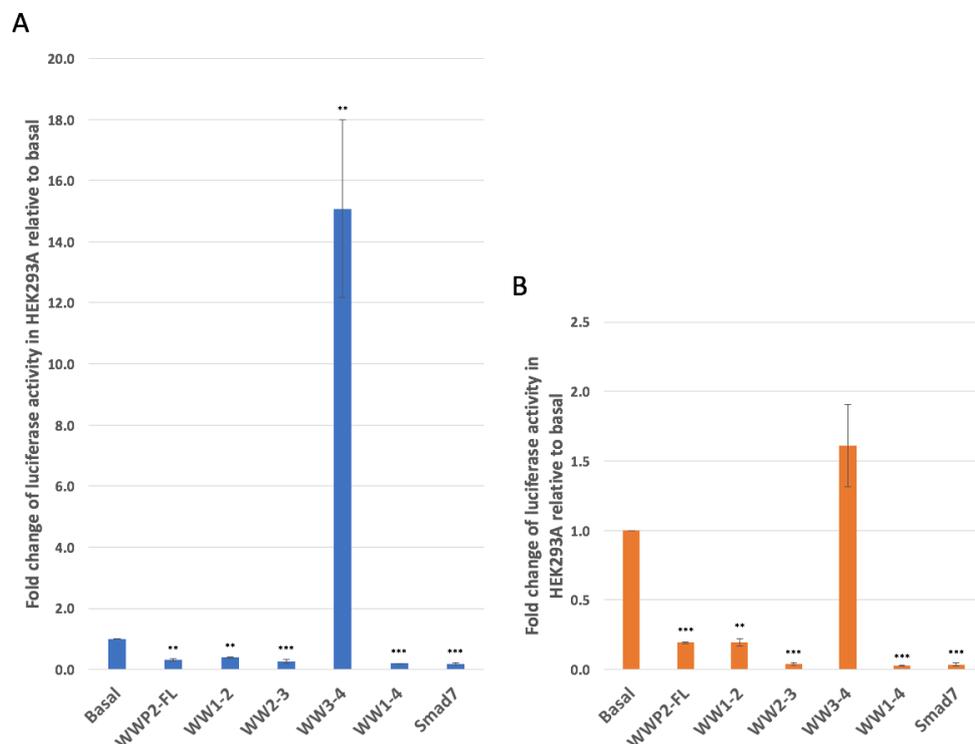
are shown in fig 3.2.1. Bands demonstrated the expression of proteins at the correct size. The constructs were then co-transfected into HEK293A and A375 cells with p-CAGAC12-luc Smad reporter vector and pRSV- $\beta$ -galactosidase construct. The effects of TGF $\beta$  was also tested in the same experiment with an 18 hr TGF $\beta$  stimulation to investigate its effects on transfected tandem constructs and Smad3-dependent TGF $\beta$  gene expression. Cells were harvested on day 3 post transfection using the standard manufacturers protocol, and Luciferase assay was performed using the Promega Luciferase assay kit. The Promega  $\beta$ -Glo assay system was used to normalize transfection efficiency.



**Figure 3.2.1- Protein expression of WWP2 WW tandem domains and Smad7 in HEK293A cells.** Western blot analysis using primary anti-HA and anti-Smad7 antibodies to detect WWP2-FL, WW1-2, 2-3, 3-4, 1-4 tandem domain HA-tagged constructs and Smad7 transfected into HEK293A. The negative control used for the anti-HA blot was a Flag-tagged construct whilst lysates transfected with WWP2 was used as the negative control in the anti-Smad7 blot. (Experimental replicate n=1)

Results of tandem constructs transfected into HEK293A is shown in fig 3.2.2 comparing TGF $\beta$  stimulation (b) and the non-stimulated control (a). Smad7 transfected constructs was used as a control to demonstrate its effect on Smad3-dependent gene expression. In non-stimulated cells, a fold change of less than 1 was seen in all constructs except for WW3-4 suggesting that there was a decrease in Smad3-dependent TGF $\beta$  gene expression (FL p=  $\leq$ 0.01,

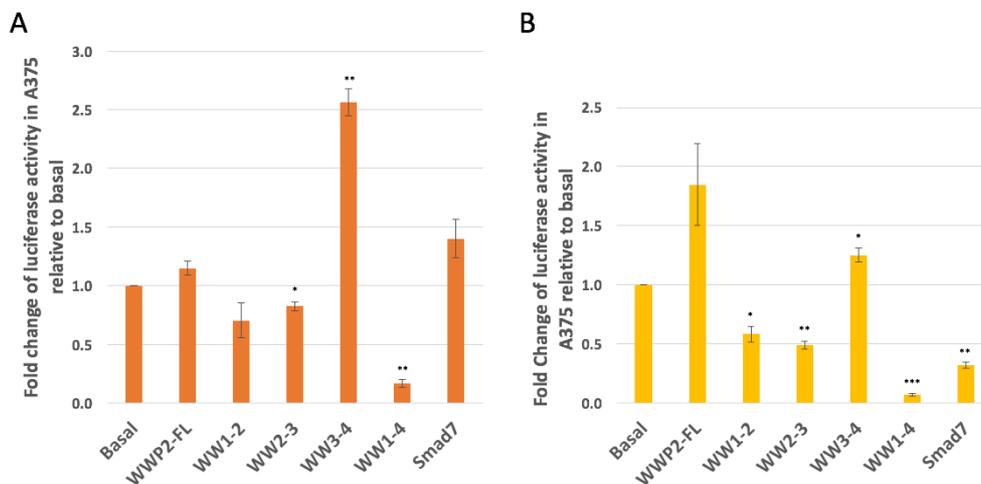
WW1-2  $p = \leq 0.01$ , WW2-3  $p = \leq 0.001$ , WW1-4  $p = \leq 0.001$ , Smad7  $p = \leq 0.001$ ). The 15-fold increase induced by WW3-4 transfection ( $p = \leq 0.01$ ) suggests that this tandem domain causes a massive increase in levels of Smad3 present in the cells. However, in TGF $\beta$  stimulated cells, the increase in Smad3-dependent TGF $\beta$  gene expression caused by overexpression of WW3-4 was found to be insignificant. Furthermore, it is also notable that TGF $\beta$  seem to decrease Smad3-dependent promoter activity across all six samples (FL  $p = \leq 0.001$ , WW1-2  $p = \leq 0.01$ , WW2-3  $p = \leq 0.001$ , WW1-4  $p = \leq 0.001$ , Smad7  $p = \leq 0.001$ ).



**Figure 3.2.2- The effects of TGF $\beta$  on Smad3-dependent gene expression of WWP2 WW tandem domain transfected HEK293A.** The WWP2-FL construct, Smad7 and other WWP2 WW tandem domains were transfected into HEK293A cells along with the p-CAGAC12-luc reporter vector. A) shows results from cells that were unstimulated whilst B) shows results from cells stimulated with TGF $\beta$  overnight. The fold change in luciferase activity was calculated relative to the basal activity of the empty vector. Results highlighted the increase in Smad3-dependent gene expression in WW3-4 transfected cells independent of TGF $\beta$  stimulation (Experimental replicate  $n=9$ . Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

As well as conducting this experiment with non-cancerous HEK293A cells, A375 melanoma cells were also used to replicate the above experiment to explore the different effects WWP2 WW binding domains have on different cell types. These results are shown in fig 3.2.3 which also shows the effect

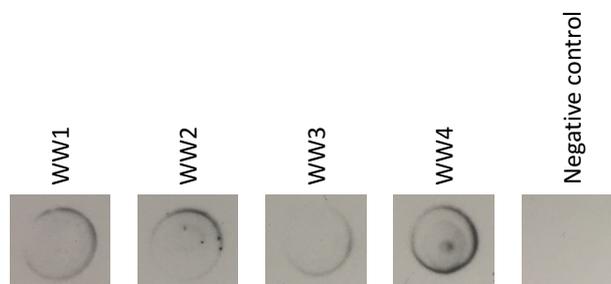
of TGF $\beta$  stimulation on Smad3-dependent TGF $\beta$  activity. Interestingly, the increase in fold change from WW3-4 is mimicked in this cell line in both TGF $\beta$  positive ( $p= \leq 0.05$ ) and negative ( $p= \leq 0.01$ ) environments however, the increase is only at 2.5-fold in comparison to the 15-fold shown in HEK293A control cells. A major difference shown between HEK293A and A375 cells is that WWP2-FL seem to induce elevated levels of Smad3 activity in A375 rather than decrease levels as shown in fig 3.2.2. Although WWP2-FL seem to cause the highest increase in Smad3-dependent gene expression in TGF $\beta$  stimulated cells, the high standard error means it is statistically insignificant, and therefore WW3-4 domain is responsible for the highest increase in gene expression in both stimulated and unstimulated cells. Results also show that WW1-4 causes consistently low levels of Smad3-dependent gene expression in both cell lines (A375 TGF $\beta$  -ve  $p= \leq 0.01$ , TGF $\beta$  +ve  $p= \leq 0.001$ ) while effects of Smad7 in A375 control cells seem to increase Smad3 activity as depicted in fig 3.2.3a. Other results shown proved statistically insignificant after used of statistical analysis described in section 2.4.7.



**Figure 3.2.3- The effects of TGF $\beta$  on Smad3-dependent gene expression of WWP2 WW tandem domain transfected A375 cells.** Luciferase assay of WWP2 constructs including WW tandem domains in A375 cells with A) showing data from unstimulated cells and B) showing data from cells stimulated with TGF $\beta$  overnight. Results are calculated as fold change relative to basal activity of empty pRK5-HA vector and highlighted an increase in Smad3-dependent activity, once again as a result of WW3-4 overexpression in unstimulated cells (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

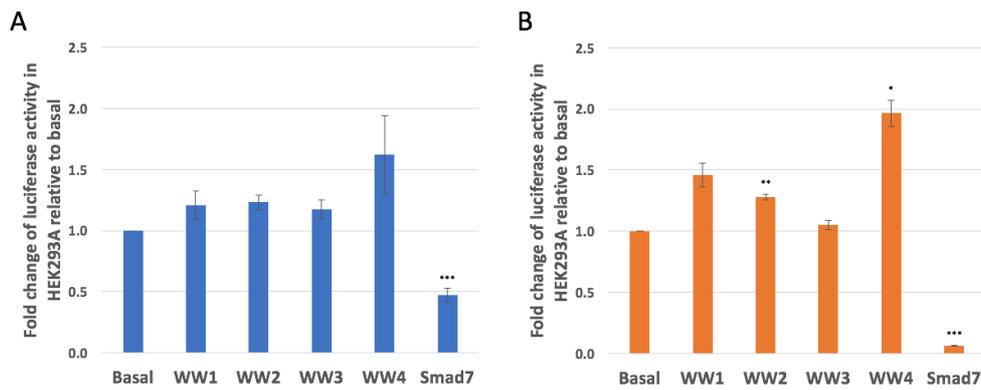
3.2.1.2 The effects of individual recombinant WWP2 WW domains on Smad3-dependent TGF $\beta$  gene expression

WWP2 WW1, WW2, WW3 and WW4 recognition domains were cloned into pRK5-HA to produce individual WW constructs for use in luciferase assays. These domains individually are 4 kDa, 4.1 kDa, 3.9 kDa and 3.9 kDa respectively, therefore were too small to be detected using standard western blotting protocols. To confirm its protein expression, dot blots using primary anti-HA antibodies were consequently performed (Shown in fig 3.2.4). These constructs were co-transfected into HEK293A and A375 cells with p-CAGAC12-luc vector and pRSV- $\beta$ -galactosidase. TGF $\beta$  stimulation was also performed to investigate their effects on cellular Smad3 expression. Fig 3.2.5 shows that in HEK293A, although all transfected individual WW domains seem to increase levels of Smad3 activity, statistical analysis showed that the results were insignificant. This was similar in cells stimulated with TGF $\beta$  with the exception of WW2 and WW4 overexpression which showed significant increase in Smad3-dependent TGF $\beta$  gene expression.

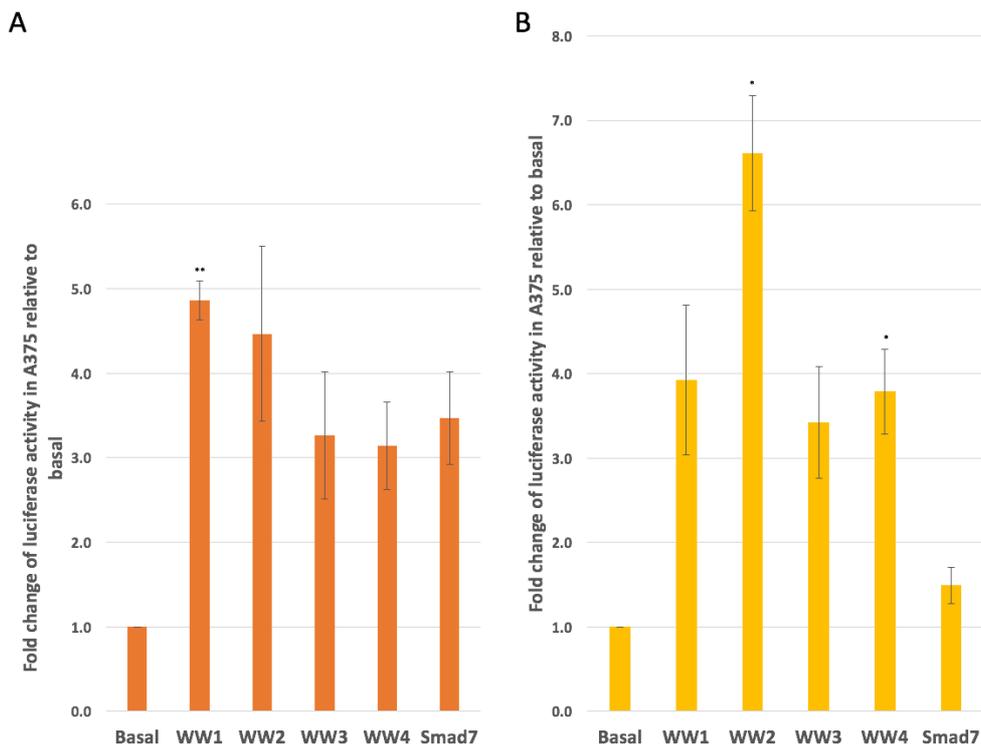


**Figure 3.2.4- Protein expression of WWP2 WW individual domains in HEK293A.** A Dot blot using primary anti-HA antibodies to identify the expression of WW1 (4 kDa), WW2 (4.1 kDa), WW3 (3.9 kDa), and WW4 (3.9 kDa). A Flag tagged construct was used as negative control.

In a repeat experiment using A375 cells shown in fig 3.2.6, transfected individual WW domains all appear to have no significant effect on the level of Smad3 gene expression except for the overexpression of WW1 ( $p \leq 0.01$ ), which caused a near 5-fold increase in non-stimulated cells and WW2 ( $p \leq 0.05$ ) which caused the largest fold change in TGF $\beta$  stimulated A375 cells.



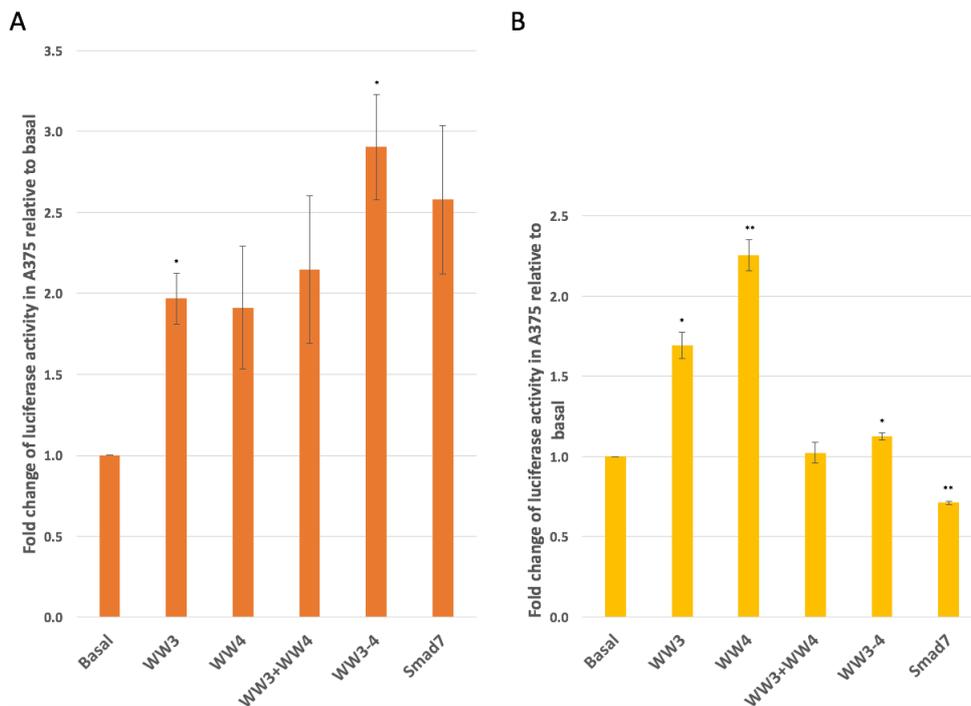
**Figure 3.2.5- The effects of TGF $\beta$  on Smad3-dependent gene expression of WWP2 WW individual domain transfected HEK293A.** The individual WW domains were cloned into pRK5-HA vector, and transfected into HEK293A epithelial cells along with the p-CAGAC12-luc binding vector. Data from luciferase assays normalized to  $\beta$ -galactosidase are shown as fold changes in A) unstimulated cells and B) cells stimulated overnight with TGF $\beta$ . Results from unstimulated cells showed no significant change except for the decrease caused by Smad7 whilst WW2 and WW4 overexpression both significantly increased Smad3-dependent gene expression in TGF $\beta$  stimulated cells. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)



**Figure 3.2.6- The effects of TGF $\beta$  on Smad3-dependent gene expression of WWP2 WW individual domain transfected A375.** Individual WW domain luciferase assay was repeated in A375 cells and results were calculated as fold change relative to basal activity of the empty pRK5-HA vector. A) shows results from non-stimulated A375 cells whilst B) is of cells that were stimulated with TGF $\beta$  overnight. WW1 overexpression increased Smad3-dependent expression in unstimulated conditions whilst WW2 and WW4 increased Smad2-dependent expression under TGF $\beta$  stimulation. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

### 3.2.1.3 Investigating the positive effect WWP2 WW3-4 has on levels of Smad3-dependent gene expression

Results from luciferase assays using WWP2 tandem WW domains in section 3.2.1 showed that overexpression of WW3-4 causes a substantial increase in Smad3 activity reflected by the 15-fold increase in HEK293A cells. To investigate the role of WW3 and WW4 in the dramatic increase in luciferase reading, the following combinations of WW3 and WW4 constructs were transfected into A375 cells with results compared using luciferase assay; WW3 only, WW4 only, WW3 and WW4 co-transfection, and the original WW3-4. Luciferase results were normalized to  $\beta$ -galactosidase levels and fold changes were calculated against blank pRK5-HA plasmid readings. Fig 3.2.7 suggests that in the absence of TGF $\beta$ , WW3-4 produced the highest luciferase reading ( $p \leq 0.05$ ), whilst in the presence of TGF $\beta$ , WW4 alone caused the highest levels of Smad3 activity ( $p \leq 0.01$ ). In non-stimulated cells, WW4, WW3+WW4 and Smad7 showed no significant difference whilst in TGF $\beta$  stimulated cells, although WW4 produced the most significant increase, both WW3 and WW3-4 also produced a significant increase in Smad3-dependent gene expression.

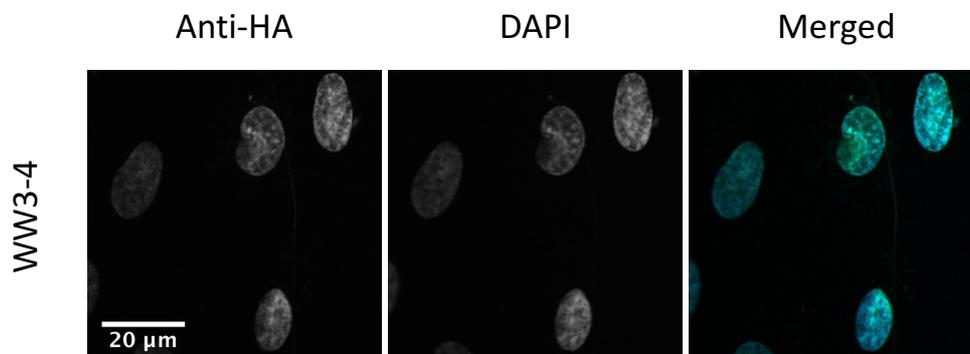


**Figure 3.2.7- The effects of TGF $\beta$  stimulation on WW3 and WW4 domain transfected A375 cells.** To investigate the substantial increase in luciferase activity produced by the transfected WW3-4 domain, a combination of WW3, WW4, WW3+4, WW3-4 tandem domains were transfected into A375 cells stimulated and unstimulated by TGF $\beta$ , B) and A), respectively. Results were normalized using  $\beta$ -galactosidase and fold change was calculated to basal activity of empty pRK5-HA vector. In unstimulated cells, WW3-4 overexpression remains the construct to increase Smad3-dependent expression the most. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

#### 3.2.1.4 Cellular localization of WWP2 WW3-4 tandem domain in ARPE-19

The WWP2 protein contains the C2 lipid binding domain, four WW recognition domains and a catalytic HECT domain. The WW3-4 tandem domain is a 73 AA region of WWP2. To understand the localization of this construct would help clarify the mechanism of function and could explain results shown in the Smad3-dependent TGF $\beta$  luciferase activity discussed previously. Due to the lack of other functional domains in the WW3-4 recombinant construct, namely the C2 and HECT domain, this study may also give an insight into the function and binding properties of the absent domains. For example, the WW3-4 construct may show little localization to the cell membrane which would highlight the lipid binding function of the C2 domain.

ARPE-19 cells transfected with HA-tagged WW3-4 were labelled with an anti-HA antibody and a secondary Alexa fluor 488 antibody. These cells are commonly used in fluorescence microscopy as they grow in a flat monolayer making it easier for analysis (Dunn *et al*, 1996). Interestingly as shown in fig 3.2.8, the HA-tagged peptide, is only present in the nucleus, and is absent from the membrane and cytoplasm. The potent nuclear translocation event demonstrated suggests the presence of a nuclear localization signal (NLS) within the WW3-4 peptide, but may also be due to the lack of potential nuclear export signals (NES) that may be present in regions absent from the recombinant WW3-4 construct such as the C2 and the HECT domains. Furthermore, the lack of WW3-4 membrane localization observed is due to the absence of C2 lipid binding domain suggesting that the primary function of this domain is the binding of lipid molecules in the cellular membrane and therefore without this domain, the WWP2 WW3-4 construct is unable to localize at the membrane.

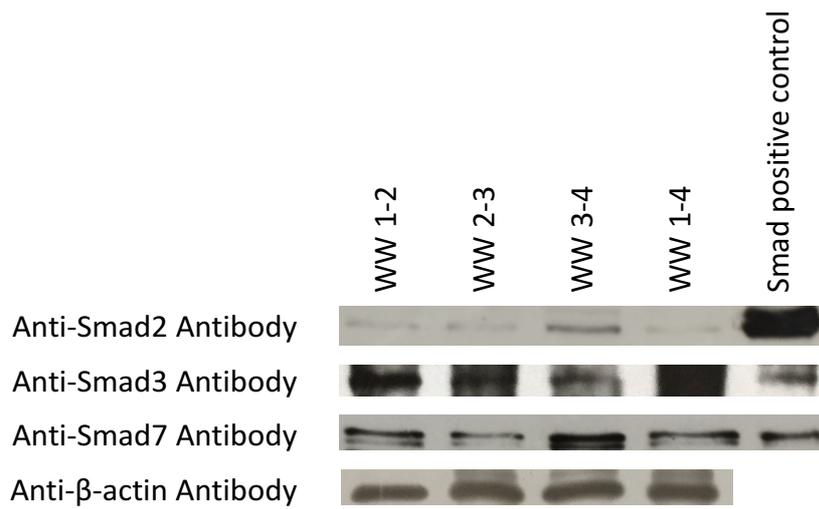


**Figure 3.2.8- Cellular localization of WWP2 WW3-4 tandem domain in ARPE-19.** Unstimulated ARPE-19 cells transfected with the WWP2 WW3-4 domain, detected via the cloned HA-Tag. Anti-HA primary antibody and secondary A488 antibody was used to identify the WW3-4 shown on the left whilst the blue channel was used to detect the DAPI nuclei staining and the merged image on the right shows both the green channel and the blue DAPI staining. Together the results suggest that WW3-4 localises in the nucleus exclusively. The images shown are representative of all cells examined.

### 3.2.1.5 Exploring the effects of transfected WWP2 WW domains on levels of Smad expression using western blots

The Smad3-dependent TGF $\beta$  gene expression was investigated in section 3.2.1, 3.2.2 and 3.2.3 by performing luciferase assays on transfected WWP2 WW domains. The aim in this segment is to decipher the effects transfected WWP2 WW tandem domains have on intracellular Smad levels. To do this, HEK293A cell lysates transfected with various WWP2 WW domains and Smad positive controls were loaded onto 10% SDS PAGE gels and transferred onto nitrocellulose membrane. These were then probed with Anti-Smad2, Smad3 and Smad7 antibodies, whilst the anti- $\beta$ -Actin antibody was used for normalization. The investigation into this binding may provide an explanation for the effects these transfected domains had on the Smad3-dependent TGF $\beta$  activity in previous luciferase assays.

Results from four different western blots using Smad2, 3, 7 and  $\beta$ -actin antibodies were collated and displayed in fig 3.2.9. Firstly, the result from Smad2 antibody probing suggests that the levels of Smad2 are higher on transfection with WW3-4 in comparison to WW1-2, 2-3 and 1-4, although Smad2 was present in all four samples. This may indicate that Smad2 has preferential binding to WWP2 WW3-4 tandem domain or that WW3-4 somehow increases the levels of Smad2 in the cell. Although results are not clear in the Smad3 western blot due to the high levels of background smears, it appears that Smad3 was present in samples transfected with all four of the tandem constructs and is arguably at the highest level of expression in the WW1-2 transfected HEK293A lysates. The highest levels of Smad7 detected in the Anti-Smad7 western blot was in WW3-4 transfected cells. Interestingly, although Smad7 was detected in all protein samples, WW2-3 seem to contain the least amount of Smad7. The high levels of Smad7 detected in the WW3-4 sample, and the contrasting low levels in WW2-3 may indicate that the WW4 domain is responsible for the increase of Smad7 levels, potentially due to WW4-Smad7 binding.

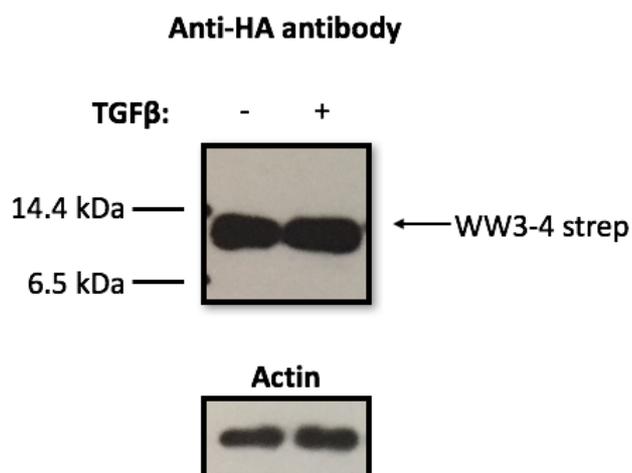


**Figure 3.2.9- Protein expression of Smad 2, 3, and 7 in WW tandem domain transfected HEK293A cells.** Western blots using Anti-Smad2, Smad3 and Smad7 antibodies to probe for differences in levels of Smad expression in HEK293A cells transfected with WW tandem domains. Results show that the highest levels of Smad2 and Smad7 detected was due to the overexpression of WW3-4. Lysates transfected with Smad 2 (52 kDa), 3 (48 kDa) and 7 (46 kDa) constructs were used as positive controls and were loaded onto gels respective to the antibody used. Anti- $\beta$ -actin antibodies were used to normalize levels of protein across different samples. (Experimental replicate n=1)

### 3.2.2 The study of WWP2 WW3-4 tandem domain and associated proteins using Quantitative Mass spectrometry

To further investigate the effects WWP2 WW3-4 has on Smad3-dependent TGF $\beta$  activity as shown in fig 3.2.2, quantitative mass spectrometry was carried out by the John Innes Proteomics facility on pull downs of HEK293A cells transfected with the streptavidin and HA tagged WWP2 WW3-4 construct using Strep-Tactin Sepharose. Using such quantitative method should allow the identification of peptides associated to the WW3-4 protein and the comparison between levels of these proteins in TGF $\beta$  and control samples.

To begin with, an 8 aa strep tag (WSHPQFEK) was cloned onto the 3' end of the HA-tagged WW3-4 construct using a combination of Phusion PCR, *XhoI/HindIII* digestion and T4 ligation. This construct was then transfected into HEK293A cells using standard PEI protocol and harvested 3 days post transfection. Whole cell lysates were then pulled down using Strep-Tactin Sepharose suspension overnight leaving only strep-tagged WW3-4, and any proteins bound to this construct after multiple washing steps (as explained in section 2.4.5). To check the success of pull down and the expression of WW3-4, a western blot was performed on the pull-down sample using HA-tag antibodies which showed the WW3-4-HA-Strep construct at the correct size of 9.2kDa (as shown in fig 3.2.10).



*Figure 3.2.10- The protein expression of WWP2 WW3-4 in HEK293A from a streptavidin pull-down experiment. A streptavidin tag was cloned onto WW3-4 pRK5-HA constructs. The expression of WW3-4 pRK5-HA-strep is shown here from lysates pulled using strep-tactin beads on a western blot using anti-HA antibodies. (Experimental replicate n=1)*

The transfection of HEK293A cells with the WWP2 WW3-4 construct, subsequent TGFβ stimulation and pull down was then repeated on a larger scale. These lysates were then processed in preparation of trypsin digestion and Orbitrap/iTRAQ quantitative mass spectrometry. A summary of the findings is shown in table 3.2-1.

*Table 3.2-1- Quantitative mass spectrometry results showing effects of TGFβ on proteins associated with WWP2 WW3-4.*

<b>Proteins found in the WW3-4 transfected HEK293A cell sample</b>	<b>Abundance ratio (TGFβ:Control)</b>
Histone H2A	11.164
Histone H2B	6.802
Histone H4	4.089
Myosin heavy polypeptide 11	1.916
FUS RNA binding protein	1.678
Myosin heavy polypeptide 9	1.641
Keratin	1.581
Myosin light polypeptide 6	1.577
Methylcrotonyl CoA carboxylase	1.550
Nucleolin	1.546
Chaperonin	1.369
14-3-3 epsilon	0.708
Lupus La protein	0.66

Comparing the protein levels in WW3-4 transfected cells between TGF $\beta$  stimulated and non-stimulated cells highlights the effects of TGF $\beta$  on the expression of associated proteins. Results from the mass spectrometry suggests that the largest TGF $\beta$ -dependent increase in protein interaction was the 11-fold increase in the levels of Histone H2A detected. The stimulation of TGF $\beta$  also increased levels of interaction between WW3-4 and two other Histone proteins, Histone H2b and H4 showing a 6, and 4-fold increase, respectively. Three myosin proteins, myosin heavy polypeptide 11, 9 and light polypeptide 6 were detected in the pull down samples and showed an increase in levels by 1.9, 1.6 and 1.58-fold, respectively. Levels of other associated proteins including Fusion (FUS) protein (1.7-fold increase detected), Keratin (1.6-fold increase), Methylcrotonyl CoA carboxylase (1.5-fold increase), Nucleolin (1.5-fold increase) and Chaperonin (1.4-fold increase) also displayed an increase in binding in response of TGF $\beta$  stimulation. However, together with the knowledge that Keratin is a common contaminant found in mass spectrometry, we surmise that the levels of Keratin found was an artefact and not a protein associated with WWP2 WW3-4. Interestingly, TGF $\beta$  seem to also exert a negative effect on certain proteins interactions. The levels of 14-3-3 epsilon protein and Lupus La protein interaction with WW3-4 decreased by 30% to a ratio of  $\sim 0.7$  suggesting that TGF $\beta$  not only positively affect proteins associated with WWP2 WW3-4, but can also decrease the levels of some proteins associated with the WW3-4 peptide.

### 3.3 Discussion

The focus of this chapter was to explore the biological roles of the WWP2 WW tandem domains and the possible effects TGF $\beta$  has on the interactions between WW domains and their potential binding partners. The effects of WWP2 WW domain on Smad3-dependent TGF $\beta$  gene expression was investigated using the p-CAGAC12-luc construct in luciferase assays whilst the binding of other proteins to WW3-4 domain was explored using quantitative mass spectrometry. The knowledge gained from these experiments combined with the results from the following chapters highlighting the expression, function and regulation of WWP2 isoforms should allow us to achieve a broader understanding of the interplay between WWP2, TGF $\beta$  and cancer.

#### 3.3.1 Investigating WWP2 WW domain and Smad3-dependent gene expression

Previous studies conducted by Soond *et al* suggested that the WWP2 WW1 domain binds preferentially to the PPGY motif of the linker region within residue 180-240 of Smad3 whilst the WW4 domain binds exclusively to the PPXY motif of Smad7 (Soond and Chantry, 2011). Additionally, recent NMR studies showed that the WW4 domain actually has a higher binding affinity to phospho-Smad3 compared with native-Smad7 whilst also demonstrating low levels of Smad2 binding (Wahl, 2016). Results from luciferase reporter assays showed that WW3-4 transfected HEK293A cells increased Smad3 activity via the CAGAC12 promoter by 15-fold. This increase could be the result of several different binding mechanisms. First, we postulate that the increase in WW3-4 tandem domain lacking in HECT ligase function leads to the binding of Smad3 and therefore prevents WWP2-FL and other functional NEDD4 ubiquitin ligases from binding to and facilitating the degradation of Smad3, ultimately stabilizing the levels of intracellular Smad3 activity (shown in fig 3.3.1A). In another study, it was found that WW domains in Smurf1 and Smurf2 had the propensity to form WW dimers (Aragón *et al*, 2012), therefore it may be possible that transfected WW3-4 constructs are

binding to WWP2-FL naturally expressed in cells through the WW3-4 domain or even WW1-2, inhibiting Smad3 binding and subsequent degradation mediated by FL. However, it is not yet clear what region of the isoform is involved in the binding, or whether this binding prevents FL from binding to and facilitating the degradation of Smad3. It is for this reason that further experiments are needed to investigate the potential binding of WWP2-FL to WWP2 WW tandem domains. Co-immunoprecipitation studies using HA-tagged WWP2 WW tandem domains and the Flag-tagged WWP2-FL can be performed with anti-HA-antibodies to extract the HA-WW tandem domains and any associated proteins. To test the presence of WWP2-FL bound to the WW tandem domains, western blot protein analysis using anti-Flag antibodies can then be performed on bound fractions.

Under TGF $\beta$  stimulation, overexpression of WWP2 WW3-4 lead to a 1.5-fold increase in Smad3-dependent TGF $\beta$  gene expression which was later found to be statistically insignificant. According to the previously mentioned NMR study by Wahl, in TGF $\beta$  conditions, WW4 should have increased binding to phospho-Smad7 in comparison to Smad3 (Wahl, 2016). This means that in our experiment, we should have seen a decrease in Smad3-dependent TGF $\beta$  gene expression. Due to our result being statistically insignificant, we therefore postulate that the WW3-4 domain may still bind preferentially to Smad7 under TGF $\beta$  stimulation as proposed by Wahl, subsequently causing the stabilisation of Smad7 levels and a marked decrease in Smad3-depedent TGF $\beta$  gene expression when compared with the 35-fold increase detected in experiments without TGF $\beta$ .

**A****TGFβ Independent**

In the absence of WW3-4:



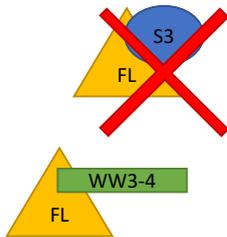
WWP2-FL binds to Smad3, mediating ubiquitin-dependent degradation subsequently causing a decrease in Smad3 levels.

↓ Smad3 levels

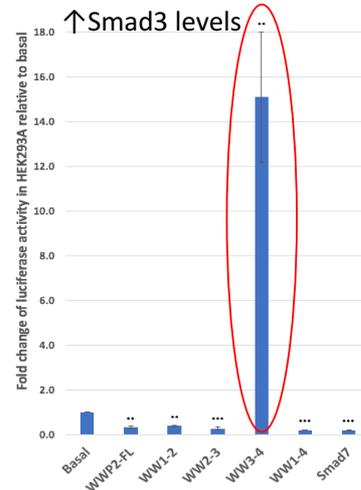
In the presence of WW3-4:



The catalytically non-functional WW3-4 binds to and stabilizes Smad3 levels.



WW3-4 may also bind to WWP2-FL to block functional binding of FL to Smad3 therefore inhibiting Smad3 degradation.

**B****TGFβ Dependent**

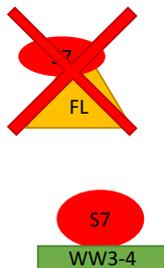
In the absence of WW3-4:



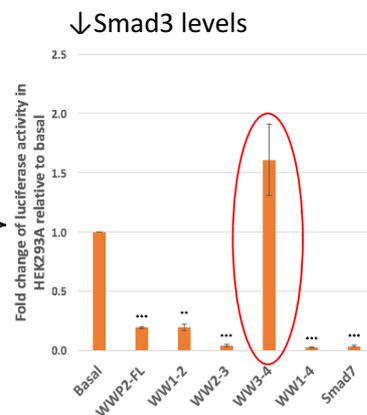
The WW4 domain within WWP2-FL binds to Smad7 leading to subsequent degradation of inhibitory Smad7 and therefore increased Smad3 levels.

↑ Smad3 levels

In the presence of WW3-4:



The catalytically inactive WW3-4 competitively binds to Smad7 causing the stabilization of Smad7 and therefore increasing degradation of Smad3. Note that the increase in Smad3 levels shown here are insignificant.



**Figure 3.3.1- Suggested mechanism of WW3-4/Smad binding.** WWP2 WW3-4 binds to Smad3, effectively blocking binding of functional ubiquitin ligases such as WWP2-FL resulting in the stabilization of Smad3 in a TGFβ independent manner (A). In the presence of TGFβ, (B), The increase in Smad3 levels are insignificant and therefore we postulate that the catalytically inactive WW3-4 may be able to bind to and subsequently stabilise levels of Smad.

Also seen from the luciferase assay of transfected HEK293A cells is the negative effect transfected Smad7 has on the levels of Smad3 binding which can be explained by the natural inhibitory effects Smad7 has on the TGF $\beta$  pathway. Similarly, the fold change induced by transfected FL is also lower than basal suggesting that WWP2-FL binds to and mediates degradation of Smad3 in both the presence and absence of TGF $\beta$ . When the Luciferase assay was repeated on A375 melanoma cells, it was found that once again, WW3-4 induced the highest fold change in Smad3 gene expression in both TGF $\beta$  stimulated and unstimulated cells. This suggests that the WWP2-Smad binding are similar in both normal and malignant cancer cells.

To investigate the increase in Smad3 levels seen in cells transfected with WW3-4 and identify the WW domain(s) responsible for the increase in Smad3-dependent gene expression, the same experiment was repeated on A375 cells transfected with WW3, WW4 domains individually, WW3+4 domains co-transfection in comparison with cells transfected with the WW3-4 tandem domain. Comparing these results should allow us to pinpoint the key WW recognition domain that was responsible for the significant increase in Smad3-dependent TGF $\beta$  gene expression. Results showed that in TGF $\beta$  negative cells, WW3-4 tandem domain transfections produced the highest level of Smad3 activity, whilst in the presence of TGF $\beta$ , transfected WW4 gave the highest luciferase reading. However, this is not due to an increase in luciferase results from WW4 transfected cells in response to TGF $\beta$  stimulation, but rather the decrease in Smad3-dependent gene expression of WW3-4 transfected cells. Once again results from the binding affinity study by Wahl can be used to explain the changes detected in the luciferase study. In unstimulated cells, WWP2 WW3-4 preferentially binds to Smad3 and therefore due to the lack of catalytic domain in the WW tandem domains, Smad3 levels are stabilized. However, when cells are stimulated with TGF $\beta$ , Smads become phosphorylated and the binding affinity of WW4 domain within WW3-4 shifts to favour phospho-Smad7 and therefore acts to increase Smad7 and subsequent Smad7-dependent inhibition of Smad3 (Wahl, 2016). Ultimately, results from the Smad3-

dependent gene expression luciferase study suggests that after investigating all the possible combinations of WW3 and WW4 domains, the WW3-4 tandem domain, containing the linker region between WW3 and WW4 domains, still causes the highest Smad3-dependent gene expression. Due to the low Smad3-dependent expression from WW3 and WW4 domains that were co-transfected in one sample (that does not contain the WW3-4 linker region), we can speculate that the 6 aa linker region between these WW domains holds some importance in the positive effect WWP2 WW3-4 has on Smad3-dependent gene expression. A recent study has highlighted the importance of the linker region between WWP2 WW2 and WW3 recognition domains in the auto-inhibition of WWP2 by the phosphorylation of two tyrosine residues within this region (Chen *et al*, 2017). Although the linker region between WW2 and WW3 is 35 aa longer than the linker region between WW3 and WW4 domains, it is reasonable to surmise that the WW3-4 linker region has a role to play in the increase of Smad3-dependent gene expression observed in the luciferase assays and therefore may be involved in WWP2 regulation, similar to the WW2-3 linker region.

To hone in on specific WW domain function, individual WW domains were cloned and transfected into HEK293A cells with p-CAGAC12-Luc reporter to study Smad3-dependent TGF $\beta$  gene activity. Results from WW1, 2 and 3 transfected cells showed no significant difference relative to basal activity however, WW4 induced a fold change of over 1.5 in absence of TGF $\beta$  and just under 2 in TGF $\beta$  stimulated cells. This can be explained once again using the binding affinities of WW4 domain with different Smad molecules. Using NMR, it was found that the binding affinity of phospho-Smad7 to WW4 was the highest, with Native-Smad3/WW4 binding next, then native-Smad7/WW4 binding third (Wahl, 2016). This therefore suggests that in cells unstimulated with TGF $\beta$ , WW4 binds to native-Smad3 over native-Smad7 therefore stabilizing levels of Smad3 leading to the increase in luciferase activity. In TGF $\beta$  stimulated cells, WW4 has preferential binding to phospho-Smad7 over native-Smad3. This could potentially inhibit the process of functional E3 ligase recruitment such as NEDD4L and WWP2-FL by Smad7,

which can act as an adapter protein to consequently send Smad3 for E3-dependent degradation (Yan *et al*, 2016). The competitive binding to WW4 to Smad7 therefore blocks the binding of functional E3s, ultimately leading to the stabilization of Smad3-dependent gene expression.

Interestingly, when the experiment was repeated in A375 melanoma cell line, the effects of TGFβ on Smad7 was similar to that of HEK293A cells. However, in TGFβ negative cells, WW1 transfected cells produced the highest Smad3-dependent gene expression whereas WW2 transfected cells induced the highest Smad3 activity in TGFβ stimulated A375s. Previously, it was found that the WW1 domain was likely responsible for the binding of Smad2/3 (Soond and Chantry, 2011) which suggests that here, in unstimulated A375s, WW1 domain binds to and stabilizes levels of Smad3 reflected by the near 5-fold increase in luciferase activity. In a study by Ganji *et al*, it was found that TAT-fused Smurf2 WW2/WW3 domain constructs lead to an intracellular increase in levels of phospho-Smad3 (Ganji *et al*, 2015). The WW2 domain of Smurf2 spans from amino acid 251-284 whilst WWP2 WW2 spans amino acid 330-363. Although there is only a 36.4% sequence similarity in these two 33 amino acid domains (shown below in fig 3.3.2 using ClustlW), it can be suggested that the increase in levels of Smad3 in WWP2 WW2 transfected cells are due to the binding and therefore the stabilization of Smad3.

```

Smurf2 WW2      PDLPEGYEQRTTQQGQVYFLHTQTGVSTWHDPRV
WWP2    WW2      RPLPPGWEKRTDPRGRFYVDHNTRTTTWQRPTA
                ** *:*:** :*:.*::. :* .:***: * .

```

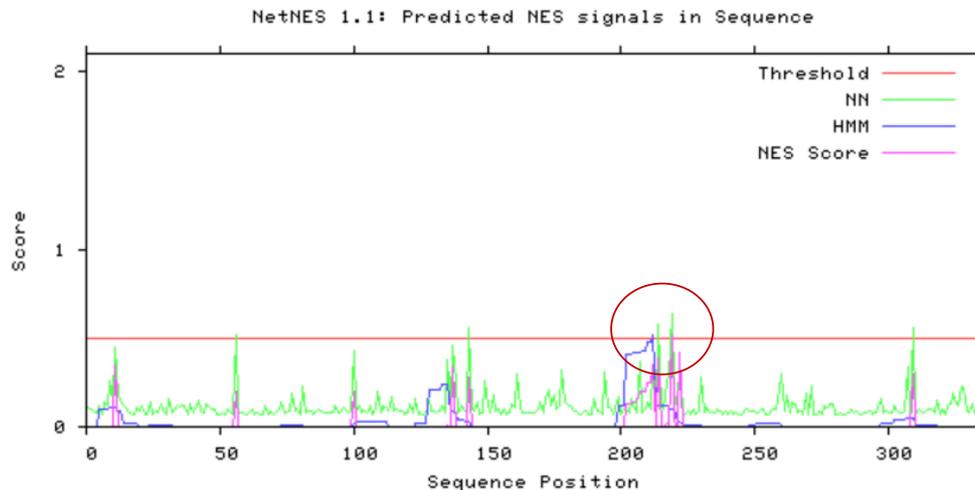
**Figure 3.3.2** A sequence alignment of the WW2 domain within WWP2 and Smurf2. The comparison between WW2 domain of WWP2 and Smurf2 using ClustlW ENSEMBL-EMBI alignment tool to show 36% sequence similarity.

To help further understand results from the WWP2 WW tandem domain luciferase analysis, western blots using anti-Smad antibodies were performed to investigate effects transfected WWP2 WW tandem domains have on the levels of Smad 2, 3, and 7 proteins. Results showed that although all three Smads were present in the cell lysates transfected with

WW1-2, WW2-3, WW3-4 and WW1-4 tandem constructs, the highest levels of Smad7 were detected in the WW3-4 sample. The increase in Smad7 levels is due to the binding of overexpressed WW3-4 to Smad7 which interferes with the binding of the catalytically functional WWP2-FL to Smad7. This subsequently means that the level of Smad7 ubiquitin-dependent degradation is lowered, as the non-functional WW3-4 peptide binds to and stabilizes the levels of Smad7. Also shown in the western blot analysis, WW2-3 sample detected levels of Smad7 that were significantly lower in comparison to samples transfected with WW3-4. The difference in levels of Smad7 between the overexpression of these two tandem constructs, both containing WW3, suggests that the increase in Smad7 levels is most likely due to the WW4 domain rather than the WW3 domain. To support this finding, data from another study has suggested that Smad7 has preferential binding to WW4 domain in WWP2 meaning that WW4 within the transfected WW3-4 is able to bind to and stabilize the levels of intracellular Smad7 (Soond and Chantry, 2011). Interestingly, the highest levels of Smad2 in the western blot study was found in the WW3-4 transfected sample whilst Smad3 levels were the highest in WW1-2 transfected cells. The latter result is supported by previous studies highlighting the propensity WWP2 WW1 domain has, to bind preferentially to Smad3 therefore inhibiting the formation and subsequent degradation of WWP2-FL/Smad3 complex, which explains the increase in Smad3 levels detected in WW1-2 transfected cells (Soond and Chantry, 2011). However, the same study shows evidence that conflict with findings from the western blots shown in fig 3.2.9 which shows expression of Smad2 is the highest in the WW3-4 sample. According to Soond *et al*, WW1 rather than WW3-4 has preferential binding to Smad2 which should have reflected in the western blot study by the increase in Smad2 levels in the WW1-2 sample, not WW3-4. Therefore, using the information gathered, it can be speculated that Smad2 may bind to regions of the WW3-4 tandem domains as well as the WW1 domain within the WWP2 protein.

In addition to the experiments mentioned above, immunolabelling was also carried out on ARPE-19 cells transfected with WWP2 WW3-4 in attempts to locate the intracellular localization of this tandem domain. Results from the cellular sub-localization study suggests that the transfected WW3-4 domain is localized exclusively in the nucleus and not in the cytoplasm nor at the membrane suggesting that the WW3-4 peptide was transported into the nucleus post translation in the cytoplasm. There are currently no known NLS signals found in the WW domains (Sudol *et al*, 2001), suggesting that the WW3-4 peptides were chaperoned into the nucleus by associating with another protein that may contain a NLS signal. Furthermore, we propose that the nuclear retention of WW3-4 might be due to its binding with Smad7 and subsequent nuclear retention which has been shown to be mediated by p300-dependent acetylation of Lys64 and Lys70 in Smad7 (Xu, 2006). The acetylation at these residues prevents the ubiquitination and subsequent translocation into the cytoplasm where degradation of Smad7 would have taken place. This therefore suggests that the ubiquitination of a protein is a potent signal for nuclear export and therefore when WW3-4 binds to Smad7, effectively blocking the binding of functional WWP2, Smad7 is not ubiquitinated and therefore stays in the nucleus, whilst bound to WW3-4, hence the exclusive nuclear localization of WW3-4 seen in the immunolabelling study. Moreover, the absence of a functional C2 lipid binding domain in the WW3-4 peptide further supports the lack of membrane localization of the WW3-4 peptide. In addition, the WW3-4 tandem construct that lacks the HECT catalytic domain also indicates a possibility of a nuclear export signal (NES) in HECT therefore suggesting that without HECT, the tandem construct is unable to carry out nuclear export. It was shown in a recent study that truncated WWP2 constructs lacking the HECT domain resulted in increased nuclear accumulation, therefore providing further evidence to suggest that the WWP2 HECT domain may contain a potent NES (Zhu *et al*, 2017). Using netNES, the amino acid sequence of WWP2 HECT domain was probed for potential NES sequences. The software combines the Artificial neural network and Hidden Markov model scores to calculate the likelihood of NES signals within the input

sequence. Results shown in fig 3.3.3 suggests that although NES activity is low within the HECT domain, there is one region circled in red that might contain a NES signal. Specifically, the Glutamic acid residue at position 219 of the HECT domain, (res. 754 of the WWP2-FL protein) might be involved with nuclear export activities (La Cour *et al*, 2004).



**Figure 3.3.3- Prediction of NES sequences in WWP2 HECT domain.** Using netNES, the Artificial neural network (NN) and Hidden Markov model (HMM) scores were combined to give an overall NES score which correlates to the prediction of a NES sequence. The higher the score is above the threshold, the more likely a NES is present. Results suggest the likelihood of an NES present at position 219 of the WWP2 HECT domain.

### 3.3.2 Effects of TGF $\beta$ on WWP2 WW3-4 protein associations

Using a Strep-Tactin pulldown method, the recombinant strep-tagged WWP2 WW3-4 construct and other proteins found associated with WW3-4 in the pulldown sample were purified and samples ran using OrbiTrap/iTRAQ quantitative mass spectrometry. By analysing and comparing two different samples, one stimulated with TGF $\beta$  and the other a negative control, any effects TGF $\beta$  had on the proteins associated with WW3-4 could therefore be detected. Results showed that the largest increase caused by the stimulation of TGF $\beta$  were three different histone proteins; Histone H2A, Histone H2B and Histone H4, which are three of the four core histone proteins found in the nucleosome (Boyer *et al*, 2000), exhibiting a 11, 6 and 4-fold increase respectively. Histones are proteins bound to DNA in the nucleus and plays an important role in transcriptional gene regulation. Studies have shown

that in cells experiencing TGF $\beta$ -induced apoptosis, levels of ubiquitinated histone H2A (uH2A) falls dramatically (Marushige and Marushige, 1995). The Deubiquitination of uH2A is needed for the activation of gene transcription therefore suggesting that ubiquitinated H2A has a role in gene repression. The decrease in uH2A caused by TGF $\beta$  induced apoptosis therefore leads to the mis-regulation and overexpression of certain genes ultimately contributing to different disease formations including tumour development (Vissers *et al*, 2008; Han *et al*, 2012). The cause of the increase in H2A detected in the WW3-4 pulldown mass spectrometry under TGF $\beta$  conditions is yet unclear, however we speculate that this increase could represent H2A binding to WW3-4 domains within the novel WWP2- $\Delta$ HECT isoform, which contain a non-functional HECT catalytic domain, therefore lowering levels of uH2A. The increased association of H2A to WW3-4 in the presence of TGF $\beta$  might therefore represent the mechanism which causes the lowering of uH2A levels via binding to WW3-4 domain of the catalytically inactive WWP2- $\Delta$ HECT during the event of TGF $\beta$ -dependent apoptosis.

Whilst uH2A is most frequently associated with gene repression, ubiquitinated H2B (uH2B) is most often correlated with transcription activation. Although no specific binding between WWP2 and H2B has yet been found, RNF20 is an E3 ubiquitin ligase that has been shown to interact with H2B. Evidence from a RNF20 depletion study found that the lack of RNF20-dependent mono-ubiquitination of H2B lead to the increase in levels of certain proto-oncogenes (Shema *et al*, 2008). In addition, another E3 ubiquitin ligase that binds to and mono-ubiquitinates H2B is RNF40 which has also been shown to exert anti-oncogenic properties in breast cancer cells (Prenzel *et al*, 2011). Therefore, result from the mass spectrometry analysis might suggest WWP2 isoforms as novel binding partners of H2B. Furthermore, the increase in levels of H2B as a result of TGF $\beta$  stimulation in HEK293A cells reflects on the tumour suppressive nature of TGF $\beta$  in non-cancerous cells as we postulate that WW3-4 within the functional WWP2-FL binds to and facilitates the mono-ubiquitination of H2B therefore activating the transcription of potential tumour suppressors.

Another protein found in the cell lysate pulldowns transfected with WWP2 WW3-4 was myosin heavy chain 9 non-muscle protein encoded by the MYH9 gene. The levels of this protein associated with WW3-4 increased by 1.6-fold in response to TGF $\beta$  stimulation. This protein has previously been shown to act as a tumour suppressor in squamous cell carcinomas (SCC), similar to Smad7 which can also act as a tumour suppressor by inhibiting the TGF $\beta$  pathway. It was found that MYH9 knockouts induced invasive tissue formation in SCC which may be explained by the role myosin IIA has in the post-translational regulation and stabilization of the p53 tumour suppressor (Schramek *et al*, 2014). We therefore speculate that the binding between WW3-4 and MYH9 might reflect a mechanism by which this tumour suppressor is degraded as part of its normal life cycle via binding to WW3-4 domain of the functional WWP2-FL in non-cancerous HEK293A cells. However, upon TGF $\beta$  stimulation which has been shown to possess both tumour-suppressive and oncogenic properties and the ability to increase EMT potential, WWP2-FL binding of MYH9 via the WW3-4 domain is increased therefore supporting the increased degradation of the MYH9 tumour suppressor which aids oncogenesis.

FUS RNA binding protein is a protein found in the mass spectrometry results to have increased in association to WWP2 WW3-4 by 1.7-fold post TGF $\beta$  stimulation. This protein is known to be implicated in many RNA regulatory process such as transcription, transport and translation. More interestingly, it has been suggested that FUS plays an important role in RNA splicing (Lagier-Tourenne *et al*, 2010). The increase in FUS protein found as a response to TGF $\beta$  might therefore imply that TGF $\beta$  stimulation increases the production of truncated protein isoforms that result from alternate splicing events. The FUS protein has also been suggested to be involved in DNA damage response as it was found that similar to some E3 ubiquitin ligases, FUS is recruited in response to DNA damage (Mastrocola *et al*, 2013).

Another protein found to have increased interaction with WW3-4 in the presence of TGF $\beta$  was Nucleolin, a protein involved in transcription, RNA assembly and modification in the nucleolus (Ginisty *et al*, 1999). More recently, Nucleolin was found to be able to translocate from the nucleolus to the membrane of endothelial cells, a process that negatively regulates angiogenesis. Intriguingly, MYH9, a protein also detected in our mass spectrometry analysis, was identified as a Nucleolin binding protein responsible for the membrane translocation (Huang *et al*, 2006). The TGF $\beta$ -dependent increase in levels of Nucleolin associated with WW3-4 therefore might act as evidence to support the tumour-suppressive role TGF $\beta$  has by negatively regulating angiogenesis through the increase of Nucleolin. To further highlight the anti-oncogenic role of TGF $\beta$ , studies have found that Nucleolin directly binds to Mdm2, an E3 ubiquitin ligase, to inhibit the ubiquitin-dependent proteasomal degradation of p53 (Saxena *et al*, 2006). This subsequently causes the stabilization of tumour-suppressor p53 and therefore decreases the oncogenic potential within the cell. However, note that the above speculation regarding the tumour-suppressive role of TGF $\beta$  reflected by the increase in Nucleolin only holds truth if Nucleolin binds to WW3-4 domain within the non-functional WWP2- $\Delta$ HECT isoform. If the results reflect the TGF $\beta$ -dependent increase in Nucleolin bound to the catalytically functional WWP2-FL, then the subsequent proteasomal degradation would lead to the decrease in levels of Nucleolin.

The final protein found to have increased interaction with WW3-4 following the stimulation of TGF $\beta$  is Chaperonin, for which we detected a 1.4-fold increase compared to the negative control. Chaperonin is a protein responsible for the ATP-dependent folding of proteins from their basic primary structure (Spiess *et al*, 2004). The proteasomal degradation of chaperonin-containing t-complex polypeptide 1 (CTT) has been suggested to be a ubiquitin-dependent process following a study showing a marked increase in levels of CTT in a mutant cell line expressing non-functional E1 ubiquitin activating enzyme (Yokota *et al*, 2000). The presence of

Chaperonin in our WW3-4 pulldown sample may therefore support the finding that CTT degradation is dependent on the ubiquitination process.

Intriguingly, protein expression found in the mass spectrometry results did not always show an increase in response to TGF $\beta$  stimulation. For example, the La protein, also known as Sjogren syndrome antigen B (SS-B) exhibited a decrease in abundance ratio of 0.66. The involvement of the La protein in transcription and mRNA translation has been well established in literature, however, only recently has its role in oncogenesis been identified. In a study by Sommer *et al*, it was found that the La protein is a potent oncogene possessing the ability to increase cell growth, migration and invasion in metastatic hypopharyngeal squamous cell sarcoma (Sommer *et al*, 2011). In a separate study, evidence was provided suggesting that the oncogenic potential of the La protein derives from its ability to upregulate the expression of Mdm2, a protein that suppresses levels of the p53 tumour-suppressor (Sommer *et al*, 2011). The decrease in the La oncogenic protein in our mass spectrometry might therefore support the tumour-suppressive effects of TGF $\beta$  in non-cancerous cells.

The 14-3-3 protein was another protein to show decreased interactions with WWP2 WW3-4 following TGF $\beta$  stimulation. This protein, encoded by the YWHAE gene, has a wide range of binding partners which allows the mediation of a many cellular process including regulation of cell cycle, signal transduction and cell death (Fu *et al*, 2000). Recent studies have shown the role 14-3-3 plays in oncogenesis, suggesting that the decrease in 14-3-3 leads to an increase in oncogenic potential, specifically the increased levels of EMT (Raychaudhuri *et al*, 2016). The decrease in this tumour-suppressor gene following TGF $\beta$  stimulation might therefore suggest that TGF $\beta$  can be favourable to oncogenesis. Ultimately, the evidence provided above may support the dual role TGF $\beta$  has in oncogenesis, however, without further, more specific experiments targeting the effects of TGF $\beta$  on the proteins discussed above, the results from our experiment cannot act as solid evidence to prove the role of TGF $\beta$  on oncogenesis.

The key finding in this chapter is arguably the increase in Smad3-dependent TGF $\beta$  activity caused by the overexpression of WWP2 WW3-4 domain and the consequent hypothesis suggesting the potential importance of the 6 aa linker region (-QGMIQE-) to the function of the WW3 and WW4 domain. Moreover, it is important to note the ability of WWP2 WW tandem domains in competitively binding to Smad mediators and subsequently inhibiting the formation of functional WWP2/ Smad complexes. This therefore leads to the stabilization of Smad levels and could ultimately alter signal mediation in the TGF $\beta$  pathway.

Chapter 4 Prediction and validation of novel and  
existing WWP2 isoforms

### 3.4 Introduction

Ubiquitination is involved in many biological regulatory activities, most known is its role in protein degradation. The E3 ubiquitin ligase is an integral part of the process due to its involvement in the recognition of target substrates and the facilitation of ubiquitin transfer from the E2 ubiquitin-conjugating enzyme (Soond and Chantry, 2011). E3 ubiquitin ligases are categorized into 2 main groups and the WWP2 NEDD4 ubiquitin ligase falls under the HECT umbrella as it contains intrinsic catalytic activity in the form of its C-terminal HECT domain. Like many other mammalian proteins, NEDD4 ubiquitin ligases can produce truncated variants known as isoforms. Generation of these isoforms commonly result from alternate splicing events including intron retention and splicing at alternative splice sites which could lead to retention of putative promoters or premature stop codons (Wang *et al*, 2015; Soond *et al*, 2013). The function of these isoforms usually differs from the full-length proteins due to the difference in length, structure and domain composition. The existence of isoforms increases the range of protein functions thus increasing the diversity of a protein. An example of a NEDD4 protein that functions differently to its isoform is WWP1. An increased expression of this HECT containing ubiquitin ligase has been detected in human prostate cancers with further investigations linking its overexpression to the inhibition of TGF $\beta$ -dependent tumour-suppressant activities. However, although WWP1 displays strong oncogenic properties, the presence of a truncated WWP1 isoform in both normal prostate cell lines and prostate cancer cell lines suggest that this isoform is not involved in oncogenesis (Chen *et al*, 2007).

WWP2 is amongst 8 other NEDD4 proteins and currently 3 established isoforms are known; WWP2-FL (full-length), and two truncated versions, WWP2-N and WWP2-C. The full-length isoform contains the N-terminal C2 lipid binding domain, WW 1-4 substrate recognition domains and the C-terminal catalytic HECT domain. The C-terminal isoform, WWP2-C comprises of the WW4 and HECT domain, whilst the N-terminal isoform, WWP2-N

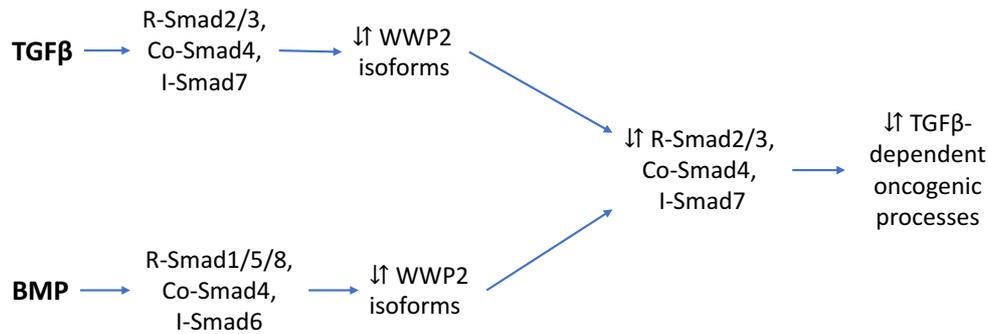
contains only the N-terminal C2 and WW1 domain. The transcription of the WWP2-C isoform is likely the product of the putative promoter P2 within intron 10/11 as shown in fig 1.3.2, whilst the generation of WWP2-N is thought to be due to a stop codon present in the retained intron 9/10 as a consequence of an alternate splicing event (Chantry, 2011).

The different combinations of WW recognition domains present in these WWP2 isoforms suggest different preferred interactions between target substrate molecules. Due to the interactions between WWP2 isoforms and Smad mediators, the role of WWP2 in TGF $\beta$ -dependent oncogenesis is a highlight in our investigations. Furthermore, the dual function of TGF $\beta$  in melanoma development, where it can act as both a tumour-suppressor and an oncogenic factor (Piek and Roberts, 2001), suggests that the interactions between WWP2 and TGF $\beta$  mediators at different stages of oncogenesis can result in very different outcomes. Previous studies on WWP2-N and WWP2-C suggests that regulatory-Smad2/3 binds preferentially to WW1 domain whilst the Inhibitory-Smad7 binds specifically to the WW4 domain (Soond and Chantry, 2011). WWP2-C is therefore a potential oncogene as it is the only isoform with both the intact WW4 and HECT catalytic domain. This means that during the later stages of tumorigenesis, when TGF $\beta$  switches from having tumour suppressant properties to becoming an oncogenic factor, WWP2-C could potentially bind to and block the natural negative regulation by I-Smad7 in the oncogenic signalling of the TGF $\beta$  pathway. On the contrary, WWP2-N has been suggested to act as a tumour suppressor due to its specificity for R-Smad2/3 which leads to the decrease in levels of EMT, a process driven by TGF $\beta$  (Soond *et al*, 2013). The TGF $\beta$  signalling pathway is therefore investigated in relation to WWP2 due to its importance in cancer development.

The relationship between TGF $\beta$  and WWP2 is a complex feedback loop in which the effects of TGF $\beta$  stimulation on the levels of WWP2 expression is facilitated by R-Smad2/3, Co-Smad4 and I-Smad7 where R-Smad/Co-Smad complexes can act as potential transcription activators or repressors of the

WWP2 gene after nuclear translocation. The resulting change in WWP2 isoform levels could then affect Smad activity via binding through WWP2 WW recognition domains as a feedback response. Ultimately, this change in Smad activity resulting from WWP2 binding will alter levels of TGF $\beta$ -dependent molecular processes such as EMT, represented in a simple diagram in fig 4.1.1. Studying the levels of WWP2 isoforms under TGF $\beta$  stimulation will therefore allow a deeper understanding of not only the role TGF $\beta$  has in oncogenesis but also validate the oncogenic or tumour suppressant properties any WWP2 isoforms may possess.

BMP is another growth factor highly implicated in melanoma that was used in our WWP2 studies. It has been shown to increase the oncogenic potential of cells by promoting processes such as cell migration and invasion (Rothhammer *et al*, 2007). Although binding between BMP mediators Smad1, 4, 5, 6, 8 and WWP2 has not yet been identified, the effect of BMP stimulation on WWP2 isoform expression could suggest BMP mediators as potential transcription activators or repressors of the WWP2 gene which could contribute to the levels of TGF $\beta$ -dependent oncogenic processes (also demonstrated in fig 4.1.1). As mentioned before, many studies have shown that BMP mediates a negative effect on oncogenic processes whilst others have suggested an oncogenic role. The results from our WWP2 expression study may produce evidence supporting BMP as either an oncogenic factor or a tumour suppressant.



**Figure 3.4.1- A simplified representation of the effects of TGFβ and BMP on levels of oncogenic processes.** TGFβ signalling is facilitated by Smad 2/3, 4 and 7 which alter levels of WWP2 expression by acting as transcriptional activators and repressors. The change in WWP2 levels then adjusts the levels of Smads by 1) non-functional WWP2 isoforms stabilizing Smad levels, and 2) functional WWP2 isoforms mediating proteasomal degradation of Smads. The change in Smad activity then affects TGFβ-related oncogenic gene transcription. Similarly, BMP stimulation can alter levels of WWP2 gene transcription via BMP Smad mediators, however, WWP2 is not yet known to bind to BMP Smads therefore suggesting that the altered WWP2 levels will affect TGFβ-dependent Smad2, 3, 7 oncogenic processes rather than BMP Smad1, 4, 5, 6, and 8 related oncogenic processes.

In this chapter, novel isoforms unseen from current literature were predicted and validated using a range of different experiments to investigate the levels of novel and existing WWP2 isoforms present in an array of melanoma cell lines and patient samples. Investigating the expression of these isoforms and the potential effects TGFβ and BMP have on WWP2 could increase the understanding of their roles in melanoma growth and progression. In a collaboration with the Nieto lab at the Institute of Neuroscience Alicante, the expression of WWP2 isoforms in epithelial and mesenchymal cells were explored. Due to the potential ability of WWP2 isoforms in modifying effects of the TGFβ pathway via Smad binding, we surmise that these isoforms may possess the potential to alter TGFβ-dependent EMT. To investigate expression patterns of WWP2 in epithelial and mesenchymal cells would allow the identification of possible key players in TGFβ-dependent EMT. The potential for the development of a prognostic tool in the future, adds to the importance of profiling the many WWP2 isoforms in the disease of melanoma (Soond *et al*, 2013).

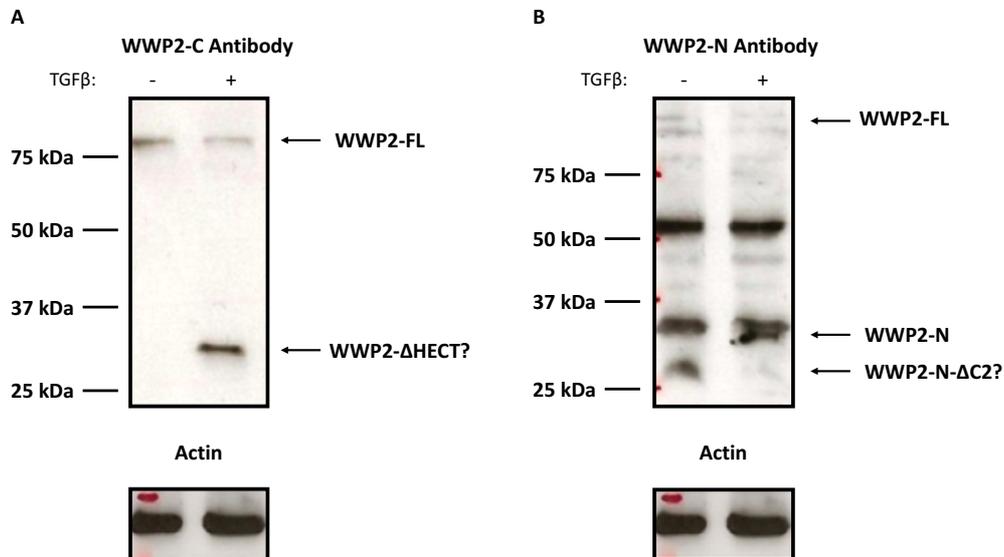
## 3.5 Results

### 3.5.1 Identification of novel isoforms using bioinformatics

There are three recognized WWP2 isoforms that are well established in literature; WWP2-FL, WWP2-N and WWP2-C. These are depicted in the WWP2 amino acid sequence shown in fig 1.3.3. The WWP2-N isoform starts at the methionine at position 1 (circled in orange) and ends with glycine at position 335 (also circled in orange). This 335-amino acid isoform therefore contains the C2 domain highlighted in pink, and the WW1 domain highlighted in grey. Note that the WW1 and WW2 domains overlap by 4 amino acids at position 330-333, meaning that WWP2-N also contains the first 4 amino acids of the WW2 domain. The C isoform on the other hand, starts with a methionine residue at position 440 (circled in purple) and is a 431-amino acid isoform that ends at the last amino acid of the full-length protein (Glutamic acid circled in purple). This suggests that the C isoform contains that WW4 domain highlighted in red and the complete HECT catalytic domain highlighted in green.

Using a WWP2-C terminal antibody (which targets a 14-22 amino acid sequence around the C terminal region of WW4 and the subsequent linker region) and a N terminal specific antibody that probes a region between residue 200 and 250 (corresponding to the linker region between the C2 and WW1 domain) in western blotting, bands corresponding to these established isoforms were identified at the correct sizes (as shown in fig 4.2.1 below). In addition to these bands, bands of a lower molecular weight were also highlighted suggesting either the presence of breakdown fragments or the expression of other potential truncated isoforms. Shown in fig 4.2.1A a western blot using WWP2-C terminal specific antibodies, a band of around 30kDa was observed induced by TGF $\beta$  whilst a band of around 25kDa, negatively affected by TGF $\beta$  stimulation is visible in fig 4.2.1B, a western blot using N-terminal specific antibodies. The effects of TGF $\beta$  on the expression of these lower molecular weight bands suggests that they are

TGF $\beta$  regulated WWP2 isoforms rather than breakdown fragments. To help identify these novel isoforms, a bioinformatics approach was taken to find potential novel isoforms that match the size and positions suggested by the western blot results.

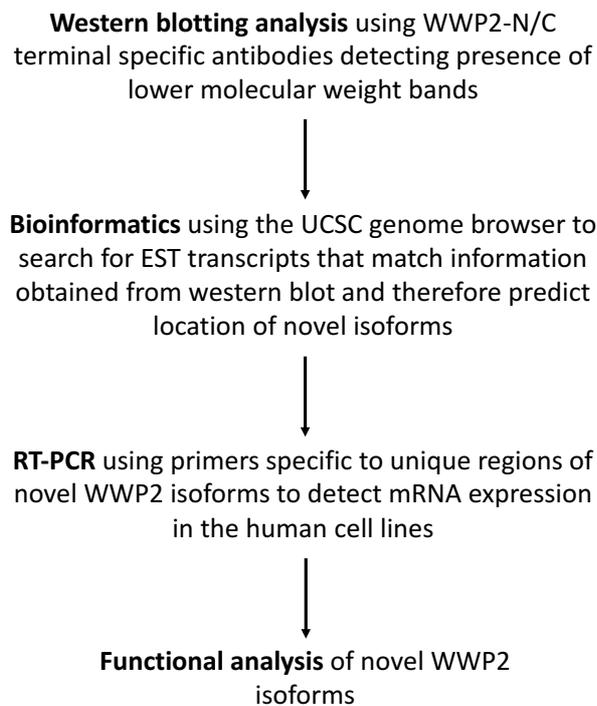


*Figure 3.5.1- The protein expression of WWP2 isoforms previously identified in western blots of HEK293A cells using primary WWP2-N and WWP2-C terminal antibodies. The level of proteins was normalized using a primary actin antibody and detected potential novel truncated isoforms labelled as WWP2- $\Delta$ HECT and N- $\Delta$ C2 (From S.Soond in the Chantry lab).*

The UCSC genome browser was used to detect any EST transcripts that included intronic sequences in the human WWP2 gene which might have resulted from an alternate splicing event. Shown in fig 4.2.2A, DC341937.1 was an EST transcript found containing regions of the retained intron 9/10 which suggests the potential for a new C-terminal isoform. Together with the finding shown in fig 4.2.2B, highlighting the presence of an EST transcript that ends in intron19/20 (BX471495.1), we therefore predict that this novel isoform spans exon 10-19 which suggest the presence of WW3, WW4 and a partial HECT catalytic domain (as shown in fig 4.2.6). This predicted isoform WWP2- $\Delta$ HECT will overlap the WWP2-C terminal specific antibody binding region and have a molecular weight of 38.3kDa which coincides with findings in the western blot shown in fig 4.2.1A.



The same methodology which is summarized in fig 4.2.3 was employed in the identification of WWP2-N-ΔC2 which initially presented itself as a 25kDa band on the western blot in fig 4.2.1B. Searching through the UCSC genome browser, we were able to find an EST transcript that began in intron 3/4 (DC396665.1 shown in in fig 4.2.4A) which would allow the translation to begin at the next start codon present in exon 5. Combined with another EST transcript which ends in a partially retained intron 9/10 (DB213134.1 shown in fig 4.2.4B), there is reason to believe that WWP2-N-ΔC2 may span exon 5-9 which would contain the WW1 domain and parts of the linker region prior making it 22.8kDa in weight and therefore corresponding to both the size and location within the WWP2 locus suggested by the Western blot results and antibody epitopes.



*Figure 3.5.3- A flow chart summarizing the detection, prediction and validation of WWP2 novel isoforms.*

## WWP2-N-ΔC2

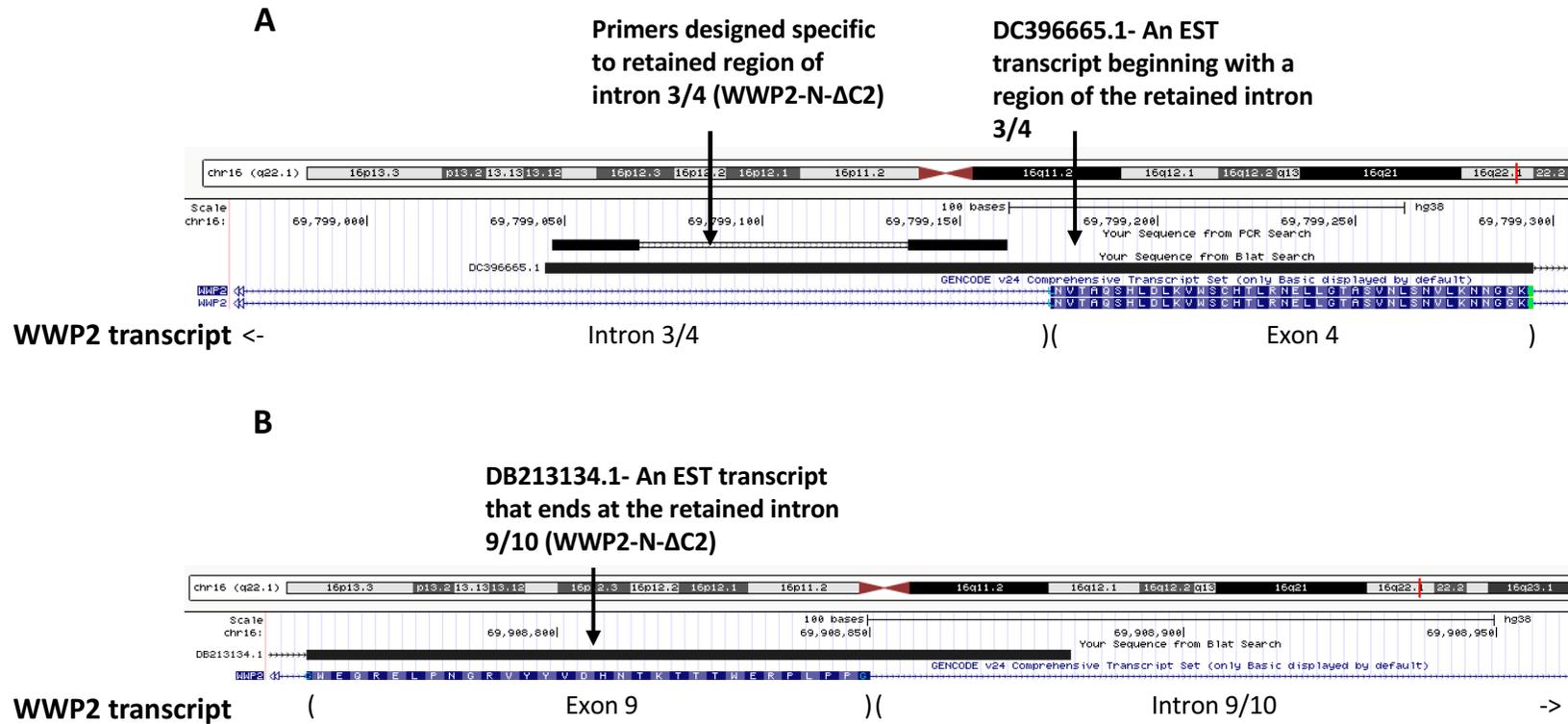
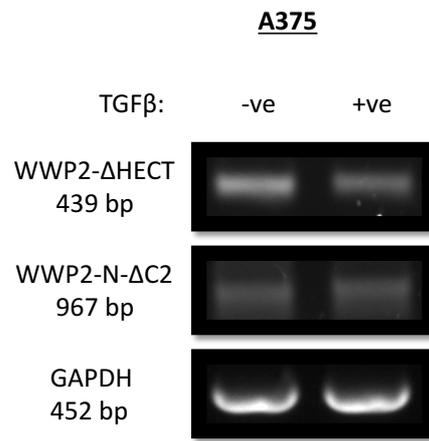
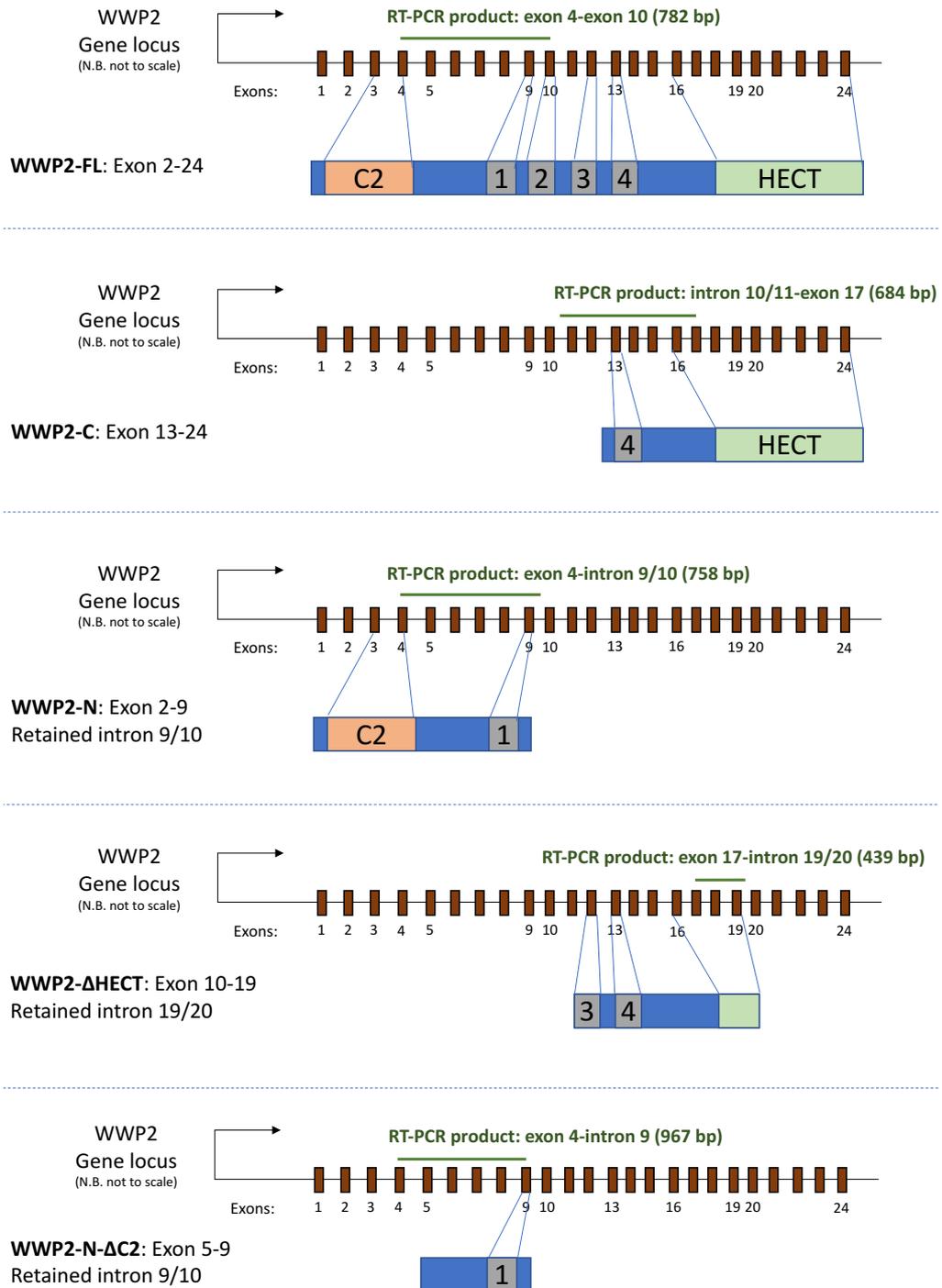


Figure 3.5.4- Whole genome analysis resulting in predicted regions of WWP2-N-ΔC2. A schematic derived from the UCSC genome browser highlighting A) the EST transcript responsible for the predicted isoform WWP2-N-ΔC2 and the location of the qPCR primer designed to detect the novel isoform. B) shows the EST transcript containing regions of intron 9/10 responsible for the predicted WWP2-N-ΔC2.

To confirm expression of these two predicted isoforms, RNA from melanoma cell line A375 stimulated with TGF $\beta$  was extracted and converted into cDNA using GoScript PCR. These were then used in Go-Taq RT-PCR (results shown in fig 4.2.5), along with specifically designed primers (highlighted in green on the WWP2 gene locus in fig 4.2.6) that recognize unique regions of these predicted isoforms to probe for the transcript expression of WWP2- $\Delta$ HECT and WWP2-N- $\Delta$ C2. Results in fig 4.2.5 confirm the presence of these novel isoforms at the correct predicted size. Interestingly, the effects of TGF $\beta$  stimulation in A375 using RT-PCR differ from the original western blot results of HEK293A cells shown in fig 4.2.1. The difference in band intensity was measured using Image J and showed that whilst TGF $\beta$  seem to induce expression of  $\Delta$ HECT in the HEK293A western blot, the reverse was observed in RT-PCR of the A375 melanoma cell line as the control sample showed 73% density compared to the TGF $\beta$  stimulated samples at 27% density. The difference of isoform expression between cancerous and non-cancerous cell lines in response to TGF $\beta$  stimulation thus suggest that more research is needed to verify the effects of TGF $\beta$  on WWP2 isoform expression as TGF $\beta$  may act as a key regulatory component of WWP2 expression in oncogenesis through the different effects it has on isoform expression in cancerous and non-cancerous cell lines. Following the validation of these novel isoforms using RT-PCR, specific cloning PCR primers were designed and used to clone the predicted novel isoforms from WWP2-FL construct template. These cloned isoforms were used in subsequent characterization experiments. The amino acid sequence of the novel isoforms WWP2- $\Delta$ HECT and WWP2-N- $\Delta$ C2 is highlighted in fig 4.2.7.



**Figure 3.5.5- mRNA expression of novel WWP2 isoforms.** An agarose gel image showing the presence of predicted novel WWP2 isoforms using RT-PCR of A375 cells and isoform specific primers. Cells were subjected to TGFβ stimulation and results were normalized to GAPDH levels. (Experimental replicate n=1)



*Figure 3.5.6- Novel and established WWP2 isoforms FL, N, C, ΔHECT, N-ΔC2 and FL-ΔC2 isoforms highlighting the domains included and the transcription region.*

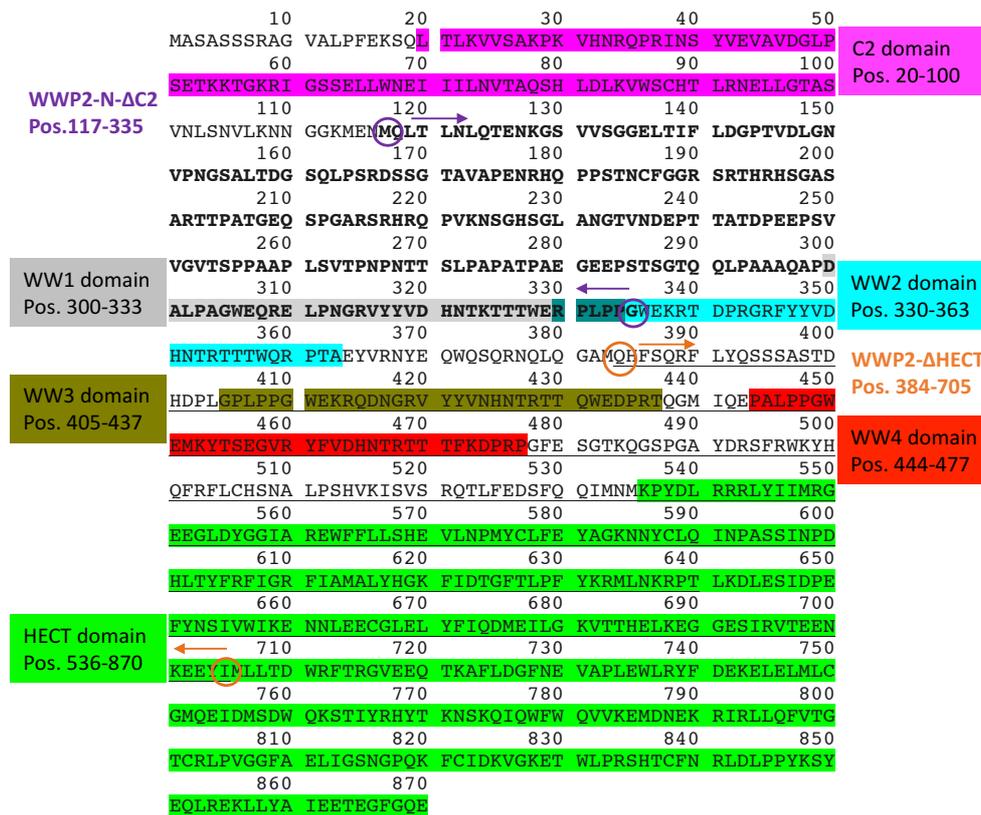


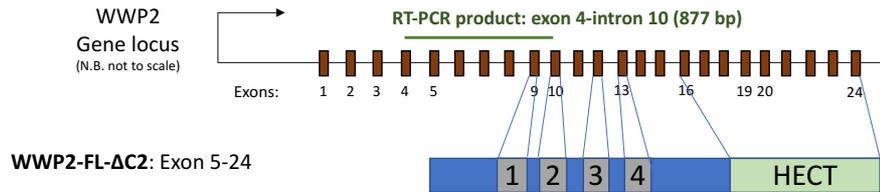
Figure 3.5.7- WWP2 novel isoforms highlighted in the WWP2-FL amino acid sequence. The WWP2 amino acid sequence highlighting the start and end positions of novel WWP2-N-ΔC2 is circled in purple (in bold letters), whilst WWP2-ΔHECT (underlined) is circled in orange. The locations of the C2 domain (pink), WW1 (grey), WW2 (blue), WW3 (dark green), WW4 (red) and HECT domain (green) are also highlighted.

### 3.5.1.1 Investigating the expression of WWP2-FL-ΔC2

WWP2-FL-ΔC2 is a C-terminal continuation of the N-ΔC2 isoform containing the WW1, 2, 3, 4 and HECT catalytic domain. Fig 4.2.8 is a schematic showing the domains present in FL-ΔC2 and also highlights its position within the amino acid sequence. This isoform, initially known as WWP2 variant 4, was recently highlighted in a study by Choi *et al* using Western blotting analysis and a genome database search, however, the expression of this isoform was never confirmed (Choi *et al*, 2015). Due to its similarities with predicted N-ΔC2 isoform, we decided to test the mRNA expression in a few of the cell lines used previously, and compare the effects of TGFβ and BMP stimulation. Unfortunately, when run on 1% agarose gels (fig 4.2.9), the products produced smears that prevented any accurate observations to be made. However, whilst only bands larger or smaller than the predicted size of

877 bp were detected in A375 and SKMEL28, M229 arguably shows the presence of WWP2-FL- $\Delta$ C2 in the unstimulated control and the BMP sample. MCF-7, a breast cancer cell line that was used in some experiments as an epithelial non-melanoma comparison also showed the potential expression of this novel isoform in all three of the different conditions.

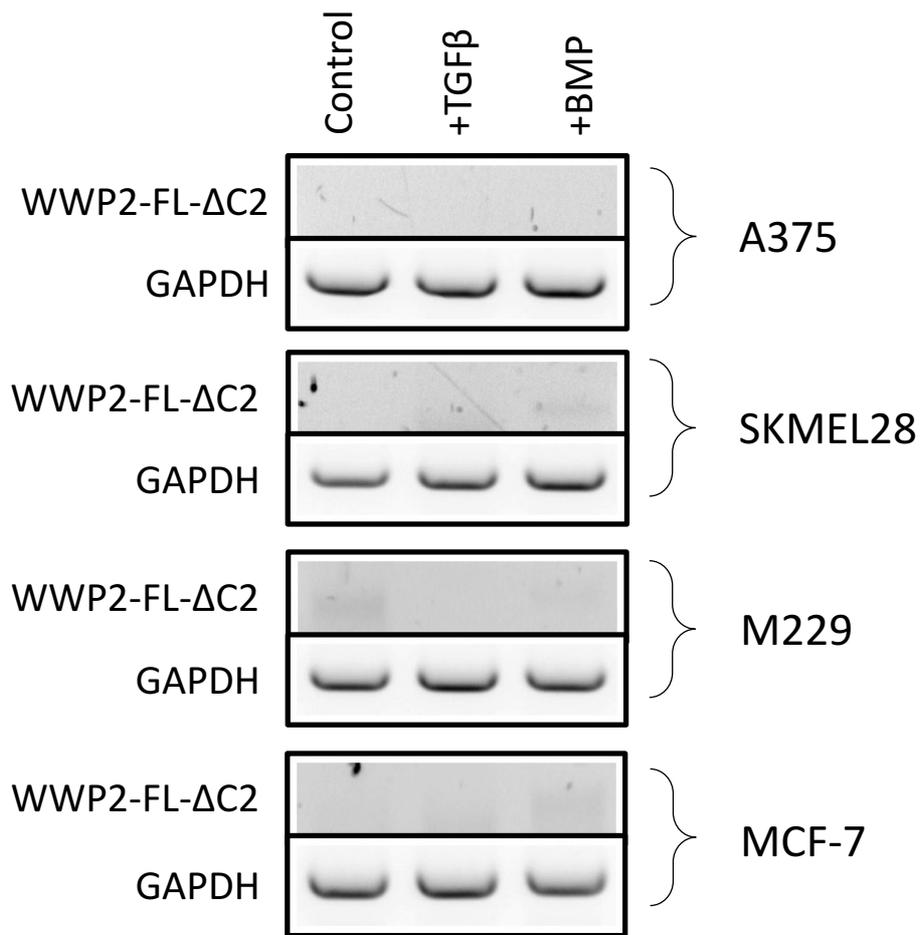
A



B

	10	20	30	40	50	
	MASASSSRAG	VALPFEKSQL	TLKVVSAPK	VHNROPRINS	YVEVAVDGLF	C2 domain Pos. 20-100
	60	70	80	90	100	
	SETKKTGKRI	GSSELLWNEI	IILNVTAQSH	LDLKVWSCHT	LRNELLGTS	
	110	120	130	140	150	
WWP2-FL- $\Delta$ C2 Pos. 117-870	VNLSNVLKNN	GGKMENMOLT	LNLTENKGS	VVSGGELTIF	LDGPTVDLGN	
	160	170	180	190	200	
	VPNGSALTDG	SQLPSSDSSG	TAVAPENRHQ	PPSTNCFGGR	SRTHRHSGAS	
	210	220	230	240	250	
	ARTTPATGEO	SPGARSRHQ	PVKNSGHSGL	ANGTVNDEPT	TATDPEEPSV	
	260	270	280	290	300	
WW1 domain Pos. 300-333	VGVTSPPAAP	LSVTPNPNTT	SLPAPATPAE	GEEPSTSGTQ	QLPAAAQAPD	WW2 domain Pos. 330-363
	310	320	330	340	350	
	ALPAGWEQRE	LPNGRVYYVD	HNTKTTWER	PLPPGWEKRT	DPRGRFYYVD	
	360	370	380	390	400	
WW3 domain Pos. 405-437	HNTRTTTWQR	PTAEYVRNIE	QWQSQRNQLQ	GAMQHFSQRF	LYQSSSASTD	WW4 domain Pos. 444-477
	410	420	430	440	450	
	HDPLGPLPPG	WEKRODNGRV	YYVNHNTRTT	QWEDPRTQGM	IQEPALPPGK	
	460	470	480	490	500	
	EMKYTSEGVR	YFVDHNTRTT	TFKDPREGF	SGTKQGSPPA	YDRSFRWKYH	
	510	520	530	540	550	
	QFRFLCHSNA	LPSHVKISVS	RQTLFEDSFQ	QIMNMKPYDL	RRRLYIIMRG	
	560	570	580	590	600	
	EGLDYGGIA	REWFLLSHE	VLNPMYCLFE	YAGKNNYCLQ	INPASSINPD	
	610	620	630	640	650	
HECT domain Pos. 536-870	HLTYFRFIGR	FIAMALYHGK	FIDTGFTLPF	YKRMLNKRPT	LKDLESIDPE	
	660	670	680	690	700	
	FYNSIVWIKI	NNLEECGLEL	YFIQDMEILG	KVTHELKEG	GESIRVTEEN	
	710	720	730	740	750	
	KEEYIMLLTD	WRFTRGVVEQ	TKAFLDGFNE	VAPLEWLRYF	DEKELEMLC	
	760	770	780	790	800	
	GMQEIDMSDW	OKSTIYRHYT	KNSKQIQWFV	QVVKEMDNEK	RIRLLQFVTG	
	810	820	830	840	850	
	TCRLPVGGA	ELIGSNGPOK	PCIDKVGKET	WLPRSHTCFN	RLDLPPYKSY	
	860	870				
	EQLREKLLYA	IEETEGFGQE				

Figure 3.5.8- A summary of the novel WWP2-FL- $\Delta$ C2 isoform domain composition and amino acid sequence. A) shows the correlation between domains and exon location within the WW2 gene locus and the RT-PCR product detected by specific primers shown in green. B) depicts the start and end of FL- $\Delta$ C2 (in bold) highlighted with orange circles in the full-length WWP2 amino acid sequence whilst also highlighting the C2, WW1, WW2, WW3, WW4 and HECT domain in grey, blue, dark green, red and green, respectively.



*Figure 3.5.9- mRNA expression of WWP2-FL-ΔC2 in melanoma cell lines. RT-PCR showing the expression of novel WWP2-FL-ΔC2 (877bp) in M229 and MCF-7 but not in A375 and SKMEL28. Levels of cDNA were optimized to GAPDH housekeeping gene levels. (Experimental replicate n=1)*

### 3.5.2 Characterizing the expression pattern of existing and novel WWP2 isoforms

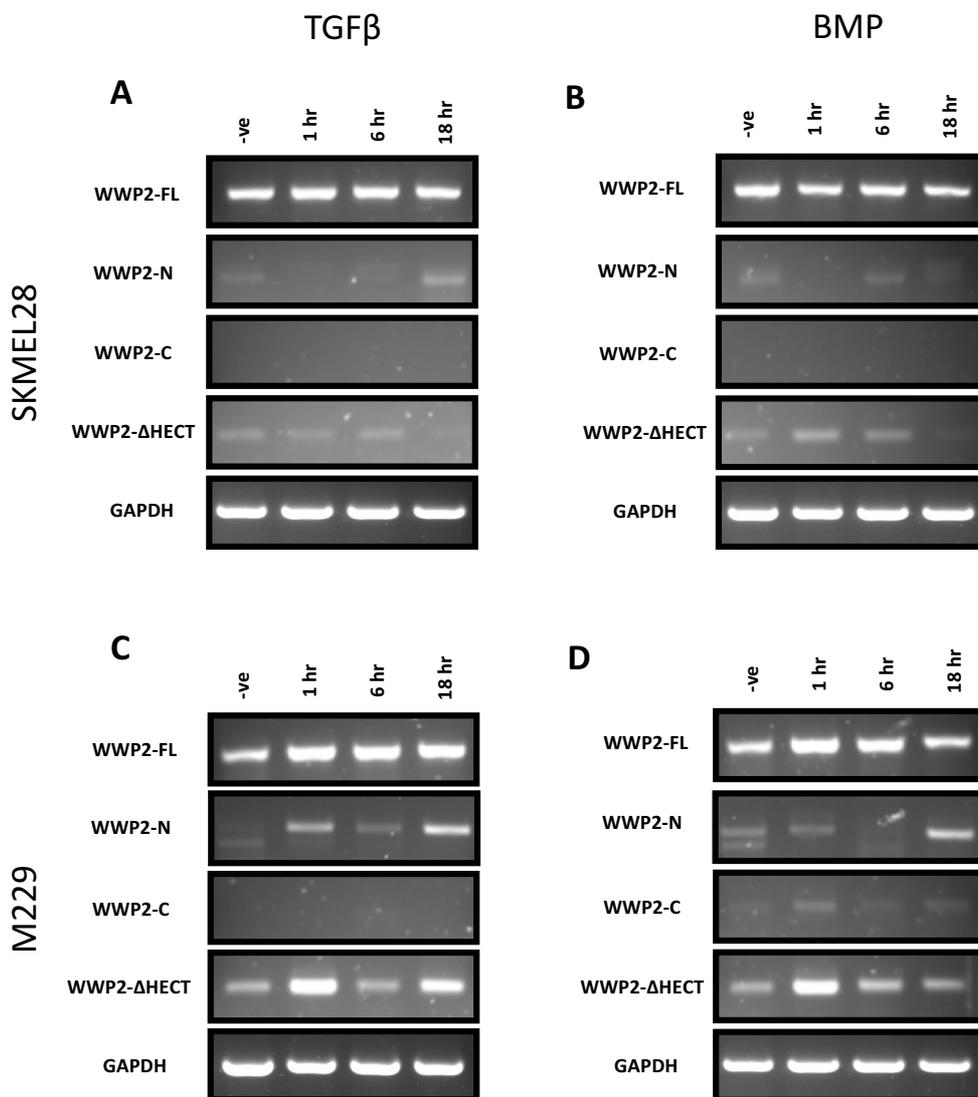
#### 3.5.2.1 Studying the expression of WWP2 isoforms in melanoma cell lines using RT-PCR and qPCR

To investigate the expression of WWP2 isoforms at the RNA level, a combination of RT-PCRs and qPCRs were performed using mRNA extracted from melanoma cells stimulated with either TGF $\beta$ 1 or BMP4. This was converted into cDNA using GoScript PCR (Promega) then combined with WWP2 isoform-specific primers in a GoTaq reaction (Promega) to produce results showing the expression of WWP2 isoforms. Levels of mRNA expression were normalized using the GAPDH housekeeping gene in RT-PCR. The combination of 18S and UBC reference genes were used in qPCR after performing geNORM qPCR on a range of samples to determine the most stable reference gene. All qPCR results were analysed using two-tailed paired t-test to determine their significance.

##### 3.5.2.1.1 Studying the mRNA expression of WWP2 isoforms using RT-PCR

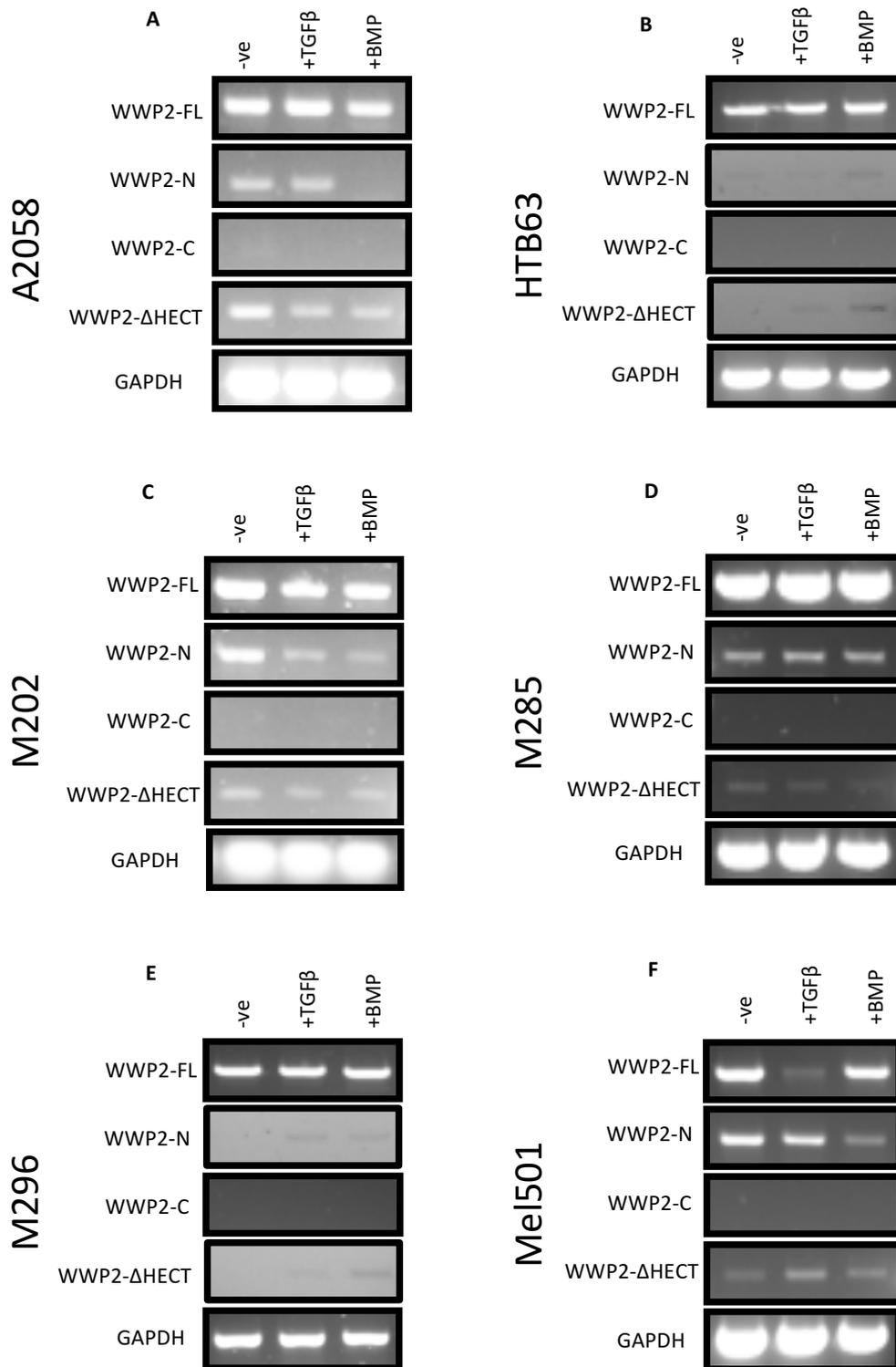
While RT-PCR is a semi-quantitative method, it is a sufficient tool capable for showing the presence or absence of certain isoforms and can provide an approximate idea of the levels of expression in different cell lines. RT-PCRs were originally performed on nine cell lines of which three, showed WWP2-C expression. A375 was chosen as the model cell line due to the high expression of WWP2-C and was used in follow-up experiments including TGF $\beta$  and BMP time-course stimulations and qPCR analysis (fig 4.2.13 and fig 4.2.14). SKMEL28 and M229 also showed C isoform expression, and therefore these were used in TGF $\beta$  and BMP time-course experiments using RT-PCR (fig.4.2.10). Although WWP2-C expression in SKMEL28 on agarose gel images are unclear, there were visible bands seen in both TGF $\beta$  and BMP stimulated cells under UV lights. The WWP2-C expression in M229 cells (shown in fig 4.2.10D) however, seems to be induced by BMP stimulation. Using Image J, bands on the agarose gel were quantified showing a 31% increase in density between the negative control and the sample stimulated

with BMP for 1 hr. Interestingly, M229 cells stimulated by both TGF $\beta$  and BMP showed an increase in  $\Delta$ HECT expression depicted by a 25% and 29% increase in band density respectively, 1 hr post stimulation. Nevertheless, without further experiments, we are unable to predict the effects of the change in isoform levels to their function as it is not yet clear whether the incomplete HECT domain in  $\Delta$ HECT has retained catalytic function. Expression of other isoforms were detected however the level of change in those isoforms are not significant enough to form conclusions using semi-quantitative RT-PCR.



**Figure 3.5.10- RT-PCR analysis of WWP2 isoform expression in SKMEL28 and M229 stimulated with a TGF $\beta$  and BMP time course of 1, 6, and 18 hr. (Experimental replicate n=1)**

Shown in fig 4.2.11 are images of RT-PCRs of the remaining six different melanoma cell lines ran on 1% agarose gels. The levels of WWP2-FL were relatively unaffected by TGF $\beta$  and BMP stimulation in all the melanoma cell lines tested except for in MEL501 (fig 4.2.11f) where TGF $\beta$  stimulation seemed to decrease the expression of FL by 47% in band density when compared to the negative control. The levels of WWP2-N in A2058 (fig 4.2.11a), M202 (fig 4.2.11c) and MEL501 (fig 4.2.11f) all seem to decrease significantly showing a change of 53%, 70% and 37% respectively in band density when stimulated with BMP. Most notable however, is the complete obliteration of N expression in A2058 when stimulated with BMP showing a 53% decrease in band density. In M296 (fig 4.2.11f e), the N isoform was not detected in unstimulated cells however, increase in expression was highlighted by a 45% band density increase with TGF $\beta$  stimulation and 40% increase with BMP stimulation. WWP2-C, an isoform containing only the WW4 recognition domain and the HECT catalytic domain, was predicted to be oncogenic due to its potential in binding and subsequent degradation of the Smad7 inhibitor in TGF $\beta$ -mediated oncogenesis. Interestingly, this isoform was not detected in any of the cell lines shown in fig 4.2.11 suggesting that its expression in different cell lines may hold significance due to its nature as a predicted oncogene. Finally, the presence of WWP2- $\Delta$ HECT was detected in all six of the melanoma cell lines shown in fig 4.2.11 with varying effects of TGF $\beta$  and BMP stimulation on levels of expression. Whilst TGF $\beta$  and BMP seem not to influence  $\Delta$ HECT expression in HTB63 (fig 4.2.11b) and M202 (fig 4.2.11c), the levels of this isoform showed a decrease of 39% in band intensity in A2058 (fig 4.2.11a) when stimulated with TGF $\beta$  and 40% in BMP samples. A 11% decrease was detected in band intensity in TGF $\beta$  stimulated samples whilst BMP resulted in a 22% decrease in M285 (fig 4.2.11d). In MEL501 (fig 4.2.11f) however, TGF $\beta$  seemed to increase the expression of WWP2- $\Delta$ HECT shown by 27% increase in band intensity detected by Image J.

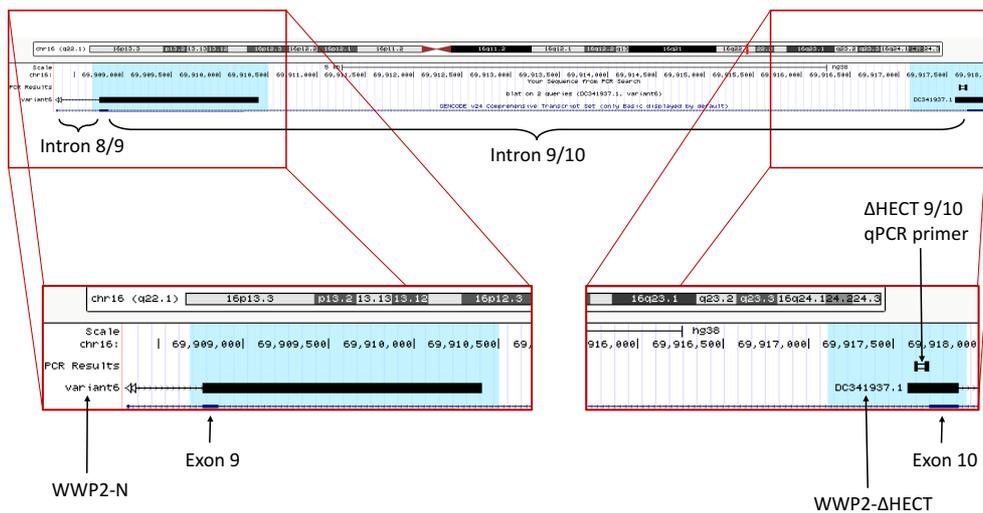


**Figure 3.5.11- The effects of TGFβ and BMP stimulation on mRNA expression of WWP2 isoforms in melanoma cell lines.** Images of RT-PCRs ran on 1% agarose gels showing the expression of WWP2 isoforms when stimulated with TGFβ and BMP compared to non-stimulated cells in A) A2058, B) HTB63, C) M202, D) M285, E) M296 and F) Mel501. Expression was normalized using GAPDH. (Experimental replicate n=1 for all except A2058 WWP2-N where n=2)

#### 3.5.2.1.2 Quantitative mRNA expression study of WWP2 isoforms

Upon review of the RT-PCR results from nine different melanoma cell lines, A375 was chosen for further experiments where cells were subjected to a TGF $\beta$ /BMP stimulation time course, and the mRNA extracted were analysed using both RT-PCR and qPCR for direct comparison. This cell line was chosen as it showed the highest levels of the proposed oncogene WWP2-C expression and also due to the interesting pattern of isoform expression when subjected to TGF $\beta$  and BMP stimulation. Whilst RT-PCR lacks in the quantitative aspect, qPCR although fully quantitative, probes for target sequences irrespective of the product size. It is for this reason that both types of PCR were performed to try and obtain a better understanding of the levels of expression of WWP2 isoforms.

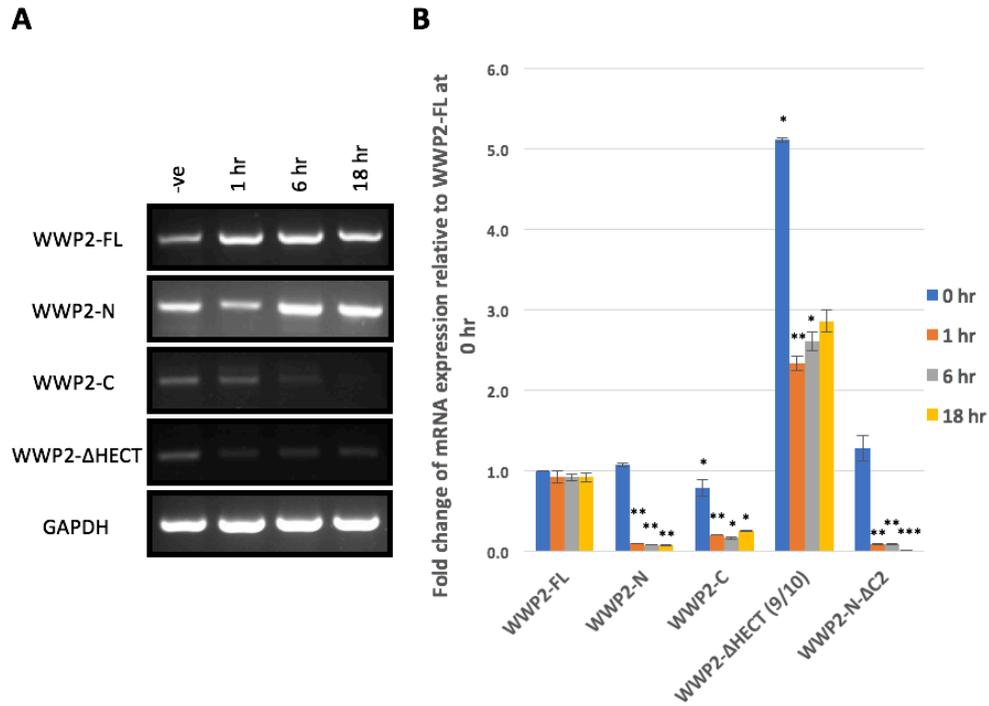
Primers used in qPCR were designed and produced by Primerdesign, and were based on the following accession numbers; WWP2-FL (NM\_007014), WWP2-N (NM\_199423) and WWP2-C (NM\_199424). Probes designed for the novel isoforms were based on predicted transcription boundaries. A set of qPCR primers were designed to detect the  $\Delta$ HECT isoform by recognizing a sequence in the retained intron 9/10. Although the WWP2-N isoform also contains the retained intron 9/10, fig 4.2.12 using the UCSC genome browser shows that the gene sequence of WWP2-N ends more than 15 Kbp before the primer site and the beginning of predicted  $\Delta$ HECT and therefore it is unlikely that the  $\Delta$ HECT 9/10 primer will detect the N isoform. It is also notable that the WWP2-N- $\Delta$ C2 probe may detect levels of FL- $\Delta$ C2 due to the overlap of sequence similarity, however, this is an unlikely problem in A375 cells as the WWP2-FL- $\Delta$ C2 was not detected using RT-PCR. Unfortunately, RT-PCRs using WWP2-N- $\Delta$ C2 specific primers produced smeared bands with a lot of background that were inconclusive, and therefore were not included in the following analyses.



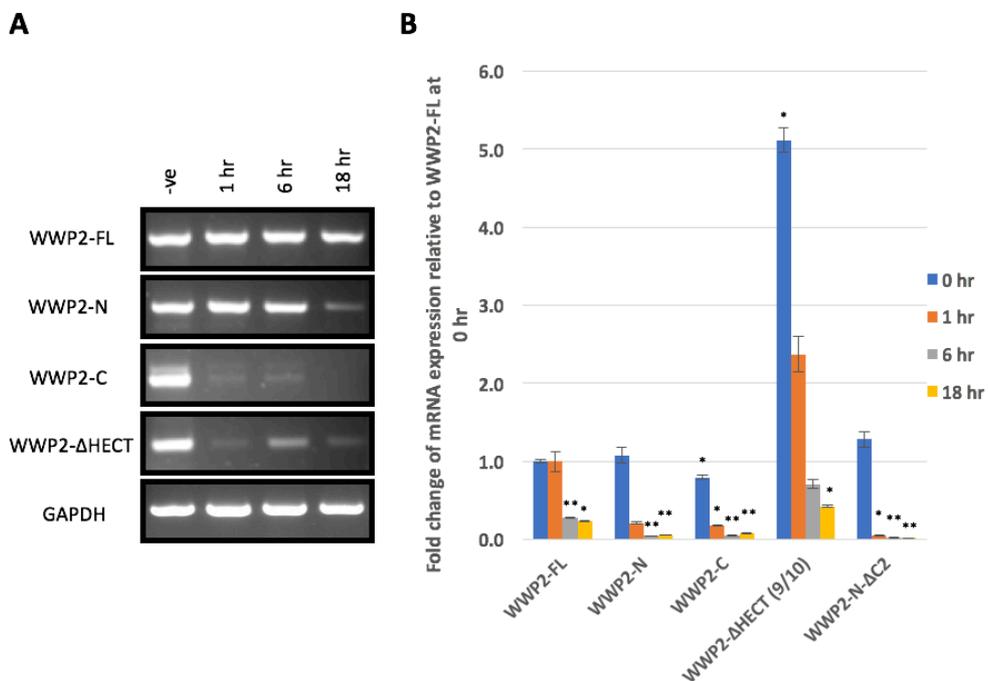
**Figure 3.5.12- A bioinformatic approach to compare WWP2-N and ΔHECT.** Comparing the WWP2-N and ΔHECT transcript to verify that ΔHECT qPCR primer detects only the ΔHECT isoform and not WWP2-N. An annotated screenshot from the UCSC genome browser showing the WWP2-N transcript (Variant 6), the retained intron 9/10 that leads to the transcription of ΔHECT (DC341937.1) and the qPCR primer designed to detect the intronic region of ΔHECT overall suggesting that it is unlikely the primer would detect the N isoform.

Results from fig 4.2.13 were obtained using mRNA extracted from A375 cells that were stimulated with TGFβ for 1, 6 and 18 hrs. The RT-PCR (A) suggests that whilst levels of FL showed little or no difference in expression, WWP2-N showed an increase of 5% in band intensity in response to TGFβ stimulation, and WWP2-C and ΔHECT experienced a decrease in levels of expression reflected by the 50% and 23% respective decreases in band intensity when compared to the unstimulated negative control. Results from the qPCR (B) are presented as fold change relative to expression levels of the WWP2-FL unstimulated control and supports the finding that TGFβ has little effect on the levels of FL expression. It also supports the likelihood that levels of WWP2-C and WWP2-ΔHECT expression decreases dramatically with TGFβ stimulation (WWP2-C TGFβ stimulation 1 hr  $p = \leq 0.01$ , 6 hr  $p = \leq 0.05$ , 18 hr  $p = \leq 0.05$ , WWP2-ΔHECT TGFβ stimulation 1 hr  $p = \leq 0.01$ , 6 hr  $p = \leq 0.05$ ). On the other hand, results from qPCR show a decrease in WWP2-N levels as a result of TGFβ stimulation which is the opposite of the increase detected by RT-PCR. Due to the semi-quantitative nature of RT-PCR which could be a result of limited dNTPs, primers or other essential reaction components, we predict that the qPCR more accurately measures the levels of isoform in

response to TGF $\beta$  stimulation. When directly comparing effects of TGF $\beta$  stimulation, the decrease shown by WWP2-C from 0 hr to 1 hr showed a p-value of  $p = \leq 0.05$ . Interestingly in unstimulated cells, there is 5-fold of  $\Delta$ HECT detected by the  $\Delta$ HECT 9/10 probe compared with FL ( $p = \leq 0.05$ ). Overall, TGF $\beta$  stimulation seemed to negatively regulate all the WWP2 isoforms significantly with the exception of WWP2-FL. Similar findings were shown in fig 4.2.14 where A375 cells were subjected to a BMP time course. The stimulation of BMP decreased expression of all WWP2 isoforms which can be seen in both the RT-PCR (A) and qPCR (B). Semi-quantitative analysis of the results from RT-PCR (fig 4.2.14A) using Image J showed that WWP2-N band intensity decreased by 23% in samples stimulated with TGF $\beta$  for 18 hr, whilst WWP2-C decreased by 16% and WWP2- $\Delta$ HECT showed a negative change of 54%. When comparing levels of isoform expression between unstimulated samples and samples stimulated with BMP for 18 hr, decrease in FL showed a p-value of  $\leq 0.05$ , whilst all other isoforms showed  $p = \leq 0.01$ . The levels of unstimulated  $\Delta$ HECT once again is shown to be higher than FL ( $p = \leq 0.05$ ). Interestingly, the negative effects of BMP are much more pronounced than that of the TGF $\beta$  in FL and  $\Delta$ HECT as increased BMP stimulation caused a direct and proportionate decrease in expression levels with time of stimulation.

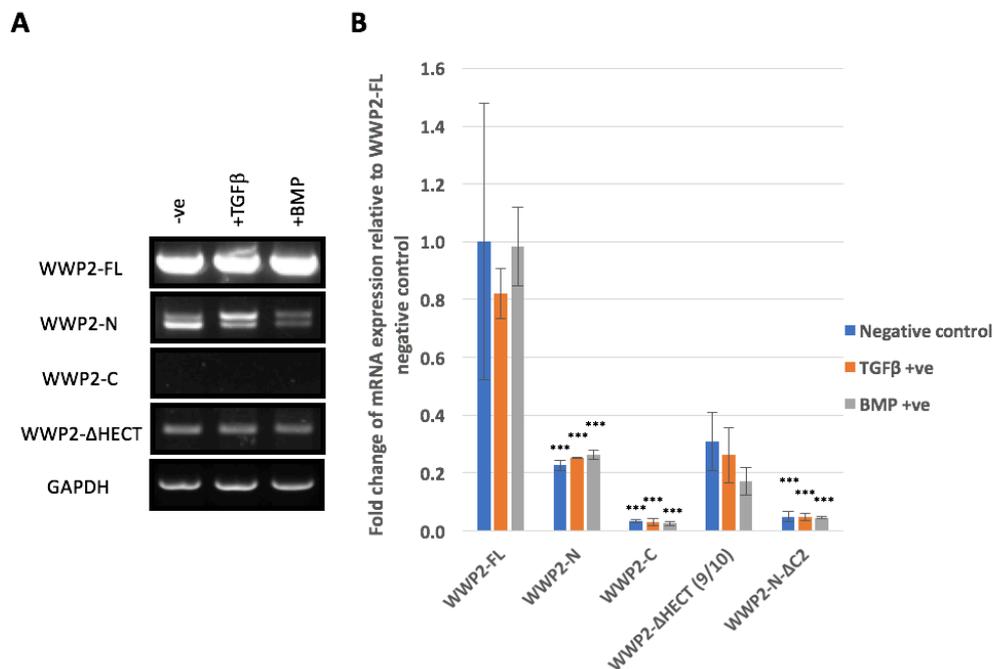


**Figure 3.5.13- The mRNA expression of WWP2 isoforms in A375 cells stimulated with TGF $\beta$  for 1, 6 and 18 hours.** This is shown by (A) RT-PCR normalised to GAPDH (experimental replicate  $n=1$ ) and (B) qPCR which shows the fold change relative to WWP2-FL expression in unstimulated cells (Experimental replicate  $n=2$ . Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)



**Figure 3.5.14- The mRNA expression of WWP2 isoforms in A375 cells stimulated with BMP for 1, 6 and 18 hours.** A RT-PCR (A) (experimental replicate  $n=1$ ) and qPCR (B) analysis of the expression of WWP2 isoforms using mRNA extracted from A375 cells stimulated by BMP for 1, 6 and 8 hrs. RT-PCR was normalized to GAPDH whilst qPCR was optimized to 18S and UBC. (Experimental replicate  $n=2$ . Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

As a non-melanoma, epithelial cell comparison, MCF-7 was used in the TGF $\beta$ /BMP stimulation WWP2 expression study (fig 4.2.15). Cells were stimulated for 18 hr with either TGF $\beta$  or BMP. mRNA was extracted as before and RT-PCR (A) and qPCR (B) were performed using the same protocol as previously mentioned. WWP2-N- $\Delta$ C2 and the proposed oncogenic WWP2-C isoform are unlikely to be expressed in this cell line due to the PCR results shown in figure 4.2.15. All three results from stimulated and non-stimulated samples showed significantly lower levels of the C isoform compared with FL ( $P = \leq 0.001$ ). All levels of truncated isoform expression are lower than the FL basal with no significant effect of TGF $\beta$  or BMP stimulation detected. An observation made in the RT-PCR results from the MCF-7 cell line which was not detected in previous expression studies is the presence of a WWP2-N doublet. It could be suggested that the detection of these double bands is due to a new previously undetected cell line-specific isoform or simply an artefact due to experimental variations.



**Figure 3.5.15- The effects of TGF $\beta$  and BMP on WWP2 mRNA expression in MCF-7 cells.** The mRNA extracted from cells stimulated with TGF $\beta$  and BMP were used to determine WWP2 isoform expression in A) RT-PCR (experimental replicate n=1) and B) qPCR. Results from the RT-PCR were normalized to GAPDH whilst data from qPCR was calculated as fold change using WWP2-FL expression in non-stimulated cells as baseline. (Experimental replicate n=2. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

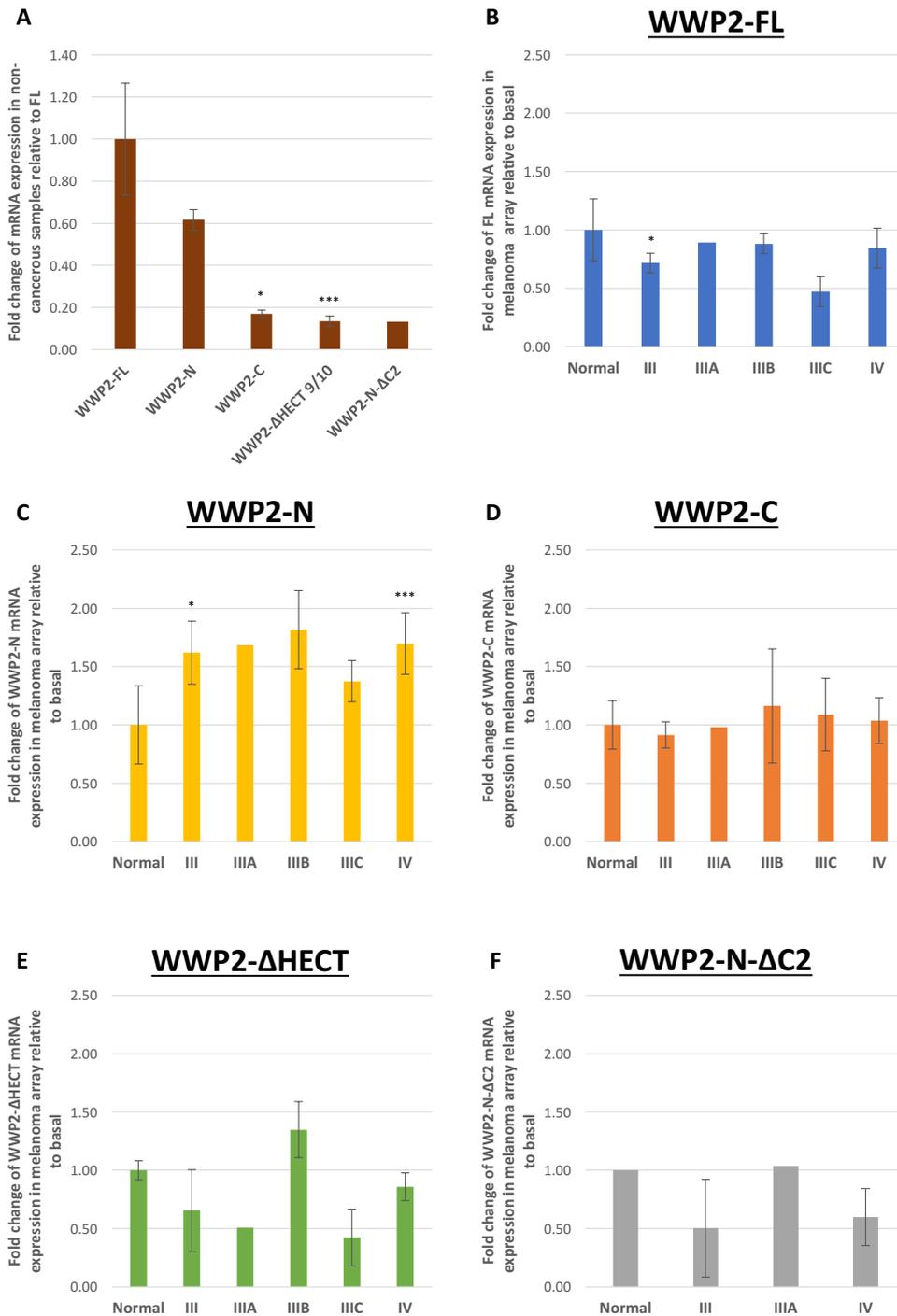
### 3.5.2.2 Expression of WWP2 isoforms in melanoma samples

The increased expression levels of various WWP2 isoforms in melanoma highlighted in established research has led to the hypothesis that this ubiquitin ligase may be involved with TGF $\beta$ -dependent oncogenesis in melanoma (Soond *et al*, 2013). In the previous section, the effects of TGF $\beta$  and BMP on expression of WWP2 isoforms was investigated in multiple melanoma cell lines. Due to the proposed oncogenic role TGF $\beta$  has in cancer, it can be suggested that TGF $\beta$  stimulation in our studies mimics the role it plays in favouring oncogenesis. However, the use of melanoma cell lines does not provide information about WWP2 expression pattern during different stages of cancer progression. It is for this reason that results in this section focus on the expression of WWP2 isoforms in different disease stages of melanoma. To pinpoint the levels of WWP2 isoforms in different melanoma stages could potentially identify isoforms as oncogenes, tumour suppressors or potential prognostic markers. Specific WWP2 Taqman primers previously used in qPCRs are used here in conjunction with TissueScan melanoma cDNA arrays supplied by Origene (MERT502). Each array contains 3 cDNA samples from normal healthy tissue, 9 from Stage III melanoma, 1 from stage IIIA, 3 from IIIB, 2 from IIIC and 22 samples from stage IV patients. All five WWP2 isoforms, FL, N, C,  $\Delta$ HECT, and N- $\Delta$ C2 were tested with results normalized to  $\beta$ -actin levels.

Firstly, expression of different isoforms were compared in normal healthy tissue samples. Fold changes shown in fig 4.2.16A were calculated relative to levels of WWP2-FL and shows that in normal tissue, WWP2-FL expression is the highest of all the isoforms. The levels of WWP2-N was the second highest at 0.4 fold less than FL expression and WWP2-C ( $p \leq 0.05$ ),  $\Delta$ HECT ( $p \leq 0.001$ ) and N- $\Delta$ C2 levels were all very low in comparison to FL in healthy tissues. In fig 4.2.16B, the levels of WWP2-FL expression is compared across different melanoma stages with results shown relative to FL levels in normal tissue samples. The expression of FL decreases in melanoma with the largest decrease seen in stage IIIC samples where levels of FL dropped by more than half. The most significant decrease however, was in stage III where  $p \leq 0.05$ .

The expression of WWP2-N is shown in (C) where there is an apparent increase of isoform levels in all stages of melanoma by an average of 1.6 fold compared with healthy tissue samples (the most significant being in stage III and IV where  $p \leq 0.05$  and  $\leq 0.001$  respectively). Fig 4.2.16D shows the expression of WWP2-C in stages of melanoma relative to healthy samples. Interestingly, there was no significant change in the levels of this predicted oncogene which could reflect its strong oncogenic properties that might be independent of melanoma aggressiveness.

The expression of WWP2- $\Delta$ HECT is shown in fig 4.2.16E. The levels of this isoform seem to be lower in diseased samples than normal healthy tissue in all melanoma stages except for IIIB where the levels of  $\Delta$ HECT is 1.3 fold higher than in normal cells. Lastly, the results shown in fig 4.2.16F are representative of the expression of WWP2-N- $\Delta$ C2 in all melanoma stages. This isoform was undetected in all 3 stage IIIB samples, and 2 stage IIIC samples. In addition to this, 2/3 normal samples had levels of N- $\Delta$ C2 which were too low to detect, 6/9 from stage III and 8/22 from stage IV were also undetected. The remaining results show that levels of this isoform is lower in stage III and IV compared to normal samples and only slightly increased in the stage IIIA sample.



**Figure 3.5.16- The mRNA expression of WWP2 isoforms in melanoma tissue samples using TissueScan cDNA arrays purchased from Origene.** This array included 3 normal cDNA samples, 9 stage II melanoma patient samples, 1 stage III samples, 3 stage IIIB samples, 2 IIIC samples and 22 Stage IV melanoma patient samples. A) show the fold change in expression of WWP2 isoforms relative to FL levels in normal healthy tissue samples whilst B, C, D, E, and F shows the expression of WWP2-FL, N, C, WWP2-ΔHECT and WWP2- N-ΔC2, respectively in different melanoma disease stages relative to normal healthy tissue. Note that the lack of error bar in IIIA samples is due to there being only one IIIA sample in this array. The lack of error bar on WWP2-N-ΔC2 in (A) and normal bar in (F) was due to there being only one viable reading. The levels of WWP2-N-ΔC2 were undetected in 2/3 normal samples. (Experimental replicate n=1. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

### 3.5.2.3 Exploring the EMT potential of melanoma cell lines and the expression of WWP2 isoforms in epithelial/ mesenchymal cells

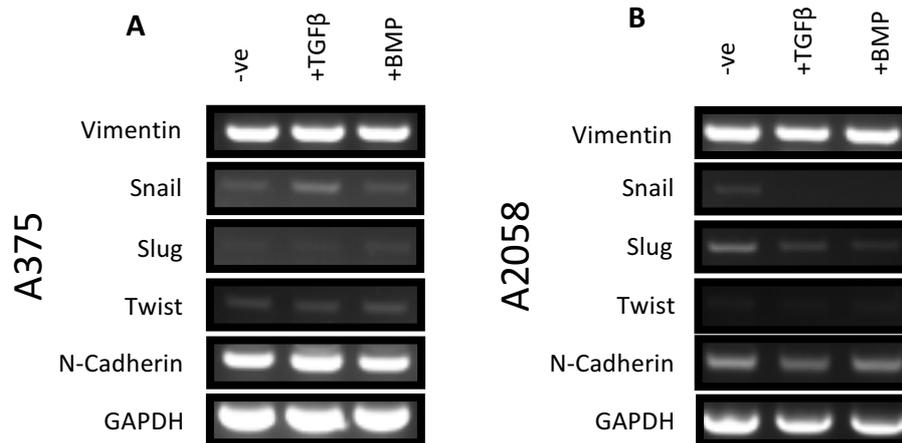
EMT is the process of epithelial to mesenchymal transition which allows cancer cells to escape their primary location and metastasize to a secondary location in the body. It is also an event that is often linked to TGF $\beta$ . This is a crucial process in the progression of cancer and can be determined by the change in levels of several EMT markers such as Vimentin, SNAIL, SLUG and Twist. The expression of these EMT markers in A375 and A2058 melanoma cell lines were profiled to not only further our understanding of the oncogenic potential these cell lines possess, but also to confirm the regulation of EMT by TGF $\beta$  and BMP. Furthermore, by looking at the difference in WWP2 isoform expression in epithelial and mesenchymal cell lines, we could be able to predict the roles of these isoforms in EMT and subsequently oncogenesis.

#### 3.5.2.3.1 Expression of EMT markers in melanoma cell lines

The six major EMT markers that are investigated here are Vimentin, SNAIL, SLUG, Twist and N-Cadherin. According to literature, a major change that happens during EMT is the increase of N-cadherin and decrease in E-cadherin, mediated by the increase in E-cadherin repressors SNAIL and SLUG (Kim *et al*, 2013). Another E-cadherin repressor that is involved in EMT is Twist which has been shown to increase in expression during this oncogenic process (Wendt *et al*, 2012). A key marker of EMT is Vimentin which has been suggested to contribute to the shape of mesenchymal cells and the increase in motility of cells undergoing EMT (Mendez *et al*, 2010). In this section, the expression of EMT markers were studied to identify the EMT potential of A375 and A2058, melanoma cell lines used in previous experiments. Furthermore, the effects of TGF $\beta$  and BMP on EMT marker expression levels was also investigated to explore the regulatory roles these growth factors have on EMT within these melanoma cell lines.

Over the years, significant evidence has shown that TGF $\beta$  and other members of its family can have oncogenic properties by inducing EMT

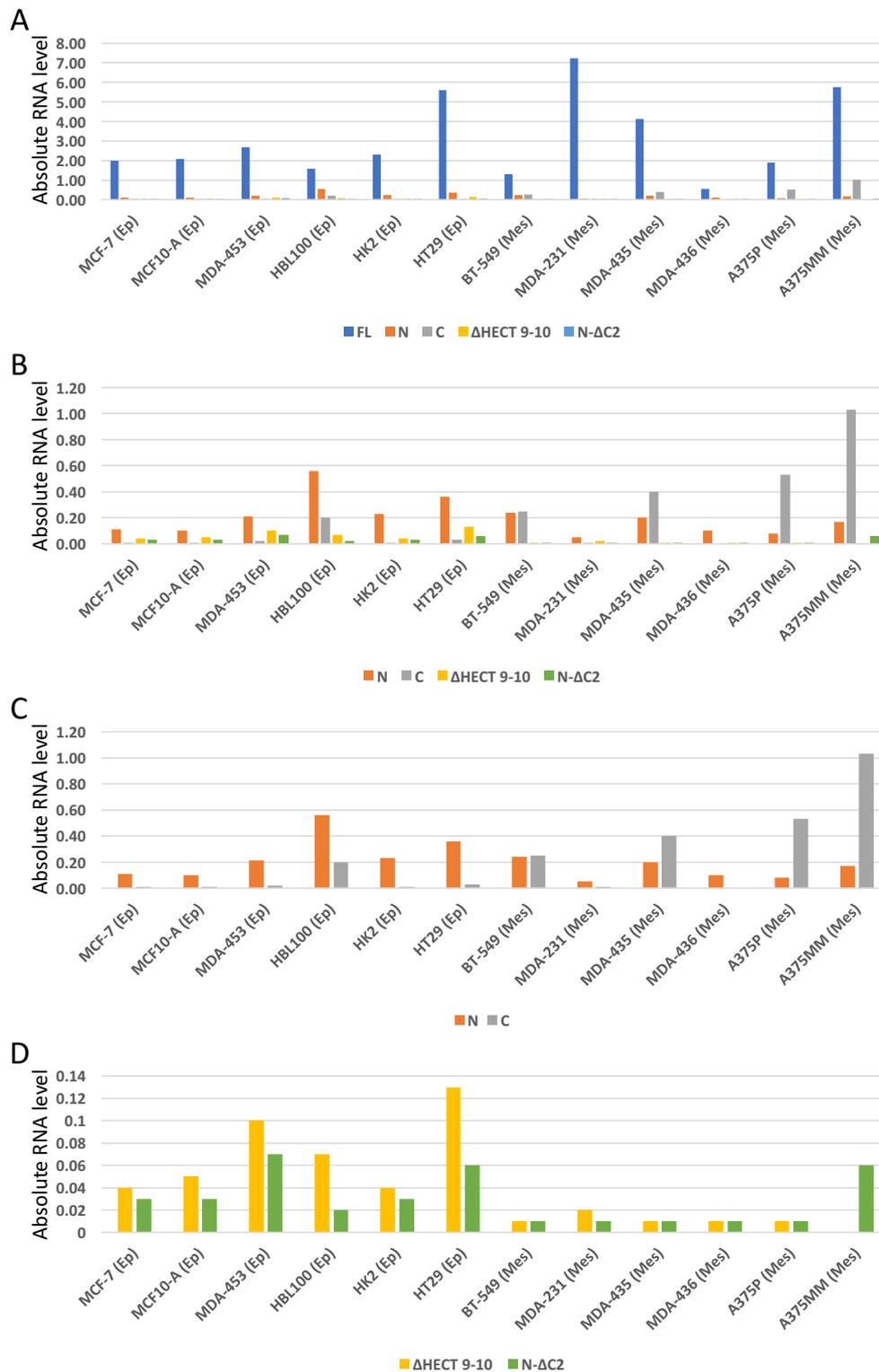
(Wendt *et al*, 2012). Therefore, A375 and A2058 melanoma cells lines were stimulated with TGF $\beta$  and BMP to mimic the effects of EMT on EMT markers to examine their regulatory role in this oncogenic process. These cells were then harvested using the Promega SV total RNA isolation system and cDNA was made using GoScript Reverse transcriptase PCR. Using specific primers, GoTaq PCR was performed to detect levels of EMT markers in A375, and A2058 stimulated with TGF $\beta$  and BMP individually. Results from RT-PCR using agarose gel electrophoresis are shown in fig 4.2.17. Note that results are not fully quantitative and must only be used as a rough guide of expression levels. Shown across both melanoma cell lines, expression of Vimentin seems to be unaffected by TGF $\beta$  and BMP stimulation. The growth factors seem to also have no effect on levels of N-Cadherin in all cell lines. In A375, TGF $\beta$  stimulation increased SNAIL expression reflected by a 33% increase in band intensity, whilst in A2058, the levels of SNAIL were inhibited by TGF $\beta$  and BMP by more than 95% band intensity. Expression of SLUG was unaffected by growth factor stimulation in A375, however, TGF $\beta$  and BMP stimulation had a negative effect on expression in A2058 show by the decrease in band intensity of 37% and 45%, respectively. Results from RT-PCR to detect Twist expression in these melanoma cell lines suggest that TGF $\beta$  and BMP stimulation had little or no effect on the levels of Twist detected. Overall, the changes in levels of EMT marker expression observed suggests that the melanoma cell lines A375 and A2058 could undergo various EMT changes such as change in morphology and cell-cell interaction characteristics, in the presence of TGF $\beta$  and BMP stimulation.



**Figure 3.5.17- The mRNA expression of EMT markers in melanoma cell lines.** RT-PCR ran on 1% agarose electrophoresis showing the expression of 6 different EMT marker in RNA extracted from a) A375 and b) A2058 under 18 hr TGF $\beta$  and BMP stimulation. (Experimental replicate n=1)

#### 4.2.2.2 WWP2 isoform expression in epithelial and mesenchymal cells

In a close collaboration with the Nieto lab at the Institute of Neuroscience Alicante, a panel of twelve epithelial and mesenchymal cell lines were probed using qPCR analysis to help identify any patterns of isoform expression. Results shown in fig 4.2.18 suggests that FL is consistently the highest expressed isoform across all twelve cell lines regardless of whether they are of an epithelial or mesenchymal origin. Overall, the levels of WWP2 N and C were higher than the expression of  $\Delta$ HECT and N- $\Delta$ C2 which can be seen in fig 4.2.18B. Interestingly, as shown in fig 4.2.18C, a direct comparison between the N and C isoforms, expression of N is higher in epithelial cells whilst expression of the C isoform is higher in mesenchymal cells. When comparing the expression of  $\Delta$ HECT and N- $\Delta$ C2 in these cell lines, there is a notably higher expression of both novel isoforms in epithelial cells as opposed to mesenchymal cells with the exception of A375MM.



**Figure 3.5.18- qPCR analysis of WWP2 isoform expression in cancerous and non-cancerous cell lines to investigate the expression in epithelial and mesenchymal cells.** A) shows the RNA expression levels of all WWP2 isoforms whilst B) excludes the FL isoform allowing focus on the lower levels of WWP2-N, C, ΔHECT and N-ΔC2. C) highlights the expression of WWP2 N and C whilst D) compares the RNA expression of WWP2-ΔHECT and N-ΔC2. Overall, results suggest that WWP2-N and ΔHECT are expressed at higher levels in epithelial cells whilst WWP2-C is expressed at a higher level in mesenchymal cells. Note that these results were generated by the Nieto lab as part of a collaboration project. (Experimental replicate n=6)

### 4.3 Discussion

The main focus in the chapter was the prediction, validation and detection of WWP2 novel isoforms. Firstly, the novel WWP2- $\Delta$ HECT and N- $\Delta$ C2 isoforms were predicted and identified. Then specific PCR primers were designed to allow the mRNA expression levels of these novel and established isoforms to be explored in a range of melanoma cell lines and patient cDNA using different PCR techniques. Additionally, the expression of these isoforms was investigated in a panel of epithelial and mesenchymal cell lines whilst RT-PCRs were performed to probe the expression of EMT markers thereby confirming the ability of TGF $\beta$  to induce EMT in melanoma cells. The overall aim of this chapter was to validate the expression of novel WWP2- $\Delta$ HECT and N- $\Delta$ C2 isoforms whilst exploring the role TGF $\beta$  and BMP has on the expression of WWP2 existing isoforms in multiple cancerous and non-cancerous cell lines, and in the context of EMT.

#### 4.3.1 Identification and validation of novel WWP2 isoforms

Currently, there are three WWP2 isoforms established in literature: WWP2-FL, WWP2-N and WWP2-C. Through an initial Western blot investigating the protein expression of these isoforms, bands of smaller molecular weights appeared suggesting expression of other truncated WWP2 isoforms. A bioinformatics approach was then adopted and the UCSC genome browser was used to search for possible transcription of WWP2 that resulted from retained intronic sequences, or alternative promoters. Using this method, combined with the knowledge gained from original western blot, two novel WWP2 isoforms were identified that matched the molecular weight, and contained the target regions of binding which corresponded to the antibodies used. WWP2- $\Delta$ HECT is a 38.3 kDa isoform that spans exon 10-19 and is a result of a putative promoter present in intron 9/10 and a stop codon in retained intron 19/20 which indicates the presence of WW3, WW4 and an incomplete HECT catalytic domain in this truncated isoform. The second novel isoform is the WWP2-N- $\Delta$ C2 which is a 22.8 kDa protein containing only the WW1 domain. The transcription of this isoform is

predicted to begin at exon 5 and end at exon 9 resulting from a putative promoter at intron 3/4 and a stop codon in intron 9/10. Finally, the FL- $\Delta$ C2 isoform was highlighted in a recent paper by Choi *et al* (Choi *et al*, 2015) with transcription also beginning at exon 5 but ending at exon 24, the end of the FL isoform. In the study by Choi *et al*, the expression of FL- $\Delta$ C2 was hypothesized by the presence of a band at the correct molecular weight in western blotting protein analysis. To eliminate the possibility that smaller molecular weight bands seen in the original western blot using WWP2-N and C terminal antibodies were degraded fragments of existing WWP2 isoforms rather than the predicted novel isoforms, and to verify the expression of FL- $\Delta$ C2, the expression of novel isoforms was confirmed by the performing a range of expression studies. However, due to the unique combinations of domains these isoforms possess, it is difficult to predict accurately the binding and function of these isoforms therefore we have conducted a string of experiments in attempts to further characterize WWP2 isoforms.

#### 4.3.2 Characterizing expression of WWP2 isoforms

To explore the mRNA expression of WWP2 isoforms, RT-PCR was used as a semi-quantitative measure alongside qPCR allowing detection of amplification at the correct size. To get a more comprehensive idea of WWP2 expression patterns, cDNA from epithelial, mesenchymal and melanoma cell lines were used as well as patient melanoma cDNA array. TGF $\beta$  and BMP were used to stimulate cells used for PCR probing techniques to elucidate the effects of growth factors on WWP2 isoform expression.

##### 4.3.2.1 mRNA expression of WWP2 isoforms in melanoma cell lines

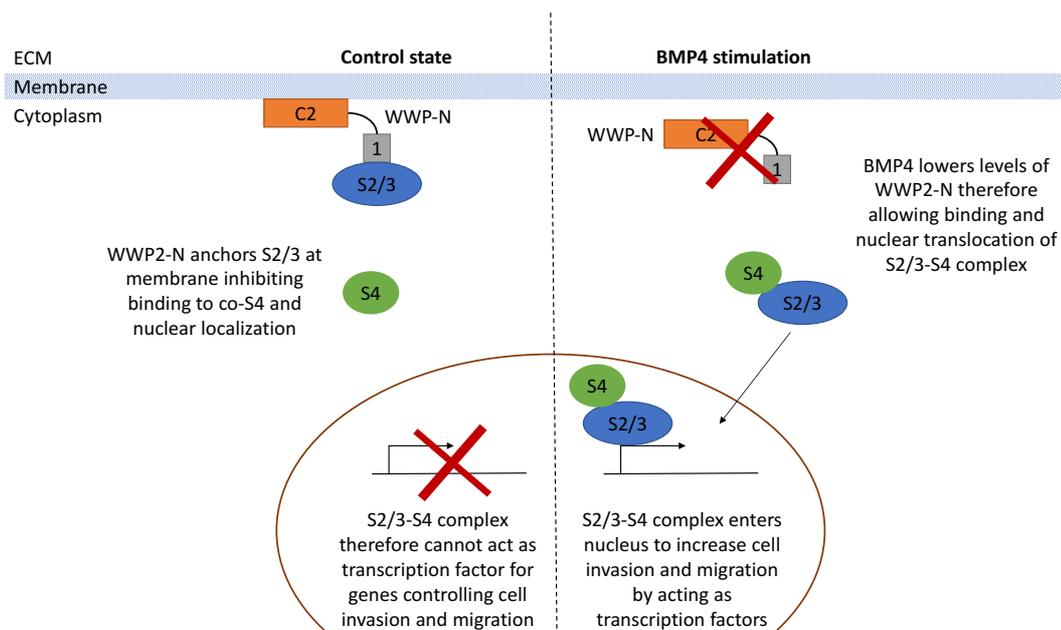
Investigating the expression levels of the WWP2 isoforms using the techniques explained previously has produced many interesting findings. From the mRNA expression analyses of melanoma and breast cancer cell lines, only 3 out of the 9 melanomas, A375, M229 and SKMEL28 expressed the WWP2-C isoform. Although the expression of this oncogene is mostly very low compared to the basal expression of FL in qPCRs, the presence is undeniable from the evidence shown in the RT-PCRs. Intriguingly, after

studying the characteristics of cell lines used in these experiments, it was found that all 3 cell lines expressing WWP2-C shared the presence of a homozygous BRAF<sup>V600E</sup> mutation. As shown in table 4.3-1, other cell lines that did not express the C-isoform either had a wild type or heterozygous BRAF<sup>V600E</sup> genotype of the kinase. BRAF is a kinase involved in the MAPK signalling cascade which is mostly stimulated by EGF and ultimately signals for cellular process such as proliferation. In the BRAF<sup>V600E</sup> mutant, a nucleotide mutation at position 1799 from a thymine to an adenine causes the coding of a glutamic acid in place of a valine. This mutation causes constitutive activation of BRAF<sup>V600E</sup> in the absence of ligand binding, leading to the increased cell proliferation and oncogenesis in melanoma (Ascierto *et al*, 2012). It has also been suggested that this BRAF mutant induces senescence and other oncogenic processes such as metastasis and contractility by affecting normal function of proteins e.g. PDE5A and LKB1 (Maurer *et al*, 2011). The presence of this BRAF mutation in all cell lines expressing WWP2-C may suggest a potential cause or provide an explanation to the regulation of this oncogenic isoform. Previously, BRAF<sup>V600E</sup> has been shown to increase both TGFβ secretion and Smad-dependent transcription in thyroid cancer (Riesco-Eizaguirre *et al*, 2009). This could suggest that WWP2-C is a product of BRAF<sup>V600E</sup>-dependent increase in TGFβ. However, in qPCR studies, the expression of this isoform decreased with TGFβ stimulation in A375 which argues against the above hypothesis.

*Table 4.3-1- BRAF status of Melanoma cell lines used in WWP2 expression study*

<b>Cell Line</b>	<b>BRAF Status</b>
A2058	V600E
A375	HOMO V600E
HTB63	V600E
M202	WT
M229	HOMO V600E
M285	WT
M296	WT
Mel501	WT
SKMEL28	HOMO V600E

Another intriguing observation found in RT-PCR analysis is that BMP4 stimulation of A2058 cells results in the complete obliteration of WWP2-N expression. In addition, BMP4 stimulation was also shown to decrease expression of WWP2-N in M202 and Mel501. As mentioned before, studies have shown the involvement of BMP4 in the increase of cell invasion and migration in oncogenesis (Rothhammer *et al*, 2005). This supports the hypothesis that WWP2-N is a tumour-suppressor as BMP4 stimulation inhibits expression of WWP2-N to subsequently aid tumour development. We postulate that the decreased levels of WWP2-N results in decreased levels of Smad2/3 retention at the membrane via binding with the WW1 domain of WWP2-N therefore allowing Smad2/3 nuclear translocation and ultimately affecting transcription of genes that alter cell invasion and migration. This is depicted in a simple illustration shown in fig 4.3.1. Together with further analysis of the results in the following functional and localization studies, we hope to clarify the role of BMP4 has in melanoma development.



**Figure 4.3.1- A simple diagram to illustrate effects of BMP4 on cell invasion and migration.** In the control state without BMP stimulation, WWP2-N is able to bind to Smad2/3 shown in blue via its WW1 domain (highlighted in grey) and subsequently act as an anchor to the membrane using its C2 lipid binding domain (shown in orange). Smad2/3 is therefore inhibited from binding with Co-smad4 (green) and nuclear translocation leading to the inability of Smad complex to act as a tissue factor and aid oncogenesis. Under BMP4 stimulation, levels of WWP2-N decrease significantly allowing Smad2/3 to form a complex with Co-smad4 and therefore localize into the nucleus where they act as transcription factors to aid cell invasion and migration.

Another isoform that was not found in most of the cell lines tested was WWP2-FL- $\Delta$ C2. Here, the expression of WWP2-FL-C2 was only identified in MCF-7 and potentially M229 suggesting that this isoform may have preferential expression patterns in different cell types. Interestingly, compared to A375 which is an epithelial cell line with strong mesenchymal properties, MCF-7 is an epithelial cell line with strictly epithelial properties. This hints that the expression of WWP2-FL- $\Delta$ C2 isoform may be favoured in epithelial cell lines before the occurrence of EMT.

In the majority of cell lines tested using RT-PCR and qPCR, WWP2-FL remained the isoform with the highest expression when compared to other truncated versions. However, qPCR study of A375 detected levels of  $\Delta$ HECT using the 9/10 probe that were 5-fold higher than the full length equivalent in unstimulated samples. Due to the lack of knowledge in the regulation of this novel isoform, it is still unclear what causes the increased promoter activity seen in A375, however, the most logical explanation might be that

TGF $\beta$  mediator Smads might act as transcriptional activators or repressors to WWP2 gene regulators, or the WWP2 gene itself.

Results from qPCR analysis suggests that the effects of growth factor stimulation are similar within cell lines. For example, the levels of WWP2-FL, N, C,  $\Delta$ HECT and N- $\Delta$ C2 decreased in response to TGF $\beta$  stimulation in A375. Whilst the effects of TGF $\beta$  and BMP was noticeable in melanoma cell lines, it is obvious that the effect were not so clear in MCF-7. It is suggested that TGF $\beta$  and WWP2 both control cell proliferation in MCF-7 (Mazars *et al*, 1995) (Jung *et al*, 2014), however, the levels of WWP2 expression in this cell line is not affected by the levels of TGF $\beta$  stimulation. The lack of response in expression levels of WWP2 from growth factor stimulation is not unusual as it was shown that TGF $\beta$  stimulation does not affect the levels of WWP2 in chondrocytes (Tardif *et al*, 2013).

4.3.2.2 mRNA expression of WWP2 isoforms in melanoma tissue cDNA array  
Combining the results from WWP2 expression levels in response to TGF $\beta$  stimulation in melanoma cell lines, with results on expression levels in patient cDNA samples allows important links between three key factors to be made; the effects of TGF $\beta$ , the progression of melanoma, and the role WWP2 isoforms have in this disease. Although results from this section do not include stage I or II patient samples, stage III represents the crucial point in melanoma where cancerous cells metastasize to the lymphatic system (www.bad.org.uk, 2013) which is not possible without EMT, a process heavily linked to TGF $\beta$ .

In normal non-cancerous cDNA samples, WWP2-FL and N were expressed at a higher level than all other isoforms. These two isoforms have been suggested to form a complex in the absence of TGF $\beta$  and auto-activate leading to the degradation of Smad-3 (Soond and Chantry, 2011). This would support the higher levels of expression of FL and N in healthy cells where these isoforms are active as tumour-suppressors. Furthermore, a decrease in the levels of FL and N were also seen with dose-dependent TGF $\beta$

stimulation in previous qPCR results for 2 out of 3 melanoma cell lines suggesting that TGF $\beta$  stimulation might correlate with melanoma progression.

The results from cDNA array showed that WWP2-FL in melanoma samples of stage III to IV showed a general decrease in expression when compared with normal healthy cDNA. In stage IV melanoma, cancer metastasizes from primary tumour and nearby lymph nodes to form a secondary metastasis in the body ([www.cancerresearchuk.org](http://www.cancerresearchuk.org), 2015b) which might explain the low levels of FL most notably in stage IIIC. This suggests that the FL acts as a tumour suppressor which is expressed at its lowest level allowing the increase of TGF $\beta$  signalling in preparation of high levels of EMT in the metastasis of cancerous cells in stage IV to secondary locations.

WWP2-N has previously been suggested to act as a tumour-suppressor due to the negative effects it has on levels of the EMT marker vimentin in presence of TGF $\beta$  (Soond *et al*, 2013). The general increase of WWP2-N transcript levels in melanoma compared with non-cancerous cDNA suggests that this may be part of the body's natural defence against melanoma triggered by the progression of this disease. The dip in expression levels of N seen in stage IIIC mimics that of FL expression and might also be a result of increased TGF $\beta$  signalling allowing increased EMT in preparation for secondary metastasis in stage IV melanoma.

Another significant isoform investigated in these experiments is WWP2-C. The notion that this isoform may act as an oncogene derives from the fact that it contains the intact HECT catalytic domain and the WW4 domain allowing it to bind to and facilitate the degradation of natural inhibitory Smad7. The increase in levels of WWP2-C seen in fig 4.2.16D from stage IIIB onwards might suggest a key point in the increase of aggressiveness of the melanoma. The difference between stage IIIA and IIIB is that cases from the latter may involve ulcerations, enlarged lymph nodes or metastasis of tumour to skin area surrounding the primary site

([www.cancerresearchuk.org](http://www.cancerresearchuk.org), 2015a). The increase in WWP2-C levels in stage IIIB could therefore mask the inhibitory effects of smad-7 allowing TGF $\beta$  signalling leading to increased EMT and ultimately formation of ulcerations and metastasis.

WWP2- $\Delta$ HECT was one of the novel isoforms investigated in these experiments. The presence of the WW4 domain suggests that it binds to Smad7 (Soond and Chantry, 2011), however, due to an incomplete HECT domain, it is not possible to pinpoint its exact function. Therefore,  $\Delta$ HECT could either act as a tumour suppressor or an oncogene. Although this is the case, the results from melanoma cDNA array highlights an increase in  $\Delta$ HECT levels at stage IIIB which was also shown in WWP2-C expression, and a decrease of more than half at stage IIIC, mimicking expression of FL and N. The presence of characteristics akin to FL, N and C suggests the possibility that  $\Delta$ HECT may have a dual function in melanoma acting as both an oncogene and a tumour-suppressor interchangeably which could be controlled by levels of TGF $\beta$  signalling or other similar factors. Lastly, WWP2-N- $\Delta$ C2 was investigated using this same melanoma cDNA array. This isoform contains only the WW1 domain which theoretically allows it to bind to and stabilize levels of Smad-2/3 therefore potentially acting as an oncogene by enhancing TGF $\beta$  signalling and its oncogenic effects such as EMT. The levels of this isoform were too low to be detected in many samples including all IIIB, IIIC and the majority of normal samples. The remaining results, showed high variability as demonstrated by the standard error bars in fig 4.2.16F. The lack of results unfortunately allows no real trend to be predicted.

#### 4.3.2.3 The EMT potential of WWP2 isoforms and melanoma cell lines

TGF $\beta$  has long been associated with the induction of EMT, a process crucial to the progression of cancer. Using specific EMT marker RT-PCR primers, the levels of RNA expression of Vimentin, SNAIL, SLUG, N-cadherin and Twist in A375 and A2058 melanoma cell lines were explored. Effects of TGF $\beta$  and BMP used to mimic cancer progression was also investigated by performing 16 hr stimulations on cells prior to RNA extraction. Results showed that in

A375, TGF $\beta$  and BMP stimulation either had positive or no effect on the expression of EMT markers. Due to previous studies on EMT markers showing SLUG, SNAIL, Twist, Vimentin and N-Cadherin as proto-oncogenes that increase in expression during oncogenesis, it is reasonable to surmise that TGF $\beta$  and BMP stimulation in our experiments successfully mimicked effects of EMT and cancer (Kim *et al*, 2013; Mendez *et al*, 2010; Wendt *et al*, 2012).

A2058 on the other hand, showed a decrease in EMT markers following TGF $\beta$  and BMP stimulation. Most notably, SNAIL and SLUG expression were both negatively impacted by these growth factors which interestingly might be due to a homozygous PTEN deletion present in A2058 but not A375. PTEN is a tumour suppressor that has been shown to be a key player in regulation of tumorigenesis in melanoma (Stahl *et al*, 2003). The loss of PTEN therefore naturally contributes to the highly invasive properties of A2058. However, the response to TGF $\beta$  and BMP stimulation seen in A2058 was the decrease in both SLUG and SNAIL EMT markers suggesting a reflected decrease in oncogenic potential. This may therefore indicate a tumour suppressive role in TGF $\beta$  that have previously been described as the dual role of TGF $\beta$  in cancer. Although literature has explored the negative effects of TGF $\beta$  and SNAIL/ SLUG EMT markers on PTEN in cancer and other diseases, there has been no evidence to offer an explanation on the negative effect TGF $\beta$  and BMP has on the expression of SNAIL and SLUG in the presence of a PTEN deletion (Chow *et al*, 2008; Lauzier *et al*, 2016), therefore more work is needed in this area to further understand the role of TGF $\beta$  in oncogenesis in regards to expression of different EMT markers. All in all, results obtained from EMT marker expression studies confirm the ability of TGF $\beta$  and BMP to induce EMT in melanoma cell lines.

To further study the role of WWP2 in EMT, the expression of WWP2 isoforms were probed in a panel of cancer and normal cell lines with epithelial and mesenchymal characteristics. Results from qPCR conducted on six epithelial and six mesenchymal cell lines showed that the expression of WWP2-N and

$\Delta$ HECT was higher in epithelial cell lines suggesting that these WWP2 isoforms may act as guardians for the epithelial phenotype. The reasoning for this is because N binds to and localizes Smad2/3 to the membrane via its C2 lipid binding domain (shown in cell localization studies in section 5.2.4.1) effectively disrupting its nuclear localization and therefore inhibiting TGF $\beta$  signals for EMT. We propose that the increased expression of the catalytically inactive WWP2- $\Delta$ HECT in epithelial cells allows the binding and subsequent stabilization of Smad7 which inhibits phosphorylation of Smad2/3 and ultimately negatively regulates TGF $\beta$ -dependent EMT. The negative regulation of these isoforms on TGF $\beta$ -dependent EMT suggests that they work in guarding the epithelial phenotype. On the other hand, a higher expression of WWP2-C was detected in mesenchymal cell lines. Keeping in mind that the WWP2-C is a catalytically active isoform with preferential binding to Smad7, it can be hypothesized that the increase in C expression suggests its role in being the mesenchymal phenotype guardian by lowering levels of the natural TGF $\beta$  signal inhibitor and in turn supporting TGF $\beta$ -dependent EMT. Consequently, we postulate that WWP2-C has an oncogenic role by preventing MET, the reversed process of EMT, whilst WWP2-N and  $\Delta$ HECT have tumour-suppressant properties by acting as guardians of the epithelial phenotype.

Overall in this chapter, the expression of three novel isoforms WWP2- $\Delta$ HECT, N- $\Delta$ C2 and FL- $\Delta$ C2 has been validated in different cancerous and non-cancerous cell lines. TGF $\beta$  and BMP has also been used to stimulate cells allowing the investigation of their effects on the expression of novel and established WWP2 isoforms. The results from these expression studies might help the identification of WWP2 isoforms as potential oncogenes or tumour-suppressors. Furthermore, the V600E BRAF mutation has also been suggested to play a role in WWP2-C expression whilst the WWP2-N and  $\Delta$ HECT novel isoforms have been proposed as guardians of the epithelial phenotype.

Chapter 5 Deciphering the regulation and  
functional roles of WWP2 isoforms

## 5.1 Introduction

WWP2 ubiquitin ligase-mediated degradation of Smads is a vital regulatory mechanism of the TGF $\beta$  pathway. Due to the established effects TGF $\beta$  has on cancer development, it can therefore be suggested that WWP2 also has an important role in oncogenesis. Following expression studies of novel and established WWP2 isoforms in the previous chapter, the next logical step is the study of transcription mechanisms involved in generating WWP2 truncated isoforms, and to investigate other aspects such as function and cellular sub-localization of these isoforms.

Although there has been previous research carried out to investigate the generation of WWP2 isoforms, there are still no established mechanisms explaining the transcription of these isoforms. Epithelial Splicing Regulatory Proteins (ESRPs) have been suggested to be responsible for the alternate splicing events leading to retained introns, however, there is still a lack of literature investigating the relationship between ESRPs and the expression of truncated WWP2 isoforms (Chantry, 2011). It is therefore the aim of this chapter to clarify the transcriptional mechanisms responsible for the production of WWP2-C, WWP2-N and WWP2- $\Delta$ HECT which could have the potential to positively or negatively affecting oncogenesis through interacting with mediators of the TGF $\beta$  pathway. The understanding of these mechanisms could allow us in the future to produce treatments that may interrupt transcription of potential oncogenes, or enhance production of tumour suppressors. First, luciferase assays were used in attempts to identify enhancers or repressors of the predicted promoter regions of WWP2-C and  $\Delta$ HECT. To do this, regions of the predicted promoter sequences for both isoforms were cloned into pGL4.27 expression vectors containing the luciferase reporter gene. Using a combination of restriction enzyme digestions, T4 ligations, TA and In-fusion cloning techniques, WWP2-C and  $\Delta$ HECT predicted promoter sequences were successfully cloned into pGL4.27 (For more details of cloning methods, please refer to section 2.1).

The role of the transcription factor SOX9 in the transcription of the WWP2-C isoform was also investigated. Previous studies by Zou *et al* demonstrated that SOX9 regulates the expression of WWP2 in mouse craniofacial development by showing that the overexpression of SOX9 leads to elevated expression of WWP2 in chondrogenic and mesenchymal cell lines (Zou *et al*, 2011). The SOX9 binding site, -CCTTGAG- was also found within intron 10/11 upstream of the translational start of WWP2-C promoter, therefore we hypothesize that SOX9 may be involved in the transcriptional regulation of WWP2-C. To verify the function of SOX9, a WWP2-C promoter region mutant was produced using QuikChange XL site-directed mutagenesis kit (Agilent Technologies) containing the complete deletion of the SOX9 binding sequence, and was used in subsequent luciferase reporter assays.

Following the investigation into the WWP2- $\Delta$ HECT transcription mechanism, the functional activity of  $\Delta$ HECT was also investigated. This novel isoform was cloned into pRK5 expression vector containing the HA tag, and co-transfected into HEK293A and A375 cell lines with the p-CAGAC12-luc Smad response vector. Luciferase assays were then performed on the transfected cell lysates to explore the Smad3-dependent TGF $\beta$  activity of WWP2- $\Delta$ HECT. Previously, the function of this isoform could only be postulated due to the lack of an intact HECT catalytic domain, therefore the aim of these experiments is to elucidate the effects this isoform has on Smad modulators of the TGF $\beta$  pathway, and its potential effects on oncogenesis.

The transcriptional mechanism of a third truncated isoform, WWP2-N, was also studied in this chapter. It was hypothesized that the expression of WWP2-N, resulting from the presence of a stop codon in the retained intron 9/10, is affected by the post-transcriptional gene regulation of alternate splicing events, specifically by Epithelial splicing regulatory proteins 1 and 2 (ESRP1+2). Previous studies have shown that ESRPs modulate the splicing events in proteins during EMT by having the ability to both splice out and retain exons within a gene to create translation of a truncated protein

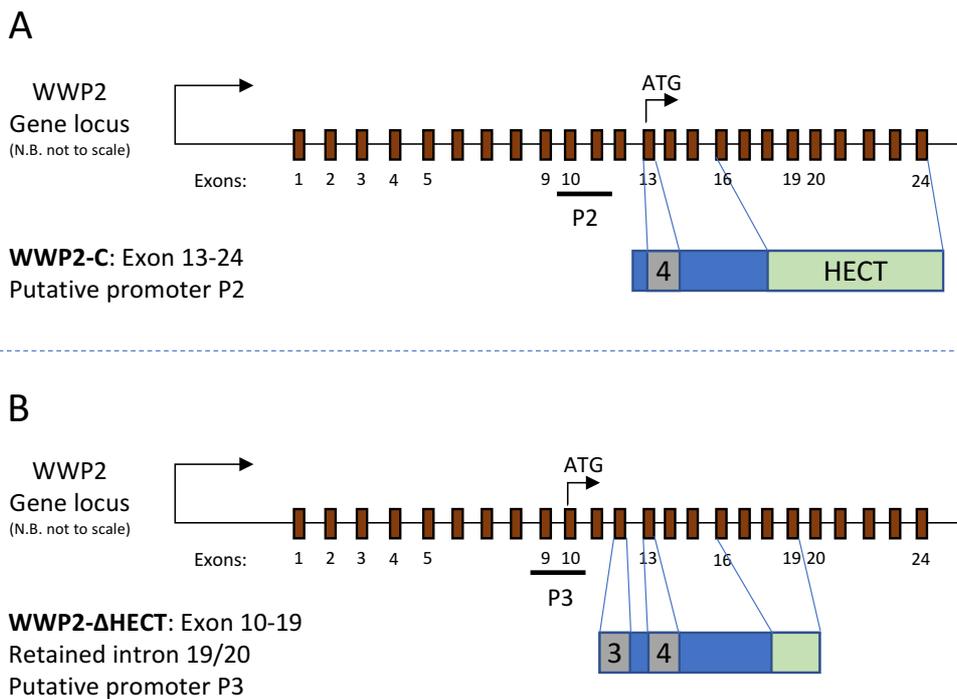
(Warzecha *et al*, 2010). This therefore suggests that the expression of WWP2-N, a splice-regulated truncated isoform shown to be involved in TGF $\beta$ -mediated EMT (Soond and Chantry, 2011), may be regulated by ESRPs in a co-transcriptional manner. To investigate the possible relationship between ESRP proteins and WWP2-N expression, RT-PCRs and qPCRs were performed using samples from HEK293A cells transfected with ESRP1+2 constructs and probed with WWP2 isoform specific primers to examine the effects of ESRPs on expression of WWP2-N. Due to the role TGF $\beta$  has in EMT and subsequently oncogenesis, the effects of TGF $\beta$  was investigated by stimulating cells with TGF $\beta$  18 hr prior to mRNA extraction. By uncovering effects of TGF $\beta$  on ESRP activity and subsequent WWP2-N expression, results should highlight the significance of ESRPs and WWP2 in TGF $\beta$ -mediated EMT.

To further assist in the hypothesis of the role WWP2 isoforms have in oncogenesis through its effect on the TGF $\beta$  pathway, the cellular sub-localization of WWP2 isoforms were examined. To do this, WWP2 isoforms were cloned into pRK5-HA vectors and transfected into ARPE-19 cells. These cells were stimulated with TGF $\beta$  then labelled with a primary anti-HA antibody and secondary Alexafluor antibody allowing the visualization of transfected proteins in both TGF $\beta$  stimulated cells and control cells without stimulation. Performing these immunolabelling experiments should highlight potential TGF $\beta$ -mediated translocation of WWP2 isoforms within the cell and ultimately help the understanding of the overall biological roles of WWP2 isoforms, and perhaps lead to a better appreciation of their putative roles in oncogenic mechanisms.

## 5.2 Results

### 5.2.1 Transcriptional regulation of WWP2-C and WWP2- $\Delta$ HECT

Although WWP2-C is a well-established isoform of WWP2, mechanisms that result in its expression are still unclear. In this section, attempts will therefore be made to identify regulation of the predicted promoter region and transcriptional mechanism of WWP2-C and the WWP2- $\Delta$ HECT novel isoform to increase our understanding of WWP2 and its role in oncogenesis. The schematic in fig 5.2.1 summarizes the WWP2-C (A) and  $\Delta$ HECT (B) isoforms in the WWP2 gene locus and highlights the relevant domains present in each isoform. The P2 promoter shown in fig 5.2.1 displays the predicted location of WWP2-C promoter in intron 10/11 whilst the promoter P3 in fig 5.2.1B illustrates the location of the putative promoter responsible for WWP2- $\Delta$ HECT.



*Figure 5.2.1- A schematic representation of the WWP2-C and  $\Delta$ HECT isoforms on the WWP2 gene locus. The predicted locations of P2 and P3 putative promoters are also highlighted respectively.*

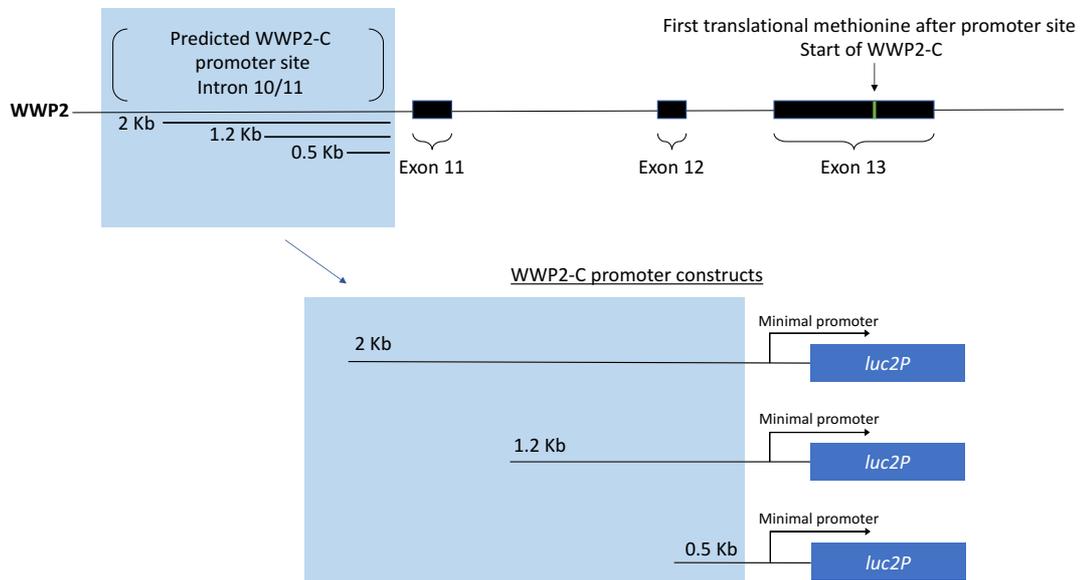
#### 5.2.1.1 Identification of potential enhancer/repressors of WWP2-C

promoter activity and its regulation by SOX9 transcription factor

Recent studies have suggested that the promoter region of WWP2-C is located at intron 10/11 of the WWP2 gene in a region upstream of the WWP2-C EST transcript (NM\_199424.2) (Soond and Chantry, 2011). To identify potential enhancers or repressors of WWP2-C transcription, the predicted regions were cloned into the pGL4.27 vector containing the *luc2P* luciferase reporter gene for the use in luciferase assays. Regions of 0.5 kb, 1.2 kb and 2 kb located at the end of intron 10/11 where transcription of WWP2-C begins were cloned using genomic DNA extracted from HEK293A cells, into the reporter vector producing three separate predicted C-promoter constructs (shown in fig 5.2.2). The reason a maximum of 2 kb region of the predicted WWP2-C promoter was chosen is highlighted in fig 5.2.3, a bioinformatics approach comparing the conserved elements between species using phastCons and phyloP in the UCSC genome browser. This analysis emphasized the high levels of predicted conserved elements within this 2kb region therefore indicating a high possibility that the promoter region lies within. The Log odd score (lod) scores presented with each predicted conserved element are log odd score which represents the confidence in each prediction. Within this 2 kb region, the predicted conserved elements with the highest lod scores are circled in red, orange and blue, and located in the 1.2 kb region, 0.5 kb region and an overlap between exon 11 and the 0.5 kb WWP2-C predicted promoter region, respectively.

Previous studies have suggested that SOX9 transcription factor regulates expression of WWP2 (Zou *et al*, 2011). The presence of the -CCTTGAG- SOX9 binding site in the predicted C-promoter region shown in fig 5.2.4, along with the conservation of this binding site shown in fig 5.2.5 using the Multiz alignment of 20 mammals, provides further evidence of the role SOX9 might have in WWP2-C expression. The predicted C-promoter constructs were therefore transfected into HEK293A, A375 and MCF-7 cells along with a SOX9 expression vector and  $\beta$ -galactosidase using a standard PEI

transfection protocol and harvested day 3 post-transfection. Finally, 100  $\mu$ l of luciferase substrate (Promega) was added to 30  $\mu$ l of harvested cleared cell lysates and results were normalized to  $\beta$ -galactosidase readings.



**Figure 5.2.2-** A diagrammatic representation of the WWP2 gene sequence between intron 10/11 and 13/14 and the WWP2-C promoter constructs in pGL4.27. The predicted promoter region of WWP2-C is located at intron 10/11 as suggested by the EST transcript of WWP2-C. The translation of WWP2-C however does not start until the next methionine in exon 13. Three promoter constructs were cloned into pGL4.27 reporter vector including the penultimate 2 Kb, 1.2 Kb and 0.5 Kb of the intron 10/11 sequence.



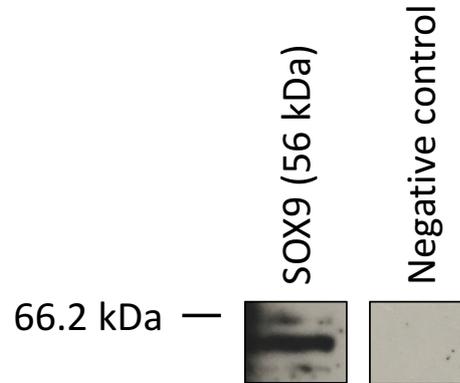
WWP2-C predicted promoter sequence

AAAGGCAAGGATT **CAGA**AAAGGGCTGATCAGTATTT **TCTG**ATCCTAGTTCAGTGCTAGTCAG  
 Start of 2 kb sequence GGAGACTGTCAGTAAAGACCTTTTGG **TCTG**TTGATTGTAGAATTAAGCAAAGATGACCA  
 TGCCCTGTGGAAGCTGGGTGGTACCAGCTGGTAGCTGATTTCCAGGAGAACCCTGAACG  
 CTTTA **CAGA**TGATCTACTGACTCCAAAACAAGACCTGAAGCAGCAAAACACTTTTACACAAA  
 TCGCTCTCTTTTAAATGTCTTAAATAACCTTTTCTATAGTCCTTCCCTTTTGAAGGTCAC  
 TGCTGTAGCTCAGCACTTACCAGTGCC **TCTG**GCTCCATGATGTAGAATACGTGTGTTTCT  
 CATTTGTTCTTTGCC **TCTG**TTGCCGAATGTTTTTAAATGTGTAACTACAAA **CAGA**GTTGG  
 AGGAACTTGGAATTTATGAGATGCTGTACCTATCTCCCCAACCCCCAACACA **CAGA**CAC  
 ACGCCCCATGACTTGGAATTCGGGAGATTTTTTTCAGTAATGATTAACGTGATTTGTGAT  
 GGAGAATTTTAGGCACAGCTTGAGATCAATGATAAGGGGCCATAATCAAGGTGAGCTTCCC  
 CCACCTCCGTGCACTTGCA **TCTG**TGTGGCAGGAGGCCCTGCCCTGAAAGGGGGCTCCGTC  
 CTGTCAAGAGCCTATGGTTTCAAGACTTAAACATCTTTGTTGGAACCTTGCTTGAAGAGGGG  
 CCTGTGTTGCCACCTTGGAAGAAAGTTTT **TCTG**TCCCTTTCAGCTGGGGTGGAGG **GAGG** Start of 1.2 kb sequence  
 TGTGAAGGGGGTGTGGCAGTTGAGTTCACCAATGTTTTCTCAAGCCGACACTGAGCC  
 CCATTTCATCTCTGCCATGTGTGGCTCCGGCACCCCTAGACGGGGCCAGCCAGCAGCCT  
 GAAGTCCGCATTTAGAACTTGATCCCTCT **TATAA**CGATTAATTTTAGTAAGTAAAATAAA  
 AATGCTT **CAGA**AAGAAAAAAGAAAGGAAGAAAGAGTGAACAATACTGTTTTATTTAAC  
 GAC **CAGA**ACTTTTGGACAACACCCCTTCCCCGAGGATATGAGAATGCCTGGTCCACTGC  
 AGTCCATCCCAGGGACCTCGAAG **CAGA**ATTTGGCTGGAGGGGTGTGGGGGGATCTAAT  
 GTAACAGTGACCCTGAAGCCCTGGAGCCAGTCCCCCTTCTTCCCCTGTCTCCCCT **TCTG**  
 TCTTGTGCCCCACCCAGCCCCAACACCCCCACCCCCACCCCCACCCCGAGACAAA  
 CA **CAGA**ACACTGAGCAACTTCAGGTTGAGGCAGGGGAGGAATAAAGGTGCTTTGTAAAGGGG  
 AAGGAAAACATTCCTGGGGGAGGTGAGGGCTTGGGGCATGAATGTGCCCTCAG **TCTG**GGTGG  
 ACGCTGCACACC **CAGA**TGGGAGGTTGGGGGGACGCCGAGGTGGAGGGAGGAGGGCCGGGCC  
 Start of 0.5 kb sequence CCAGCCAGGAA **TCTG**GGCGTGGGAGGGCTTGTGGGAATGATTTTCATTGGAAGGCCTGC  
 GATATTTTTTCCCTCCTGCGTGTGTTCTTGGAGAAAGTTGGAGGTGGTGGTATTTTCAGT  
 CGCCTTGGCCG **CCTGAG**CCGAGCTGAGCGGAGGCACTGGGCCGAGCCTGCTTCCGGGCC  
 TTCCATCCATGCCAGGCTGCTCCCTGCCCTCCGCCACCCTGGCACACCTTACCCGCGTACC  
 GCCTCCTCCCGTCGC **TCTG**CCTTTTCCAAAACACTTGGGCCCTCCGTGCGCAGGGTTCT  
 TTTTGGTTTT **TCTG**TAAAAATCAAAACAAAAA **CAGA**GACTTTTGAAGAGG **CAGA**TGCCA  
 CCTAAAGTCCCACTGCATTCCTGCAAAGCGCTCAAATGTGAAGCCAGTCATTGGCATT  
 TATTTTTTATGATTGATTGATTTTTTACCAGTGGCTTTTTGTAACC **TCTG**GT **TCTG**CTG  
 TGTTTCTGTGTTTAG

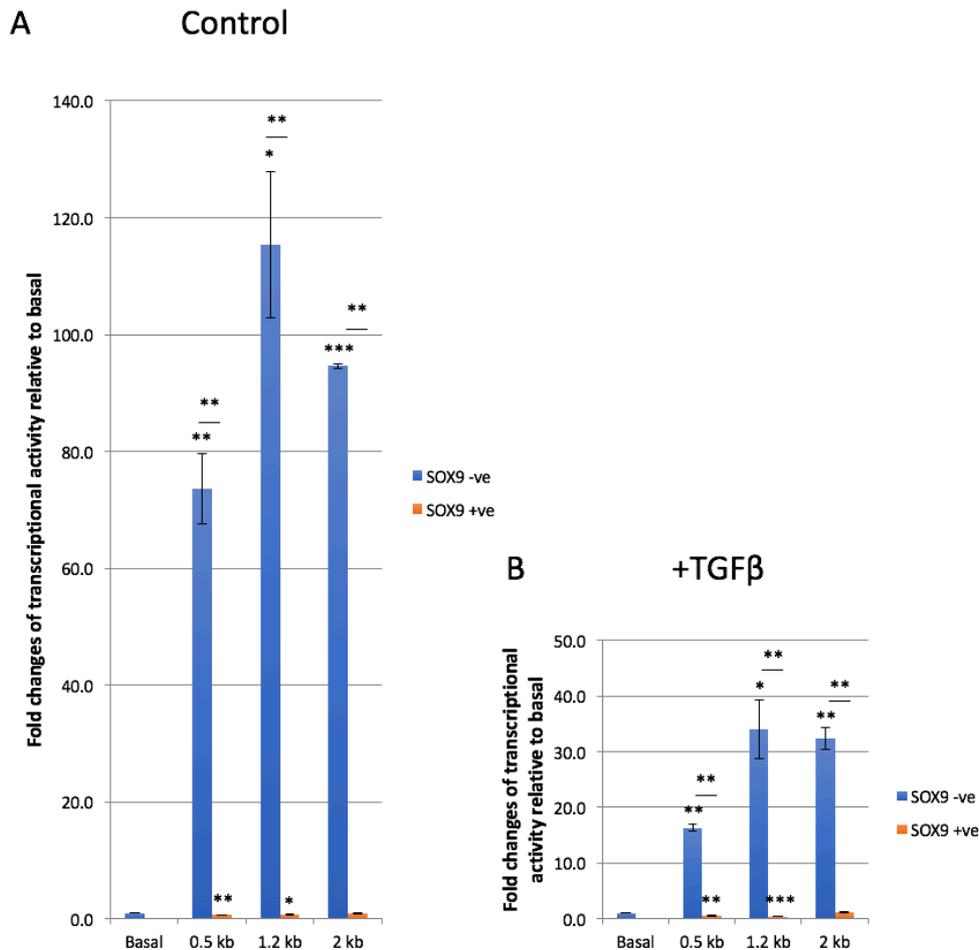
Figure 5.2.4- WWP2-C predicted promoter region nucleotide sequence. The 2 kb nucleotide sequence of the predicted WWP2-C promoter regions within intron 10/11 with the sequence in brown highlighting the beginning section of the WWP2-C transcript. The start of the 1.2 kb promoter sequence is circled in purple whilst the start of the 0.5 kb sequence is circled in green. Several other regions are featured in this sequence including the SOX9 binding site highlighted in yellow, the Smad binding sites highlighted in blue, and the TATA box highlighted in grey.



Most notably, an enormous decrease in promoter activity in the presence of SOX9 suggested that SOX9 negatively regulates WWP2-C promoter activity. These decreases seen in both cells with and without TGF $\beta$  stimulation all show a significance value of  $p \leq 0.01$ .



*Figure 5.2.6- A western blot using primary anti-Flag antibodies to detect SOX9 protein expression in HEK293A cells. A non-transfected HEK293A cell lysate sample was used as negative control. (Experimental replicate n=1)*

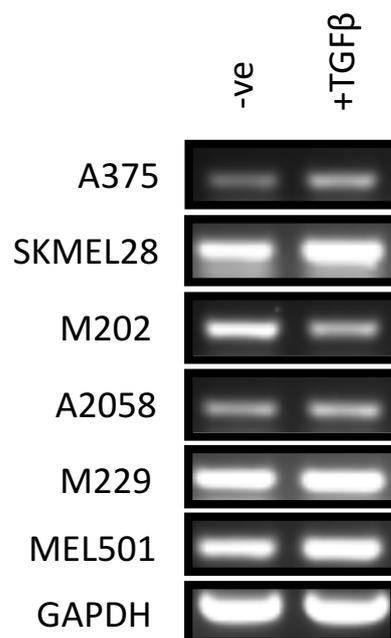


**Figure 5.2.7- The effects of SOX9 and TGFβ on WWP2-C predicted promoter activity.** The 0.5 kb, 1.2 kb and 2 kb predicted WWP2-C promoter regions were cloned into the pGL 4.27 luciferase reporter vector and transfected into HEK293A cells together with SOX9. (A) shows results from unstimulated cells whilst (B) shows results from cells stimulated with TGFβ overnight. The transfection efficiency was first normalized to β-galactosidase, then fold changes were calculated relative to readings from cells transfected with the empty pGL 4.27 control. (Experimental replicate n=6. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

### 5.2.1.2 Effects of TGFβ on SOX9 transcription factor expression in melanoma cell lines

An important aspect to bear in mind is that in previous studies on the WWP2-C promoter, the SOX9 construct was transfected into cells. To validate relevant hypotheses, the expression of naturally occurring SOX9 must be confirmed in cells not transfected with recombinant SOX9. To do this, RT-PCR using RNA extracted from six melanoma cell lines were amplified with SOX9 RT-PCR primers which confirmed SOX9 mRNA expression as shown in fig 5.2.8. Investigating the effects of TGFβ on the

expression of SOX9 transcription factor is also important as SOX9 is a crucial element in the transcriptional regulation of WWP2-C. To find out the effects TGF $\beta$  has on SOX9 expression could shed light on the knock-on effects it has on WWP2-C transcription which could subsequently make clearer the relevance of SOX9 in WWP2-dependent oncogenesis and EMT. RNA from six melanoma cell lines stimulated with TGF $\beta$  were extracted and converted into cDNA using GoScript PCR. RT-PCR for these samples allowed the comparison of SOX9 expression between control and TGF $\beta$  stimulated cells. The results from this experiment can also support data regarding WWP2-C transcription and potentially explain the negative effects SOX9 has on WWP2-C transcription mechanism. As shown in fig 5.2.8, SOX9 is expressed in all six melanoma cell lines tested, furthermore, TGF $\beta$  stimulation seem to have little or no effect on SOX9 expression in all melanoma cell lines except M202 which decreased by 22% in band density when analysed using Image J.



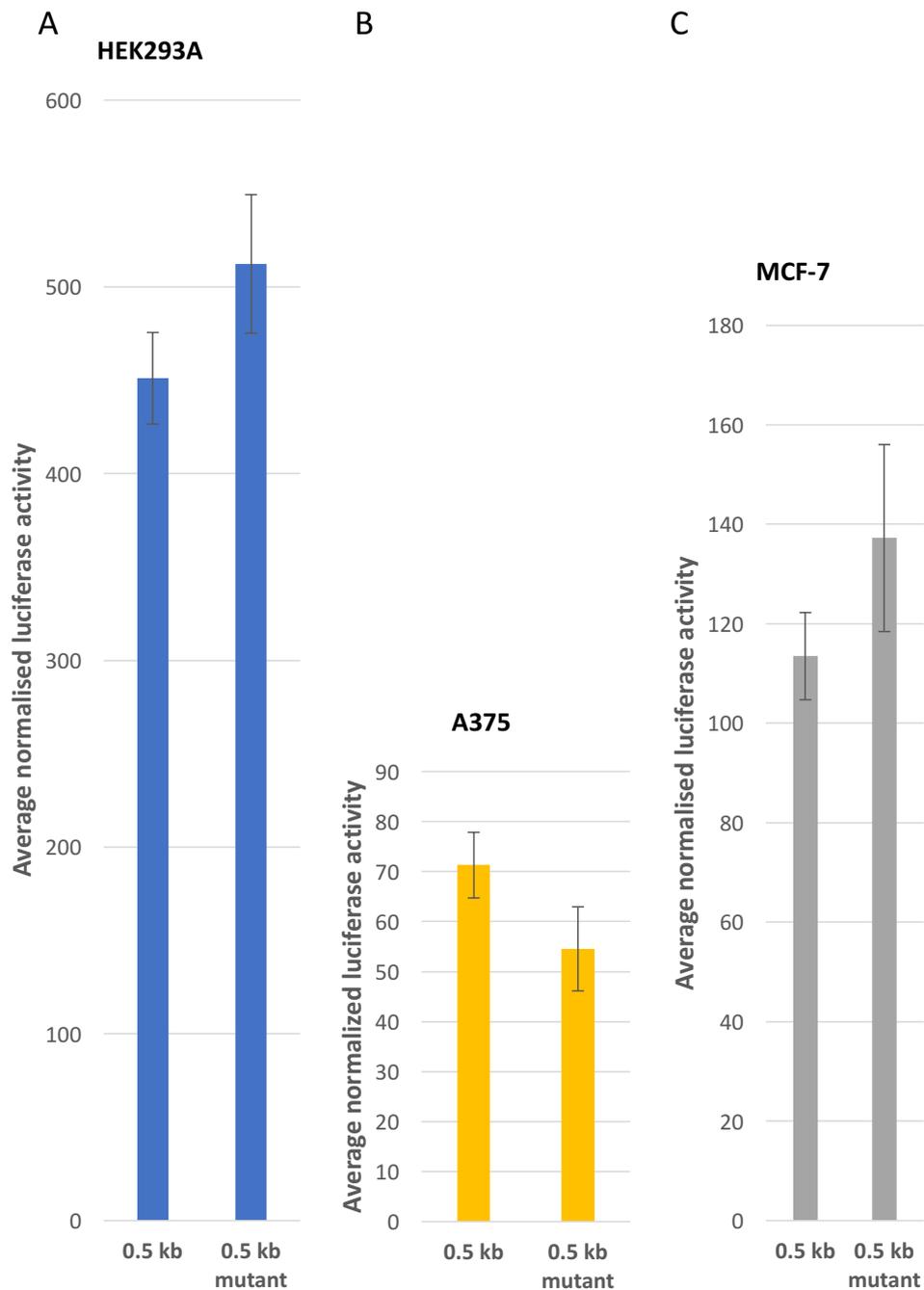
*Figure 5.2.8- The effect of TGF $\beta$  stimulation on mRNA expression of SOX9 in melanoma cell lines. Results from RT-PCR of melanoma cell lines showing the expression of SOX9 in the presence and absence of TGF $\beta$ . Results suggests that TGF $\beta$  stimulation seem to have no significant effect on expression of SOX9 except for the decrease seen in M202 cell line. (Experimental replicate n=1)*

### 5.2.1.3 Investigating the role of SOX9 in WWP2-C transcription using SOX9 binding site mutant constructs

To investigate the SOX9-mediated negative regulation on WWP2-C transcriptional activity, the predicted SOX9 binding site -CCTTGAG- within the 0.5 kb promoter region is deleted to produce a mutant C-promoter construct using the Agilent QuikChange XL site mutagenesis kit. Luciferase reporter assay was performed using the WWP2-C promoter mutant construct and results were compared with the wild type C-promoter construct allowing the study of the predicted SOX9 binding site and its regulatory activities in WWP2-C transcription. All results were normalized to  $\beta$ -galactosidase. The wild type and mutant constructs were compared in three different cell lines in this section to allow a broader understanding of WWP2-C transcription in different cell types.

Fig 5.2.9A shows the luciferase assay analysis from HEK293A epithelial cells transfected with wild type and mutant 0.5 kb c-promoter constructs in pGL4.27. The result compares the average luciferase activity normalized to  $\beta$ -galactosidase of the wild-type WWP2-C promoter with the SOX9 mutant C-promoter. The previous experiment shown in fig 5.2.7 highlighted the negative effects SOX9 has on WWP2-C transcriptional activity. The hypothesis for this mutant study is therefore the increase in luciferase activity caused by transfected 0.5 kb mutant C-promoter construct containing a SOX9 binding site deletion, in comparison to the wild type C-promoter construct. This predicted increase in luciferase activity is due to the inability of SOX9 to bind to the promoter sequence in the mutant construct, subsequently the negative regulation on C-promoter activity is inhibited. Unfortunately, the increase in transcriptional activity seen in HEK293A and MCF-7 transfected with the SOX9 mutant construct (fig 5.9.2A and C) were statistically insignificant, however, these results may highlight a trend that may support the negative effect SOX9 has on WWP2-C transcriptional activity.

The same experiment was repeated in an epithelial melanoma cell line expressing mesenchymal characteristics, the SOX9 binding site deletion failed to significantly affect the SOX9 mediated negative regulation of WWP2-C transcriptional activity reflected in fig 5.2.9B. Altogether, the results suggest that the SOX9 mediated regulation of WWP2-C transcriptional activity may be dependent on cell types and epithelial/mesenchymal phenotypic differences. Alternatively, the results might also suggest that the SOX9 effects on WWP2-C transcription may not be directly mediated via the predicted SOX9 binding site, and could be due to other mechanisms such as associations with other proteins to indirectly assert its effects on WWP2-C transcriptional activity.



**Figure 5.2.9- WWWP2-C wild-type and mutant promoter activity in HEK293A, A375 and MCF-7.** Results from luciferase analysis of WWWP2-C transcriptional activity comparing the effects of naturally occurring SOX9 on the wild-type predicted promoter to the mutant promoter construct containing a SOX9 binding site deletions. A) shows luciferase results from the HEK293 epithelial cell lines whilst B) and C) shows results from A375 and MCF-7 cell lines. Although trends can be predicted using data above, no significant differences were detected. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

#### 5.2.1.4 Identification of enhancers and repressors of WWP2- $\Delta$ HECT transcription

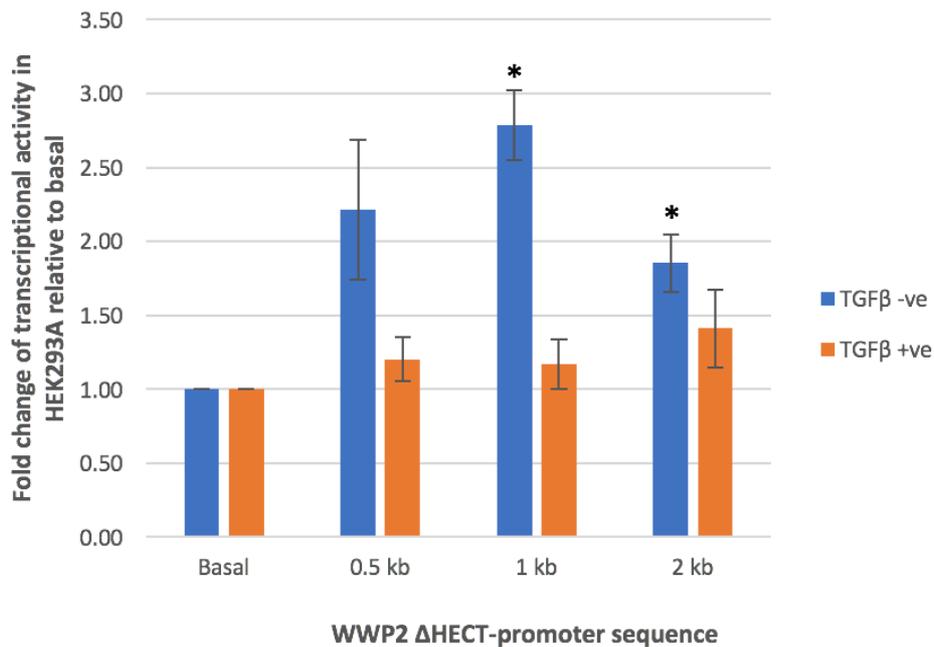
Although the function of WWP2- $\Delta$ HECT is not established, it is important to try and identify the regulation involved in the transcription mechanism of  $\Delta$ HECT due to its potential role in the development of a prognostic or therapeutic tool in melanoma. Translation of WWP2- $\Delta$ HECT begins at exon 10 therefore logically, the promoter sequence was predicted in intron 9/10. Similarly to the WWP2-C transcriptional regulation experiments, three clones were made from different lengths of the predicted  $\Delta$ HECT promoter region and used in luciferase reporter assays to probe for maximum transcriptional activity. These clones were then transfected into HEK293A cells at a concentration of 3000 ng/ml following standard PEI transfection protocol and harvested 3 days post transfection. The harvested cell lysates were then added to 100  $\mu$ l of luciferase and  $\beta$ -galactosidase substrate separately for the subsequent use in reporter assays.

Three constructs were made using different lengths of predicted  $\Delta$ HECT-promoter sequence within intron 9/10. The 0.5, 1, and 2 kb regions were cloned into pGL4.27 luciferase reporter vector using a combination of TA, T4 and In-fusion cloning techniques. The nucleotide sequence of the predicted WWP2- $\Delta$ HECT promoter is shown in fig 5.2.10 and highlights presence of CAGA TGF $\beta$  binding elements in blue, EGF AP-1 binding sequences in pink, SOX9 binding site in yellow and the TATA box in grey. Luciferase results in fig 5.2.11 shows that all three of the predicted constructs enhance the promoter activity when compared to basal level of empty vector (both 1 kb and 2 kb constructs produced results that were statistically significant ( $p = <0.05$ ) with the negative effects of TGF $\beta$  highlighted by the decrease in activity in all three constructs.

WWP2-ΔHECT predicted promoter sequence



Figure 5.2.10- The predicted promoter region of WWP2-ΔHECT. The 2 kb predicted WWP2-ΔHECT promoter region showing the start of the 1 kb sequence in purple and 0.5 kb in green. The EGF binding element, AP-1 binding site is highlighted in pink whilst the TGFβ response element, CAGA is in blue, SOX9 binding site in yellow and finally the TATA boxes are highlighted in grey.

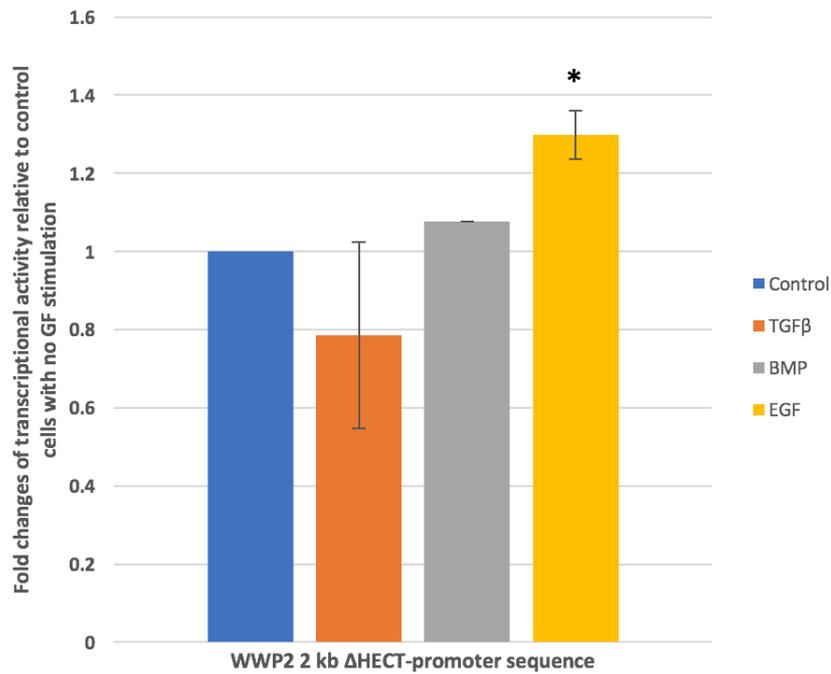


**Figure 5.2.11- The effects of TGFβ on WWP2-ΔHECT predicted promoter activity.** Fold change in luciferase activity relative to pGL4.27 basal in HEK293A cells transfected with 0.5, 1, and 2 kb WWP2-ΔHECT promoter constructs in the presence and absence of TGFβ. The overexpression of 1 and 2 kb constructs significantly increased transcriptional activity whilst other variables did not result in significant changes. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

#### 5.2.1.5 Effects of growth factor stimulation on WWP2-ΔHECT transcriptional activity

Although luciferase results show that WWP2-ΔHECT transcriptional activity is significantly enhanced by the overexpression of 1 kb and 2 kb constructs in fig 5.2.11, this activity is still relatively low when compared with the WWP2-C transcriptional activity in fig 5.2.7. This therefore prompted the investigation of the potential enhancing/repressive roles different growth factors may have in the transcriptional activity of ΔHECT. The 2 kb ΔHECT promoter construct was used in the following experiments and was subjected to TGFβ, BMP and EGF stimulation at 5 ng/ml, 10 ng/ml and 10 ng/ml, respectively. Results were calculated relative to basal empty vector readings then normalized to control results of 2 kb promoter activation from unstimulated cells and are shown in fig 5.2.12. Results suggest that whilst TGFβ and BMP had no significant effect on ΔHECT transcriptional activity,

EGF stimulation increases activity of the 2 kb  $\Delta$ HECT promoter sequence by 1.3 fold ( $p \leq 0.05$ ).



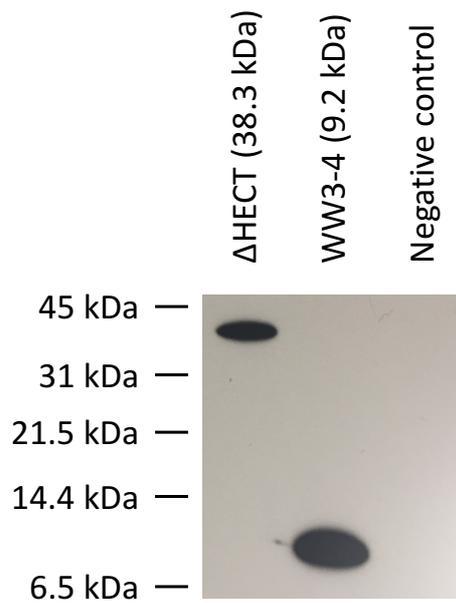
**Figure 5.2.12- The effects of TGF $\beta$ , BMP and EGF on WWP2- $\Delta$ HECT 2 kb predicted promoter activity.** The predicted 2 kb WWP2- $\Delta$ HECT promoter construct was transfected into HEK293A cells that were then stimulated with TGF $\beta$ , BMP and EGF overnight, and results were normalised to  $\beta$ -galactosidase data. The above chart shows the results as fold change relative to basal activity of untreated 2 kb construct and suggests that EGF stimulation significantly increases transcriptional activity. (Experimental replicate  $n=3$ . Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

### 5.2.2 Studying the effects of WWP2- $\Delta$ HECT on Smad-3 dependent gene expression using p-CAGAC12-luc reporter in luciferase assays

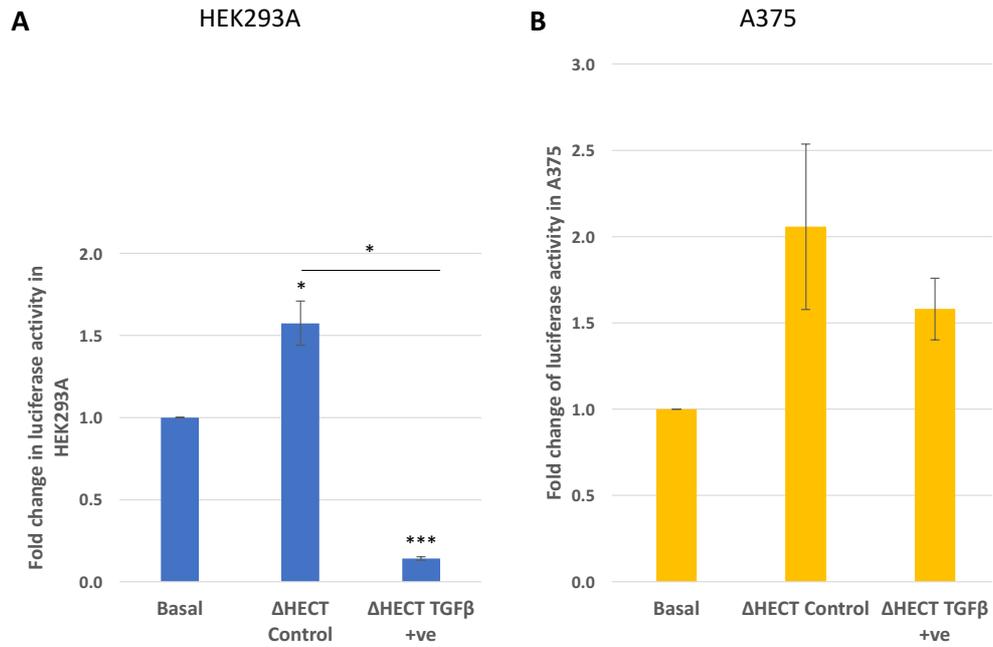
The novel WWP2- $\Delta$ HECT isoform contains only the WW3, WW4 and a partial HECT catalytic domain. It is for this reason that its function cannot be accurately predicted as it is not certain whether the incomplete HECT domain retains its catalytic activity. The combination of domains, makes it an interesting target for investigation as the WW4 domain has previously been suggested to bind to Smad7 of the TGF $\beta$  pathway. This means that if the partial HECT domain maintains its catalytic activity, the natural inhibitor Smad7 could be subjected to ubiquitin-dependent proteasomal degradation. On the other hand, if the partial HECT domain is catalytically inactive,  $\Delta$ HECT could act as an effective anti-oncogenic factor by stabilizing levels of inhibitor Smad7 and therefore disrupt the oncogenic effects of TGF $\beta$  such as EMT.

Similar to previous luciferase experiments studying the effects of tandem WW domains in chapter 3, the WWP2- $\Delta$ HECT isoform was cloned into pRK5 HA expression vector and co-transfected into HEK293A and A375 melanoma cells along with the p-CAGAC-12 Smad3 dependent TGF $\beta$  response vector and the pRSV- $\beta$ -galactosidase construct as a control. Expression of WWP2- $\Delta$ HECT was initially confirmed using western blotting with anti-HA antibodies which showed the cloned construct at the correct molecular weight of 38.3kDa (shown in fig 5.2.13). For the luciferase assays, cells were stimulated with TGF $\beta$  whilst others were left as unstimulated controls to explore the effects of TGF $\beta$  on  $\Delta$ HECT in a Smad3-dependent manner. Cell lysates were extracted 3-days post-transfection and used in luciferase assays. All results were normalized to  $\beta$ -galactosidase and an empty vector reading. Fig 5.2.14A shows the result from transfected HEK293A cells with and without TGF $\beta$  stimulation whilst fig 5.2.14B shows luciferase results from transfected A375, again, with and without TGF $\beta$  stimulation. It is clear in these results that the  $\Delta$ HECT construct increases Smad3-dependent gene transcription by over 1.5 and 2-fold in HEK293A ( $p \leq 0.05$ ) and A375,

respectively. Also notable is the repressive effect TGF $\beta$  has on  $\Delta$ HECT-mediated Smad3-dependent activity demonstrated by the decrease in fold change to 0.2 fold in HEK293A. The negative effect of TGF $\beta$  although present, at 1.6 fold, was not as significant in A375 ( $p > 0.05$ ). Another trend that was evident was the generally higher levels of Smad3-dependent gene expression in  $\Delta$ HECT transfected A375 cells in comparison to HEK293A cells suggesting functional differences of WWP2 isoforms in normal and oncogenic cells.



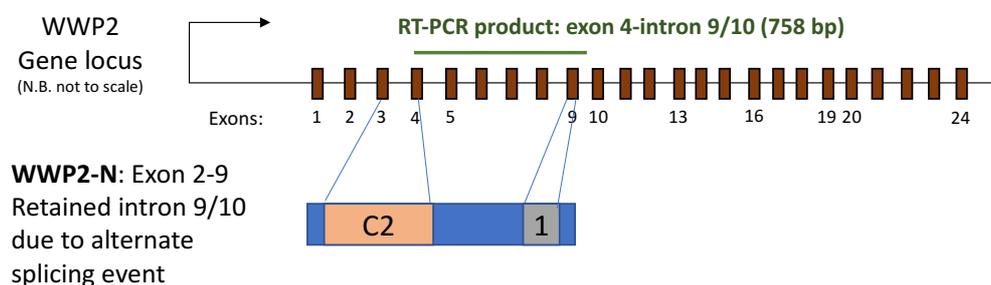
**Figure 5.2.13- Protein expression of cloned WWP2- $\Delta$ HECT construct in HEK293A.** A western blot using primary anti-HA antibodies to confirm protein expression of  $\Delta$ HECT at 38.3 kDa, with WW3-4 as a size comparison at 9.2 kDa. Cell lysate transfected with a Flag-tagged construct was used as negative control. Equal amounts of all samples were loaded onto the SDS-PAGE gels. (Experimental replicate  $n=1$ )



**Figure 5.2.14-** *The effects of WWP2-ΔHECT overexpression and TGFβ stimulation on Smad3-dependent gene expression in HEK293A and A375. HEK293A and A375 cells were co-transfected with the WWP2-ΔHECT construct with the p-CAGAC-12 Smad3-dependent gene expression vector stimulated with TGFβ for 18 hrs. Fold changes were normalized to empty vector readings and β-galactosidase reporter vector. Whilst results from A375 cell line were not significant, ΔHECT overexpression in HEK293A significantly increased Smad3-dependent gene expression whilst the overexpression of ΔHECT in the presence of TGFβ significantly decreased Smad3-dependent gene expression. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)*

### 5.2.3 Investigating ESRPs as a post-transcriptional regulator of WWP2-N expression

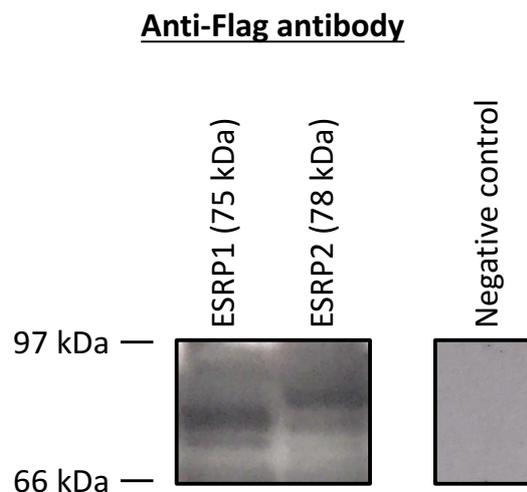
An alternate splicing event is responsible for the transcription of a stop codon within the retained intron 9/10 of WWP2 and subsequently produces a truncated WWP2-N isoform that contains only the C2 lipid binding domain and the WW1 recognition domain shown in fig 5.2.15. Although the role of WWP2-N in oncogenesis is still unclear, deciphering the transcriptional regulation of this truncated isoform can help increase the understanding in this area and allow further hypotheses to be made regarding its role in cancer. ESRP 1 and 2 has been found previously to be an integral part of the Splicing regulatory network (SRN) that is involved in the regulation of EMT proteins (Warzecha *et al*, 2010) and therefore, we hypothesize that ESRP may have a role in the post-transcriptional regulation of WWP2-N, a protein heavily implicated in TGF $\beta$  mediated EMT (Soond and Chantry, 2011).



**Figure 5.2.15- The WWP2 gene locus and the WWP2-N isoform.** A simplified schematic showing the WWP2-N isoform and the location of C2 and WW1 domains in relation to exon/ intron positions of the WWP2 gene locus. Highlighted in green is the RT-PCR product of 758 bp produced using specific RT-PCR primers that detects the WWP2-N isoform specifically.

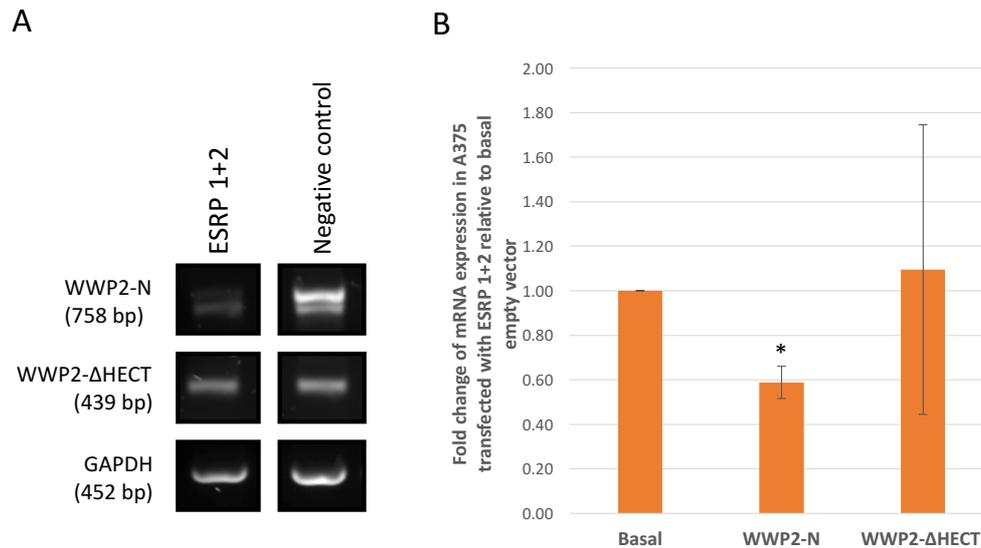
First, to test the protein expression of the flag-tagged ESRP constructs, a western blot was performed using primary anti-flag antibody, on HEK293A cells transfected with ESRP which showed both ESRP1 and 2 at the correct size (Shown in fig 5.2.16). These constructs were then transfected into HEK293A and MCF-7 cells according to a standard PEI transfection protocol and subsequently stimulated with TGF $\beta$ . RNA was harvested using the Promega SV wizard RNA extraction kit which was then made into cDNA using Promega GoScript PCR. To test the effects of ESRPs on WWP2-N expression,

and investigate the possible effects on WWP2- $\Delta$ HECT expression, the cDNA from ESRP transfected cells was used in both RT-PCR and Taqman qPCR. The reason for using both methods of PCR is that whilst qPCR provides a quantitative analysis of isoform expression, it does not account for the size of the isoforms which RT-PCR does, therefore both methods were used in the attempt to provide an extensive picture. Results from RT-PCRs were normalized to GAPDH whilst 18S and UBC were used to normalize qPCR results.



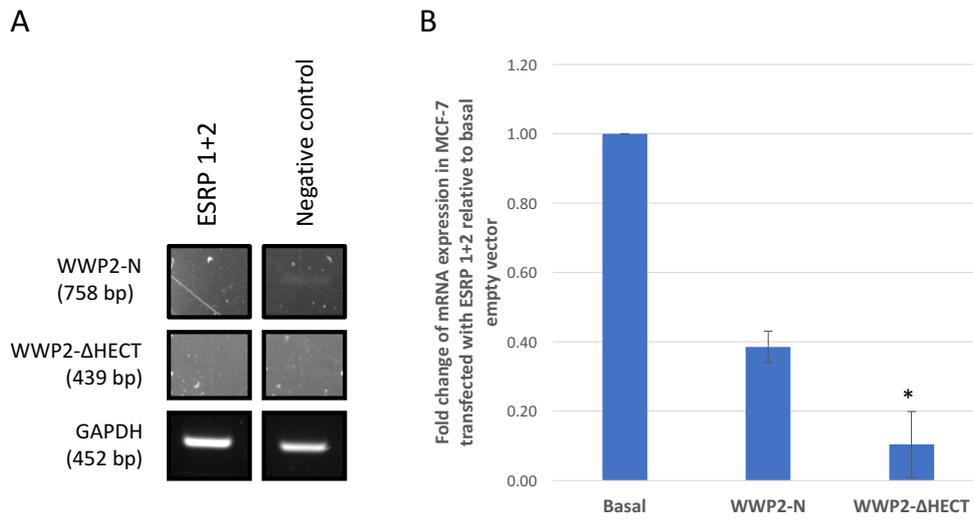
*Figure 5.2.16- Protein expression of ESRP1 and ESRP2 in HEK293A. A western blot showing the protein expression of flag tagged ESRP1 and ESRP2 using the primary anti-Flag antibody. Cell lysates from non-transfected HEK293A was used as negative control. (Experimental replicate n=1)*

First, the effects of ESRPs were tested in A375 cells as shown in fig 5.2.17. RT-PCR results in (A) suggests that ESRP1 and 2 decrease the expression of the N isoform, reflected by a decrease of 27% in band intensity when analysed using Image J. Interestingly, results from qPCR (fig 5.2.17B) also show the negative effect ESRP 1 and 2 have on WWP2-N expression by a decrease of fold change to 0.6 folds ( $p < 0.05$ ). Transfected ESRP1 and 2 also seem to have a negative effect on  $\Delta$ HECT expression as results from RT-PCR shown in (A) suggests a 12% decrease in band intensity, however, results were not supported by qPCR results as they were statistically insignificant.



**Figure 5.2.17- The effects of ESRP on the mRNA expression of WWP2 isoforms in A375.** WWP2-N and ΔHECT expression were studied using A) RT-PCR which was normalized to GAPDH where negative control was non-transfected A375 cells (experimental replicate n=1) and B) fold change relative to basal activity shown in a qPCR analysis which detected a significant decrease in WWP2-N in response to ESRP overexpression. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

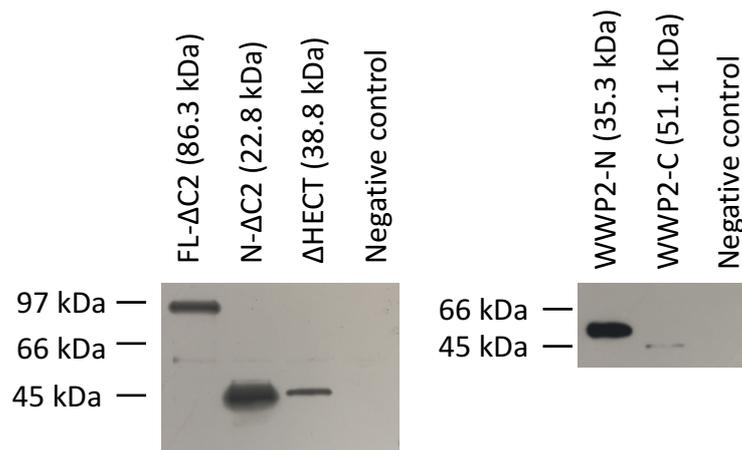
The same experiment was repeated in a non-melanoma cell line. MCF-7, a breast cancer epithelial cell line was used as a direct comparison to previous studies highlighting the difference in ESRP-regulation of WWP2-N expression in an epithelial cell line, and A375 which has increased mesenchymal properties. Results shown in fig 5.2.18 are produced from RT-PCR and qPCR analysis of MCF-7 cells transfected with ESRP1 and 2. In this cell line, the expression of WWP2-N and ΔHECT in RT-PCR (5.2.18A) are not as clear, however, the presence of a band in the WWP2-N ESRP negative control suggests that once again ESRP 1 and 2 negatively regulated the expression of this isoform. This is also highlighted in the qPCR study by a decrease of 0.6 folds when compared to the non-ESRP transfected negative control. In the qPCR analysis shown in fig 5.2.18B, the expression of WWP2-ΔHECT in the presence of transfected ESRPs decreased significantly when compared to the basal negative control ( $p < 0.05$ ). Results from mRNA expression studies of both the A375 and MCF-7 cell lines combined suggest that ESRPs may be involved in the regulation of alternate splicing events leading to the decrease in WWP2-N and ΔHECT expression.



**Figure 5.2.18- The effects of ESRP on the mRNA expression of WWP2 isoforms in MCF-7 cells.** Expression of WWP2-N and  $\Delta$ HECT shown using A) RT-PCR normalized to GAPDH levels GAPDH where negative control was non-transfected MCF-7 cells (experimental replicate n=1), and B) qPCR showing the fold change of isoforms relative to the empty vector basal activity. Note that the qPCR results highlighted a significant decrease in  $\Delta$ HECT expression in response to ESRP overexpression. (Experimental replicate n=3. Error bars show standard error. For more details of statistical analysis please refer to section 2.4.7)

#### 5.2.4 Cellular localization of WWP2 isoforms

Currently, no information has been published regarding the cellular localization of WWP2 isoforms. To identify the localization of these isoforms in the cell could aid the hypothesis of the role WWP2 isoforms has in oncogenesis and increase understanding of their normal biological roles. Furthermore, the effects of TGF $\beta$  on the WWP2 cellular sub-localization could also help reveal the location in which Smad mediators and WWP2 isoforms interact, therefore uncover the role TGF $\beta$  has in oncogenesis. Using immunolabelling, different WWP2 isoforms were located in ARPE-19 cells under TGF $\beta$  stimulation and in unstimulated cells. Note that ARPE-19 cells are epithelial retinal pigment cells that are commonly used for fluorescence microscopy due to its monolayer growth pattern that allows ease of imaging and analysis (Dunn *et al*, 1996). To begin, HA-tagged isoforms had to be cloned using Phusion DNA polymerase, *Xho*I and *Hind*III restriction enzymes and T4 ligase. Detailed protocols can be found in the methods section. A total of 5 different isoforms were cloned into pRK5-HA for these experiments with 2 additional constructs, WWP2-C and WWP2-N obtained from Jessica Watt in the Chantry lab. To confirm protein expression of WWP2 isoforms before beginning immunolabelling experiments, constructs were transfected into HEK293A cells and western blots using primary anti-HA antibodies were performed (as shown in fig 5.2.19). Note that the expression of FL was shown previously in fig 3.2.1. Constructs were transfected overnight into ARPE-19 cells using 3000 ng of DNA, then cells stimulated by TGF $\beta$  were starved overnight with media containing 0.5% FCS before standard stimulation of 2.5 ng/ml for 18 hrs. Transfected cells were fixed using methanol and blocked with 10% goat serum in PBS. Tagged proteins were then labelled with primary HA- tag antibodies and the respective Alexa Fluor 488 secondary which fluoresces green under the fluorescence microscope.



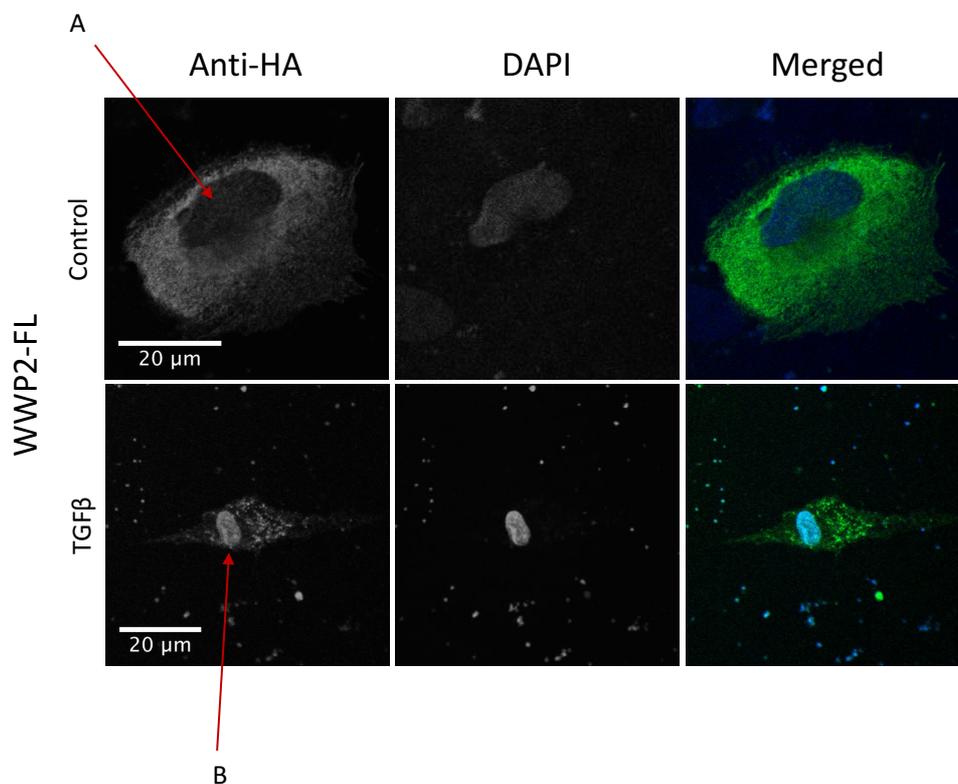
**Figure 5.2.19 - Western blots using primary anti-HA antibodies to detect the protein expression of WWP2- FL-ΔC2, N-ΔC2, ΔHECT, WWP2-N and WWP2-C.** Equal amounts of all cell lysates were loaded onto the SDS PAGE gels. Note that bands produced by N-terminal isoforms i.e. N and N-ΔC2 run higher than others. This is a pattern observed across all N-terminal isoforms. Negative controls included were non-transfected HEK293A cell lysates. (Experimental replicate n=1)

#### 5.2.4.1 Effects of TGFβ stimulation on the cellular localization of HA-tagged WWP2 isoforms

Results from cellular sub-localization studies is arguably as important as finding out the function of WWP2 protein. In many ways, the two findings will work together to give a more in depth understanding of the isoforms. First of all, the results from this study may confirm any hypotheses made regarding the C2 lipid binding domain within the isoforms and potentially suggest the effects of TGFβ on such binding. Secondly, the results may also provide us with a clearer idea of WWP2 and Smad protein interactions.

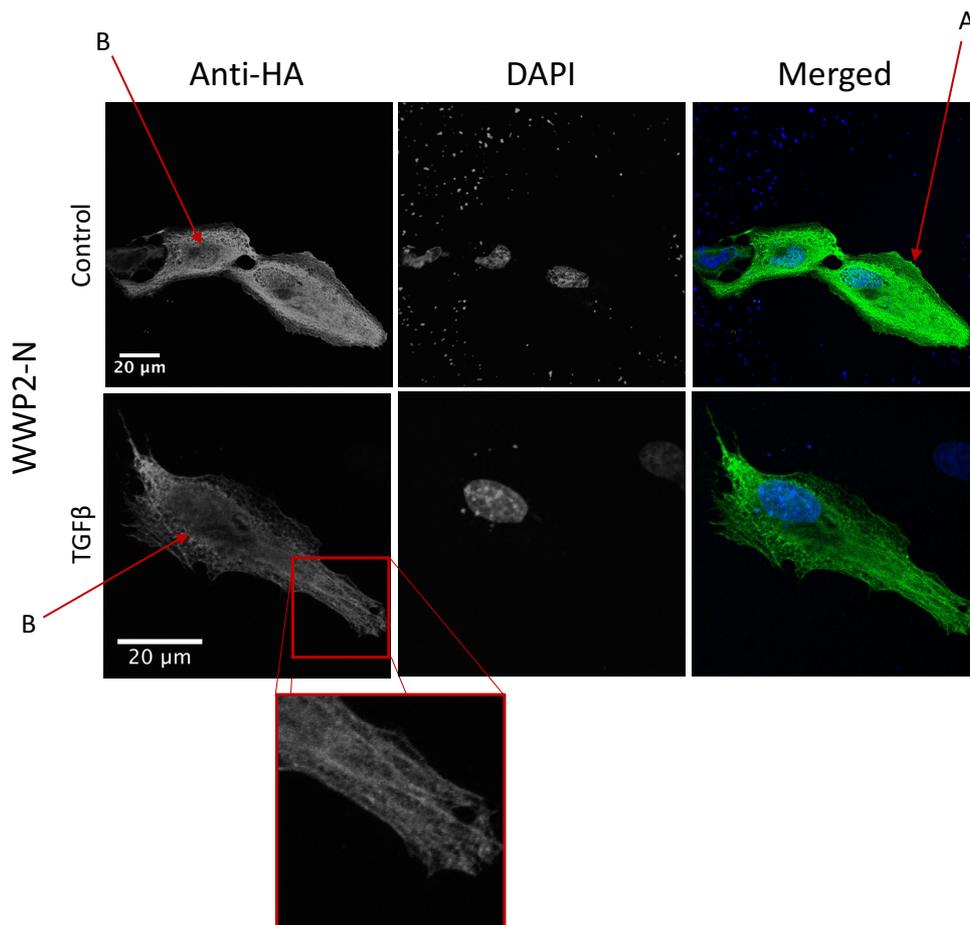
Here, ARPE-19 cells were transfected with HA-tagged WWP2 isoforms which were then labelled with primary anti-HA and secondary Alexa Fluor 488 antibodies which fluoresces green to show the protein of interest. Cells transfected with the full length WWP2 are shown in fig 5.2.20 which also compares the localization between unstimulated samples and samples stimulated with TGFβ. The images presented here are a representation of the cells observed under the fluorescence microscope.

The distribution of WWP2-FL can be seen throughout the cytoplasm with a distinct outline of the cell suggesting some membrane localization. However, without the labelling for cellular membrane or cell-cell junction proteins such as N-cad, we can only postulate its membrane localisation. When comparing the green (A488) and blue (DAPI) channel, it is also noticeable that the FL is not present in the nucleus as there is an obvious lack of staining in the nucleus (highlighted by A using red arrow). On the other hand, samples stimulated with TGF $\beta$  showed a much less defined membrane outline whilst the nucleus now displays increased staining (highlighted by B using red arrow) suggesting that membrane localization has decreased and the FL isoform has translocated into the nucleus.



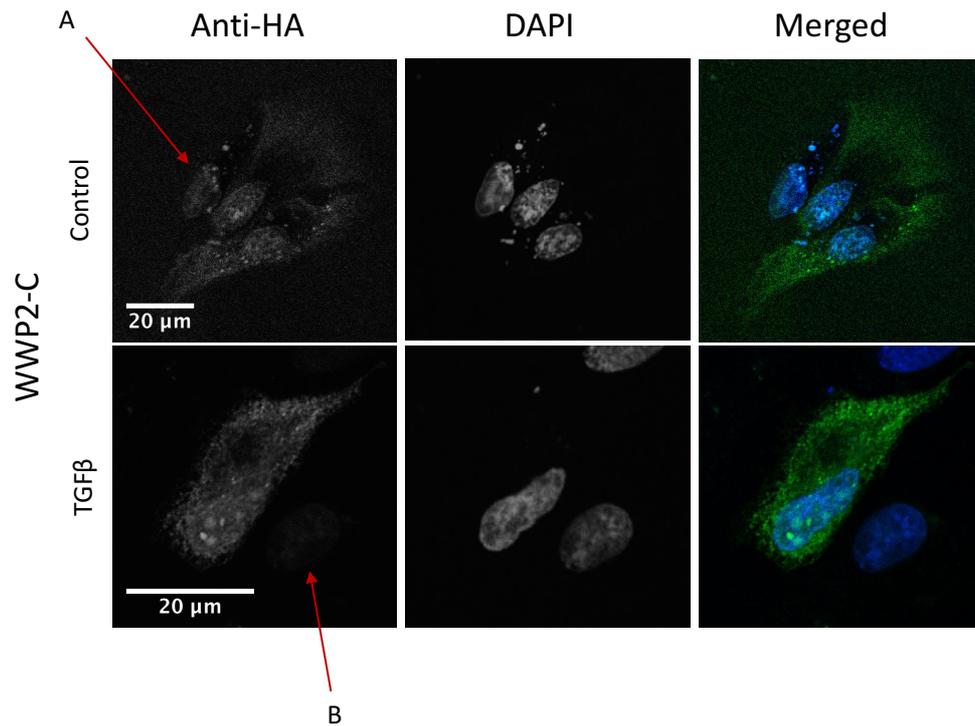
**Figure 5.2.20- Fluorescence microscopy of ARPE-19 cells expressing transfected WWP2-FL.** The full-length WWP2 isoform was HA-tagged then detected using primary anti-HA antibody and secondary rat A488 antibody which fluoresces green. The nuclei are shown in blue by the DAPI staining, and images on the right shows both the protein of interest and the nucleus together in the merge of two channels. Transfected cells were stimulated overnight with TGF $\beta$  and are compared here against unstimulated controls. A) shows the lack of nuclear staining from transfected WWP2-FL whilst B) shows the increased nuclear localisation of WWP2-FL. The results shown here are detected in all cells analysed.

The WWP2-N isoforms which contains only the C2 lipid binding domain and the WW1 recognition domain shows dense localization at the cell membrane (fig 5.2.21A) and throughout the cytoplasm. The lack of staining in the nucleus suggests that there is little or no N isoform in the nucleus in both stimulated and unstimulated cells (fig 5.2.21B). Interestingly, the N-isoform fluorescence seem to indicate a filamentous network akin to actin or microtubules which suggest a potential binding or co-localization with these filamentous proteins (fig 5.2.21 inset).



**Figure 5.2.21- Fluorescence microscopy of ARPE-19 cells expressing transfected WWP2-N.** The HA-tagged WWP2-N isoform was detected using primary anti-HA antibody and secondary rat A488 antibody which fluoresces green. The nuclei are shown in blue by the DAPI staining, and images on the right shows both the protein of interest and the nucleus together in the merge of two channels. Transfected cells were stimulated overnight with TGFβ and are compared here against unstimulated controls. A) shows the suggested membrane localisation of WWP2-N whilst B) shows the lack of nuclear localisation. Shown in the inset is the appearance of a filamentous network caused by WWP2-N localisation. Results shown were detected in all cells analysed.

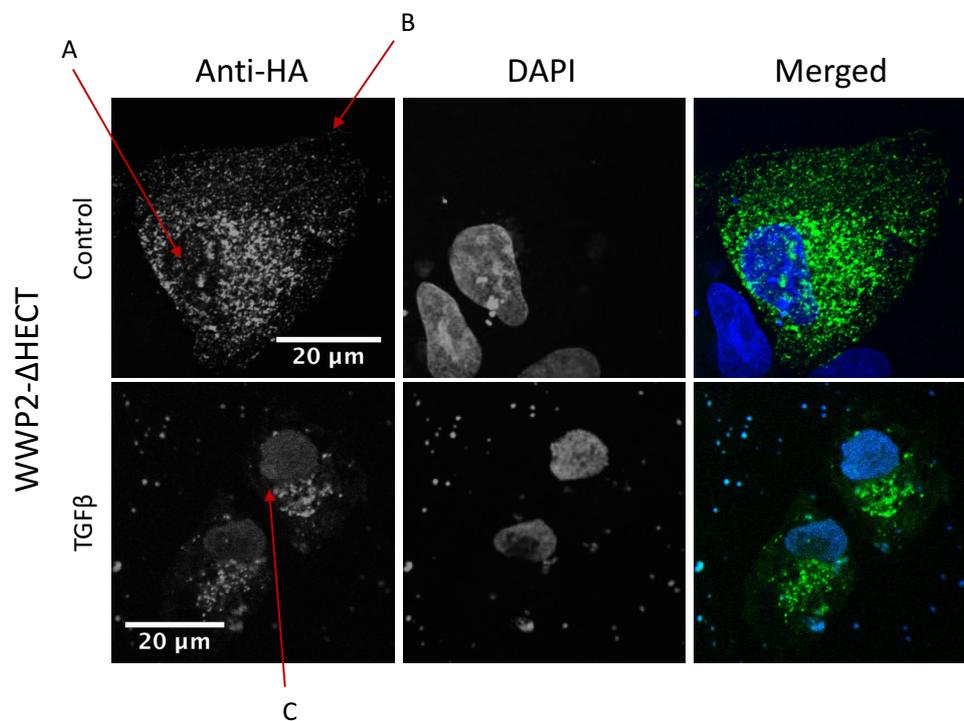
The oncogenic WWP2-C isoform containing the WW4 and HECT domain is shown to localize at the nucleus of the cells shown in fig 5.2.22A. This nuclear localization is not affected by TGF $\beta$  stimulation and interestingly, in the sample stimulated with TGF $\beta$ , the nucleus alone of a separate cell fluoresces green highlighting the absence of this HA-tagged protein in the rest of the cell (fig 5.2.22B). The reason for this trend is unclear, nonetheless this finding further supports the nuclear localization of WWP2-C.



**Figure 5.2.22- Fluorescence microscopy of ARPE-19 cells expressing transfected WWP2-C.** The WWP2-C isoform was HA-tagged and detected using fluorescence microscopy and a combination of primary anti-HA antibody and secondary rat A488 antibody which fluoresces green. The nuclei are shown in blue by the DAPI staining, and images on the right shows both the protein of interest and the nucleus together in the merge of two channels. Transfected cells were stimulated overnight with TGF $\beta$  and are compared here against unstimulated controls. A) shows nuclear staining as a result of HA-WWP2-C localisation whilst B) shows the low levels of WWP2-C localisation in the nucleus of an adjacent cell. Results shown were detected in 67% of cells analysed.

The three novel isoforms WWP2- $\Delta$ HECT, WWP2-N- $\Delta$ C2 and WWP2-FL- $\Delta$ C2 were also transfected into ARPE-19 cells and the effects of TGF $\beta$  stimulation were compared using immunolabelling. These results are shown in fig 5.2.23, 5.2.24 and 5.2.25. First, cells transfected with WWP2- $\Delta$ HECT isoform which contains the WW3, WW4 and an incomplete HECT domain showed that nuclear localization of the protein of interest is unlikely

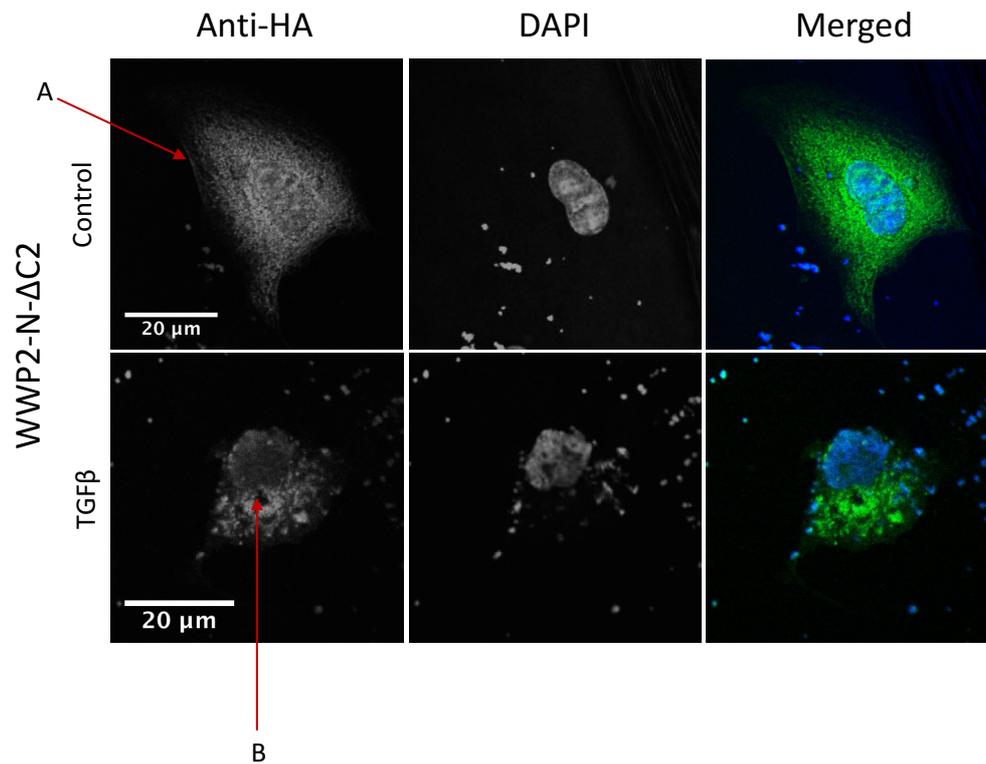
(fig 5.2.23A). This is represented by the lack of staining in the nucleus seen under the green (A488) channel. However, the faint outline of the cell (fig 5.2.23B) suggests that there may be membrane localization of the  $\Delta$ HECT protein in ARPE-19 cells. When stimulated with TGF $\beta$ , localization of protein in the membrane along with most of the cytoplasm decreases dramatically and the levels of protein detected in the nucleus substantially increases (fig 5.2.23C). This is shown by the green fluorescence overlapping with the DAPI staining in the merged image of the cells. Unlike cells transfected with other isoforms, aggregates can also be seen in cells transfected with  $\Delta$ HECT in both stimulated and unstimulated cells.



**Figure 5.2.23- Sub-cellular localization of novel WWP2- $\Delta$ HECT isoform in ARPE-19.** ARPE-19 cells transfected with the HA-tagged novel WWP2- $\Delta$ HECT shown with green Alexafluor488 staining. The nucleus was also stained using DAPI which is shown in blue. Cells were either stimulated with TGF $\beta$  overnight or kept unstimulated as the negative control. The merged pictures with both the DAPI and A488 channels are shown in the right-hand column. A) shows the lack of nuclear staining as a result of low WWP2- $\Delta$ HECT localisation whilst B) shows the suggested membrane localisation of  $\Delta$ HECT. C) highlights the increase in nuclear staining in the presence of TGF $\beta$ . Results are representative of 67% of cells analysed.

WWP2-N- $\Delta$ C2 contains only the WW1 domain and protein localization is shown in fig 5.2.24. Due to the high density of WWP2-N- $\Delta$ C2 present in the cytoplasm, the presence of the isoform cannot be confirmed in the nucleus,

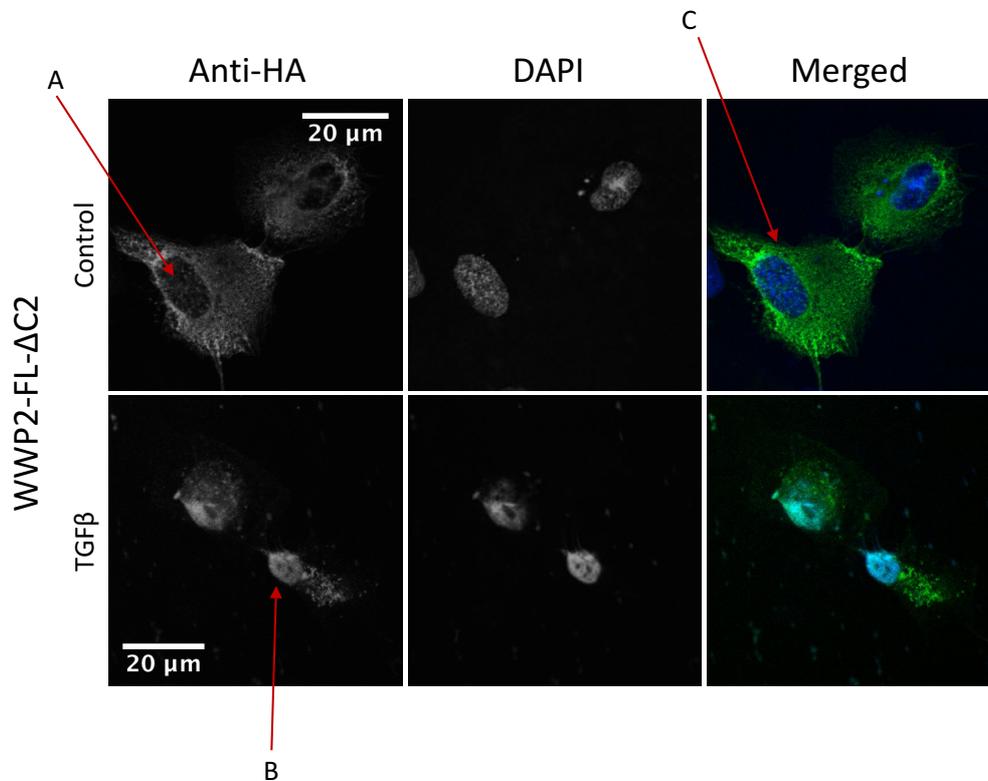
however, the faint outline of the cell suggests some membrane localization present (fig 5.2.24A). When stimulated with TGF $\beta$ , it becomes clear that nuclear localization is preferred over membrane localization as the membrane outline disappears in favour of nuclear localisation (fig 5.2.24B). It is also notable that the level of isoform decreases in the cytoplasm compared with the unstimulated cells.



**Figure 5.2.24- Sub-cellular localization of novel WWP2-N- $\Delta$ C2 isoform in ARPE-19.** The HA-tagged WWP2-N- $\Delta$ C2 isoform was transfected into ARPE-19 cells shown with green Alexafluor488 staining. The nucleus was also stained using DAPI which is shown in blue. Cells were either stimulated with TGF $\beta$  overnight or kept unstimulated as the negative control. The merged pictures with both the DAPI and A488 channels are shown in the right-hand column. A) shows a faint outline of the membrane suggesting membrane localisation of WWP2-N- $\Delta$ C2 whilst B) shows the increased nuclear localisation of WWP2-N- $\Delta$ C2 post TGF $\beta$  stimulation. Results are representative of 100% of cells analysed.

Lastly, the WWP2-FL- $\Delta$ C2 isoform which starts in the same exon as WWP2-N- $\Delta$ C2 but continues to the end of the FL isoform contains all the WWP2 domains except for the C-terminal C2 lipid binding domain. Shown in fig 5.2.25A, this isoform is not localized in the nucleus of unstimulated cells however, upon TGF $\beta$  stimulation, the levels of WWP2-FL- $\Delta$ C2 in the nucleus increases dramatically (fig 5.2.25B). Similar to the WWP2-N- $\Delta$ C2, TGF $\beta$  stimulation negatively regulates the membrane localization of this isoform

as the green outline seen in the control cells (fig 5.2.25C) disappears under TGF $\beta$  stimulation. Once again, cytoplasmic levels of this isoforms also decrease with TGF $\beta$  stimulation.

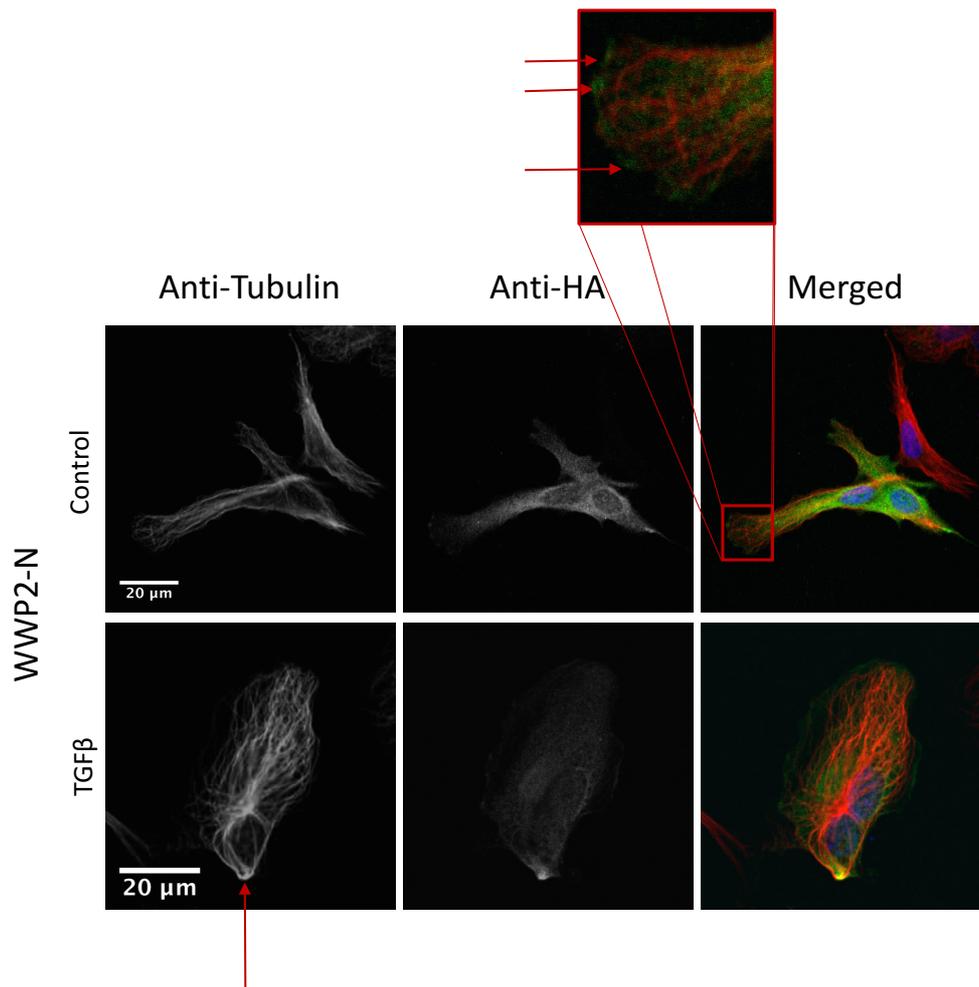


**Figure 5.2.25- Sub-cellular localization of novel WWP2-FL- $\Delta$ C2 isoform in ARPE-19.** ARPE-19 cells transfected with the HA-tagged novel WWP2-FL- $\Delta$ C2 shown with green Alexafluor488 staining. The nucleus was also stained using DAPI which is shown in blue. Cells were either stimulated with TGF $\beta$  overnight or kept unstimulated as the negative control. The merged pictures with both the DAPI and A488 channels are shown in the right-hand column. A) shows the lack of nuclear staining as a result of low WWP2-FL- $\Delta$ C2 localisation whilst B) shows the increased nuclear localisation of WWP2-FL- $\Delta$ C2. C) highlights the suggested membrane localisation of the FL- $\Delta$ C2 isoform. Results are representative of 100% of cells analysed.

#### 5.2.4.2 Investigating the co-localization of WWP2-N and tubulin

Previous results from fig 5.5.2 of the filamentous-like network of WWP2-N in transfected cells suggested a potential co-localization of the N isoform and filamentous proteins. Here, the possibility of WWP2-N binding to a microtubule network is explored by labelling both tubulin and the HA-tagged WWP2-N. Tubulin was labelled with an Alexafluor 647 antibody whilst the WWP2-N was labelled with Alexafluor488 (please refer to section 2.3 for concentrations and protocols used). In fig 5.5.4, shown in red is a clear

microtubule network highlighted by the staining of tubulin, whilst WWP2-N is shown in green. Although these images once again emphasise the lack of WWP2-N localized in the nucleus of both stimulated and unstimulated cells, the HA-tagged N isoform is unlikely to interact or bind to tubulin as there is a lack of similarity in the WWP2-N localization and the tubulin formation. However, an interesting feature shown in some of these cells is that there is an accumulation of WWP2-N at the leading edges which could suggest potential co-localisation with other filamentous proteins such as actin.



**Figure 5.2.26-** A cellular localization study that explores the relationship between WWP2-N and Tubulin. Using primary anti-HA, secondary A488 antibodies, along with an Alexafluor647 tubulin stain, images above were produced of ARPE-19 cells transfected with WWP2-N, showing WWP2-N-HA in green and tubulin in red whilst comparing cells stimulated with and without TGF $\beta$ . The merges images on the right-hand column contains the green, red fluorescence and the blue DAPI staining. Although the filamentous network of WWP2-N shown previously is not obvious here, arrows show potential localisation of WWP2-N at leading edges of the cells. Trends highlighted were observed in 80% of cells analysed.

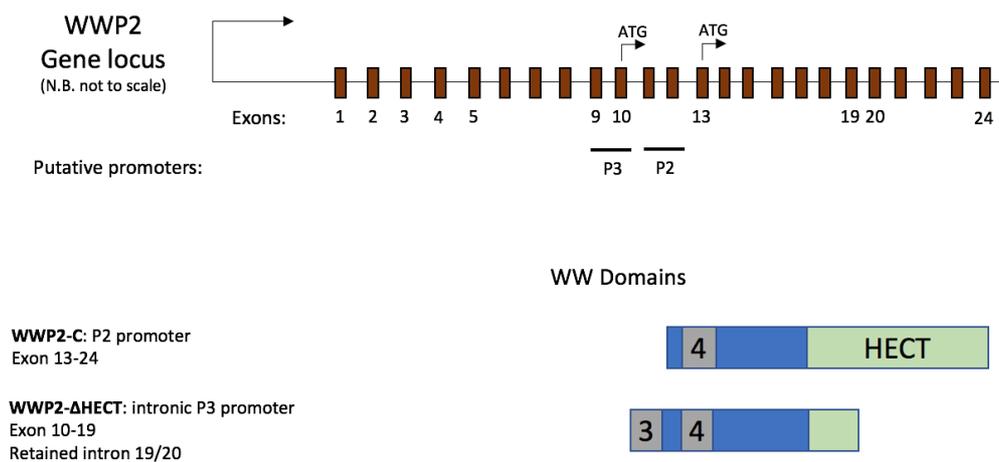
### 5.3 Discussion

In this chapter, different aspects of the transcription mechanism regulating expression of WWP2-C, N and  $\Delta$ HECT isoforms were investigated. The aim was to identify the potential enhancers and repressor of WWP2-C and  $\Delta$ HECT transcription and look at the significance of SOX9 transcription factor with regards to WWP2-C promoter activity. Thirdly, the functional aspect of the novel  $\Delta$ HECT isoforms was studied by using luciferase assays to investigate the Smad-3 dependent gene expression of this isoform and the potential effect of TGF $\beta$  on its gene activity. The regulation of intron retention that results in the transcription of WWP2-N was investigated by studying the role of ESRP splicing regulatory proteins on the mRNA expression of WWP2-N using a combination of RT-PCR and qPCR. Finally, cellular sub-localization studies were performed on WWP2 isoform transfected cells using immunolabelling techniques to visualize the cell localization of isoforms in the presence and absence of TGF $\beta$ . The overall aim of this chapter was to broaden the understanding in transcriptional regulation of WWP2 truncated isoforms which will hopefully expose the role WWP2 isoforms have in TGF $\beta$ -mediated oncogenesis.

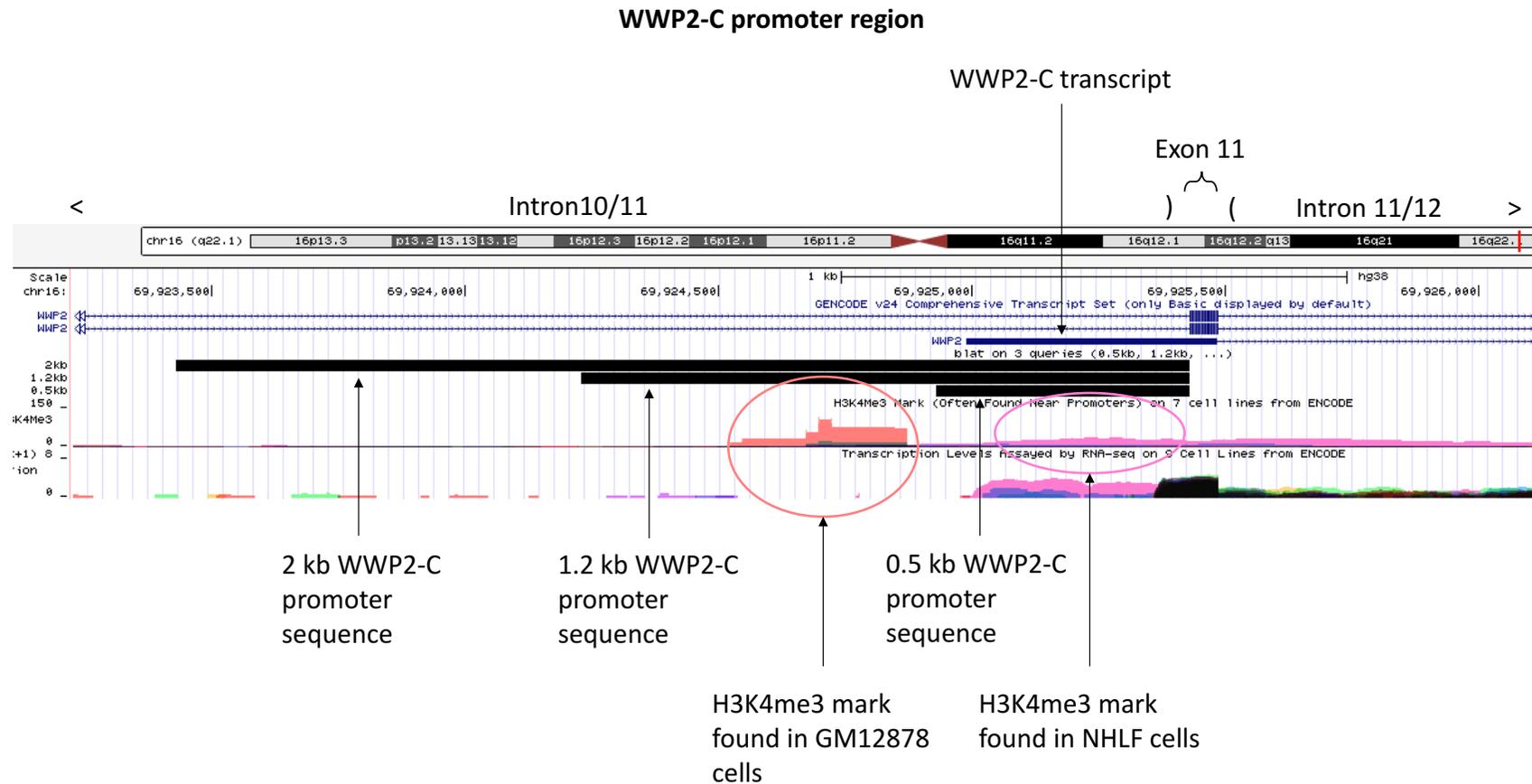
#### 5.3.1 Regulation of WWP2-C promoter region and the relevance of SOX9 transcription factor

WWP2-C is a truncated version of WWP2-FL that is considered an oncogene due to its composition of WW4 recognition domain and intact HECT ligase domain theoretically allowing the increased degradation of natural inhibitor Smad7 of the TGF $\beta$  pathway. As shown in fig 5.3.1, this isoform spanning from exon 13-24 is a result of a putative promoter (P2) predicted within intron 10/11. Three different length constructs (0.5, 1.2 and 2 kb) within the predicted promoter sequence upstream of the transcriptional start were cloned into luciferase reporter vectors and used in luciferase reporter assays. Results showed that all three constructs produced a significant fold change relative to basal that suggest a potential enhancing property of the cloned region. An analysis commonly used to detect promoter activity is the

H3K4me3 test. H3K4me3 is a modification that involves the methylation of Lys4 within histone H3 which occurs at the start of transcription often near a promoter site (Guenther *et al*, 2007). Upon analysis of H3K4me3 marks of 7 cell lines from ENCODE using the UCSC genome shown in fig 5.3.2, it was found that a region of high promoter activity circled in purple was detected in the normal human lung fibroblast (NHLF) cell line. This detected promoter activity correlates to the position of the predicted 0.5 kb WWP2-C promoter region in our studies and therefore supports the enhancing effect of the cloned region and suggests that the WWP2-C promoter might lie within the 0.5 kb region. However, circled in red, is a much higher level of promoter activity detected in GM12878 lymphoblastoid cells which might suggest the minimal promoter sequence needed for WWP2-C is in the 1 kb region rather than the predicted 0.5kb region.



*Figure 5.3.1- A schematic of the transcription mechanism involved in WWP2-C and  $\Delta$ HECT production. Above shows a brief summary of the WWP2 gene locus and the domains present in these truncated isoforms.*



*Figure 5.3.2- The depiction of predicted WWP2C-promoter sequences within intron 10/11 of the WWP2 gene locus and H3K4me3 marks using the UCSC genome browser. Circled in red and purple are high probability regions of H3K4me3 highlighted using the H2K4me3 test on 7 cell lines from ENCODE. The increased H3K4me3 circled in red is observed in the GM12878 cell line whilst the promoter activity circled in purple was found in Normal Human Lung Fibroblast (NHLF) cells.*

Within the predicted promoter region for WWP2-C isoform transcription, there lies a SOX9 transcription factor binding site (-CCTTGAG-) and due to the established links between SOX9 and melanoma development (Passeron *et al*, 2009), we decided to test the effects of SOX9 on WWP2-C transcription. To do this, a SOX9 construct was co-transfected into HEK293A cells with the predicted C-promoter constructs and cell lysates used in luciferase reporter assays. Interestingly, results from SOX9 transfected cells showed a staggering decrease in WWP2-C transcriptional activity of more than 70 fold for all three predicted promoter constructs suggesting that SOX9 binds to a binding site within the promoter sequence to decrease the promoter activity and therefore the expression of WWP2-C. According to previous studies by Passeron *et al*, the increase in SOX9 expression inhibits the progression of melanomas in both human and mice specifically by altering the cell proliferating potential of melanoma cells (Passeron *et al*, 2009). This supports the hypothesis that WWP2-C acts as an oncogene due to the negative effects SOX9 has on C expression and subsequent melanoma development, thus further highlighting the importance of the relationship between SOX9 transcription factor and WWP2-C in oncogenesis.

To confirm the expression of endogenous SOX9, and therefore the validity of the hypotheses above, RT-PCRs were performed on six different melanoma cell lines using SOX9 specific RT-PCR primers. Although TGF $\beta$  has a known dual role in oncogenesis, TGF $\beta$  stimulation was used in our study to mimic cancer progression by performing 16 hr stimulations on cells prior to RNA extraction, as malignant cells have been found to bypass the tumour-suppressive effects of TGF $\beta$  (Hussein, 2004).

Under TGF $\beta$  stimulation, the levels of SOX9 were not significantly affected in melanoma cell lines except for in M202 melanoma cells where a decrease in SOX9 expression was seen following TGF $\beta$  stimulation. Intriguingly, this cell line is the only one in the panel of six to express a Q61L mutation in N-RAS (shown in table 5.3-1). Studies have shown that N-RAS<sup>Q61R</sup> causes an overactive N-RAS, activating the AKT/P13K pathway which results in the

phosphorylation and mis-localization of FOXO transcription factor. FOXO is therefore unable to bind to its transcription co-factor Smads leading to mis-regulation in p15 and p21 which ultimately causes the N-RAS<sup>Q61R</sup> cell to become resistant to TGFβ senescence in melanoma (Lasfar and Cohen-Solal, 2010). Although the N-RAS mutation mentioned in the above study is not identical to the one expressed in M202, both mutations are amino acid substitutions that occur at position 61, altering the folding of the protein which results in the constant GTP-bound state of N-RAS (Jovanovic *et al*, 2010). We therefore postulate that N-RAS<sup>Q61R</sup> and N-RAS<sup>Q61L</sup> have similar effects on the resistance to TGFβ stimulation reflected by the resistance to TGFβ-mediated increase in SOX9 levels in the SOX9 RT-PCR expression study.

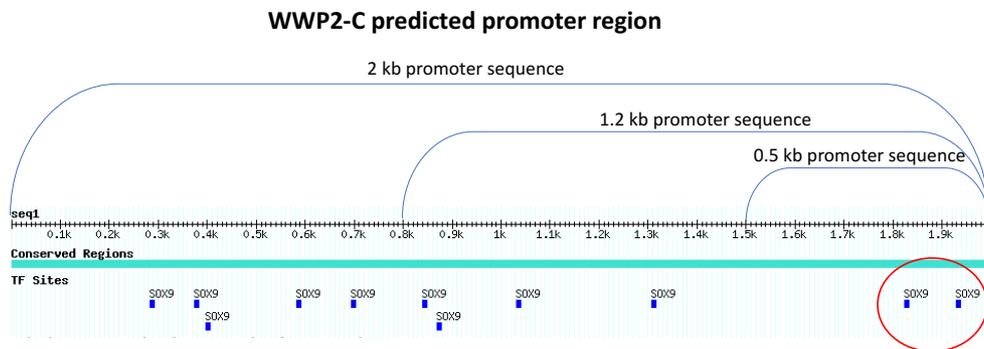
*Table 5.3-1- A table showing the NRAS statuses of melanoma cell lines used in SOX9 RT-PCR*

<b>Melanoma Cell line</b>	<b>NRAS status</b>
A2058	WT
A375	WT
SKMEL28	WT
M202	Q61L
M229	WT
MEL501	G12D

The role of TGFβ in the transcriptional activity of WWP2-C isoform was also investigated by repeating the above experiment with transfected HEK293A cells stimulated with TGFβ 18 hr prior to cell harvest. Results suggest that TGFβ stimulation negatively affects WWP2-C transcriptional activity which could be explained using the dual role of TGFβ in cancer. It has long been established that TGFβ acts interchangeably as a tumour suppressor and an oncogene dependent on the stage of tumorigenesis. In normal cells, TGFβ acts as a tumour suppressor whilst in malignant cells it has an oncogenic effect (Hussein, 2004). The decrease in the predicted WWP2-C oncogene transcriptional activity following TGFβ stimulation can therefore be explained by the tumour suppressant properties TGFβ possess in the non-malignant HEK293A cell line.

To test SOX9 binding within the predicted WWP2-C promoter region, a mutant 0.5 kb construct was created that featured a complete deletion of the SOX9 binding site (-CCTTGAG-). This construct was then transfected into HEK293A, A375 and MCF-7 cells and luciferase reporter assay was performed to compare results of wild type and mutant transcriptional activity in presence of naturally occurring SOX9 in the cell. Results from the HEK293A C-promoter mutant experiment showed that the mutant WWP2-C transcriptional activity increased compared with the wild type. Although this increase was not statistically significant, it suggests that the obliteration of the SOX9 binding site does have a negative effect on the SOX9 binding as the negative regulation of WWP2-C transcriptional activity by SOX9 was reversed by the mutant. Furthermore, results from the same experiment repeated in the epithelial breast cancer cell line MCF-7, also showed an increase in WWP2-C transcriptional activity when transfected with the SOX9 binding site mutant construct suggesting a positive trend across both cell lines indicating that the mutation of the predicted SOX9 binding site affects WWP2-C transcriptional activity via SOX9 binding inhibition. This further supports the -CCTTGAG- sequence as the SOX9 binding site, which is obliterated in the mutant and therefore reduces the SOX9-mediated negative regulation of WWP2-C. However, although an increase was observed in presence of the SOX9 binding site deletion, neither results from HEK293A nor MCF-7 were statistically significant. This perhaps suggests multiple SOX9 binding sites present, allowing SOX9 regulation through other binding sites that were not targeted through the -CCTTGAG- binding site deletion. In an article by Mertin *et al*, it was found that SOX9 not only binds to the -CCTTGAG- binding site, but -AGAACAATGG- was also found to be a SOX9 binding sequence (Mertin *et al*, 1999). Although this sequence was not found in the promoter region of the cloned 0.5 kb construct, the presence of another SOX9 binding site is possible. Using the ORCA<sub>t</sub>f tool that detects transcription factor binding sites within target sequences using phylogenetic foot printing, two other putative SOX9 binding sites, -TTATTGATT- and -AAACAAAA- were detected shown in fig 5.3.3 (Portales-Casamar *et al*, 2009). This indicates that one of the two binding sites might

therefore be responsible for the SOX9 binding that occurs in our SOX9 mutant construct. The obliteration of -CCTTGAG- in the mutant construct may therefore not have influenced SOX9 binding as there is a potential for another SOX9 binding site within the recombinant mutant promoter region.



**Figure 5.3.3-Putative SOX9 binding sites in the predicted WWP2-C promoter region.** Results from SOX9 binding site detection using ORCAtk showing that within the 0.5 kb promoter region, two SOX9 binding sites excluding -CCTTGAG- were predicted highlighted in red (Portales-Casamar et al, 2009).

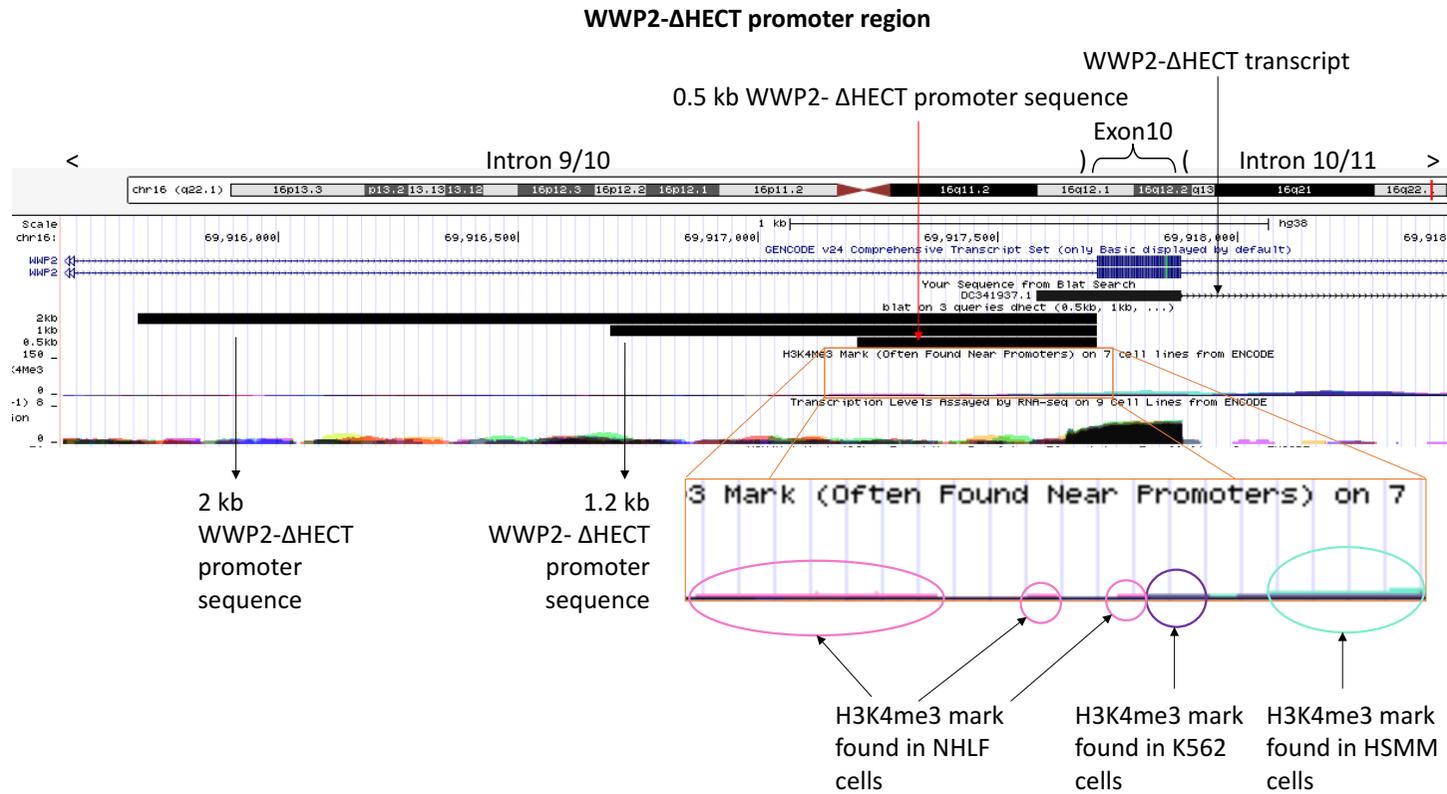
Finally, results from the WWP2-C mutant promoter luciferase analysis using A375, an epithelial melanoma cell line expressing mesenchymal properties showed a decrease in WWP2-C transcriptional activity suggesting that the negative effect of SOX9 was not affected by the binding site deletion. This result is the opposite of the findings from HEK293A and MCF-7 cells which suggests that the epithelial/mesenchymal phenotype might have a role to play in the SOX9 regulation of WWP2-C. As mentioned above, there are two other predicted SOX9 binding sites located within the 0.5 kb WWP2-C promoter region (fig 5.3.3). The negative effect SOX9 has on mutant WWP2-C transcriptional activity was only observed in A375, an epithelial cell line with strong mesenchymal characteristics. Consequently, this might imply that the use of SOX9 binding sites differ in different cell lines with the predicted SOX9 sites -TTATTGATT- and -AAACAAAAA- favoured in cells expressing a mesenchymal phenotype, and perhaps -CCTTGAG- binding site is preferred in cells with an epithelial phenotype. Although preference of SOX9 binding sites specifically have not been established, studies have shown that binding sites of serum response factor (SRF) varies in three different human cell types; a lymphoblast cell line, a neuroblastoma cell lines

and a smooth muscle cell line (Cooper *et al*, 2007). This suggests that the shift in preferences for SOX9 binding sites might be a possible mechanism of an added level in WWP2-C transcriptional regulation in oncogenesis. Alternatively, another hypothesis that might explain the inability of SOX9 binding site deletion to inhibit the negative effect of SOX9 on WWP2-C transcriptional activity, is that SOX9 might be binding to other proteins to form a complex that indirectly regulate c-promoter activity. A study by Furumatsu *et al* found that Smad3 and p300 are needed to form a functional complex with SOX9 in order to activate the transcription of chromatin (Furumatsu *et al*, 2009). The negative effect of SOX9 on WWP2-C transcription activation in our mutant study despite the deletion of SOX9 binding site may therefore be due to the association of SOX9 and Smad mediators indirectly regulating WWP2-C via binding to the Smad binding element (SBE) instead of the deleted SOX9 binding site.

### 5.3.2 Investigating the regulation of WWP2- $\Delta$ HECT promoter activity and the effects of growth factor stimulation

WWP2- $\Delta$ HECT is a truncated WWP2 isoform that arises due to the putative intronic promoter P3 shown in fig 5.3.1. The transcription of this isoform starts at exon 10 and ends at exon 19 due to a stop codon present in the retained intron 19/20. Similar to the WWP2-C transcription regulation study, three constructs containing parts of the predicted promoter region were initially cloned into the luciferase reporter vector pGL4.27 and used in luciferase reporter assays. However, the 0.5, 0.7 and 1 kb constructs that were cloned upstream from 3' region of intron 9/10 did not seem to produce high transcriptional activity therefore a 2 kb predicted  $\Delta$ HECT promoter constructs was cloned and used in further luciferase assays (shown in fig 5.2.10). Results suggests that all four constructs including the most recently cloned 2 kb construct, did not induce a high basal luciferase reporter activity indicating that the enhancing properties of these promoter regions are low. The H3K4me3 levels were analysed from ENCODE using the UCSC genome browser. High levels of H3K4me3 modification indicates transcription start sites related to possible promoter activity (Guenther *et al*, 2007). Results

taken from the UCSC analysis of the predicted  $\Delta$ HECT promoter region shown in fig 5.3.4 suggests that although H3K4me3 levels are low, indicating low promoter activity, H3K4me3 was detected in three of the seven cell lines from ENCODE including NHFL (in pink), K562 leukaemia cell line (in purple) and Human skeletal muscle cell line HSMM (in turquoise). Interestingly, TGF $\beta$  stimulation on HEK293A cells showed a decrease in transcriptional activity. Due to established research showing the tumour-suppressive effects TGF $\beta$  has in non-cancerous cells, our result might indicate that if these constructs were indeed able to enhance  $\Delta$ HECT transcriptional activity, then like WWP2-C, the negative effect TGF $\beta$  has on this isoform in a normal cell line may indicate the oncogenic role of WWP2- $\Delta$ HECT (Hussein, 2004).



**Figure 5.3.4- The predicted WWP2-ΔHECT promoter region and highlighted H3K4me3 marks.** A schematic derived using the UCSC genome browser showing the predicted WWP2-ΔHECT promoter regions within intron 9/10 of the WWP2 gene locus and the WWP2-ΔHECT transcript. H3K4me3 analysis from ENCODE (shown in the magnified section) also suggest promoter activity in the 0.5 kb region of the predicted promoter in NHFL cells circles in pink, K562 cells circled in purple and HSMM cells circled in turquoise.

Due to the lack of effect WWP2- $\Delta$ HECT cloned promoter regions had on transcription activity under control conditions and TGF $\beta$  stimulation, the role of other growth factors was then investigated by repeating the above experiments and stimulating transfected cells with EGF and BMP 18 hr before harvest. Results show that BMP stimulation has minimal effect on  $\Delta$ HECT transcriptional activity which reflects the lack of GC-rich Smad binding elements (GC-SBE) needed for BMP signal transduction (Morikawa *et al*, 2011) in the predicted  $\Delta$ HECT promoter region. TGF $\beta$  stimulation on the other hand did not show significant changes to luciferase activity.

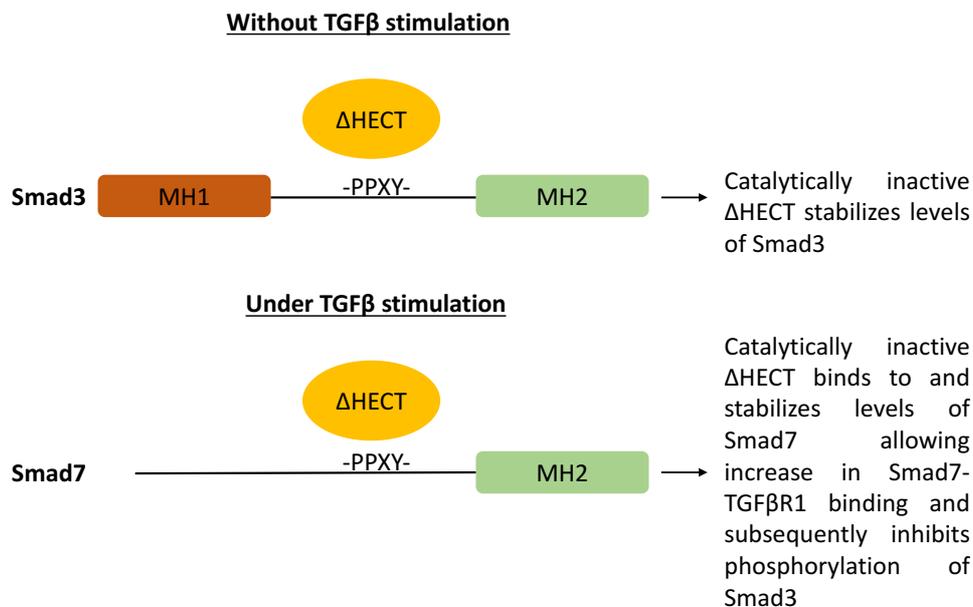
Finally, EGF was used to stimulate cells to study the effects of this growth factor on WWP2- $\Delta$ HECT transcriptional activity. Results show an increase of transcriptional activity compared to the control samples using non-stimulated cells suggesting that the EGF signalling pathway may have a role in the induction of WWP2- $\Delta$ HECT. According to Akimov *et al*, this is not at all unlikely as it was found in their study that EGF regulates WWP2 production among three other E3 ligase family members (Akimov *et al*, 2011). Moreover, studies have also shown the transcriptional activation of the CYP11A1 promoter by EGF can be mediated by the binding of c-Jun to AP-1 like sites (Pestell *et al*, 1995). Due to the large amount of AP-1 binding sites present in the predicted  $\Delta$ HECT promoter region (shown in fig 5.2.10), it can be surmised that EGF maybe able to stimulate  $\Delta$ HECT transcriptional activity via its binding to AP-1 sites.

### 5.3.3 The effects of WWP2- $\Delta$ HECT overexpression on Smad3-dependent gene expression

The novel WWP2- $\Delta$ HECT isoform containing the WW3, WW4 and an incomplete HECT catalytic domain was cloned and along with the p-CAGAC12-luc Smad3-dependent gene transcription reporter, it was transfected into HEK293A and A375 cells. These were then used in luciferase assays and results showed that in the absence of TGF $\beta$ , transfected  $\Delta$ HECT increases levels of Smad3 activity in both cell lines whilst in the presence of TGF $\beta$ , Smad3 activity dramatically decreases in HEK293A cells. Due to the

lack of a complete HECT domain, it is not certain whether this isoform has intact ligase activity. However, although WW4 domain was suggested to bind to Smad7, it was found in a recent NMR study that the WW4 domain also binds to Smad2 and Smad3 (Wahl, 2016) therefore we argue that the increase in Smad3-dependent activity caused by transfected  $\Delta$ HECT in absence of TGF $\beta$  can be explained by the binding and stabilization of Smad3 by  $\Delta$ HECT. WWP2- $\Delta$ HECT therefore becomes a competitive inhibitor of functional WWP2-Smad7 binding and subsequent ubiquitin-dependent degradation of Smad7.

In the presence of TGF $\beta$ , luciferase results of cells transfected with  $\Delta$ HECT show a significant decrease in Smad3-dependent gene activity. According to the NMR study mentioned above, the binding affinity of WW4 to phospho-Smad7 is much higher than Smad3 therefore suggesting that this novel isoform containing the WW4 domain binds to and stabilizes levels of phospho-Smad7 allowing for the decrease in levels of Smad3-dependent gene activity by binding to TGF $\beta$ R1 and inhibiting the phosphorylation of Smad2/3. Together with western blot analysis using anti-Smad antibodies highlighting the expression of both Smad3 and Smad7 in cells overexpressed with WW3-4 (shown in fig 3.2.9), it can be surmised that the incomplete HECT domain within the  $\Delta$ HECT isoform is catalytically inactive due to the binding and stabilization of phospho-Smad7 in TGF $\beta$  positive conditions, and the stabilization of Smad3 in control conditions shown in fig 5.3.5.



*Figure 5.3.5- The suggested preferential binding partners of  $\Delta$ HECT in presence and absence of TGF $\beta$  and its effects on Smad3 dependent gene transcription.*

In A375 melanoma cells, a similar increase in Smad3 dependent gene activity was detected in the absence of TGF $\beta$ , however, when subjected to TGF $\beta$  stimulation, the Smad3 dependent activity increased by 1.5 fold instead of the expected decrease. A reasonable explanation for this may be to do with the cell origin. A375 are a melanoma cell line which differ from HEK293A due to its malignancy. It has also been suggested that melanoma cell lines themselves produces increased levels of TGF $\beta$  meaning that the levels of TGF $\beta$  present in the cell medium may have been a lot higher than in the HEK293A cells (Medrano, 2003). Consequently, the now higher levels of Smad2/3 resulting from the naturally higher levels of TGF $\beta$  secretion in melanoma masks the effects of lower levels of transfected  $\Delta$ HECT therefore the stabilized levels of Smad7 is out-competed by the high levels of Smad2/3, leading to the phosphorylation and subsequent increase in Smad3-dependent activity. Alternatively, the increased levels of Smad3-dependent activity post TGF $\beta$  stimulation of melanoma cells can be explained by the dual role of TGF $\beta$  in cancer. In normal cells, TGF $\beta$  acts as a tumour suppressor which is demonstrated in our result by the negative effect it has on Smad3-dependent activity. However, in oncogenic cells, TGF $\beta$  acts as an oncogene which is also reflected in our result as TGF $\beta$  stimulation in A375 no longer reduce Smad3-dependent activity.

#### 5.3.4 Studying the role of ESRP1 and 2 as a post-transcriptional mechanism in WWP2-N production

WWP2-N is an isoform containing the N-terminal C2 domain and the WW1 recognition domain. The production of this isoform is a result of a premature stop codon present in the retained intron 9/10, however the regulation of this alternate splicing event is yet unclear. ESRPs are proteins that regulate splicing events and has been shown to be heavily implicated in EMT-related proteins (Warzecha *et al*, 2010). Due to the role WWP2 isoforms have in TGF $\beta$ -mediated EMT, we therefore hypothesize that alternate splicing activities that result in truncated WWP2 isoforms are regulated by ESRPs. To investigate this hypothesis, RNA was extracted from A375 and MCF-7 cells transfected with ESRP1 and 2. These were then converted into cDNA using GoScript PCR and used in RT-PCR and qPCR.

In A375 melanoma cells, ESRP was shown to down-regulate the expression of WWP2-N reflected by both the RT-PCR and the qPCR data. RT-PCR was a semi-quantitative method used to detect the expression of isoforms whilst qPCR was used as a quantitative measure. The combination of these two PCR methods will provide an accurate prediction of isoform mRNA expressions. Another interesting finding is the decrease of WWP2- $\Delta$ HECT expression in response to transfected ESRP. This decrease may suggest the role ESRPs play in the expression of  $\Delta$ HECT by regulating the retention of intron 19/20. Similarly in MCF-7, a breast cancer epithelial cell line, RT-PCR and qPCR results both showed that WWP2-N expression was down-regulated by ESRP. Furthermore, the expression of  $\Delta$ HECT also showed a decrease in response to transfected ESRP, although not shown in RT-PCR results. The regulation of splicing events by ESRPs can either have a positive or negative effect on the production of isoforms as ESRP has the ability to either silence an exon or enhance the inclusion of an exon therefore favouring transcription of the full-length proteins (Warzecha *et al*, 2009b). This provides evidence suggesting that ESRPs might play a role in negatively regulating the transcription of WWP2-N and  $\Delta$ HECT by reducing the

retention of intron 9/10 and 19/20 respectively, leading to the decline in levels of isoform mRNA detected as shown in the PCR results.

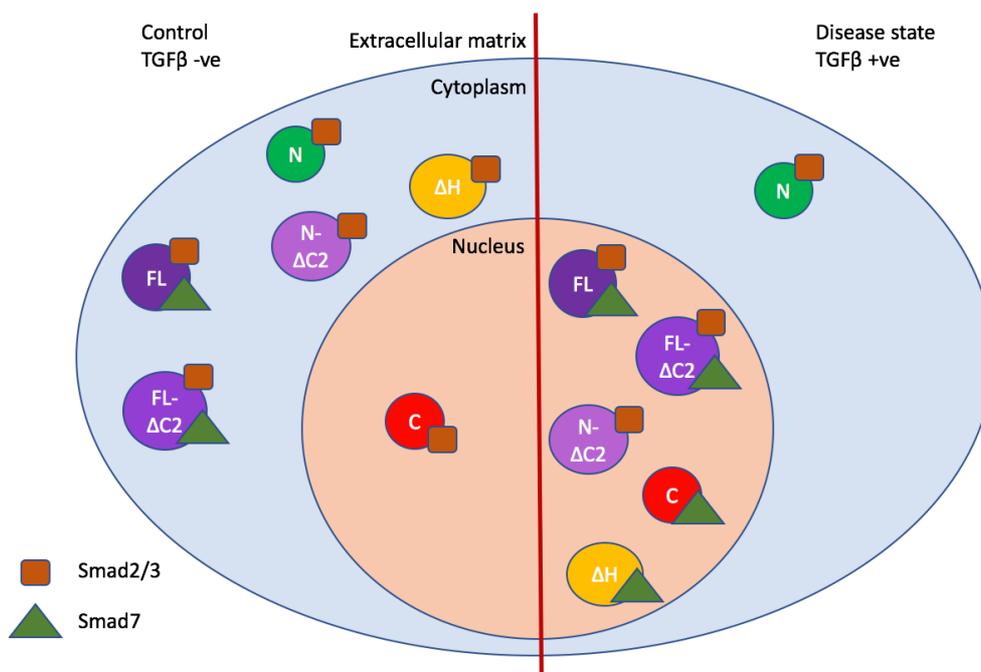
#### 5.3.5 Cellular localization of WWP2 isoforms and the hypothesized Smad interactions

Studying the cellular localization of these isoforms in depth allowed us to hypothesize their relationships with Smad proteins using existing knowledge on structural properties regarding the isoforms. Firstly, it is important to bear in mind that from previous studies (Soond and Chantry, 2011), it was found that the WW1 domain in WWP2 binds preferably to Smad2/3 whilst the WW4 domain binds to Smad7. Recently, it has been suggested that WW4 not only has preferential binding to phospho-Smad7 but was also shown to have a greater binding affinity to native-Smad3 over native-Smad7 (Wahl, 2016). The effect of TGF $\beta$  on the localization of isoforms can be explained using two factors; the presence of the C2 lipid binding domain and the combination of WW recognition domains expressed.

In a study on Smad expression in pulmonary fibrosis, Bleomycin was used to induce inflammation and fibrosis. It was found that the administration of Bleomycin increases TGF $\beta$ 1 levels causing the increased translocation of phosphorylated Smad2/3 into the nucleus which then induces transcriptional activity. It was also found that the cytosolic Smad3 levels dropped in the disease state (caused by the increase of TGF $\beta$  levels from administration of Bleomycin) whilst levels of Smad3 in the nucleus increased. Studies on the regulatory role of Smad7 have suggested that Smad7 binds to Smad2/3 in the cytosol to prevent its translocation into the nucleus where it can aid transcription (Venkatesan *et al*, 2004).

Using our knowledge on Smad localization in the above disease state, we can postulate that the stimulation of TGF $\beta$  in our study mimics the effects it has in melanoma. Figure 5.3.6 is a simplified schematic of the findings in the WWP2 localization study paired with the previously hypothesized Smad interaction. Firstly, WWP2-FL was seen to translocate from the membrane

to the nucleus when stimulated with TGF $\beta$  and due to presence of all the WW domains in the full-length isoform, it can bind to both Smad 2/3 and Smad7. However, it was found in previous studies that TGF $\beta$  stimulation shifts preferential binding of FL to Smad7 instead of Smad2/3 (Soond and Chantry, 2011). Together, this suggests that WWP2-FL is a tumour suppressor in the control state as it binds to and aids the degradation of Smad2/3 whilst in the disease state, mimicked by the stimulation of TGF $\beta$ , FL acts as an oncogenic factor by binding to Smad7 to disrupt the inhibitory properties of Smad7. The formation of this complex also stops Smad7 from binding to DNA which is another mechanism in which Smad7 acts to negatively regulate the TGF $\beta$  pathway (Yan *et al*, 2009).



*Figure 5.3.6- A simplified schematic representation showing the cellular localization of WWP2 isoforms and the respective Smad binding in normal and disease states.*

The expression of the C isoform was seen predominantly in the nucleus, regardless of TGF $\beta$  stimulation. The lack of membrane localization is supported by the fact that this isoform lacks the C2 lipid binding domain. Its role as an oncogene is highlighted as the expression of only the WW4 and HECT domain allows it to bind to and inhibit Smad7 in the nucleus when stimulated with TGF $\beta$ . However, in the control state, the WW4 domain present in WWP2-C should have increased affinity to native-Smad3 suggesting that it may have a TGF $\beta$ -dependent dual function in oncogenesis.

A shift from cytoplasmic to nuclear localization of WWP2- $\Delta$ HECT was seen in response to TGF $\beta$  stimulation. Due to an incomplete HECT domain expressed in  $\Delta$ HECT, it is not clear whether it can carry out its catalytic activities and aid proteasomal degradation of Smad7. However, if our previous predictions are correct and that this isoform is indeed catalytically inactive, then the preferential binding to native-Smad3 in absence of TGF $\beta$  could stabilize and aid TGF $\beta$ -dependent transcription whilst the  $\Delta$ HECT preferential binding to phospho-Smad7 in the presence of TGF $\beta$  could lead to the stabilization of Smad7 therefore inhibiting the oncogenic effect of TGF $\beta$ . It is for this reason that  $\Delta$ HECT might possess both oncogenic and tumour suppressive properties in oncogenesis.

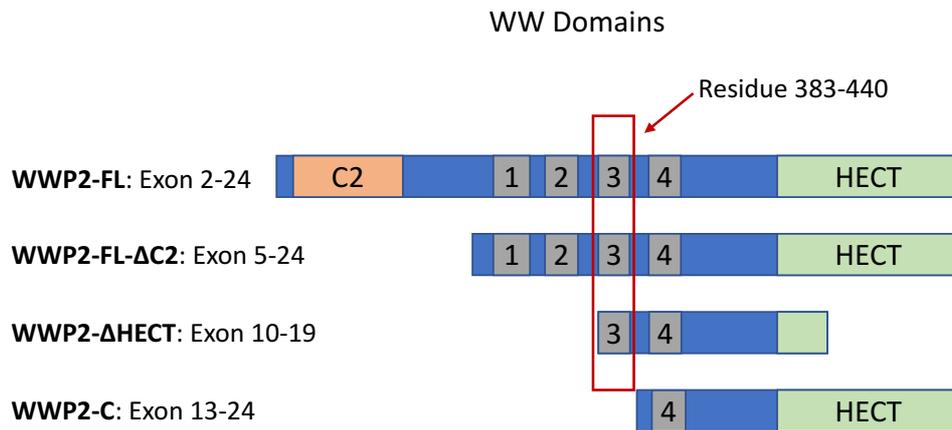
WWP2-N- $\Delta$ C2 and FL- $\Delta$ C2 were both expressed at very low levels at the membrane which could be due to the lack of C2 domain. Both these isoforms demonstrated an increase in nuclear translocation from the cytoplasm but may have very different effects in oncogenesis due to the difference in domain composition. The N- $\Delta$ C2 isoform lacks the HECT catalytic domain therefore will bind to and stabilize the levels of Smad2/3 which suggests that it has a potential oncogenic effect. On the other hand, WWP2-FL- $\Delta$ C2 has all the WW recognition domains which means that it can bind to Smad2/3 and Smad7. Together with an intact HECT domain, this isoform has the power to transport Smads for degradation. The WWP2 isoform can aid proteasomal degradation in both the nucleus and the cytoplasm as proteasomes are expressed throughout the cells (Wójcik and DeMartino, 2003). Currently, the preferential binding of FL- $\Delta$ C2 to Smads has not been discovered, however, if TGF $\beta$  affects preferential binding of FL- $\Delta$ C2 in the same way as it does to WWP2-FL, then we can postulate that it is a tumour suppressor protein.

The WWP2-N localization study produced a few interesting findings. First, the lack of staining in the nucleus shows the lack of WWP2-N expression in the nucleus regardless of TGF $\beta$  stimulation, and because of its C2 domain, there was a very dense membrane localization shown by the clear outline of the cell. Results showing localization of N in the cytoplasm could mean that

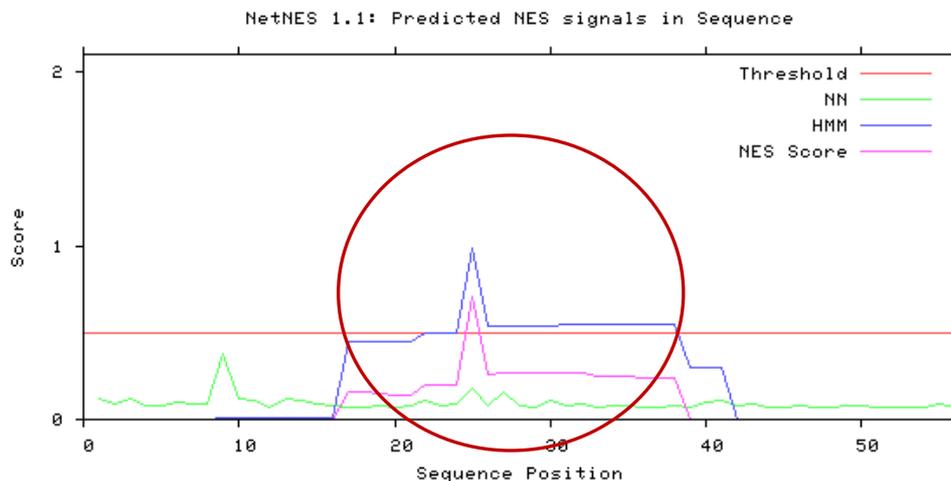
the binding of WWP2-N and Smad2/3 via the WW1 domain leads to the retention of Smad2/3 in the cytoplasm and the prevention of translocation into the nucleus therefore suggesting that WWP2-N is a tumour suppressor. However, there is also a possibility that WWP2-N can act as an oncogene. The phosphorylation of Smad2/3 and subsequent binding with Smad4 occurs in the C-terminal MH2 domain of Smad2/3 (Yan *et al*, 2009). This means that WWP2-N binding with PPXY motif in the linker region is unlikely to hinder the formation of Smad2/3-4 complex therefore allowing its translocation and ultimately gene transcription in the nucleus. Although this is true, the lack of N isoform nuclear localization shown in the immunolabelling experiments suggest that this novel isoform binding with Smads is inhibiting at least a proportion of activated co-Smad/R-Smad complexes from entering the nucleus, subsequently favouring its role as an anti-oncogenic protein. Another exciting finding that emerged from the WWP2-N localization study was the appearance of a filamentous network produced by the N expression. Previous studies have shown the binding of Smad2/3 and microtubules in several epithelial and endothelial cell lines (Dong *et al*, 2000), so here, we hypothesized that WWP2-N binds to Smad2/3 which co-localizes with tubulin. Unfortunately, after labelling both the WWP2-N isoform and tubulin (fig 5.5.4), the filamentous network previously shown with N expression was no longer clear in comparison to the tubulin network.

In addition, a recent immunolabelling study conducted by Zhu *et al* visualized the localization of WWP2 FL and WWP2 mutants similarly to the experiments carried out here. Dependent on the domain composition of WWP2 isoforms, they were able to postulate certain effects that specific domains had on nuclear translocation (Zhu *et al*, 2017). One interesting conclusion that was found suggested that the HECT catalytic domain of WWP2 contains a region that negatively regulates its translocation into the nucleus. This means that all isoforms containing the HECT domain (including the FL, C, FL- $\Delta$ C2 and potentially the  $\Delta$ HECT isoform) should produce a higher fluorescence in the cytoplasm rather than the nucleus. When comparing to

the results of our immunolabelling studies, it was found that FL, FL- $\Delta$ C2 and  $\Delta$ HECT indeed was absent or had lower levels of localisation in the nucleus, reflected by the lack of staining. Interestingly, the absence of  $\Delta$ HECT from the nucleus could also suggest that the sequence responsible for this negative regulation on nuclear translocation, possibly a nuclear export sequence (NES) is located between residue 1-171 of the HECT domain, the region present in the  $\Delta$ HECT isoform. Previous results from fluorescence microscopy of the WW3-4 tandem domain in chapter 3 highlighted a dense localization within the nucleus. The contrast in cellular localization between the HECT containing isoforms and the WW3-4 domain lacking the HECT domain further supports the presence of a NES in the HECT domain (Zhu *et al*, 2017). However, previous analysis of the WWP2 HECT domain using the netNES program (La Cour *et al*, 2004) to detect leucine-rich nuclear export signals (NES) shown in fig 3.3.3, predicted an absence of NES sites within residue 1-171 of the HECT domain. The nuclear localization of WWP2-C containing the HECT domain shown in fig 5.5.2 further suggests the lack of NES sites in residue 1-171 of the WWP2 HECT domain. Thus, we hypothesize that the lack of nuclear localization of WWP2-FL, FL- $\Delta$ C2 and  $\Delta$ HECT, but not WWP2-C might be due to a potent NES in residues 383-440 of the WWP2 protein, a region overlapping the WW3 domain that is present in WWP2-FL, FL- $\Delta$ C2 and  $\Delta$ HECT but not WWP2-C (shown in fig 5.3.7). When the amino acid sequence of this region was inputted into the netNES program (fig 5.3.8), results detected a region between aa 15-37 of the probed sequence (residue 398-420 of the full length protein) highlighted in red, and more specifically the leucine residue at position 25 (residue 408 in the full length WWP2) that might contain NES activity. The presence of this predicted NES in the beginning of the WW3 domain might therefore be responsible for the lack of WWP2-FL, FL- $\Delta$ C2 and  $\Delta$ HECT nuclear localization seen in the cellular sub-localization study.



**Figure 5.3.7- Prediction of NES signal within the WWP2 WW3 domain region.** Residue 383-440 highlighted in red is a region overlapping the WW3 domain which is present in WWP2-FL, FL- $\Delta$ C2 and  $\Delta$ HECT but not the WWP2-C isoform is predicted to contain a NES signal.



**Figure 5.3.8- The prediction of Nuclear export signals (NES) in residue 383-440 of WWP2 using the netNES program.** The predictions are generated using a combination of the Artificial neural network (NN) score and the Hidden Markov model (HMM) score and the region most likely to contain an NES is circled in red.

Another interesting trend observed in our immunolabelling study is that under TGF $\beta$  stimulation, the hypothesized lack of nuclear localization due to the presence of the predicted NES in the WW3 domain previously observed in WWP2-FL, FL- $\Delta$ C2 and  $\Delta$ HECT transfected cells is no longer detected. Cells that were stimulated by TGF $\beta$  showed nuclear translocation of FL, FL- $\Delta$ C2 and  $\Delta$ HECT into the nucleus regardless of the presence of WW3 domain suggesting that TGF $\beta$  is able to counteract or suppress the negative

effects of the NES signal within the WW3 domain in nuclear translocation, which could be the results of TGF $\beta$ -dependent Smad-WWP2 binding.

Overall in this chapter, a few key ideas regarding the transcriptional regulation of WWP2 isoforms and cellular localization were developed. From the research carried out, we predict that the 0.5 kb region in intron 10/11 containing the predicted WWP2-C promoter enhances the transcriptional activity of WWP2-C and that its activity is negatively regulated by the presence of SOX9. Luciferase results also produced significant data suggesting that a 2 kb region within intron 9/10 of the WWP2 locus enhances the activity of  $\Delta$ HECT transcription which is also promoted by EGF stimulation. Furthermore, the functional analysis of WWP2- $\Delta$ HECT lead to the hypothesis that the  $\Delta$ HECT containing an incomplete HECT domain is catalytically non-functional. Additionally, the effects of ESRP proteins on expression of WWP2-N was investigated and revealed that ESRPs not only negatively regulate the transcription of WWP2-N but also the  $\Delta$ HECT isoform. Lastly, many predictions were made using results from the immunolabelling experiments, however, the most notable is arguably the hypothesized NES within the WWP2 WW3 domain.

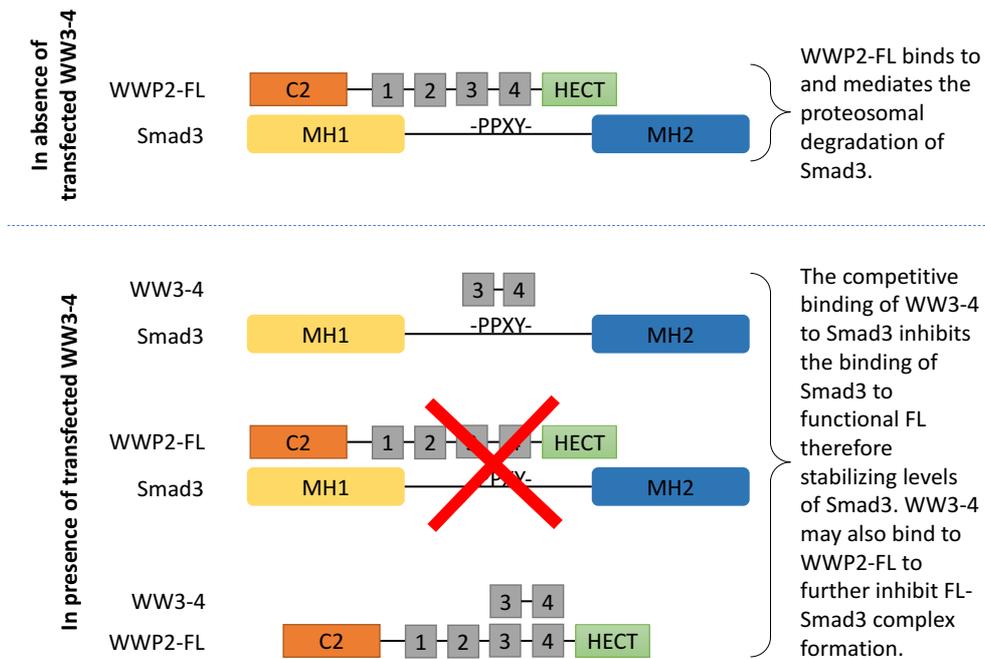
## Chapter 6 Discussion

## 6.1 Discussion

WWP2 is an E3 ubiquitin ligase that works by binding to the PPXY motif of target substrates via its WW domains, to carry out proteasomal degradation of specific proteins. This E3 ligase is highly implicated in the regulation of the TGF $\beta$  pathway by binding to and facilitating the degradation of Smad mediator proteins. Furthermore, the complex relationship between TGF $\beta$  and cancer suggest that the function of WWP2 isoforms might play an important role in oncogenesis. Therefore, in this thesis efforts have been made in attempts to connect these three fundamental players and find out the effects of TGF $\beta$  and WWP2 isoforms on the development of melanoma. In particular, three key areas of the WWP2 isoforms were explored to investigate the cause and effect including; Chapter 1) the biological implications and binding of WWP2 WW domains, chapter 2) the prediction and validation via expression studies of WWP2 existing and novel isoforms in melanoma cell lines and during EMT, and chapter 3) the transcriptional mechanism regulating WWP2 isoform expression. The knowledge of the above combined will provide us with a clearer picture of the characteristics of each isoform therefore increasing our understanding of this ligase and its relationship with TGF $\beta$  signalling and consequently oncogenesis. Together these may act as stepping stones to the potential development of diagnostic and therapeutic tools for melanoma in the future.

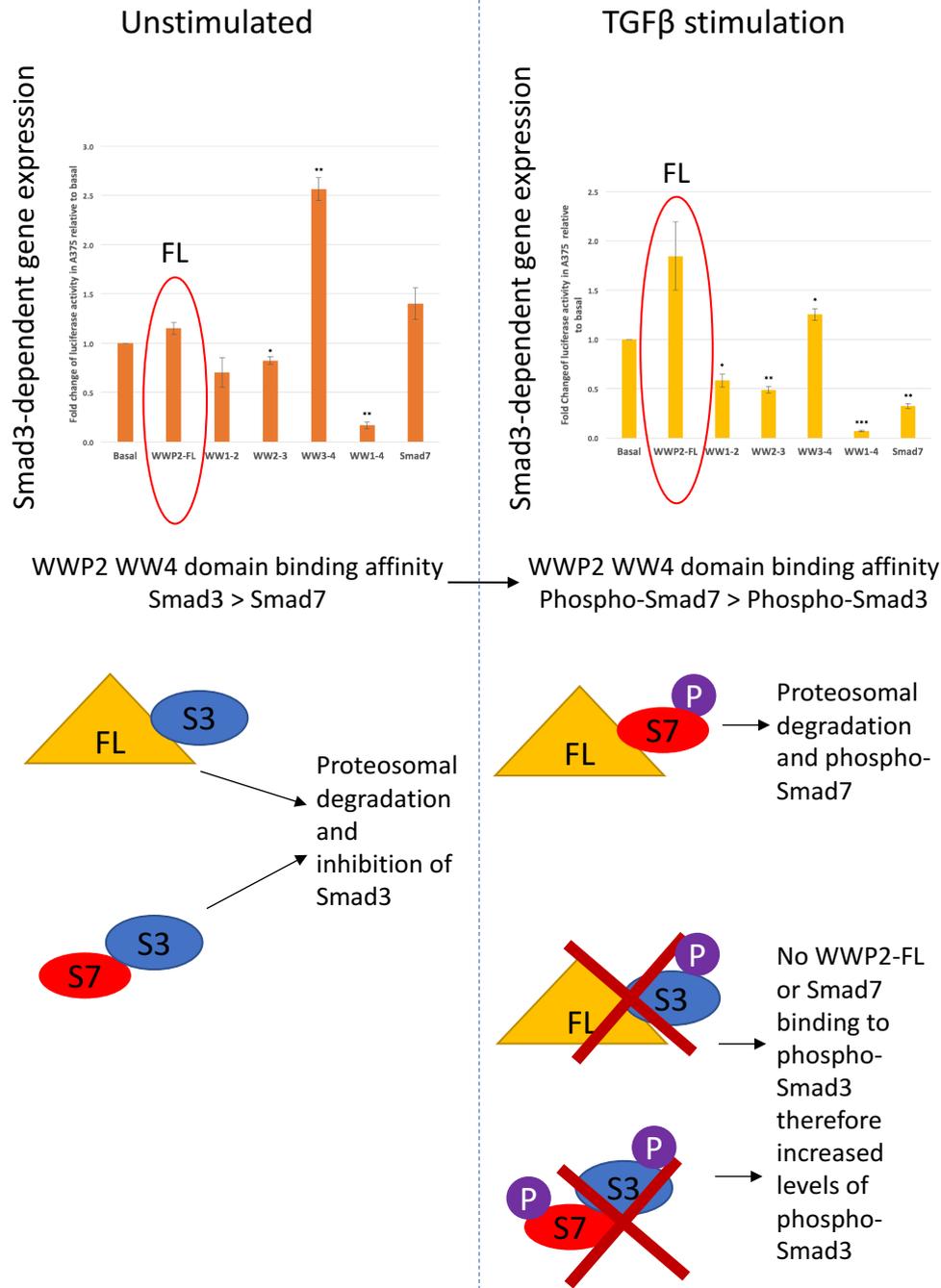
## 6.2 Functional implications of WWP2 WW domains

To begin with, luciferase assays were performed using WW tandem domains and the CAGAC reporter sequence to investigate the Smad3-dependent TGF $\beta$  activity to different regions of the WWP2 isoforms. In non-cancerous HEK293A cells, it was found that transfected WWP2 WW3-4 induced a significant increase in Smad3-dependent activity irrespective of TGF $\beta$  stimulation. Out of the six established and novel WWP2 isoforms, only the FL,  $\Delta$ HECT and FL- $\Delta$ C2 contain WW3-4. The increase in Smad3-dependent activity in response to WW3-4 expression suggests that WW3-4 binds to and stabilizes Smad3 levels, inhibiting the formation of functional FL-Smad3 complex which would lead to the degradation of Smad3. It is also suggested that WW3-4 can bind directly to FL forming a heterodimer to once again inhibit Smad3 binding and subsequent degradation (explained in fig 6.2.1). When this experiment was repeated in A375 melanoma cells, WW3-4 induced a high Smad3-dependent TGF $\beta$  gene expression in cells unstimulated with TGF $\beta$ . However, under TGF $\beta$  stimulation, WWP2 FL induced a much higher reading in comparison WW3-4 suggesting that the transfected FL construct had an increased positive effect on Smad3-dependent activity under TGF $\beta$  stimulation. The binding responsible for this observed pattern may be to do with the preferential binding of WW4 domain to native/phospho-Smad7. Shown in a previous NMR study, the WW4 domain has an increased binding affinity to phospho-Smad7 in comparison to native-Smad7 or Smad3 (Wahl, 2016). This means that under TGF $\beta$  stimulation, where native Smads are phosphorylated, the WW4 domain within the transfected WWP2-FL binds preferentially to and mediates degradation of phospho-Smad7 leading to the increase in Smad3 levels as WWP2-FL and Smad7 are both unable to bind to and inhibit Smad3. This suggested mechanism causing the change in Smad3-dependent gene expression is summarized in fig 6.2.2.



**Figure 6.2.1-** A diagram depicting the proposed WW3-4 binding from luciferase data. The catalytically functional WWP2-FL binds to Smad3 leading to its subsequent degradation as suggested by the decrease in Smad3-dependent activity in luciferase assays. Transfected WW3-4 binds competitively to Smad3 and inhibits Smad3-WWP2-FL binding therefore stabilizing Smad3 levels as it lacks the catalytic HECT domain. WW3-4 has also been suggested to bind directly to FL to further inhibit Smad3-FL binding.

## A375



**Figure 6.2.2-** A schematic representation of the suggested mechanism responsible for the Smad3-dependent gene expression in WWP2-FL transfected A375 cells. The increase in gene expression seen in TGFβ stimulated cells may be due to the preferential binding affinity of WWP2 WW4 domain to phospho-Smad7 in TGFβ conditions and Smad3 in unstimulated conditions. In unstimulated A375 cells, the WW4 domain within WWP2-FL preferentially binds to Smad3 therefore facilitates the proteasomal degradation of Smad3 in addition to the Smad7 inhibition of Smad3. In A375 cells stimulated with TGFβ, phosphorylation of Smads occur therefore shifting WW4 preferential binding within WWP2-FL to phospho-Smad7 rather than Smad3. Levels of free phospho-Smad7 and WWP2-FL subsequently decreases leading to a decreased level of phospho-Smad3 inhibition and degradation.

To investigate the significant increase WWP2 WW3-4 had on Smad3-dependent gene expression, luciferase assay was performed using A375 cells transfected with WW3, WW4 individually, WW3-4 tandem domains and a co-transfection of WW3 + WW4. Results suggested that the transfected WW3-4 tandem domain produced the highest reading for Smad3-dependent activity similar to HEK293A. However, in the presence of TGF $\beta$ , transfected WW4 produced the highest level of Smad3 expression. The Smad3-dependent gene expression in transfected WW4 cells between TGF $\beta$  stimulated and non-stimulated cells did not change significantly, however WW3-4 transfected cells did decrease in gene expression under TGF $\beta$  stimulation. According to results from previously discussed NMR study, the WW4 domain has increased binding affinity to phospho-Smad7 over Smad3 (Wahl, 2016) which supports our finding regarding the decrease in gene expression of WW3-4 transfected cells under TGF $\beta$  stimulation. We postulate that in non-stimulated cells, WWP2 WW3-4 binds preferentially to Smad3 as suggested by the NMR study and therefore the catalytically non-functional tandem WW domain stabilizes the levels of Smad3. However, as cells are stimulated by TGF $\beta$ , Smads are phosphorylated becoming phospho-Smads and therefore according to the results from the binding affinity study, WW3-4 subsequently shifts preferential binding to phospho-Smad7 and therefore stabilizing Smad7 activity, ultimately allowing the increased inhibition of Smad3 and Smad3-dependent gene expression. Furthermore, the different combinations of WW3 and WW4 domains transfected into cell samples allowed the identification of the 6 aa linker region between WW3 and WW4 as an important feature responsible for the increased Smad3-dependent gene expression. This is because the WW3 and WW4 domains were co-transfected (without the presence of the linker region) into cells and results did not reflect an increase in gene expression suggesting that the linker region (-QGMIQE-) is an important factor in the upregulation of Smad3-dependent gene expression. We therefore surmise that the linker region is needed in addition to the WW3 or WW4 domain to produce the highest up-regulation of Smad3-dependent gene expression. Recently, two tyrosine phosphorylation sites in the WW2-3 linker region were identified as

crucial WWP2 regulatory components which controlled WWP2 auto-inhibition (Chen *et al*, 2017). Although this linker region is much longer than the 6 aa linker region between WW3-4, it provides evidence suggesting linker regions within WWP2 and other NEDD4 ubiquitin ligases are part of the WWP2 regulatory mechanism. Furthermore, sequence analysis of the WW3-4 linker regions between WWP2, WWP1 and ITCH NEDD4 ligases in fig 6.2.3 shows high conservation within the NEDD4 family at 50 % sequence similarity. This similarity might further highlight the importance of WW3-4 linker region in not only WWP2 regulation but possibly WWP1 and ITCH.

```

WWP2_WW2-3_linker      QGMIQE
WWP1_WW2-3_linker      QGLQNE
ITCH_WW2-3_linker      QGQLNE
                        **  :*

```

*Figure 6.2.3- An amino acid sequence alignment showing the WW3-4 linker regions of WWP2, WWP1 and ITCH.*

To investigate the potential difference in effects of transfected WWP2 tandem domains in non-malignant and cancerous cell lines, the Smad3-dependent gene expression luciferase study was repeated in A375 cells where WWP2-FL, WW1-2, WW2-3, WW3-4, and WW1-4 were transfected into the melanoma cell line. Results showed that WW3-4 induces the highest Smad3-dependent gene expression regardless of TGF $\beta$  stimulation and cell type. Moreover, another interesting trend detected whilst comparing results from HEK293A cells and A375 cells was that independent of TGF $\beta$  stimulation, WWP2-FL decreased levels of Smad3 gene expression in HEK293A relative to basal whilst transfected FL increased levels of Smad3 gene expression in A375. We therefore propose that in normal HEK293A cells, WWP2-FL favours binding of Smad3, subsequently facilitating the degradation of this Smad protein, whilst in malignant A375 cells, where TGF $\beta$  levels are naturally higher, FL has increased preferential binding and therefore mediates proteasomal degradation of phospho-Smad7 (Kong *et al*, 1995), leading to the increase in Smad3-dependent gene expression. Although no established studies can support our hypothesis regarding WWP2-FL preferential binding in non-malignant and cancerous cells, the previously mentioned study by Soond *et al* highlights the preferential

binding of WWP2-FL to Smad3 in non-stimulated cells and Smad7 in TGF $\beta$  conditions (Soond and Chantry, 2011). It can therefore be speculated that the preferential binding of WWP2-FL to phospho-Smad7 in A375 cells might be the direct result of higher TGF $\beta$  levels present in cancerous cell lines.

Western blots using Smad2, 3, and 7 antibodies were performed on WW tandem domain transfected HEK293A cells to probe for the effects WW domains have on Smad expression. Results suggest that out of the transfected WW1-2, 2-3, 3-4 and 1-4 domains, the highest expression of Smad3 was observed in the WW1-2 sample whilst the WW3-4 exhibited the highest Smad7 levels. Both observations are supported by previous studies by Soond *et al* indicating that WW1 domain is likely responsible for Smad3 binding whereas WW4 domain binds preferentially to Smad7 (Soond and Chantry, 2011). However, results from the Smad2 western blot suggested that WW3-4 transfected cells showed the highest levels of Smad2. This does not correlate with the findings in the study by Soond *et al*, as their observations showed that Smad2 bound to WWP2-FL and N (both containing WW1) but not WWP2-C, which contains only the WW4 domain, thus suggesting that Smad2 binds preferably to the WW1 domain. A study of Smurf2, another E3 ligase, identified aa 244-328, a region between WW2 and WW3 to be responsible for Smad2 binding (Lin *et al*, 2000). After performing an alignment between this region and the WWP2 WW2-3 sequence, a 32 aa sequence present in aa 404-435 of WWP2 which is a region within the WW3 domain, and aa 296-327 of Smurf2 showed 69% sequence similarity (fig 6.2.4). To further support the binding of Smad2 via the WW3 domain, shown in both our western blot analysis and the Smurf2 study by Lin *et al*, the 32 aa highly similar regions in WWP2 and Smurf both contain the double tryptophan residues, a feature crucial to the target recognition function of WW domains. We therefore surmise that aa 404-435 within the WW3 domain of WWP2 might be linked to Smad2 binding.

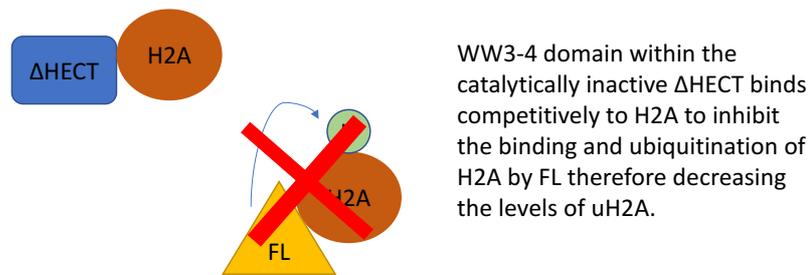
```

Smurf2      LGPLPPGWEIRNTATGRVYFVDHNNRRTTQFTDP
WWP2        LGPLPPGWEKRQD-NGRVYYVNHNRTRTTQWEDP
            ***** *: .*****:*:***.*****: **

```

*Figure 6.2.4- A sequence alignment of Smurf2 and WWP2 WW2 domains. The comparison was performed using MUSCLE ENSEMBL-EBI alignment tool to show the 69% similarity between aa 404-435 of WWP2 and aa 296-327 of Smurf2, the predicted WW3 regions responsible for Smad2 binding.*

To further understand our binding studies, a strep-tagged WWP2 WW3-4 tandem domain was cloned and used in a pull-down experiment and quantitative mass spectrometry. Performing mass spectrometry analysis allowed the detection of a comprehensive list of proteins found in the samples transfected with WW3-4. Furthermore, by stimulating samples with TGF $\beta$ , the effects of TGF $\beta$  on the levels of the protein associated with WW3-4 could be examined. Although TGF $\beta$  mediator Smads were not detected in the mass spectrometry analysis, it was interesting to uncover other proteins that may be associated with the WWP2 WW domain. Results from quantitative mass spectrometry showed that the protein associated with the WW3-4 pulldown sample that was most affected by TGF $\beta$  is histone H2A. Histones are proteins bound by DNA and plays a crucial role in DNA regulation. Whilst ubiquitinated H2A is an active gene repressor, uH2B is more commonly associated with transcription activation (Vissers *et al*, 2008). Studies have shown that TGF $\beta$  negatively regulates the uH2A during TGF $\beta$ -mediated apoptosis (Han *et al*, 2012), whilst results from Mass spectrometry showed an increase in H2A in the presence of WW3-4 under TGF $\beta$  stimulation. We therefore speculate that the increase in H2A could represent the binding of WW3-4 within the non-functional catalytic  $\Delta$ HECT isoform to decrease levels of uH2A seen in TGF $\beta$ -stimulated apoptosis due to the lack of an intact HECT domain (shown in fig 6.2.5).



*Figure 6.2.5- A simplified diagram depicting the proposed mechanisms resulting in the WW3-4-dependent increase in H2A binding in mass spectrometry. WW3-4 within the inactive  $\Delta$ HECT binds to H2A competitively, effectively inhibiting the binding of FL and therefore causing a decrease in levels of uH2A.*

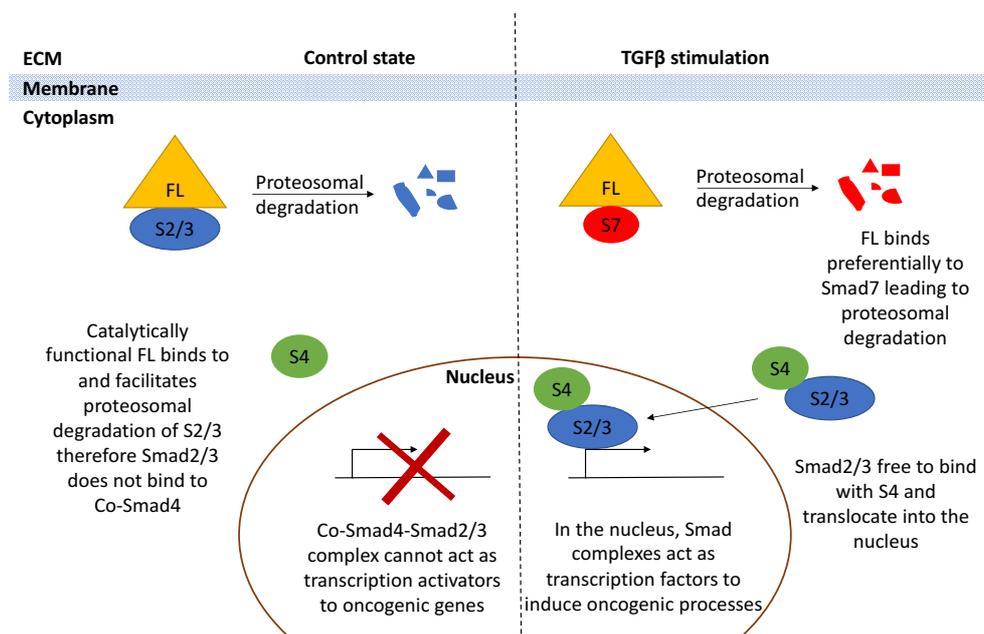
An interesting finding that arose from the quantitative mass spectrometry analysis was the increase in myosin heavy peptide 9 non-muscle, a tumour suppressor encoded by the MYH9 gene (Schramek *et al*, 2014), found associated with WWP2 WW3-4 in response to TGF $\beta$  stimulation. Furthermore, it has been shown that TGF $\beta$  represses the expression of myosin (Wang *et al*, 2003). We therefore propose myosin heavy peptide 9 to be a novel WWP2 binding partner and that the presence of myosin in our results represents the binding of WW3-4 in the catalytically functional WWP2 isoforms to myosin. The increase in levels of myosin shown associated with WW3-4 under TGF $\beta$  stimulation might therefore reflect the mechanism by which TGF $\beta$  negatively regulates the MYH9 tumour suppressor via WWP2-dependent degradation to ultimately aid TGF $\beta$ -mediated oncogenesis. Proteasomal degradation via WWP2-dependent ubiquitination of the myosin tumour suppressor might therefore be one of the mechanisms contributing to the oncogenic effects of TGF $\beta$ .

### 6.3 Characterization of novel and established WWP2 isoforms

The WWP2 protein has two truncated isoforms besides its full-length version, that are well established in literature. The N and C isoforms both contain only one of the four WW recognition domains and either the C2 lipid binding domain or HECT catalytic domain, respectively. Through the use of western blots and whole genome bioinformatics analysis, we were able to predict and confirm the expression of two extra novel isoforms that matched the size and positions suggested by the results of the western blots. These isoforms are known as the WWP2- $\Delta$ HECT and WWP2-N- $\Delta$ C2. Whilst the latter only contains the WW1 domain,  $\Delta$ HECT contains both WW3 and WW4 domains with an incomplete HECT domain. Additionally, recent studies by the Mund lab has highlighted the WWP2-FL- $\Delta$ C2, an isoform similar to the N- $\Delta$ C2 isoform with translation beginning in exon 5 (Mund *et al*, 2015). The WWP2-FL- $\Delta$ C2 transcript, however, continues onto the end of the FL protein. This 86 kDa isoform was previously confirmed only by an indication through western blotting. We have now further validated the expression of this isoform using specific primers in RT-PCR analysis. Furthermore, WWP2- $\Delta$ HECT, N- $\Delta$ C2 and FL- $\Delta$ C2 were successfully cloned and expressed in ARPE-19 cells using the WWP2-FL construct as a template, allowing the localization study of these novel isoforms to be performed. Together, this suggests that rather than the original three WWP2 isoforms reported, there are indeed six transcribed WWP2 proteins in the human genome altogether. The expression of these novel isoforms were then confirmed using RT-PCR on multiple cell lines.

It is vital that we look at all six isoforms as their function within the TGF $\beta$  regulation mechanism may differ due to the difference in the combination of domains they each contain. Furthermore, it is important to note that in some experiments, the stimulation of TGF $\beta$  was used to mimic disease stage as TGF $\beta$  ultimately exerts an oncogenic effect in cancer development.

Results from qPCR showed that WWP2-FL had the highest levels of expression in normal cells in comparison to other isoforms. This combined with the decrease in levels shown in stage IIIC patient samples, also found in another melanoma cDNA panel by Soond *et al* (Soond *et al*, 2013) suggest that FL acts as a tumour suppressor. The decrease in expression levels of the hypothesized WWP2-FL tumour suppressor at stage IIIC allows the increase in TGF $\beta$  signal transduction which mediates EMT and metastasis seen in stage IV melanoma. However, due to the preferential binding of WWP2-FL to Smad2/3 in the control state and Smad7 under TGF $\beta$  stimulation (Soond and Chantry, 2011), we can hypothesize that WWP2-FL acts as a tumour suppressor by mediating degradation of Smad2/3 in the control state whilst acting as an oncogene in the TGF $\beta$  disease state by binding to Smad7. This suggested that similar to TGF $\beta$ , WWP2-FL has a dual role in cancer development dependent on TGF $\beta$  stimulation. Fig 6.3.1 demonstrates the dual role WWP2-FL has in oncogenesis as described above.



**Figure 6.3.1-** A simplified diagram to show the predicted dual role of WWP2-FL in oncogenesis. In the control state without TGF $\beta$  stimulation, FL binds preferentially to Smad2/3 and facilitates proteasomal degradation. The reduced levels of Smad2/3 leads to the subsequent decrease in formation and nuclear translocation of Smad2/3-Smad4 complex. The lowered levels of this transcription factor mean that less oncogenic genes are transcribed therefore suggesting that FL in the control state is a tumour-suppressor. In presence of TGF $\beta$  stimulation, FL binds preferentially to and subsequently facilitates the proteasomal degradation of Smad7. The lowered levels of this natural inhibitor of the TGF $\beta$  pathway allows the levels of TGF $\beta$  signal transduction to increase as Smad2/3-Smad4 complexes translocate into the nucleus and acts as transcription factors for oncogenic genes.

The WWP2-N isoform containing the C2 and WW1 domains occur due to a stop codon present in the retained intron 9/10 which has been suggested to be linked to the expression of ESRPs (Chantry, 2011). WWP2-N is another isoform that was expressed at higher levels in normal cells compared with melanoma cells. Like WWP2-FL, the expression of WWP2-N dipped at stage IIIC (also supported by Soond *et al*, 2013), both suggesting WWP2-N as a tumour suppressor. Furthermore, results from localization study showed that the N isoform was localized in the cytoplasm irrespective of TGF $\beta$  stimulation. Additionally, a strong membrane localization of WWP2-N can be seen which reinforces the function of the C2 lipid binding domain within the N isoform. The binding and stabilization of Smad2/3 by WWP2-N via the WW1 domain and the subsequent inhibition of Smad2/3 translocation into the nucleus further suggests that WWP2-N is a tumour suppressor as Smad2/3 is inhibited from binding with Co-Smad4 and therefore cannot act as a transcription factor in the nucleus to aid oncogenesis (explained in fig 5.3.6 in chapter 5).

Another interesting finding was that WWP2-N expression highlighted a filamentous network within cells suggesting that Smad2/3 bound to WWP2-N was co-expressed with a filamentous protein such as actin or tubulin. Although it has been shown that Smad2/3 binds to microtubules in some cell lines (Dong *et al*, 2000), our study did not provide conclusive evidence supporting this. Moreover, the transcriptional regulation that results in the expression of WWP2-N was also examined in an attempt to gain the full picture regarding the role WWP2 has in oncogenesis. ESRP is a splicing regulatory protein linked to the regulation of many EMT related proteins (Warzecha *et al*, 2010). Due to the potential involvement WWP2 isoforms have in TGF $\beta$  mediated-EMT, it was therefore hypothesized that ESRP might have a role in the intron retention that leads to the expression of WWP2-N. Results from our mRNA expression study show that transfected ESRP1 and 2 lead to a marked decrease in WWP2-N expression in both A375 and MCF-7 cell lines suggesting that ESRP is a key player in the negative regulation of the transcriptional activity of WWP2-N. Research has shown

that in Non-transformed mouse mammary gland (NMUMG) epithelial cells, TGF $\beta$  reduces levels of ESRP expression as ESRP antagonizes TGF $\beta$ -mediated EMT (Horiguchi *et al*, 2012). In our mRNA expression studies, TGF $\beta$  was found to increase WWP2-N in the epithelial MCF-7 cell line. We hypothesize that the downregulation of ESRP by TGF $\beta$  therefore reduces levels of ESRP-dependent negative regulation of WWP2-N which is reflected by the increase in WWP2-N expression and subsequently consolidating the idea that ESRP is indeed involved in the negative regulation of WWP2-N. The antagonistic properties ESRP possess during TGF $\beta$ -mediated EMT, namely the ability to restore epithelial phenotypes by regulating splicing programs, suggests that ESRP is a tumour suppressor. However, if our prediction of WWP2-N also being a tumour-suppressor is correct, then this would mean the negative effect ESRP has on this potential tumour suppressor makes ESRP a possible oncogene. Although ESRP is most known for its tumour-suppressive role in TGF $\beta$ -mediated EMT, it has also been shown that ESRP may have a dual role in cancer development, providing evidence to further supporting our hypothesis (Hayakawa *et al*, 2016).

In our studies, evidence was provided supporting the notion that WWP2-C acts as an oncogene due to the presence of the WW4 and intact HECT domain allowing the degradation of Smad7. Results from the qPCR of melanoma cDNA array showed that levels of WWP2-C increased after stage IIIB. This suggests that the increased expression of this oncogene may be needed to drive the aggression of melanoma by inhibiting Smad7, subsequently allowing increased ulcerations in stage IIIB. Smad7, which is predominantly localized in the nucleus (Zhang *et al*, 2007), binds to and is degraded by WWP2-C, also localized in the nucleus, allowing the efficient inhibition of the inhibitory Smad activity on both DNA and in the context of R-Smad binding (Yan *et al*, 2009). Additionally, expression studies on melanoma cell lines showed that the expression of WWP2-C was exclusive to cell lines that contained the homozygous BRAF<sup>V600E</sup> mutation. This means that cells with this homozygous mutation will only have the ability to express the mutant BRAF kinase and not the wild type. This mutation in the MAPK

signalling pathway kinase means that it is constitutively active even in the absence of EGF stimulation, resulting in increased cell proliferation in melanoma (Ascierto *et al*, 2012). This mutation has also been shown to be responsible for increased TGF $\beta$  secretion and Smad-dependent transcription suggesting that the increase in WWP2-C oncogene may be an effect of BRAF<sup>V600E</sup>-dependent TGF $\beta$  increase. However, there is no solid evidence supporting this hypothesis therefore it is still not clear whether the BRAF mutation is a cause or effect of the C-isoform expression, or indeed merely a coincidence.

To investigate the transcription mechanism of WWP2-C, several predicted promoter sequences were cloned into luciferase reporter vectors. Results from luciferase assays showed that the penultimate 0.5 kb of intron 10/11 has enhancing properties on the transcription activity of WWP2-C. Furthermore, it was found that transfected SOX9 had a negative effect on the WWP2-C transcriptional activity which consolidates the idea of WWP2-C being an oncogene due to the finding in previous studies that detail SOX9 as an inhibitor of melanoma progression (Passeron *et al*, 2009). However, to verify the significance of our findings, the expression of naturally occurring SOX9 in the cell must be confirmed by performing RT-PCR to probe the expression of SOX9 in a panel of melanoma cell lines. Results suggested that levels of SOX9 was shown to increase in all melanoma cell lines except for M202, following TGF $\beta$  stimulation. Interestingly, M202 is the only cell line of the six to contain a Q61L N-RAS mutation. This single residue substitution leads to the altered folding and consequent constant GTP-bound state of N-RAS (Jovanovic *et al*, 2010), which has been shown in a similar mutation (N-RAS<sup>Q61R</sup>) to cause the resistance to TGF $\beta$  senescence in melanoma (Lasfar and Cohen-Solal, 2010). Therefore, we speculate that the lack of SOX9 increase in response to TGF $\beta$  in M202 could be due to the N-RAS<sup>Q61L</sup> mutation present in the cell line.

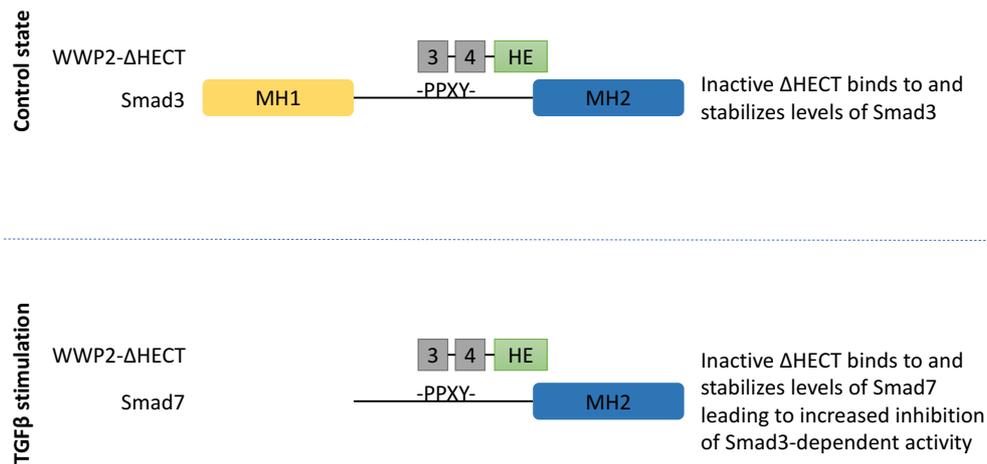
The SOX9-mediated negative regulation on WWP2-C shown above was suggested to be a result of binding through its recognition

sequence -CCTTGAG-. To investigate this predicted SOX9 binding site, WWP2-C promoter mutant constructs containing the SOX9 binding site deletion were then cloned and used in similar luciferase analysis. Results show that although SOX9-mediated negative regulation of WWP2-C transcriptional activity was reduced in HEK293A and MCF-7 epithelial cells, the deletion in the mutant failed to alter this negative regulation of WWP2-C transcriptional activity in A375, an epithelial cell line expressing strong mesenchymal characteristics. Together with the discovery of multiple predicted SOX9 binding sites within the WWP2-C promoter region, this suggests that preference to specific SOX9 binding sites is epithelial and mesenchymal tissue-specific.

The WWP2- $\Delta$ HECT isoform is the only known isoform containing an incomplete HECT domain therefore it is still unclear whether this domain has intact catalytic activities. The predicted promoter of this isoform lie in intron 9/10, and luciferase reporter assay of promoter sequence up to 2 kb, 3' of this intron showed an increase in luciferase activity compared to the basal reading suggesting that the cloned sequence from intron 9/10 might enhance the transcriptional activity of WWP2- $\Delta$ HECT. However, when compared with the massive increase seen in WWP2-C transcriptional activity, the increase in  $\Delta$ HECT activity is relatively low therefore the effects of other growth factors were tested to probe a change in  $\Delta$ HECT transcriptional activity. Whilst the stimulation of TGF $\beta$  and BMP both had no significant effect on WWP2- $\Delta$ HECT transcriptional activity, EGF stimulation increase the activity of the putative  $\Delta$ HECT promoter through the suggested binding at AP-1 like sites present in the promoter sequence. In a separate study, the effects of ESRP splicing regulatory protein on expression of  $\Delta$ HECT was also investigated using a combination of RT-PCR and qPCRs. Results suggest that ESRP is part of a negative regulatory mechanism of WWP2- $\Delta$ HECT expression shown in both A375 and MCF-7 cell lines.

In the melanoma expression study, it was found that the levels of  $\Delta$ HECT increased at stage IIIB similar to WWP2-C, and decreased at stage IIIC which

was a pattern also observed in WWP2-FL/N expression. Its similarity to both the oncogene and tumour suppressors suggest that this isoform may have a dual role in oncogenesis. Likewise, results from the cellular sub-localization study showed that  $\Delta$ HECT translocated from the cytoplasm to the nucleus under TGF $\beta$  stimulation similar to FL suggesting that once again, it may be involved in both suppressing of tumours and oncogenesis. However, without knowing the full potential of the incomplete HECT domain, we are unable to accurately hypothesize the function of  $\Delta$ HECT as it could either work to stabilize Smad7 levels, or inhibit Smad7 by mediation of degradation. To try and further understand the biological role of this novel isoform, WWP2- $\Delta$ HECT was also cloned and transfected into HEK293A cells. Results from luciferase assays show that in the absence of TGF $\beta$  stimulation, an increase of Smad3-dependent TGF $\beta$  activity was detected whilst a decrease was seen in the presence of TGF $\beta$ . Using data on WW4 domain binding affinities (Wahl, 2016), we hypothesize that in TGF $\beta$  negative conditions, WW4 domain in  $\Delta$ HECT binds to and stabilizes Smad3, and that under TGF $\beta$  stimulation,  $\Delta$ HECT binds to and stabilizes Smad7 causing the degradation of Smad3 (as depicted in fig 6.3.2). The results from this experiment also suggests that the incomplete HECT domain within  $\Delta$ HECT is catalytically inactive due to its inability to facilitate proteolytic degradation of Smad3 in control conditions and Smad7 under TGF $\beta$  stimulation. It can therefore be speculated that the WWP2- $\Delta$ HECT isoform is part of the TGF $\beta$ -mediated feedback mechanism to regulate the functional WWP2-dependent degradation of Smad proteins through the competitive binding of  $\Delta$ HECT to Smad mediators.

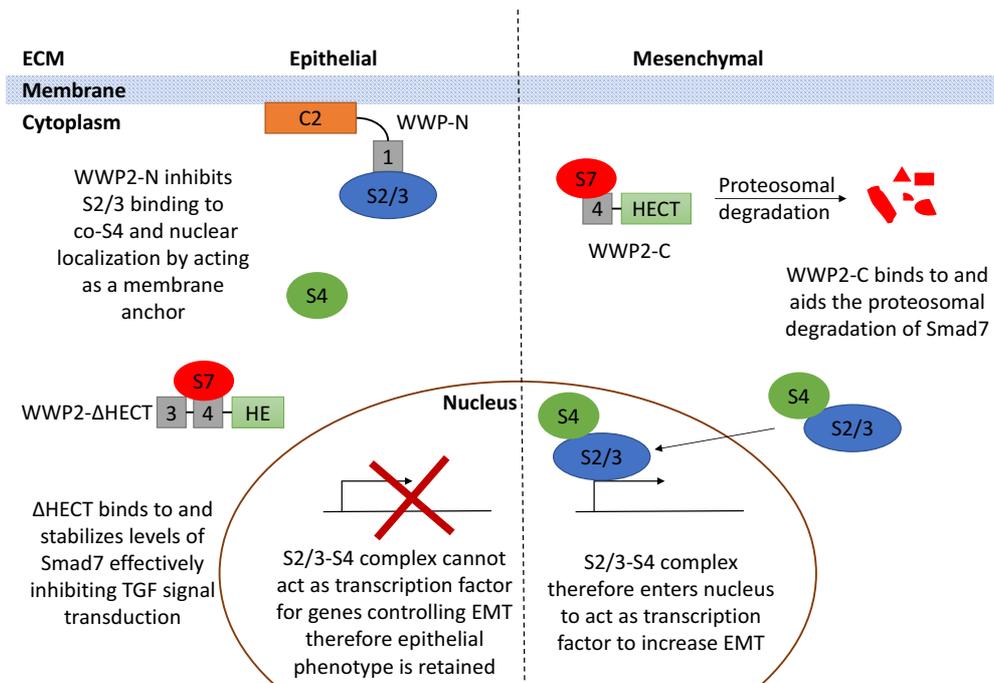


*Figure 6.3.2- A diagram showing the proposed binding from ΔHECT luciferase assays. Without TGFβ stimulation, ΔHECT binds to and stabilizes levels of Smad3 whilst in the presence of TGFβ, ΔHECT stabilizes levels of Smad7 which leads to the decrease in Smad3-dependent TGFβ activity.*

The expression study on WWP2-FL-ΔC2 and WWP2-N-ΔC2 did not provide any significant results. However, results from cellular localization study provided several interesting points. Both isoforms showed a translocation from the cytoplasm to the nucleus following TGFβ stimulation similar to WWP2-FL, nevertheless due to the difference in domain composition, the function of the isoforms may be very different. WWP2-N-ΔC2 contains only the WW1 domain suggesting binding and stabilization of Smad2/3 therefore is a proposed oncogene. On the other hand, WWP2-FL-ΔC2 contains all the domains except for the C2 lipid binding domain which suggests it could bind to both Smad2/3 and Smad7. There is no solid evidence supporting the following notion, however if this isoform expressed preferential binding similar to FL, where binding to Smad2/3 is preferred in absence of TGFβ and Smad7 binding is preferred in TGFβ conditions (Soond and Chantry, 2011), then WWP2-FL-ΔC2 may have a dual role in oncogenesis. This would mean that WWP2-FL-ΔC2 would bind to and inhibit Smad2/3 in normal cells but Smad7 in disease state effectively inhibiting the natural inhibitory mechanism of the TGFβ pathway. Ultimately, the profiling of WWP2 isoform expression at different stages of melanoma has the potential to aid development of a prognostic tool or a marker for melanoma progression. The knowledge of WWP2 isoform function and Smad binding may in the future prove useful as part of a therapeutic tool in cancer.

## 6.4 The role of WWP2 isoforms in EMT

Previously, the importance of TGF $\beta$  and BMP in EMT has been highlighted. We proposed that the TGF $\beta$ -mediated change in WWP2 isoform levels act as a feedback loop as WWP2 isoforms are able to bind to Smad proteins to ultimately alter levels of TGF $\beta$ -dependent oncogenic processes such as EMT. TGF $\beta$  has been shown to increase EMT which can be reflected by the increase in expression of EMT markers. In our collaboration with the Nieto lab at the Institute of Neuroscience Alicante, it was found that the WWP2-N and  $\Delta$ HECT isoforms were expressed at a higher level in epithelial cells whilst WWP2-C was expressed in mesenchymal cells. We therefore put forward the N and  $\Delta$ HECT isoforms as guardians of the epithelial phenotype. This means that functionally, N is able to anchor Smad2/3 to the membrane effectively inhibiting nuclear translocation and activation of genes controlling EMT. On the other hand, the catalytically inactive  $\Delta$ HECT binds to and stabilizes the levels of Smad7 allowing increased inhibition of the TGF $\beta$  signal transduction and ultimately decreases levels of EMT-related gene transcription. Due to the increased expression shown in mesenchymal cells, WWP2-C, has therefore been predicted as not only an oncogene, but also the guardian of the mesenchymal phenotype. The intact catalytic function of the C-isoform suggest that it is able to bind to and inhibit Smad7 activity. This means that Smad2/3 can continue to act as transcription factors in the nucleus for genes associated with EMT. A simple diagram illustrating the predicted role of WWP2-N and  $\Delta$ HECT as guardians of the epithelial/mesenchymal phenotypes along with effects of WWP2-C in EMT is shown in fig 6.4.1.



**Figure 6.4.1- Diagram explaining WWP2-C, N and ΔHECT as guardians of EMT.** WWP2-N and ΔHECT has been suggested as guardians of the epithelial phenotype by inhibiting TGFβ signal transduction that leads to EMT. The C2 lipid binding domain (in orange) binds to the membrane whilst the WW1 domain binds to Smad2/3. This inhibits the nuclear translocation of Smad2/3 and therefore interferes with EMT-related TGFβ signalling. WWP2-ΔHECT on the other hand binds to and stabilizes levels of Smad7 via the WW4 domain in grey. The increased levels of Smad7 allows the inhibition of TGFβ signal transduction again decreasing levels of TGFβ-mediated EMT. In mesenchymal cells, functional WWP2-C binds to and facilitates the proteosomal degradation of Smad7 in turn increasing levels of Smad transduction and TGFβ-dependent EMT.

## 6.5 Overall effects of TGF $\beta$ and BMP on oncogenesis in melanoma

TGF $\beta$  and BMP stimulations have been used throughout experiments conducted for this thesis. The aim of this was to identify the roles these growth factors have on oncogenic processes such as EMT and therefore suggest whether they have oncogenic or tumour suppressing properties. TGF $\beta$  is known to have a dual role in oncogenesis in which tumour suppressor properties can be detected in healthy cells and oncogenic effects seen in malignant cells (Hussein, 2004). A few key findings in our research have supported the dual effect of TGF $\beta$  in cancer. Firstly, in malignant A375 cells, it was found that TGF $\beta$  reverses the negative effect SOX9 has on the transcriptional activity of the oncogenic WWP2-C. This suggests that in oncogenic cells, TGF $\beta$  has the ability to increase transcription of the WWP2-C oncogene implying that TGF $\beta$  has an oncogenic role. Interestingly, it was also found that in HEK293A cells, TGF $\beta$  stimulation decreases levels of  $\Delta$ HECT transcriptional activity. Due to the tumour suppressing nature of  $\Delta$ HECT under TGF $\beta$  stimulation, (as  $\Delta$ HECT binds preferably to Smad7 under TGF $\beta$  stimulation and stabilizes its levels to ultimately decrease TGF $\beta$  signal transduction), the negative effects TGF $\beta$  has on  $\Delta$ HECT transcription therefore suggest that TGF $\beta$  has oncogenic properties even in HEK293A cells. On the other hand, results from functional  $\Delta$ HECT studies showed a significant decrease in  $\Delta$ HECT-mediated Smad-3 dependent activity in HEK293A as a result of TGF $\beta$  stimulation. We argue that in this normal cell line,  $\Delta$ HECT binds preferentially to Smad7 in response to TGF $\beta$  stimulation and therefore stabilizing levels of Smad7 ultimately reducing Smad3-dependent activity such as EMT. This suggests that TGF $\beta$  has a tumour suppressing role in normal cells. However, in malignant A375 cells, the stimulation of TGF $\beta$  does not decrease the levels of  $\Delta$ HECT-mediated Smad3-dependent activity therefore suggesting that TGF $\beta$  has lost its tumour suppressing abilities in favour of oncogenic properties. In the quantitative mass spectrometry experiment, the effects of TGF $\beta$  on proteins associated with the transfected WWP2 WW3-4 were investigated and one

of the proteins which saw the largest increase as a result of TGF $\beta$  stimulation was histone H2B. Studies have previously shown the tumour-suppressive nature of uH2B (Prenzel *et al*, 2011) suggesting that in the mass spectrometry study using HEK293A cells, the TGF $\beta$ -dependent increase in H2B, potentially bound to WW3-4 of the catalytically functional WWP2-FL, may be an indication of the tumour-suppressive role TGF $\beta$  has in non-cancerous epithelial cells.

Although conflicting evidence show that BMP can be both oncogenic and tumour suppressing, most research has identified BMP4 as oncogenic and have pinpointed its role in invasion and migration (Rothhammer *et al*, 2005). Key findings in our research has supported the oncogenic role BMP4 has, most notably in the WWP-N expression studies in melanoma cells. It was found that in five out of nine melanoma cell lines, BMP4 stimulation decreased levels of the WWP2-N tumour suppressor. Most significantly, the expression of this isoform was obliterated by BMP stimulation in A2058. Due to the ability of WWP2-N to anchor Smad2/3 to the membrane and inhibit nuclear translocation, we surmise that the suppression of this isoform by BMP4 could reflect its oncogenic tendencies. In addition, BMP4 stimulation has also shown an increase in EMT markers such as SLUG, SNAIL, Vimentin, N-cadherin and Twist suggesting that it may have a positive role in the up-regulation of EMT.

Overall, we surmise that the effects of TGF $\beta$  on WWP2 expression, and the ability of WWP2 isoforms to alter TGF $\beta$  signalling through regulation of Smad mediators, creates a feedback loop that can prove crucial to the study of oncogenesis in melanoma and perhaps other types of cancers. For this reason, further research should be carried out in this area to clarify the relationship between WWP2 isoforms, the TGF $\beta$  signalling pathway and melanoma.

## 6.6 Future work

WWP2 has a complex relationship with TGF $\beta$  and oncogenesis therefore might prove useful as a prognostic or therapeutic target in cancer. However, without the complete understanding of the WWP2 isoforms and their effects on the TGF $\beta$  pathway and subsequently oncogenesis, it is difficult to design or develop any type of prognostic or therapeutic measures. It is therefore important to continue the studies to clarify the role WWP2 has in oncogenesis. First of all, the transcriptional mechanism of WWP2 isoforms must be verified. ESRPs have been shown to down-regulate levels of WWP2-N and  $\Delta$ HECT in our studies however, this relationship could be verified by performing knock-down studies to investigate the effects lack of ESRP might infer onto the production of these isoforms. Furthermore, ESRP1 and 2 were used simultaneously in our studies, therefore in future experiments, these splicing proteins could be separated to identify their individual roles in WWP2 isoform transcriptional regulation.

In chapter 3 of this thesis, it was hypothesized that the WWP2-FL, similar to the WWP2 WW4 domain has preferential binding to Smad mediators dependent on TGF $\beta$  stimulation and cell type. Co-immunoprecipitation followed by western blot experiments using anti-Smad antibodies can be performed in the future using non-malignant and cancerous cell lines stimulated with TGF $\beta$  to study the levels of Smads interacting with WWP2-FL under different conditions. Results could show the potential preferential binding of WWP2-FL to Smads and prove or disprove our hypothesis. Also in this chapter, it was concluded that the linker region between WWP2 WW3 and WW4 held significance in the increase in Smad3-dependent gene expression detected in the luciferase analysis. To identify the exact region needed for optimal up-regulation of Smad3 gene expression, constructs containing the WW3+linker region, WW4+linker region and the linker region alone can be cloned and used in luciferase assays as before. The results would validate the hypothesis regarding the 6 aa linker region. To pinpoint

the region responsible for mediating the TGF $\beta$  signal may prove useful as a potential tool in inhibiting the TGF $\beta$  pathway.

Likewise, it is also important to characterize the novel  $\Delta$ HECT isoform and identify whether the incomplete HECT domain retains its catalytic function. The binding of  $\Delta$ HECT can also be explored by performing mass spectrometry to identify binding partners implicated in cancer cell lines and show the effects of TGF $\beta$ . To do this, a strep tag would be cloned onto the WWP2- $\Delta$ HECT isoform, transfected into various cell lines and used in a streptavidin pull-down in preparation for quantitative mass spectrometry, similar to the WW3-4 mass spectrometry performed in this thesis. Understanding the characteristics and abilities of this isoform may be a crucial step to unlocking its potential as a tool in the prognosis and treatment of cancer.

Further immunolabelling and fluorescence microscopy should also be carried out to investigate the filamentous network highlighted by WWP2-N to explore the co-expression of N to Smad2/3 and potentially actin. The cellular sub-localization studies performed during this research period utilized ARPE-19 cells, however, this experiment can be repeated on melanoma cell lines to investigate the accurate effects of TGF $\beta$  on the translocation of isoforms in cancer and thus provide comparison to localization of isoforms in normal cells. Finally, to study the function of WWP2 isoforms, ubiquitination assays can be performed to probe levels of ubiquitination by different isoforms and subsequently identify catalytically functional and non-functional WWP2 isoforms. By combining these results with results of Smad binding studies, it may be possible to provide a more accurate prediction of the effects WWP2 has on the TGF $\beta$ -mediated oncogenesis.

## References

- Akimov, V., Rigbolt, K. T. G., Nielsen, M. M. and Blagoev, B. (2011) 'Characterization of ubiquitination dependent dynamics in growth factor receptor signaling by quantitative proteomics', *Molecular BioSystems*. The Royal Society of Chemistry, 7(12), p. 3223.
- Al-Salihi, M. A., Herhaus, L., Macartney, T. and Sapkota, G. P. (2012) 'USP11 augments TGF $\beta$  signalling by deubiquitylating ALK5.', *Open biology*, 2(6), p. 120063.
- Amerik, A. Y. and Hochstrasser, M. (2004) 'Mechanism and function of deubiquitinating enzymes', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1695(1–3), pp. 189–207.
- Amerik, A. Y., Swaminathan, S., Krantz, B. A., Wilkinson, K. D. and Hochstrasser, M. (1997) 'In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome', *EMBO Journal*, 16(16), pp. 4826–4838.
- An, H., Krist, D. T., Statsyuk, A. V., Qiu, X., Lu, K., Fu, Y., Li, H., Xing, G., Li, D., Peng, R., He, F., Zhang, L., Sermet-Gaudelus, I., Rowe, S. M., Dong, Q., Rodriguez, S., Yen, K., Ordonez, C. and Elborn, J. S. (2014) 'Crosstalk between kinases and Nedd4 family ubiquitin ligases', *Molecular BioSystems*. The Royal Society of Chemistry, 10(7), p. 1643.
- Annes, J. P., Munger, J. S. and Rifkin, D. B. (2003) 'Making sense of latent TGF $\beta$  activation.', *Journal of cell science*, 116(Pt 2), pp. 217–24.
- Ascierto, P. A., Kirkwood, J. M., Grob, J.-J., Simeone, E., Grimaldi, A. M., Maio, M., Palmieri, G., Testori, A., Marincola, F. M. and Mozzillo, N. (2012) 'The role of BRAF V600 mutation in melanoma.', *Journal of translational medicine*. BioMed Central, 10, p. 85.
- Ben-Saadon, R., Zaaroor, D., Ziv, T. and Ciechanover, A. (2006) 'The Polycomb Protein Ring1B Generates Self Atypical Mixed Ubiquitin Chains Required for Its In Vitro Histone H2A Ligase Activity', *Molecular Cell*, 24(5), pp. 701–711.

- Berndsen, C. E. and Wolberger, C. (2014) 'New insights into ubiquitin E3 ligase mechanism.', *Nature structural & molecular biology.*, 21(4), pp. 301–7.
- Bogenrieder, T. and Herlyn, M. (2003) 'Axis of evil: molecular mechanisms of cancer metastasis.', *Oncogene*, 22(42), pp. 6524–36.
- Boyer, L. A., Shao, X., Ebright, R. H. and Peterson, C. L. (2000) 'Roles of the histone H2A-H2B dimers and the (H3-H4)<sub>2</sub> tetramer in nucleosome remodeling by the SWI-SNF complex.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 275(16), pp. 11545–52.
- Cancer Research UK (2014) *Skin cancer incidence statistics | Cancer Research UK*. Available at: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/skin-cancer/incidence#heading-Zero> (Accessed: 19 December 2017).
- Cao, J. and Yan, Q. (2012) 'Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer.', *Frontiers in oncology*, 2(March), p. 26.
- Chai, J., Wu, J. W., Yan, N., Massagué, J., Pavletich, N. P. and Shi, Y. (2003) 'Features of a Smad3 MH1-DNA complex: Roles of water and zinc in DNA binding', *Journal of Biological Chemistry*, 278(22), pp. 20327–20331.
- Chantry, A. (2011) 'WWP2 ubiquitin ligase and its isoforms: New biological insight and promising disease targets', *Cell Cycle*.
- Chen, C., Sun, X., Guo, P., Dong, X. Y., Sethi, P., Zhou, W., Zhou, Z., Petros, J., Frierson Jr., H. F., Vessella, R. L., Atfi, A. and Dong, J. T. (2007) 'Ubiquitin E3 ligase WWP1 as an oncogenic factor in human prostate cancer', *Oncogene*, 26(16), pp. 2386–2394.
- Chen, W., Jiang, X. and Luo, Z. (2014) 'WWP2: A Multifunctional ubiquitin ligase Gene', *Pathology and Oncology Research* 20(4) 799-803.
- Chen, Y. G., Hata, a, Lo, R. S., Wotton, D., Shi, Y., Pavletich, N. and Massagué, J. (1998) 'Determinants of specificity in TGF-beta signal

- transduction.', *Genes & development*, 12(14), pp. 2144–2152.
- Chen, Z., Jiang, H., Xu, W., Li, X., Dempsey, D. R., Zhang, X., Devreotes, P., Wolberger, C., Amzel, L. M., Gabelli, S. B. and Cole, P. A. (2017) 'A Tunable Brake for HECT Ubiquitin Ligases', *Molecular Cell*, 66(3), pp. 345–357.e6.
  - Choi, B. H., Che, X., Chen, C., Lu, L. and Dai, W. (2015) 'WWP2 is required for normal cell cycle progression', *Genes Cancer* 6(9-10), pp. 371–377.
  - Chow, J. Y. C., Dong, H., Quach, K. T., Van Nguyen, P. N., Chen, K. and Carethers, J. M. (2008) 'TGF-beta mediates PTEN suppression and cell motility through calcium-dependent PKC-alpha activation in pancreatic cancer cells.', *American journal of physiology. Gastrointestinal and liver physiology*, 294(4), pp. G899-905.
  - Cohen-Solal, K. A., Merrigan, K. T., Chan, J. L. K., Goydos, J. S., Chen, W., Foran, D. J., Liu, F., Lasfar, A. and Reiss, M. (2011) 'Constitutive Smad linker phosphorylation in melanoma: A mechanism of resistance to transforming growth factor- $\beta$ -mediated growth inhibition', *Pigment Cell and Melanoma Research*, 24(3), pp.512-24
  - Cook, J. C. and Chock, P. B. (1992) 'Isoforms of mammalian ubiquitin-activating enzyme', *Journal of Biological Chemistry*, 267(34), pp. 24315–24321.
  - Cooper, S. J., Trinklein, N. D., Nguyen, L. and Myers, R. M. (2007) 'Serum response factor binding sites differ in three human cell types.', *Genome research*. Cold Spring Harbor Laboratory Press, 17(2), pp. 136–44.
  - La Cour, T., Kiemer, L., Mølgaard, A., Gupta, R., Skriver, K. and Brunak, S. (2004) 'Analysis and prediction of leucine-rich nuclear export signals', *Protein Engineering, Design and Selection*, 17(6), pp. 527–536
  - Dong, C., Li, Z., Alvarez, R., Feng, X.-H. and Goldschmidt-Clermont, P. J. (2000) 'Microtubule Binding to Smads May Regulate TGF $\beta$  Activity', *Molecular Cell*, 5(1), pp. 27–34.
  - Dummer, R., Hauschild, A. and Pentheroudakis, G. (2009) 'Cutaneous

- malignant melanoma: ESMO Clinical Recommendations for diagnosis, treatment and follow-up', *Annals of Oncology*. Oxford University Press, 20(Supplement 4), p. iv129-iv131.
- DUNN, K. C., AOTAKI-KEEN, A. E., PUTKEY, F. R. and HJELMELAND, L. M. (1996) 'ARPE-19, A Human Retinal Pigment Epithelial Cell Line with Differentiated Properties', *Experimental Eye Research*, 62(2), pp. 155–170.
  - Ebisawa, T., Fukuchi, M., Murakami, G., Chiba, T., Tanaka, K., Imamura, T. and Miyazono, K. (2001) 'Smurf1 Interacts with Transforming Growth Factor- $\beta$  Type I Receptor through Smad7 and Induces Receptor Degradation', *Journal of Biological Chemistry*, 276(16), pp. 12477–12480.
  - Eichhorn, P. J., Rodon, L., Gonzalez-Junca, A., Dirac, A., Gili, M., Martinez-Saez, E., Aura, C., Barba, I., Peg, V., Prat, A., Cuartas, I., Jimenez, J., Garcia-Dorado, D., Sahuquillo, J., Bernards, R., Baselga, J. and Seoane, J. (2012) 'USP15 stabilizes TGF-beta receptor I and promotes oncogenesis through the activation of TGF-beta signaling in glioblastoma', *Nat Med*, 18(3), pp. 429–435.
  - Epstein, F. H., Blobel, G. C., Schiemann, W. P. and Lodish, H. F. (2000) 'Role of Transforming Growth Factor  $\beta$  in Human Disease', *New England Journal of Medicine*. Massachusetts Medical Society , 342(18), pp. 1350–1358.
  - Eric Aragón, Nina Goerner, Qiaoran Xi, Tiago Gomes, Sheng Gao, J. and Massagué, and M. J. M. (2012) 'Structural basis for the versatile interactions of Smad7 with regulator WW domains in TGF- $\beta$  pathways', *Biophysical Chemistry*, 34(1), pp. 13–23.
  - Feng, X.-H. and Derynck, R. (2005) 'Specificity and Versatility in Tgf-B Signaling Through Smads', *Annual Review of Cell and Developmental Biology*, 21(1), pp. 659–693.
  - Ferrari, G., Cook, B. D., Terushkin, V., Pintucci, G. and Mignatti, P. (2009) 'Transforming growth factor-beta 1 (TGF-beta1) induces angiogenesis through vascular endothelial growth factor (VEGF)-mediated apoptosis.', *Journal of cellular physiology*, 219(2), pp. 449–

58.

- Flaszka, M., Gorman, P., Roylance, R., Canfield, A. E. and Baron, M. (2002) 'Alternative splicing determines the domain structure of WWP1, a Nedd4 family protein.', *Biochemical and biophysical research communications*, 290(1), pp. 431–7.
- Foot, N. J., Dalton, H. E., Shearwin-whyatt, L. M., Dorstyn, L., Tan, S. and Yang, B. (2013) 'Regulation of the divalent metal ion transporter DMT1 and iron homeostasis by a ubiquitin-dependent mechanism involving Ndfips and Regulation of the divalent metal ion transporter DMT1 and iron homeostasis by a ubiquitin-dependent mechanism involving Ndfi', *Blood*, 112(10), pp. 4268–4275.
- Fu, H., Subramanian, R. R. and Masters, S. C. (2000) '14-3-3 P ROTEINS: Structure, Function, and Regulation', *Annual Review of Pharmacology and Toxicology*, 40, pp. 617–647.
- Furumatsu, T., Ozaki, T. and Asahara, H. (2009) 'Smad3 activates the Sox9-dependent transcription on chromatin.', *The international journal of biochemistry & cell biology.*, 41(5), pp. 1198–204.
- Ganji, A., Roshan, H. M., Varasteh, A., Moghadam, M. and Sankian, M. (2015) 'The effects of WW2/WW3 domains of Smurf2 molecule on TGF- $\beta$  signaling and arginase I gene expression', *Cell Biology International*, 39(6), pp. 690–695.
- Gao, S., Alarcón, C., Sapkota, G., Rahman, S., Chen, P.-Y., Goerner, N., Macias, M. J., Erdjument-Bromage, H., Tempst, P. and Massagué, J. (2009) 'Ubiquitin ligase Nedd4L targets activated Smad2/3 to limit TGF-beta signaling.', *Molecular cell.*, 36(3), pp. 457–68.
- Ginisty, H., Sicard, H., Roger, B. and Bouvet, P. (1999) 'Structure and functions of nucleolin.', *Journal of Cell Science*, 112 ( Pt 6, pp. 761–772.
- Goto, D., Nakajima, H., Mori, Y., Kurasawa, K., Kitamura, N. and Iwamoto, I. (2001) 'Interaction between Smad anchor for receptor activation and Smad3 is not essential for TGF-beta/Smad3-mediated signaling.', *Biochemical and biophysical research communications*, 281(5), pp. 1100–5. doi: 10.1006/bbrc.2001.4489.

- Grice, G. L. and Nathan, J. A. (2016) 'The recognition of ubiquitinated proteins by the proteasome', *Cellular and Molecular Life Sciences*. 73(18), pp. 3497–3506.
- Grou, C. P., Pinto, M. P., Mendes, A. V, Domingues, P. and Azevedo, J. E. (2015) 'The de novo synthesis of ubiquitin: identification of deubiquitinases acting on ubiquitin precursors.', *Scientific reports*. , 5, p. 12836.
- Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. and Young, R. A. (2007) 'A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells', *Cell*, 130(1), pp. 77–88.
- Guo, X. and Wang, X.-F. (2009) 'Signaling cross-talk between TGF-beta/BMP and other pathways.', *Cell research*. , 19(1), pp. 71–88.
- Haas, A. L. and Rose, I. A. (1982) 'The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis.', *Journal of Biological Chemistry*, 257(17), pp. 10329–10337.
- Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P. P. and Dikic, I. (2003) 'Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation.', *Nature cell biology*, 5(5), pp. 461–466
- Han, X.-J., Lee, M.-J., Yu, G.-R., Lee, Z.-W., Bae, J.-Y., Bae, Y.-C., Kang, S.-H. and Kim, D.-G. (2012) 'Altered dynamics of ubiquitin hybrid proteins during tumor cell apoptosis', *Cell Death and Disease*, 3, p. e255
- Handley-gearharts, P. M., Stephen, A. G., Trausch-azarf, J. S., Ciechanovers, A. and Schwartzfn, A. L. (1994) 'Human Ubiquitin-activating Enzyme , E 1' *The journal of Biological Chemistry*, 269(52) pp.33171-33178.
- Harrigan, J. A., Jacq, X., Martin, N. M. and Jackson, S. P. (2017) 'Deubiquitylating enzymes and drug discovery: emerging opportunities', *Nature Reviews Drug Discovery*. Nature Publishing Group
- Hata, a, Lo, R. S., Wotton, D., Lagna, G. and Massagué, J. (1997) 'Mutations increasing autoinhibition inactivate tumour suppressors

- Smad2 and Smad4.', *Nature*, 388(6637), pp. 82–87.
- Hayakawa, A., Saitoh, M. and Miyazawa, K. (2016) 'Dual Roles for Epithelial Splicing Regulatory Proteins 1 (ESRP1) and 2 (ESRP2) in Cancer Progression', in *Advances in experimental medicine and biology*, 925, pp. 33–40.
  - Hayashi, H., Abdollah, S., Qiu, Y., Cai, J., Xu, Y.-Y., Grinnell, B. W., Richardson, M. A., Topper, J. N., Gimbrone, M. A., Wrana, J. L. and Falb, D. (1997) 'The MAD-Related Protein Smad7 Associates with the TGF $\beta$  Receptor and Functions as an Antagonist of TGF $\beta$  Signaling', *Cell*, 89(7), pp. 1165–1173
  - Herhaus, L. and Sapkota, G. P. (2014) 'The emerging roles of deubiquitylating enzymes (DUBs) in the TGF $\beta$  and BMP pathways.', *Cellular signalling*, 26(10), pp. 2186–92.
  - Hoeben, A., Landuyt, B., Highley, M. S., Wildiers, H., Van Oosterom, A. T. and De Bruijn, E. A. (2004) 'Vascular Endothelial Growth Factor and Angiogenesis', *Pharmacological Reviews*, 56(4), pp. 549–580.
  - Honda, R., Tanaka, H. and Yasuda, H. (1997) 'Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53', *FEBS Letters*, 420(1), pp. 25–27.
  - Horiguchi, K., Sakamoto, K., Koinuma, D., Semba, K., Inoue, a, Inoue, S., Fujii, H., Yamaguchi, a, Miyazawa, K., Miyazono, K. and Saitoh, M. (2012) 'TGF- $\beta$  drives epithelial-mesenchymal transition through  $\delta$ EF1-mediated downregulation of ESRP', *Oncogene*, 31(26), pp. 3190–3201.
  - Huang, Y., Shi, H., Zhou, H., Song, X., Yuan, S., Luo, Y. and Dc, W. (2006) 'The angiogenic function of nucleolin is mediated by vascular endothelial growth factor and nonmuscle myosin The angiogenic function of nucleolin is mediated by vascular endothelial growth factor and nonmuscle myosin', *Blood*, 107(9), pp. 3564–3571.
  - Huibregtse, J. M., Scheffner, M., Beaudenon, S. and Howley, P. M. (1995) 'A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase.', *Proceedings of the National Academy of Sciences of the United States of America*, 92(7), pp.

2563–2567

- Humbert, L. and Lebrun, J.-J. (2013) 'TGF-beta inhibits human cutaneous melanoma cell migration and invasion through regulation of the plasminogen activator system', *Cellular Signalling*, 25, pp. 490–500
- Hussein, M. R. (2004) 'Transforming growth factor-b and malignant melanoma: molecular mechanisms'.
- Inoue, Y. and Imamura, T. (2008) 'Regulation of TGF- $\beta$  family signaling by E3 ubiquitin ligases', *Cancer Science*, 99(11), pp. 2107–2112
- Jilaveanu, L. B., Aziz, S. A. and Kluger, H. M. (2009) 'Chemotherapy and biologic therapies for melanoma: do they work?', *Clinics in Dermatology*, 27(6), pp. 614–625.
- Johnson, E. S. (2002) 'Ubiquitin branches out', *Nat Cell Biol*, 4(12), pp. E295-8
- Jovanovic, B., Egyhazi, S., Eskandarpour, M., Ghiorzo, P., Palmer, J. M., Bianchi Scarrà, G., Hayward, N. K. and Hansson, J. (2010) 'Coexisting NRAS and BRAF mutations in primary familial melanomas with specific CDKN2A germline alterations.', *The Journal of investigative dermatology*, 130(2), pp. 618–20.
- Jung, J.-G., Stoeck, A., Guan, B., Wu, R.-C., Zhu, H., Blackshaw, S., Shih, I.-M., Wang, T.-L., Swiatek, P., Lindsell, C., Amo, F. del, Weinmaster, G., Gridley, T., Krebs, L., Xue, Y., Norton, C., Sundberg, J., Beatus, P., Joutel, A., Park, J., Li, M., Nakayama, N., Davidson, B., Eberhart, C., Mosquera, J., Sboner, A., Zhang, L., Chen, C., Sung, Y., Park, J., Chen, X., Trope, C., Davidson, B., Ie, M. S., Chen, X., Thiaville, M., Chen, L., Stoeck, A., Xuan, J., Jeong, J., Jiang, L., Albino, E., Marrero, J., Rho, H., Wilkin, M., Carbery, A., Fostier, M., Aslam, H., Mazaleyrat, S., Sakata, T., Sakaguchi, H., Tsuda, L., Higashitani, A., Aigaki, T., Chastagner, P., Israel, A., Brou, C., Jia, L., Yu, G., Zhang, Y., Wang, M., Cerami, E., Gao, J., Dogrusoz, U., Gross, B., Sumer, S., Mermel, C., Schumacher, S., Hill, B., Meyerson, M., Beroukhi, R., Alqudah, M., Agarwal, S., Al-Keilani, M., Sibenaller, Z., Ryken, T.,

- McAuliffe, S., Morgan, S., Wyant, G., Tran, L., Muto, K., MacGurn, J., Hsu, P., Emr, S., Qiu, L., Joazeiro, C., Fang, N., Wang, H., Elly, C., McGill, M., Dho, S., Weinmaster, G., McGlade, C., Oberg, C., Li, J., Pauley, A., Wolf, E., Gurney, M., Hu, S., Wan, J., Su, Y., Song, Q. and Zeng, Y. (2014) 'Notch3 Interactome Analysis Identified WWP2 as a Negative Regulator of Notch3 Signaling in Ovarian Cancer', *PLoS Genetics*, 10(10), p. e1004751.
- Kalluri, R. and Weinberg, R. a (2009) 'Review series The basics of epithelial-mesenchymal transition', *Journal of Clinical Investigation*, 119(6), pp. 1420–1428.
  - Kavsak, P., Rasmussen, R. K., Causing, C. G., Bonni, S., Zhu, H., Thomsen, G. H. and Wrana, J. L. (2000) 'Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF $\beta$  receptor for degradation', *Molecular Cell*, 6(6), pp. 1365–1375.
  - Kim, J. E., Leung, E., Baguley, B. C. and Finlay, G. J. (2013) 'Heterogeneity of expression of epithelial-mesenchymal transition markers in melanocytes and melanoma cell lines', *Frontiers in Genetics*, 31(4) 97
  - Kimura, Y. and Tanaka, K. (2010) 'Regulatory mechanisms involved in the control of ubiquitin homeostasis', *Journal of Biochemistry*, 147(6), pp. 793–798
  - Kito, Y., Bai, J., Goto, N., Okubo, H., Adachi, Y., Nagayama, T. and Takeuchi, T. (2014) 'Pathobiological properties of the ubiquitin ligase Nedd4L in melanoma', *International Journal of Experimental Pathology*, 95(1), pp. 24–28.
  - Komander, D., Clague, M. J. and Urbé, S. (2009) 'Breaking the chains: structure and function of the deubiquitinases.', *Nature reviews. Molecular cell biology.* , 10(8), pp. 550–563.
  - Kong, F. M., Anscher, M. S., Murase, T., Abbott, B. D., Iglehart, J. D. and Jirtle, R. L. (1995) 'Elevated plasma transforming growth factor-beta 1 levels in breast cancer patients decrease after surgical removal of the tumor.', *Annals of surgery*, 222(2), pp. 155–62.
  - Lagier-Tourenne, C., Polymenidou, M. and Cleveland, D. W. (2010)

- 'TDP-43 and FUS/TLS: Emerging roles in RNA processing and neurodegeneration', *Human Molecular Genetics*, 19(R1), pp. 46–64.
- Lamouille, S., Xu, J. and Derynck, R. (2014) 'Molecular mechanisms of epithelial–mesenchymal transition', *Nature Reviews Molecular Cell Biology*. , 15(3), pp. 178–196.
  - Lasfar, A. and Cohen-Solal, K. A. (2010) 'Resistance to transforming growth factor ??-mediated tumor suppression in melanoma: Are multiple mechanisms in place?', *Carcinogenesis*, 31(10), pp. 1710–1717
  - Lauzier, A., Lavoie, R. R., Charbonneau, M., Gouin-Boisvert, B., Harper, K. and Dubois, C. M. (2016) 'Snail is a critical mediator of invadosome formation and joint degradation in arthritis', *American Journal of Pathology*. , 186(2), pp. 359–374.
  - Li, W., Bengtson, M. H., Ulbrich, A., Matsuda, A., Reddy, V. A., Orth, A., Chanda, S. K., Batalov, S. and Joazeiro, C. A. P. (2008) 'Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling', *PLoS ONE*, 3(1).
  - Lin, X., Liang, M. and Feng, X. H. (2000) 'Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent degradation of Smad2 in transforming growth factor-beta signaling', *Journal of Biological Chemistry*, 275(47), pp. 36818–36822. doi: 10.1074/jbc.C000580200.
  - Maddika, S., Kavela, S., Rani, N., Palicharla, V. R., Pokorny, J. L., Sarkaria, J. N. and Chen, J. (2011) 'WWP2 is an E3 ubiquitin ligase for PTEN.', *Nature Cell Biology*. , 13(6), pp. 728–733.
  - Marushige, Y. and Marushige, K. (1995) 'Disappearance of ubiquitinated histone H2A during chromatin condensation in TGF beta 1-induced apoptosis.', *Anticancer research*, 15(2), pp. 267–72.
  - Maspero, E., Mari, S., Valentini, E., Musacchio, A., Fish, A., Pasqualato, S. and Polo, S. (2011) 'Structure of the HECT:ubiquitin complex and its role in ubiquitin chain elongation.', *EMBO reports*. , 12(4), pp. 342–349.
  - Massagué, J. (1998) 'Tgf- B Signal Transduction', *Annual Review of*

*Biochemistry*, 67, pp. 753–91.

- Mastrocola, A. S., Kim, S. H., Trinh, A. T., Rodenkirch, L. A. and Tibbetts, R. S. (2013) 'The RNA-binding protein fused in sarcoma (FUS) functions downstream of poly(ADP-ribose) polymerase (PARP) in response to DNA damage.', *The Journal of biological chemistry*, 288(34), pp. 24731–41
- Maurer, G., Tarkowski, B. and Baccharini, M. (2011) 'Raf kinases in cancer—roles and therapeutic opportunities', *Oncogene*, 30(32), pp. 3477–3488.
- Mazars, P., Barboule, N., Baldin, V., Vidal, S., Ducommun, B. and Valette, A. (1995) 'Effects of TGF- $\beta$ 1 (transforming growth factor- $\beta$ 1) on the cell cycle regulation of human breast adenocarcinoma (MCF-7) cells', *FEBS Letters*, 362(3), pp. 295–300.
- Mclaughlin, F., Finn, P. and Thangue, N. B. La (2003) 'The cell cycle, chromatin and cancer : mechanism-based therapeutics come of age', *Drug Discovery Today*, 8(17), pp. 793–802.
- Medrano, E. E. (2003) 'Repression of TGF- $\beta$  signaling by the oncogenic protein SKI in human melanomas: consequences for proliferation, survival, and metastasis', *Oncogene*, 22(20), pp. 3123–3129.
- Melanoma Research Foundation (2017) *Stages of Melanoma Diagnosis | Melanoma Research Foundation*. Available at: <https://www.melanoma.org/understand-melanoma/diagnosing-melanoma/stages-of-diagnosis> (Accessed: 19 December 2017).
- Mendez, M. G., Kojima, S.-I. and Goldman, R. D. (2010) 'Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition.', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 24(6), pp. 1838–51
- Mertin, S., McDowall, S. G. and Harley, V. R. (1999) 'The DNA-binding specificity of SOX9 and other SOX proteins.', *Nucleic acids research*. Oxford University Press, 27(5), pp. 1359–64.
- Morikawa, M., Koinuma, D., Tsutsumi, S., Vasilaki, E., Kanki, Y.,

- Heldin, C. H., Aburatani, H. and Miyazono, K. (2011) 'ChIP-seq reveals cell type-specific binding patterns of BMP-specific Smads and a novel binding motif', *Nucleic Acids Research*, 39(20), pp. 8712–8727.
- Mort, R. L., Jackson, I. J. and Patton, E. E. (2015) 'The melanocyte lineage in development and disease', *Development*, 142(7), pp. 1387–1387.
  - Mund, T., Graeb, M., Mieszczanek, J., Gammons, M., Pelham, H. R. B. and Bienz, M. (2015) 'Disinhibition of the HECT E3 ubiquitin ligase WWP2 by polymerized Dishevelled', *Open Biology*, 5(12), pp. 3–13.
  - Mund, T. and Pelham, H. R. B. (2009) 'Control of the activity of WW-HECT domain E3 ubiquitin ligases by NDFIP proteins', *EMBO reports*, 10(5), pp. 501–507.
  - Munoz, M. A., Saunders, D. N., Henderson, M. J., Clancy, J. L., Russell, A. J., Lehrbach, G., Musgrove, E. A., Watts, C. K. W. and Sutherland, R. L. (2007) 'The E3 ubiquitin ligase EDD regulates S-phase and G2/M DNA damage checkpoints', *Cell Cycle*, 6(24), pp. 3070–3077.
  - Na, Y. R., Seok, S. H., Kim, D. J., Han, J. H., Kim, T. H., Jung, H., Lee, B. H. and Park, J. H. (2009) 'Bone morphogenetic protein 7 induces mesenchymal-to-epithelial transition in melanoma cells, leading to inhibition of metastasis', *Cancer Science*, 100(11), pp. 2218–2225.
  - Nye, M. D., Almada, L. L., Fernandez-Barrena, M. G., Marks, D. L., Elswa, S. F., Vrabel, A., Tolosa, E. J., Ellenrieder, V. and Fernandez-Zapico, M. E. (2014) 'The transcription factor GLI1 interacts with SMAD proteins to modulate transforming growth factor beta-induced gene expression in a p300/CREB-binding protein-associated factor (PCAF)-dependent manner', *Journal of Biological Chemistry*, 289(22), pp. 15495–15506.
  - Passeron, T., Valencia, J. C., Namiki, T., Vieira, W. D., Passeron, H., Miyamura, Y. and Hearing, V. J. (2009) 'Upregulation of SOX9 inhibits the growth of human and mouse melanomas and restores their sensitivity to retinoic acid', *Journal of Clinical Investigation*, 119(4) 954-963
  - Perrot, C. Y., Javelaud, D. and Mauviel, A. (2013) 'Insights into the

Transforming Growth Factor- $\beta$  Signaling Pathway in Cutaneous Melanoma.', *Annals of dermatology*. Korean Dermatological Association, 25(2), pp. 135–44.

- Pestell, R. G., Albanese, C., Watanabe, G., Johnson, J., Eklund, N., Lastowiecki, P. and Jameson, J. L. (1995) 'Epidermal growth factor and c-Jun act via a common DNA regulatory element to stimulate transcription of the ovine P-450 cholesterol side chain cleavage (CYP11A1) promoter.', *The Journal of biological chemistry*, 270(31), pp. 18301–8.
- Peth, A., Uchiki, T. and Goldberg, A. L. (2010) 'ATP-dependent steps in the binding of ubiquitin conjugates to the 26S proteasome that commit to degradation.', *Molecular cell*, 40(4), pp. 671–81.
- Pickart, C. M. and Eddins, M. J. (2004) 'Ubiquitin: Structures, functions, mechanisms', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1695(1–3), pp. 55–72.
- Piek, E. and Roberts, A. B. (2001) 'Suppressor and oncogenic roles of transforming growth factor-beta and its signaling pathways in tumorigenesis.', *Advances in cancer research*, 83, pp. 1–54.
- Portales-Casamar, E., Arenillas, D., Lim, J., Swanson, M. I., Jiang, S., McCallum, A., Kirov, S. and Wasserman, W. W. (2009) 'The PAZAR database of gene regulatory information coupled to the ORCA toolkit for the study of regulatory sequences', *Nucleic Acids Research*, 37(suppl\_1), pp. D54–D60.
- Prenzel, T., Begus-Nahrman, Y., Kramer, F., Hennion, M., Hsu, C., Gorsler, T., Hintermair, C., Eick, D., Kremmer, E., Simons, M., Beissbarth, T. and Johnsen, S. A. (2011) 'Estrogen-Dependent Gene Transcription in Human Breast Cancer Cells Relies upon Proteasome-Dependent Monoubiquitination of Histone H2B', *Cancer Research*, 71(17), pp. 5739–5753
- Qin, Y., Xu, S.-Q., Pan, D.-B., Ye, G.-X., Wu, C.-J., Wang, S., Wang, C.-J., Jiang, J.-Y. and Fu, J. (2016) 'Silencing of WWP2 inhibits adhesion, invasion, and migration in liver cancer cells', *Tumor Biology*.
- Raikwar, N. S. and Thomas, C. P. (2008) 'Nedd4-2 isoforms

ubiquitinate individual epithelial sodium channel subunits and reduce surface expression and function of the epithelial sodium channel.', *Am J Physiol Renal Physiol*, 294(5), pp. F1157-65.

- Raychaudhuri, K., Chaudhary, N., Gurjar, M., D'Souza, R., Limzerwala, J., Maddika, S. and Dalal, S. N. (2016) '14-3-3 $\sigma$  gene loss leads to activation of the epithelial to mesenchymal transition due to the stabilization of c-Jun protein', *Journal of Biological Chemistry*, 291(31), pp. 16068–16081
- Riesco-Eizaguirre, G., Rodriguez, I., De la Vieja, A., Costamagna, E., Carrasco, N., Nistal, M. and Santisteban, P. (2009) 'The BRAFV600E Oncogene Induces Transforming Growth Factor Secretion Leading to Sodium Iodide Symporter Repression and Increased Malignancy in Thyroid Cancer', *Cancer Research*, 69(21), pp. 8317–8325
- Rizo, J. and Südhof, T. C. (1998) 'C2-domains, structure and function of a universal Ca<sup>2+</sup>-binding domain.', *The Journal of biological chemistry*, 273(26), pp. 15879–82.
- Rothhammer, T., Bataille, F., Spruss, T., Eissner, G. and Bosserhoff, A.-K. (2007) 'Functional implication of BMP4 expression on angiogenesis in malignant melanoma', *Oncogene.*, 26(28), pp. 4158–4170
- Rothhammer, T., Poser, I., Soncin, F., Bataille, F., Moser, M. and Bosserhoff, A.-K. (2005) 'Bone Morphogenic Proteins Are Overexpressed in Malignant Melanoma and Promote Cell Invasion and Migration', *Cancer Research*, 65(2), pp. 448-456.
- Sadowski, M., Suryadinata, R., Tan, A. R., Roesley, S. N. A. and Sarcevic, B. (2012) 'Protein monoubiquitination and polyubiquitination generate structural diversity to control distinct biological processes', *IUBMB Life*, 64(2), pp. 136–142.
- Saitoh, M. and Miyazawa, K. (2012) 'Transcriptional and Post-transcriptional Regulation in TGF-beta-mediated epithelial-mesenchymal transition', *Journal of Biochemistry*, 151(6), pp.563-571
- Saxena, a, Rorie, C. J., Dimitrova, D., Daniely, Y. and Borowiec, J. a

- (2006) 'Nucleolin inhibits Hdm2 by multiple pathways leading to p53 stabilization.', *Oncogene*, 25(55), pp. 7274–7288.
- Scheffner, M. and Staub, O. (2007) 'HECT E3s and human disease', *BMC Biochem*, 8 Suppl 1, p. S6.
  - Schramek, D., Sendoel, A., Segal, J. P., Beronja, S., Heller, E., Oristian, D., Reva, B. and Fuchs, E. (2014) 'Direct in vivo RNAi screen unveils myosin IIa as a tumor suppressor of squamous cell carcinomas.', *Science (New York, N.Y.)*, 343(6168), pp. 309–13.
  - Shema, E., Tirosh, I., Aylon, Y., Huang, J., Ye, C., Moskovits, N., Raver-Shapira, N., Minsky, N., Pirngruber, J., Tarcic, G., Hublarova, P., Moyal, L., Gana-Weisz, M., Shiloh, Y., Yarden, Y., Johnsen, S. A., Vojtesek, B., Berger, S. L. and Oren, M. (2008) 'The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression', *Genes & Development*, 22(19), pp. 2664–2676.
  - Shi, Y. and Massagué, J. (2003) 'Mechanisms of TGF-beta signaling from cell membrane to the nucleus.', *Cell*, 113(6), pp. 685–700.
  - Smith, M. P., Ferguson, J., Arozarena, I., Hayward, R., Marais, R., Chapman, A., Hurlstone, A. and Wellbrock, C. (2013) 'Effect of SMURF2 targeting on susceptibility to MEK inhibitors in melanoma', *Journal of the National Cancer Institute*, 105(1), pp. 33–46.
  - Sommer, G., Rossa, C., Chi, A. C., Neville, B. W. and Heise, T. (2011) 'Implication of RNA-binding protein La in proliferation, migration and invasion of lymph node-metastasized hypopharyngeal SCC cells', *PLoS ONE*, 6(10), pp. 1–8.
  - Sommer, Dittmann, J., Kuehnert, J., Reumann, K., Schwartz, P. E., Will, H., Coulter, B. L., Smith, M. T. and Heise, T. (2011) 'The RNA-binding protein La contributes to cell proliferation and CCND1 expression.', *Oncogene*, 30(4), pp. 434–444.
  - Soond, S. and Chantry, A. (2011) 'Selective targeting of activating and inhibitory Smads by distinct WWP2 ubiquitin ligase isoforms differentially modulates TGFb signalling and EMT', *Oncogene*, 30, pp. 2451–2462

- Soond, S. M. and Chantry, A. (2011) 'How ubiquitination regulates the TGF $\beta$ - signalling pathway: New insights and new players: New isoforms of ubiquitin-activating enzymes in the E1-E3 families join the game', *BioEssays*, 33. pp.749-758
- Soond, S. M., Smith, P. G., Wahl, L., Swingler, T. E., Clark, I. M., Hemmings, A. M. and Chantry, A. (2013) 'Novel WWP2 ubiquitin ligase isoforms as potential prognostic markers and molecular targets in cancer', *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1829, pp.2127-2135
- Spiess, C., Meyer, A. S., Reissmann, S. and Frydman, J. (2004) 'Mechanism of the eukaryotic chaperonin: Protein folding in the chamber of secrets', *Trends in Cell Biology*, 14(11), pp. 598–604
- Stahl, J. M., Cheung, M. and Sharma, A. (2003) 'Loss of PTEN Promotes Tumor Development in Malignant Melanoma Loss of PTEN Promotes Tumor Development in Malignant Melanoma 1', pp. 2881–2890.
- Stephen, A. G., Trausch-Azar, J. S., Ciechanover, A. and Schwartz, A. L. (1996) 'The ubiquitin-activating enzyme E1 is phosphorylated and localized to the nucleus in a cell cycle-dependent manner', *Journal of Biological Chemistry*, 271(26), pp. 15608–15614.
- Stopa, M., Anhuf, D., Terstegen, L., Gatsios, P., Gressner, A. M. and Dooley, S. (2000) 'Participation of Smad2, Smad3, and Smad4 in transforming growth factor  $\beta$  (TGF- $\beta$ )-induced activation of Smad7: The TGF- $\beta$  response element of the promoter requires functional Smad binding element and E-box sequences for transcriptional regulation', *Journal of Biological Chemistry*, 275(38), pp. 29308–29317
- Subbaiah, V. K., Zhang, Y., Rajagopalan, D., Abdullah, L. N., Yeo-Teh, N. S. L., Tomaić, V., Banks, L., Myers, M. P., Chow, E. K. and Jha, S. (2016) 'E3 ligase EDD1/UBR5 is utilized by the HPV E6 oncogene to destabilize tumor suppressor TIP60', *Oncogene*, 35(16), pp. 2062–2074
- Sudol, M., Chen, H. I., Bougeret, C. and Einbond, A. (1995)

- 'Characterization of a novel protein-binding module - the W W domain', *FEBS Letters*, 369, pp. 67–71.
- Sudol, M., Recinos, C. C., Abraczinskas, J., Humbert, J. and Farooq, A. (2005) 'WW or WoW: the WW domains in a union of bliss.', *IUBMB life*, 57(12), pp. 773–8
  - Sudol, M., Sliwa, K. and Russo, T. (2001) 'Functions of WW domains in the nucleus', *FEBS Letters*, 490(3), pp. 190–195.
  - Sun, L. and Chen, Z. J. (2004) 'The novel functions of ubiquitination in signaling', *Current Opinion in Cell Biology*, 16, pp.119-126
  - Sun, Y. (2006) 'E3 ubiquitin ligases as cancer targets and biomarkers.', *Neoplasia (New York, N.Y.)*. Neoplasia Press, 8(8), pp. 645–54.
  - Swatek, K. N. and Komander, D. (2016) 'Ubiquitin modifications', *Cell Research*, 26, pp.399-422
  - Tang, L.-Y., Yamashita, M., Coussens, N. P., Tang, Y., Wang, X., Li, C., Deng, C.-X., Cheng, S. Y. and Zhang, Y. E. (2011) 'Ablation of Smurf2 reveals an inhibition in TGF- $\beta$  signalling through multiple mono-ubiquitination of Smad3.', *The EMBO journal*, 30(23), pp. 4777–89.
  - Tardif, G., Pelletier, J.-P., Fahmi, H., Hum, D., Zhang, Y., Kapoor, M. and Martel-Pelletier, J. (2013) 'NFAT3 and TGF- $\beta$ /SMAD3 regulate the expression of miR-140 in osteoarthritis.', *Arthritis research & therapy*. BioMed Central, 15(6), p. R197
  - Tenno, T., Fujiwara, K., Tochio, H., Iwai, K., Morita, E. H., Hayashi, H., Murata, S., Hiroaki, H., Sato, M., Tanaka, K. and Shirakawa, M. (2004) 'Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains', *Genes to Cells*, 9(10), pp. 865–875.
  - Thrower, J. S., Hoffman, L., Rechsteiner, M. and Pickart, C. M. (2000) 'Recognition of the polyubiquitin proteolytic signal.', *The EMBO journal*, 19(1), pp. 94–102
  - Valimberti, I., Tiberti, M., Lambrughi, M., Sarcevic, B. and Papaleo, E. (2015) 'E2 superfamily of ubiquitin-conjugating enzymes: constitutively active or activated through phosphorylation in the catalytic cleft', *Scientific Reports*, 5(April), p. 14849.

- Venkatesan, N., Pini, L. and Ludwig, M. S. (2004) 'Changes in Smad expression and subcellular localization in bleomycin-induced pulmonary fibrosis', *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 287(6), pp.L1342-1347.
- Viel, S., Marçais, A., Guimaraes, F. S., Loftus, R., Rabilloud, J., Grau, M., Degouve, S., Djebali, S., Sanlaville, A., Charrier, E., Bienvenu, J., Marie, J. C., Caux, C., Marvel, J., Town, L., Huntington, N. D., Bartholin, L., Finlay, D., Smyth, M. J. and Walzer, T. (2016) 'TGF- $\beta$  inhibits the activation and functions of NK cells by repressing the mTOR pathway.', *Science signaling*, 9(415), p. ra19.
- Vijay-kumar, S., Bugg, C. E. and Cook, W. J. (1987) 'Structure of ubiquitin refined at 1.8 Å resolution', *Journal of Molecular Biology*, 194(3), pp. 531–544.
- Vissers, J. H., Nicassio, F., van Lohuizen, M., Di Fiore, P. P. and Citterio, E. (2008) 'The many faces of ubiquitinated histone H2A: insights from the DUBs.', *Cell division*, 3, p. 8.
- Wahl, L. (2016) *Decoding the structure and function of WWp2 in the TGF $\beta$  signalling pathway*. University of East Anglia.
- Wang, H., Yang, G.-H., Bu, H., Zhou, Q., Guo, L.-X., Wang, S.-L. and Ye, L. (2003) 'Systematic analysis of the TGF-beta/Smad signalling pathway in the rhabdomyosarcoma cell line RD.', *International journal of experimental pathology*, 84(3), pp. 153–163.
- Wang, J., Peng, Q., Lin, Q., Childress, C., Carey, D. and Yang, W. (2010) 'Calcium activates Nedd4 E3 ubiquitin ligases by releasing the C2 domain-mediated auto-inhibition', *Journal of Biological Chemistry*, 285(16), pp. 12279–12288.
- Wang, Y., Liu, J., Huang, B. O., Xu, Y.-M., Li, J., Huang, L.-F., Lin, J., Zhang, J., Min, Q.-H., Yang, W.-M. and Wang, X.-Z. (2015) 'Mechanism of alternative splicing and its regulation.', *Biomedical reports*, 3(2), pp. 152–158
- Warzecha, C. C., Jiang, P., Amirikian, K., Dittmar, K. A., Lu, H., Shen, S., Guo, W., Xing, Y. and Carstens, R. P. (2010) 'An ESRP-regulated splicing programme is abrogated during the

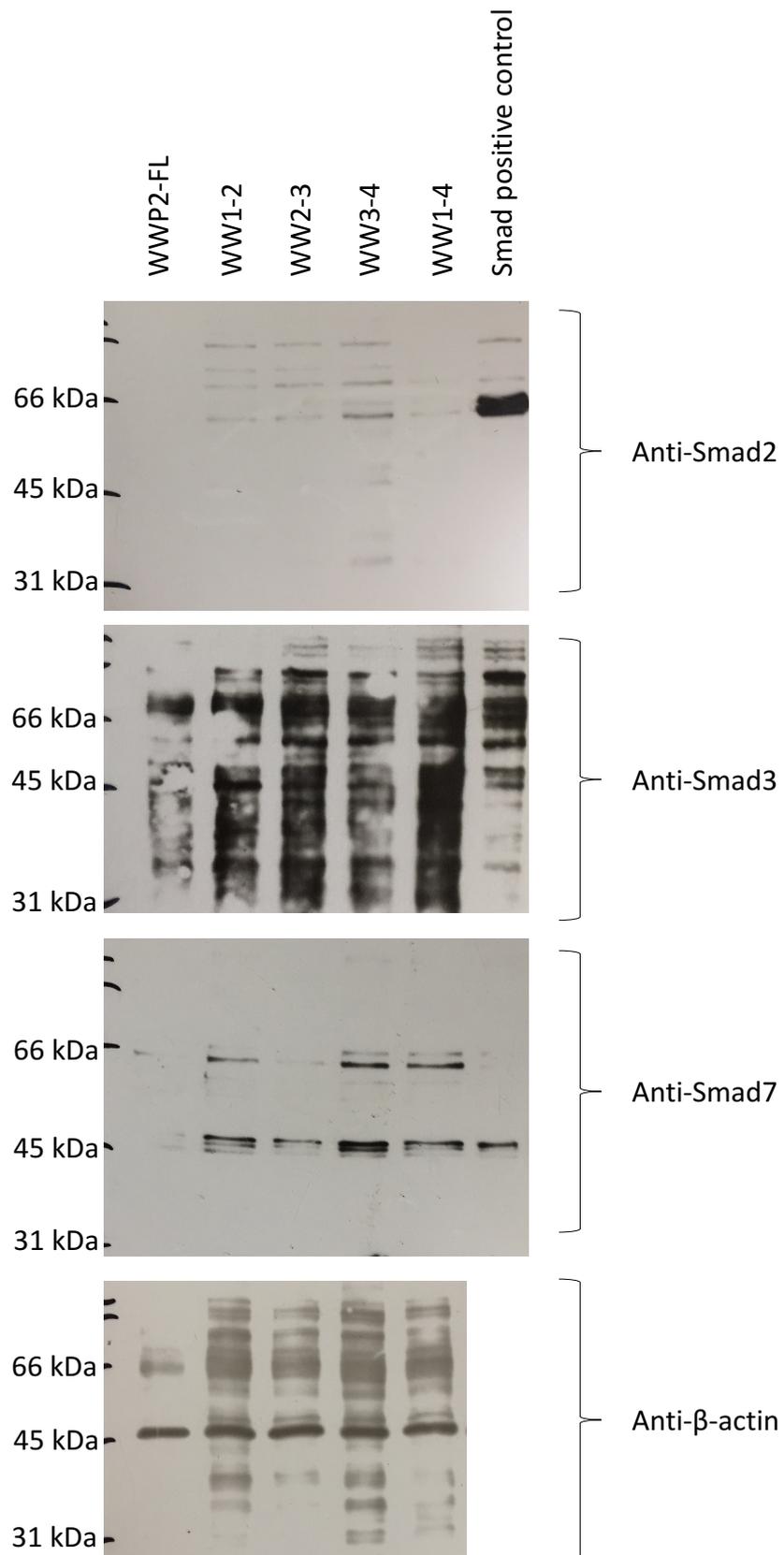
- epithelial–mesenchymal transition’, *The EMBO Journal*, 29, pp. 3286–3300
- Warzecha, C. C., Shen, S., Xing, Y. and Carstens, R. P. (2009a) ‘The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events’, *RNA biology*, 6(5), pp. 546–62.
  - Warzecha, C. C., Shen, S., Xing, Y. and Carstens, R. P. (2009b) ‘The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events.’, *RNA biology*, 6(5), pp. 546–62.
  - Wendt, M. K., Tian, M. and Schiemann, W. P. (2012) ‘Deconstructing the mechanisms and consequences of TGF-beta-induced EMT during cancer progression’, *Cell and Tissue Research*, 347, pp. 85-101
  - Wiesner, S., Ogunjimi, A. A., Wang, H. R., Rotin, D., Sicheri, F., Wrana, J. L. and Forman-Kay, J. D. (2007) ‘Autoinhibition of the HECT-Type Ubiquitin Ligase Smurf2 through Its C2 Domain’, *Cell*, 130(4), pp. 651–662
  - Wójcik, C. and DeMartino, G. N. (2003) ‘Intracellular localization of proteasomes.’, *The international journal of biochemistry & cell biology*, 35(5), pp. 579–89.
  - Wotton, D., Lo, R. S., Massagué, J., Swaby, L. C. and Massague, J. (1999) ‘Multiple Modes of Repression by the Smad Transcriptional Corepressor TGIF Multiple Modes of Repression by the Smad Transcriptional Corepressor TGIF \*’, *CELL BIOLOGY AND METABOLISM*, 274(52), pp. 37105–37110.
  - Wu, J. W., Hu, M., Chai, J., Seoane, J., Huse, M., Li, C., Rigotti, D. J., Kyin, S., Muir, T. W., Fairman, R., Massagué, J. and Shi, Y. (2001) ‘Crystal structure of a phosphorylated Smad2: Recognition of phosphoserine by the MH2 domain and insights on Smad function in TGF-beta signaling’, *Molecular Cell*, 8(6), pp. 1277–1289.
  - [www.bad.org.uk](http://www.bad.org.uk) (2013) *Stage 3*.
  - [www.cancerresearchuk.org](http://www.cancerresearchuk.org) (2015a) *Stage 3 | Cancer Research UK*. Available at: <http://about-cancer.cancerresearchuk.org/about->

- cancer/melanoma/stages-types/stage-3 (Accessed: 9 February 2017).
- [www.cancerresearchuk.org](http://www.cancerresearchuk.org) (2015b) *Stage 4 | Cancer Research UK*. Available at: <http://about-cancer.cancerresearchuk.org/about-cancer/melanoma/stages-types/stage-4> (Accessed: 9 February 2017).
  - Xiao, Z., Liu, X. and Lodish, H. F. (2000) 'Importin beta mediates nuclear translocation of Smad 3', *Journal of Biological Chemistry*, 275(31), pp. 23425–23428.
  - Xu, L. (2006) 'Regulation of Smad activities.', *Biochimica et biophysica acta.*, 1759(11–12), pp. 503–13.
  - Xu, L., Kang, Y., Çöl, S. and Massagué, J. (2002) 'Smad2 nucleocytoplasmic shuttling by nucleoporins CAN/Nup214 and Nup153 feeds TGF $\beta$  signaling complexes in the cytoplasm and nucleus', *Molecular Cell*, 10(2), pp. 271–282.
  - Xu, P., Liu, J. and Derynck, R. (2012) 'Post-translational regulation of TGF- $\beta$  receptor and Smad signaling.', *FEBS letters.*, 586(14), pp. 1871–84.
  - Yan, X., Liao, H., Cheng, M., Shi, X., Lin, X., Feng, X. H. and Chen, Y. G. (2016) 'Smad7 protein interacts with receptor-regulated Smads (R-Smads) to inhibit transforming growth factor-beta (TGF-beta)/Smad signaling', *Journal of Biological Chemistry*, 291(1), pp. 382–392.
  - Yan, X., Liu, Z. and Chen, Y. (2009) 'Regulation of TGF- signaling by Smad7', *Acta Biochimica et Biophysica Sinica*, 41(4), pp. 263–272.
  - Yang, J. and Weinberg, R. A. (2008) 'Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis', *Developmental Cell*, 14(6), pp. 818–829.
  - Yang, R., He, Y., Chen, S., Lu, X., Huang, C. and Zhang, G. (2016) 'Elevated expression of WWP2 in human lung adenocarcinoma and its effect on migration and invasion', *Biochemical and Biophysical Research Communications*, 479, pp. 146–151.
  - Yokota, S., Kayano, T., Ohta, T., Kurimoto, M., Yanagi, H., Yura, T. and Kubota, H. (2000) 'Proteasome-Dependent Degradation of Cytosolic

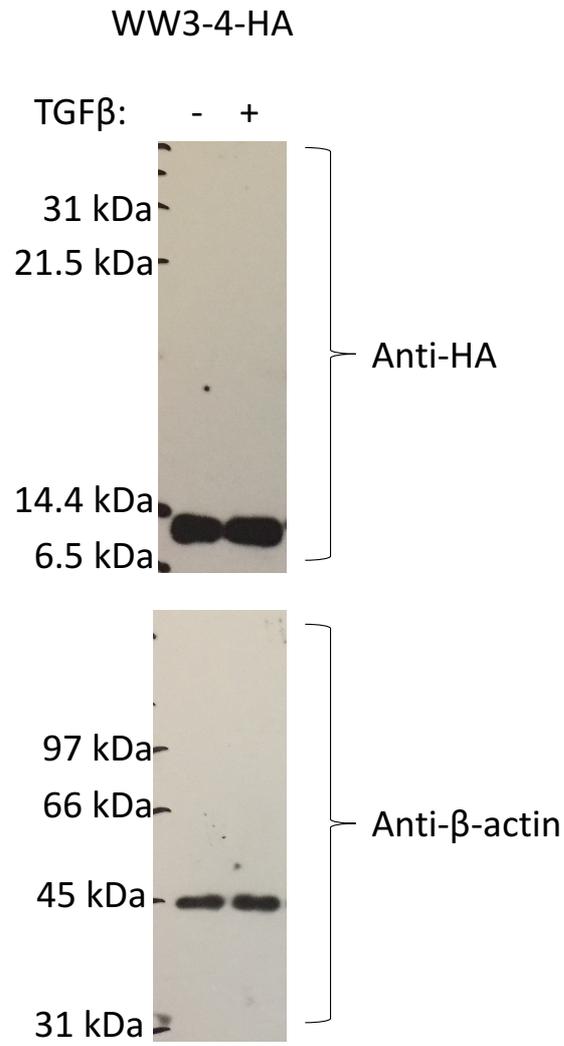
Chaperonin CCT', *Biochemical and Biophysical Research Communications*, 279(2), pp. 712–717.

- Zavadil, J. and Böttinger, E. P. (2005) 'TGF- $\beta$  and epithelial-to-mesenchymal transitions', *Oncogene*, 24(37), pp. 5764–5774.
- Zhang, L., Zhou, F., Drabsch, Y., Gao, R., Snaar-Jagalska, B. E., Mickanin, C., Huang, H., Sheppard, K.-A., Porter, J. a., Lu, C. X. and ten Dijke, P. (2012) 'USP4 is regulated by AKT phosphorylation and directly deubiquitylates TGF- $\beta$  type I receptor', *Nature Cell Biology*, 14(7), pp. 717–726
- Zhang, S., Fei, T., Zhang, L., Zhang, R., Chen, F., Ning, Y., Han, Y., Feng, X., Meng, A. and Chen, Y. (2007) 'Smad7 antagonizes transforming growth factor beta signaling in the nucleus by interfering with functional Smad-DNA complex formation.', *Molecular and cellular biology*, 27(12), pp. 4488–4499
- Zhu, H., Kavsak, P., Abdollah, S., Wrana, J. L. and Thomsen, G. H. (1999) 'A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation.', *Nature*, 400(6745), pp. 687–693.
- Zhu, W., He, X., Hua, Y., Li, Q., Wang, J. and Gan, X. (2017) 'E3 Ubiquitin Ligase WWP2 Facilitates RUNX2 Transactivation Through A Mono-ubiquitination Manner During Osteogenic Differentiation', *Journal of Biological Chemistry*, (1), p. jbc.M116.772277.
- Zou, W., Chen, X., Shim, J.-H., Huang, Z., Brady, N., Hu, D., Drapp, R., Sigrist, K., Glimcher, L. H. and Jones, D. (2011) 'The E3 ubiquitin ligase Wwp2 regulates craniofacial development through mono-ubiquitylation of Goosecoid', *Nature Cell Biology*, 13(1), pp. 59–65

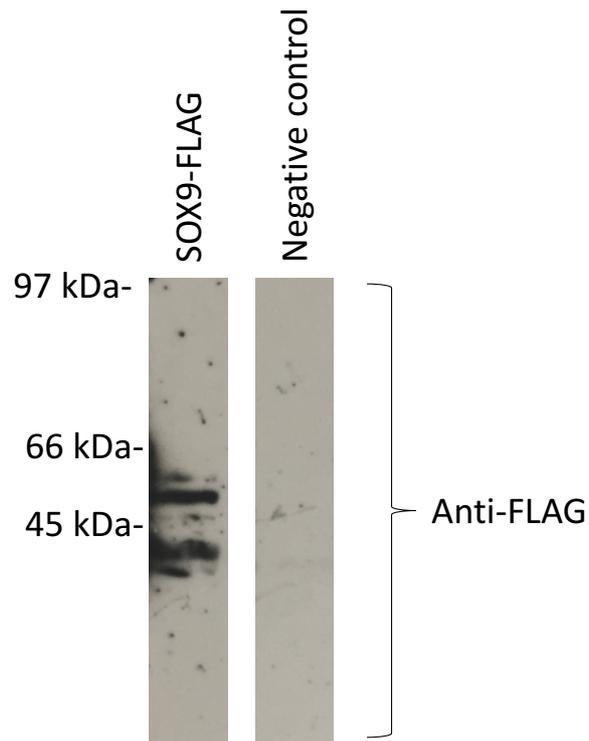
## Appendix



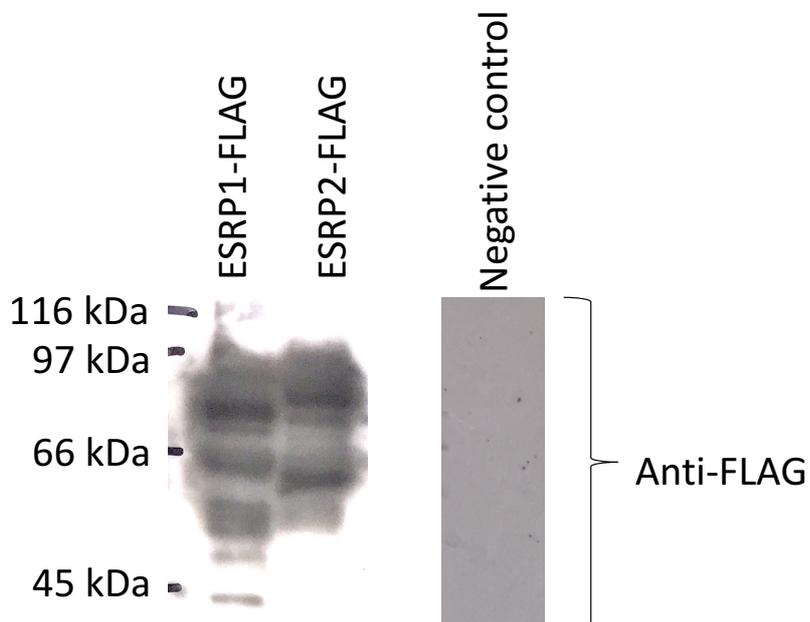
A 1- Full versions of Western blots performed using anti-Smad 2,3 and 7 antibodies to show Smad levels in HEK293A cells overexpressing WWP2 constructs as shown in figure 3.2.9. Smad 2, 3 and 7 are used in the respective blots as positive controls.



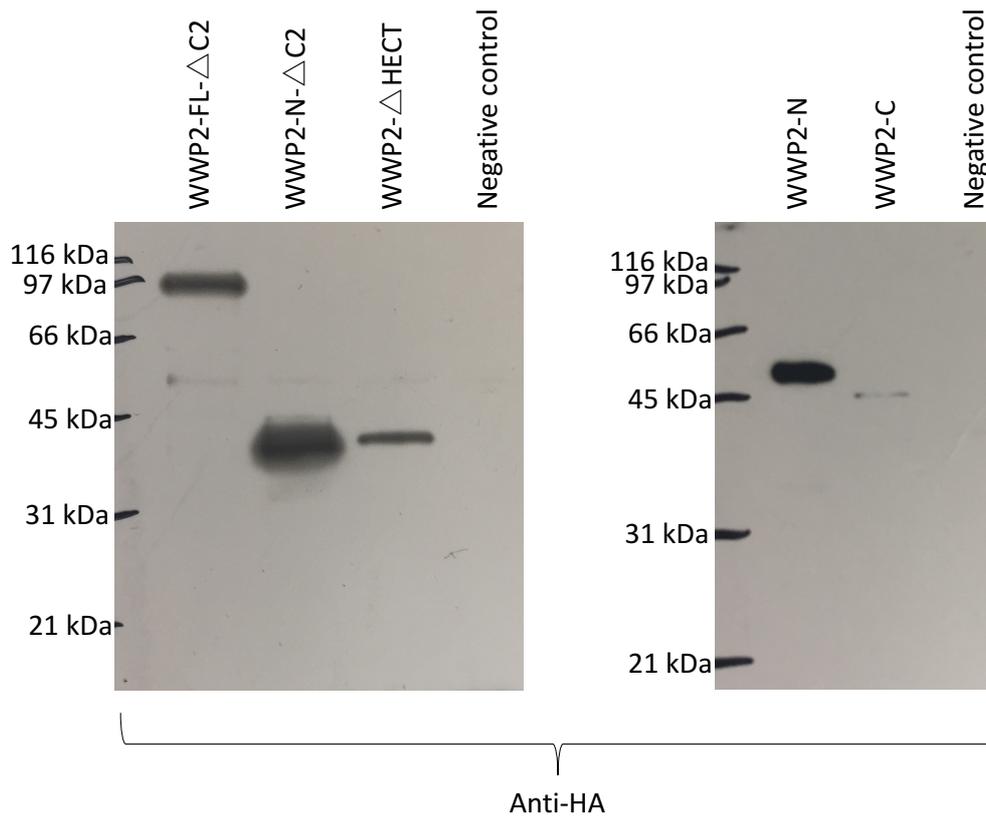
A 2- Full versions of Western blots using anti-HA and  $\beta$ -actin antibodies shown in figure 3.2.10. This western blot confirms the presence of HA-WW3-4 in HEK293A cells transfected with HA-WW3-4.



A 3- Full versions of Western blots using anti-FLAG antibodies shown in figure 5.2.6. The presence of SOX9 confirms successful transfection of SOX9 construct. Note that negative control was a HA- tagged WWP2 construct.



A 4- Full versions of Western blots using anti-FLAG antibodies shown in figure 5.2.16. The above western blot confirms the overexpression of ESRP constructs indicating a successful transfection. Negative control was an HA-tagged WWP2 construct. Note that the negative control lane derived from a different picture of the same blot therefore showing a different exposure.



A 5- Full versions of Western blots using anti-HA antibodies shown in figure 5.2.19. The expression of transfected HA-tagged WWP2 constructs were confirmed using anti-HA antibodies.