DOCTOR OF PHILOSOPHY

Brachyury in the Human Colon and Colorectal Cancer

Williams, Jason

Award date:
2018

Awarding institution:
Bangor University

PRIFYSGOL
BANGOR
UNIVERSITY

Brachyury in the Human Colon and Colorectal Cancer

Ph. D. Thesis
2017

Jason Saunders Williams

Declaration and Consent

Details of the Work

I hereby agree to deposit the following item in the digital repository maintained by Bangor University and/or in any other repository authorized for use by Bangor University.

Author Name:
Title:
Supervisor/Department:
Funding body (if any):
Qualification/Degree obtained: PhD
This item is a product of my own research endeavours and is covered by the agreement below in which the item is referred to as "the Work". It is identical in content to that deposited in the Library, subject to point 4 below.
Non-exclusive Rights
Rights granted to the digital repository through this agreement are entirely non-exclusive. I am free to publish the Work in its present version or future versions elsewhere.
I agree that Bangor University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Bangor University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.
Bangor University Digital Repository
I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web.
I understand that once the Work is deposited, the item and its metadata may be incorporated into public access catalogues or services, national databases of electronic theses and dissertations such as the British Library's EThOS or any service provided by the National Library of Wales.
I understand that the Work may be made available via the National Library of Wales Online Electronic Theses Service under the declared terms and conditions of use (http://www.llgc.org.uk/index.php). I agree that as part of this service the National Library of Wales may electronically store, copy or convert the Work to any approved medium or format for the purpose of future preservation and accessibility. The National Library of Wales is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

Statement 1:
This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless as agreed by the University for approved dual awards.
Signed: ……..
Date: ……

Statement 2:

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).
All other sources are acknowledged by footnotes and/or a bibliography.
Signed: ……
Date: ……

Statement 3:

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loan and for electronic storage (subject to any constraints as defined in statement 4), and for the title and summary to be made available to outside organisations.

Signed: ……
Date: ……

NB: Candidates on whose behalf a bar on access has been approved by the Academic Registry should use the following version of Statement 3:
Statement 3 (bar):
I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loans and for electronic storage (subject to any constraints as defined in statement 4), after expiry of a bar on access.

Signed: ……
Date: ……

Statement 4:

I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University and where necessary have gained the required permissions for the use of third party material.
In addition to the above I also agree to the following:
1. That I am the author or have the authority of the author(s) to make this agreement and do hereby give Bangor University the right to make available the Work in the way described above.
2. That the electronic copy of the Work deposited in the digital repository and covered by this agreement, is identical in content to the paper copy of the Work deposited in the Bangor University Library, subject to point 4 below.
3. That I have exercised reasonable care to ensure that the Work is original and, to the best of my knowledge, does not breach any laws – including those relating to defamation, libel and copyright.
4. That I have, in instances where the intellectual property of other authors or copyright holders is included in the Work, and where appropriate, gained explicit

permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, or that I have identified and removed that material for which adequate and appropriate permission has not been obtained and which will be inaccessible via the digital repository.
5. That Bangor University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.
6. That I will indemnify and keep indemnified Bangor University and the National Library of Wales from and against any loss, liability, claim or damage, including without limitation any related legal fees and court costs (on a full indemnity bases), related to any breach by myself of any term of this agreement.

Signature: ……
Date: ……

Summary

The T-box transcription factor Brachyury (T) has a well-established role in essential morphogenetic movements in the developing embryo, but in recent years has emerged as a cancer biomarker and promising therapeutic target. Studies in patient-derived cancer tissues and cancer cells suggest that Brachyury is associated with poor patient survival and the epithelial to mesenchymal transition. These observations suggest that Brachyury has an important function in the progression of several cancers. However, the exact mechanism by which it is involved in cancer progression is still incompletely understood.

Here, we show that Brachyury has two expressed transcript variants (TV1 and TV2) that are expressed in both normal and cancer tissues. We report the expression of Brachyury in a subset of enteroendocrine cells (EECs) of the normal human colonic crypts, marked by the EEC marker Chromogranin A and occasionally by the reserve stem cell marker Lrig1. We hypothesise that these cells are a population of reserve stem cells in the normal colonic crypt and that Brachyury contributes to their maintenance and plastic response to damage and environmental cues.

Brachyury positive EEC-like cells were observed in colorectal adenomas, primary colorectal cancers and metastases of colorectal origin. Furthermore, expression of Brachyury in EEC-like cells in primary colorectal carcinomas confers a poor survival outcome and progression-free survival in patients in which they are present.

Acknowledgements

Funding for this work was provided by Coleg Cymraeg Cenedlaethol (CCC) and National Institute for Social Care and Health Research Academic Health Science Collaboration (NISCHR AHSC). I would like to thank John Sammut who obtained tissue samples from patients undergoing colonic resection and my co-supervisor Simon Gollins who obtained colorectal tissue samples as part of the RICE trial (ISRCTN59158409 DOI 10.1186/ISRCTN59158409). Invaluable advice on histopathology was received from Dr Ian Cree, Dr Sarah Coupland, Dr Anthony Caslin and Dr Geraint Williams. Thanks to Dr Vin Sahota for advice on confocal microscopy.

Thanks to everyone in D7 lab for educating me inside and outside the lab. I will miss making and talking about coffee, lunch times will never be the same, Ellen Vernon, Natalia Gomez-Escobar, Vicente Planells-Palop, Erik Waskiewicz, Jana Jezkova and Vin.

Ali, Fayez, Ahmed, Naif, Othman, Fahad and Hosam, I will miss our colourful conversations and cultural exchanges.

Diolch arbennig i'r Athro Deri Tomos am fod yn fentor a hefyd fy efengleiddiad i achos yr iaith Cymraeg.

A special thanks to Jane Wakeman for spending hours scrutinising my work and being a wonderful supervisor throughout.

Finally, a special thanks to my family and friends for their patience and perseverance.

Jason

Abbreviations

Adult Intestinal Stem Cells	AISCs
Adult Stem Cells	ASCs
Androgen Receptor	AR
American Type Culture Collection	ATCC
Bicinchonionic acid	BCA
complementary DNA	cDNA
Crypt Base Columnar	CBC
Confidence Interval	CI
Colorectal Cancer	CRC
Computerised Tomography	CT
Deep Crypt Secretory	DCS
Deoxyribose nucleic acid	DNA
Embryonic Stem Cells	ESCs
Enteroendocrine cells	EECs
Epithelial to Mesenchymal Transition	EMT
Epidermal Growth Factor Receptor	EGFR
Familial Adenomatous Polyposis	FAP
Fetal Bovine Serum, 45	FBS
Formalin-fixed Paraffin-embedded, 50	FFPE
Fibroblast Growth Factor	FGF
Green Fluorescent Protein	GFP
Gastrointestinal	GI, 9
Haemotoxylin and Eosin, 50	H&E
Hereditary Non-polyposis Colorectal Cancer	HNCC
High Fat Diet	HFD
Horse Radish Peroxidase, 47	HRP
Human Embryonic Stem cells	HESCs
Indirect Immunofluorescence, 51	IF
Intestinal Stem Cells	ISCs
Immunohisotchemistry	IHC
Inner Cell Mass	ICM
Knockdown	KD
Kilodalton	kDa
Knockout	KO
Magnetic Resonance Imaging	MRI
mouse Embryonic Stem Cells	mESCs
Mesenchymal to Epithelial Transition	MET
Microsatellite Instability	MSI
mitochondrial DNA	mtDNA
messenger ribonucleic acid	mRNA
Multiple Cloning Site	MCS
Phosphate-buffered saline, 47	PBS
Paraformaldehyde, 51	PFA

Primordial Germ Cells,	PGCs
Positron Emission Tomography	PET
Polyvinylidene fluoride, 47	PVDF
quantitative Real-time Polymerase Chain Reaction,	qPCR
regenerating islet derived family member 4	Reg4
Ribonucleic Acid	RNA
Roswell Park Memorial Institute, 45	RPMI
small interfering Ribose Nucleic Acid, 45	siRNA
Small Molecule Inhibitors	SMIs
The American Joint Committee	AJCC
Tissue Microarrays	TMAs
Trancript Variant 1	TV1
Transcript Variant 2	TV2
Triple-negative Breast Cancer	TNBC

Table of Contents

Table of Figures

List of Tables

1 Introduction

1.1 Colorectal Cancer Progression

1.1.1 The anatomy of the human colon

The colon is the final part of the gastrointestinal (GI) tract and the digestive system in vertebrates. It functions mainly to absorb water, sodium and essential vitamins that are created by colonic bacteria (Hill et al., 1997).

The colon is preceded by the small intestine in the GI tract. The ascending colon is on the right side of the body and is the first component of the colon. The ascending colon runs upwards through the abdominal cavity toward the transverse colon. The transverse colon spans from the hepatic flexure where it meets the ascending colon, to the splenic flexure on the left side of the body where it meets the descending colon. Between the descending colon and the rectum is the sigmoid colon. After the rectum is the anus (Searle et al., 1995; Peifer, 2002 Eickhoff et al., 2010).

1.1.2 The microanatomy of the normal colon

The wall of the small and large intestine is composed of four layers: mucosa (or mucous membrane), submucosa, muscularis (or muscularis propria, and adventitia (or serosa) (Figure 1.1). The mucosa is the innermost layer and consists of glandular epithelium, lamina propria and muscularis mucosae. The glandular epithelium forms cylindrical structures called crypts. These are embedded within a reticular connective tissue known as the lamina propria. This consists also of lymphocytes (see aggregated lymphoid module in Figure 1.1), plasma cells and eosinophilic granulocytes as well as lymphatics and capillaries. The muscularis mucosae is a thin layer of smooth muscle at the boundary between the mucosa and submucosa. Beyond the submucosa is a layer of muscle known as the muscularis externa. This layer is responsible for peristalsis. The muscularis externa is lined by the serosa that consists of an inner epithelial layer of mesothelium; an outer connective tissue layer

followed by a second layer of epithelial cells called the parietal layer (Reviewed in Peifer, 2002; Schepers & Clevers, 2012).

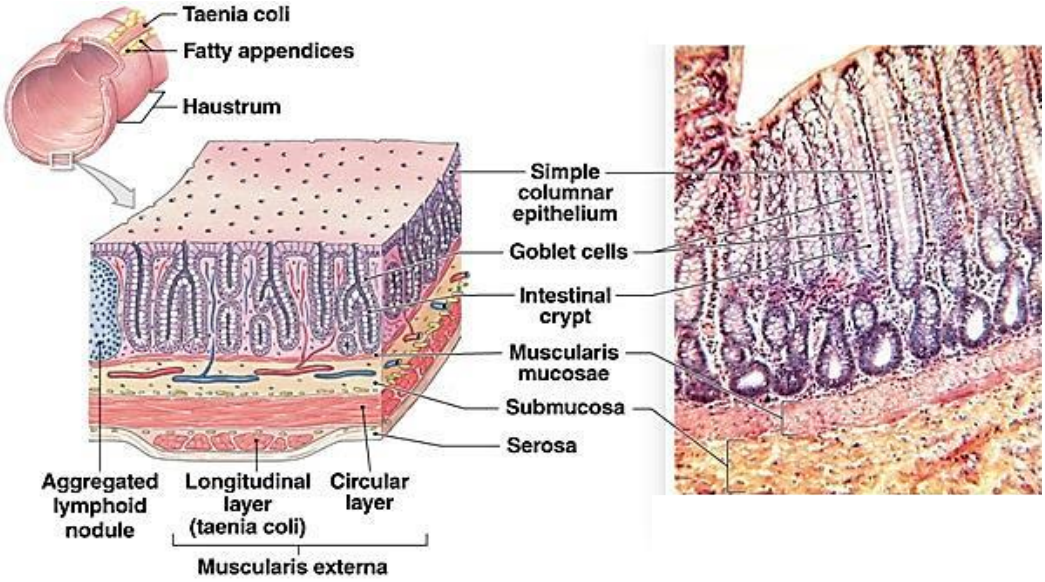

Figure 1.1 Architecture of the large intestine (colon) The digestive tract has four main layers (mucosa, submucosa, muscularis mucosa and serosa). Unlike the small intestine, the colon lacks villi. Image retrieved December 19, 2016 from http://keywordsuggest.org/gallery/245230.html

1.1.3 Colorectal Cancer (CRC)

Bowel cancer represents 11% of all cancer cases and is the fourth most common cancer in males and females in the United Kingdom (Cancer Research, UK, 2014). In 2014, 41,265 new cases were diagnosed with 15,903 deaths being reported in the same year. Bowel cancer mortality rates correlate strongly to age, with age-specific mortality rates rising sharply from around age 50-54 and the highest rates seen in the 90+ age group for both males and females. Since the early 1970s there has been an overall decrease in the mortality rates across all adult age groups.

Colorectal cancer has a long latency period, and thus is very amenable to screening and early detection. Indeed, a recent systematic review provided empirical evidence of an effect of screening on both colorectal-cancer incidence and mortality. A 14% reduction in colorectal cancer mortality and 5% reduction in incidence were observed in four trials of faecal occult blood testing. Dueto the low uptake of screening (50% of the over 50 US population, 2005) and magnitude of the decrease improved surgical technique and adjuvant chemotherapy have likely also

contributed to this dramatic decrease (Welch & Robertson, 2016). Colorectal cancer sites are distributed as follows; Caecum (15%), ascending colon (5%), transverse colon (10%), descending colon (3%), sigmoid (29%) and rectum (38%) (Davies, Miller, & Coleman, 2005).

Approximately 15% of CRCs are familial, the most common of which are familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNCC).

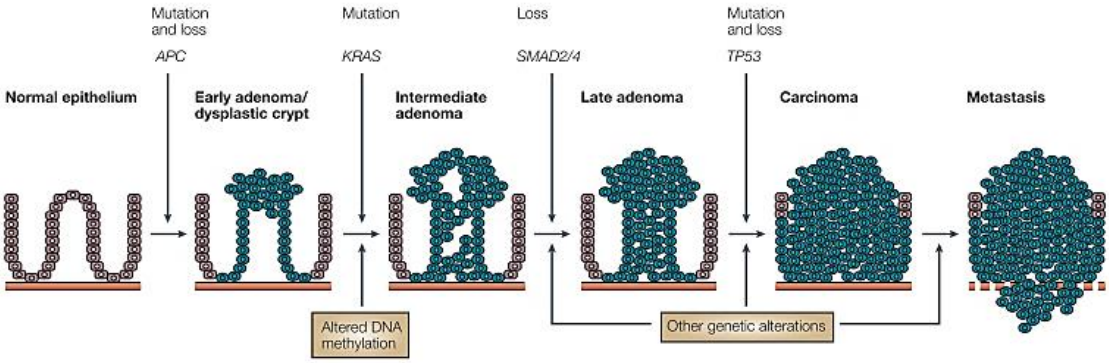

Figure 1.2 The Adenoma-carcinoma sequence. Colorectal cancer progresses through a well-defined series of mutations that manifest themselves in a morphological phenotypic change. Loss of function of the *APC* tumour suppressor gene leads to a dysplastic crypt, a subsequent mutation in *KRAS* and loss of *SMAD2/4* are observed together with intermediate and late precancerous adenomas. Further mutation and loss of *TP53* leads to carcinoma followed by invasive carcinoma and metastasis. (Davies et al., 2005)

Both inherited CRCs display an early onset and are characterised by mutations in mismatch-repair genes including *MLH1, MASH2, MSH3, MSH6, PMS1* and *PMS2*, leading to genomic instability and an increase in the length of microsatellite sequences that result in microsatellite instability (MSI). FAP is caused by an inherited mutation in the adenomatous polyposis coli *(APC)* gene (Davies et al., 2005).

Most sporadic CRCs are thought to develop from benign adenomas. The stepwise genetic alterations that accumulate and transform normal tissue from well-defined colonic crypts, to aberrant crypt foci, then adenomas, adenomatous polyps and finally carcinoma and metastasis, is known as the adenoma-carcinoma sequence (Davies et al., 2005)(Figure 1.2). This adenoma-carcinoma sequence theory was formulated using evidence from FAP. Patients of this disease have greater than 100 gastrointestinal polyps (mushroom-like adenomatous projections in the lumen of the bowel (Figure 1.3). By the age of 30 years, almost three-quarters of FAP patients develop colorectal carcinoma. This provided direct evidence that a mutation

in the tumour suppressor gene *APC* causes adenomas, adenomatous polyps leading to colorectal carcinoma (Half et al., 2009).

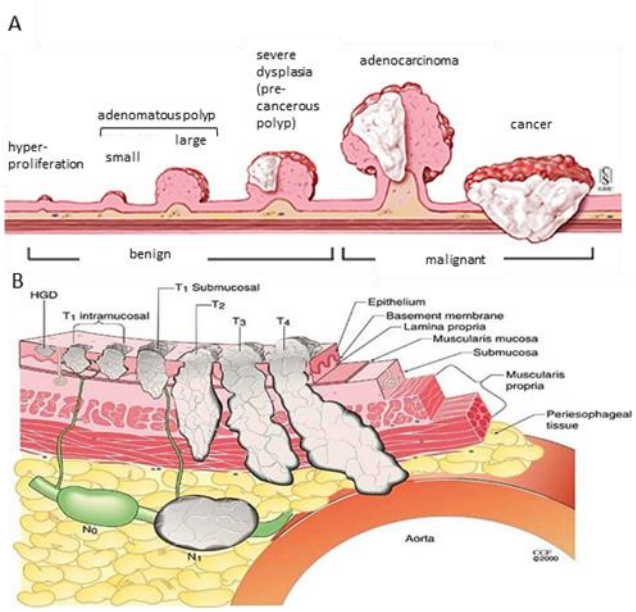

Figure 1.3 Summary of the adenoma to carcinoma progression and TNM staging. (A) Progression from hyperproliferation through dysplastic polyp to adenocarcinoma and then the invasive cancer (B) The progressive growth and invasion of the cancer and its corresponding TNM classification. Reproduced and adapted from *clinicalgate.com/colorectalcancer*.

1.1.4 Colorectal Cancer Staging

The most commonly used colorectal classification used today is the TNM system. It includes a clinical (pre-treatment) and a pathological (postsurgical histopathological) classification. Brachyury describes the size of the primary tumour and its invasion status, N describes the invasion status of local nymph nodes and M describes the presence of distant metastasis sites (Table 1.1 and Figure 1.3).

TNM classifications are formulated based on a histopathological assessment of the morphology of a post-resected biopsy tissue section stained with haemotoxylin and eosin (H&E) for the colon. So definitive staging can only be performed after surgery. However, a polypectomy taken during a colonscopy of a malignant pedunculated polyp (a mushroom-like adenomatous projection) can be analysed in the same way with minimal invasion. Endoscopic ultrasound of rectal cancers may be perfomed before surgery. Metastases may also be staged using abdominal ultrasound, MRI, CT and PET (*AJCC Cancer Staging Manual (8th Edition) 2016*).

Table 1.1 Definitions of TNM staging classifications from The American Joint Committee (AJCC) on Cancer – Cancer Staging Manual 7th edition and the International Union against Cancer (UICC).
Reproduced from AJCC Cancer Staging Manual (8th Edition) 2016

Primary Tumour (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ: intraepithelial or invasion of lamina propria
T1	Tumour invades submucosa
T2	Tumour invades muscularis propria
T3	Tumour invades through the muscularis propria into pericolorectal tissues
T4a	Tumour penetrates to the surface of the visceral peritoneum
T4b	Tumour directly invades or is adherent to other organs or structures
Regional Lymph Nodes (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in 1–3 regional lymph nodes
N1a	Metastasis in one regional lymph node
N1b	Metastasis in 2–3 regional lymph nodes
N1c	Tumor deposit(s) in the subserosa, mesentery, or nonperitonealized pericolic or perirectal tissues without regional nodal metastasis.
N2	Metastasis in 4 or more regional lymph nodes
N2a	Metastasis in 4–6 regional lymph nodes
N2b	Metastasis in 7 or more regional lymph nodes
Distant Metastasis (M)	
M0	No distant metastasis
M1	Distant metastasis
M1a	Metastasis confined to one organ or site (for example, liver, lung, ovary, nonregional node)
M1b	Metastases in more than one organ/site or the peritoneum

1.2 Stem cell function in normal and cancer tissue

1.2.1 Stem cell definitions

Stem cells are cells that have the ability to self-renew over a long period of time with different potencies, i.e. the ability to differentiate into a wide range of cell lineages during development and normal tissue homeostasis (Morrison & Spradling, 2008; Wu & Izpisua Belmonte, 2016). They exist in a hierarchy as detailed below (Figure 1.4)

Embryonic stem cells (ESCs) (Figure 1.4), fetal progenitors (multipotent in Figure 1.4 and adult stem cells (ASCs) (also multipotent in Figure 1.4) have been isolated and have proven a powerful means to model human development in a dish . These embryonic stem cells display an ability to renew virtually limitlessly in culture and can, to varying extents, differentiate into the various lineages that are observed during *in vivo* development and/or injury, repair and maintenance of normal tissue (Wu & Izpisua Belmonte, 2016).

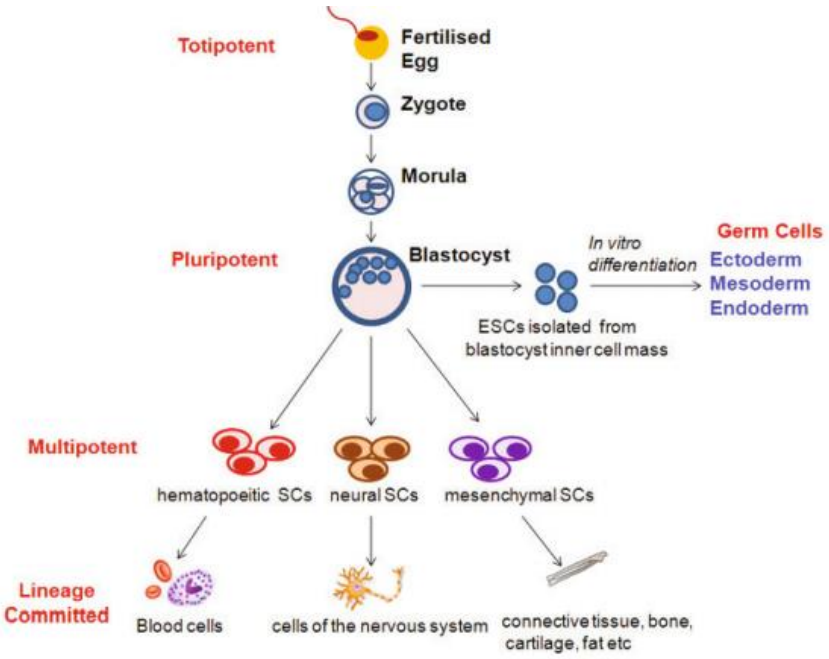

Figure 1.4 Stem cell hierarchy. As the embryo develops from a totipotent zygote to a pluripotent, multipotent, unipotent progenitors through to differentiated cells; the ability of the cell to differentiate into a wide variety of cell lineages decreases until it becomes a specialised functional differentiated cell type (Hayes et al., 2012).

Human Embryonic Stem cells (HESCs) were first derived in 1998 (Thomson et al., 1998). They are derived from the pre-implantation blastocyst (Figure 1.4), a sphere of cells consisting of an outer layer of trophoblast cells that give rise to the placenta and an inner cell mass (ICM), from where ES cells are derived. The cells of the ICM ultimately form the embryo and thus have the capacity to generate all of the tissue types of the body (Rippon & Bishop, 2004).

One of the major events of embryonic development in vertebrate animals is the specification of three embryonic germ layers: ectoderm (gives rise to skin and neural lineages), mesoderm (generates blood, bone muscle, cartilage and fat), and endoderm (contributes to the tissues of the respiratory and digestive tracts). Embryonic cells are separated into being committed to one of these three layers. This requires the sequential action of multiple gene products, it is thought that this segregation of cells in to one of these germ layers during embryonic development is an enduring event (Weissman, Anderson, & Gage, 2001). With a few exceptions, all subsequent progeny, including differentiated cells, progenitor cells and stem cells irreversibly maintain this hierarchy throughout adulthood. Furthermore, in adult tissues, homeostatic cell renewal and tissue regeneration maintain such tissue specificity. In these tissues and organs that retain stem cell dependence, rare tissue-resident stem cells generate only the differentiated cell types corresponding to their tissue of origin. They do not cross germ layer boundaries to give rise to cell types of different lineages. Although multiple lines of evidence have challenged this notion in recent years and called into question the lineage commitment of some adult stem cell populations (Herzog, Chai, & Krause, 2003; Sánchez Alvarado & Yamanaka, 2014).

Pluripotency is a stem cell state that exists transiently in the early embryo and that can be recapitulated through the derivation of ESCs or by reprogramming somatic cells to become induced pluripotent stem cells (IPSCs). The latter were first derived in 2006 by Yamanaka and co-workers.

1.2.2 Adult intestinal stem cell-driven homeostasis

Adult stem cells are the resident stem cell of a specific functional tissue. They reside in an anatomical site known as a niche that has a well-defined combination of secreted factors generated by surrounding cells. When a stem cell divides, one daughter may stay in the niche and continue its function as an adult stem cell and the other leave on a path towards differentiation and ultimately physiological functionality (Snippert, Schepers, Van Es, Simons, & Clevers, 2014). Here we will evaluate the evidence from the two main mouse models of adult intestinal homeostasis i.e. the small intestine and colon.

Adult intestinal stem cells are maintained by a niche that is dependent on the Wnt signalling cascade (Korinek et al., 1998). There exists a decreasing gradient of Wnt from the bottom to the top of the stem cell compartment in both the mouse small intestine and colon (Figure 1.5 colon) (Barker, 2014). As well as Wnt, Notch, Hedgehog, Bmp and Egfr signal gradients contribute both the stem cell niche and the sequential differentiation events that create the intestinal crypt (Figure 1.5) (Barker, 2014; Reya & Clevers, 2005).

Wnt signalling is the master switch between proliferation and differentiation (Neth et al., 2007; Schuijers & Clevers, 2012). Like the Wnt signalling cascade, the Notch pathway is also essential to maintain the crypt stem cell compartment in its undifferentiated, proliferative state (Barker, 2014; Sato et al., 2011). The rare stem cells at the base of the crypt give rise to a much larger pool of transit amplifying cells (Barker, 2014).

Evidence from the mouse colon and small intestine demonstrated that these cells will terminally differentiate into one of four main epithelial lineages of the GI tract. Goblet cells, enteroendocrine cells (both mouse colon and small intestine) and Paneth cells (mouse small intestine) or their equivalent in the mouse colon; deep secretory that belong to the secretory lineage (Barker, 2014). Absorptive enterocytes represent the fourth cell type (mouse colon and small intestine). Figure 1.5 (mouse colon) (right) shows the cell fates directed by both Wnt and Notch, Wnt directs secretory cell fate and Notch directs an absorptive cell fate (Farin et al., 2016; van der Flier & Clevers, 2009).

A gradient of Wnt decreases from the bottom of the crypt (high) to the top (low) in both the mouse small intestine and colon (Figure 1.5). The colonic stem cell niche is less well defined than the small intestine that has niche-generating Paneth cells (Sato et al., 2011). Recently however, Sasaki and co-workers identified regenerating islet derived family member 4; $Reg4^{+}$ deep crypt secretory (DCS) cells as generators of the epithelial niche of $Lgr5^{+}$ stem cells in the mouse colon (Sasaki et al., 2016). These DCS cells are interspresed with $Lgr5^{+}$ stem cells and their ablation results in loss of $Lgr5^{+}$ cells and are proposed as the Paneth cell equivalent (and generators of Wnt) in the colon. Along with epithelial sources of niche factors, the stroma provides a niche for stem cells. In the absence of an epithelial niche; indeed, Kabiri and co-workers (2014) reported that stromal sources of Wnts are sufficient to maintain small intestinal homeostasis in the mouse.

There are two models of adult stem cell identity. The 'stem cell zone model' posited by Leblond, Cheng and Bjerknes (Barker, 2014; Bjerknes & Cheng, 1999) and the '+4 model' by Potten that suggests that stem cells are present in a 16-cell ring immediately above the Paneth cells (small intestine) (Potten, 1977). Each model has struggled to reproducibly describe stem cell identity and position correctly (Barker, 2014) and instead a unified theory that incorporates aspects of both is most informative.

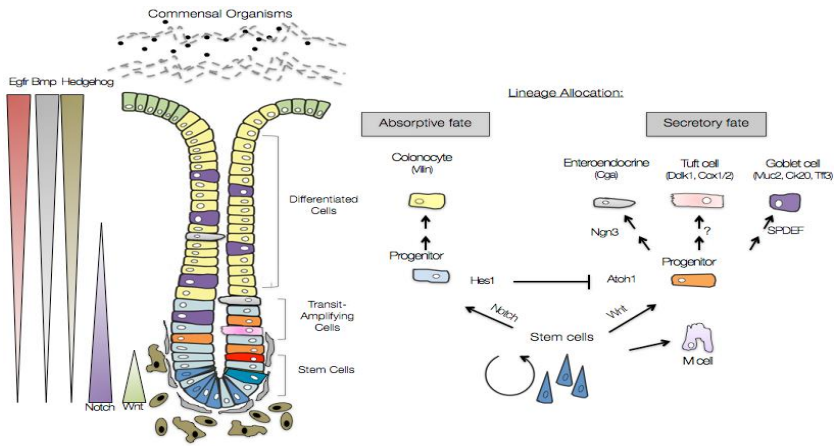

Figure 1.5 The colonic crypt stem cell niche. Stem, transit-amplifying and differentiated cells depend upon regulation of signalling cascades that inform cell division, migration and homestasis. The brown pericryptal stromal cells cross-talk with the epithelium and are a known source of Wnt. The stem cells in blue are intersperesed with deep crypt secretory cells that also generate Wnt. Gradients of these ligands and their receptors dictate the form and funciton of the crypt. (Powell et al., 2012).

The 'unitarian' theory of the origin of the four epithelial cell types' was formulated when undifferentiated, mitotically active crypt base columnar cells intercalated

between Paneth cells were identified by Cheng and Leblond in the mouse small intestine (Barker, 2014; Bjerknes & Cheng, 1999). CBC cells were described as being actively phagocytic, helping to clear dead cells from the crypt base. Taking advantage of this trait, Cheng and colleagues (1974) irradiated mice and showed that surviving CBC cells contained radiolabelled phagosomes as they engulfed neighbouring CBC cells that were killed via the incorporation of tritiated thymidine. Initially, only CBC cells were marked, but at later time points the cells from three major epithelial lineages became radiolabelled within the crypts. This suggested that CBC cells were of these lineages common ancestor. This suggested that CBC cells were of these lineages common ancestor. Cheng and co-workers (1974) next introduced inheritable somatic marks within the crypts. A small set of epithelial clones included all major epithelial lineages. This suggested that the original mutation had occurred in a common self-renewing stem cell ancestor. This led Cheng and Bjerknes (1999) to form the stem cell zone model. This model predicted that CBC cells are resident in a niche at the base of the crypt of the mouse small intestine. They proliferate and generate daughter cells that exit this niche and commit to one of the four major lineages at 'the common origin of differentiation' around position +5.

In the mouse small intestine, enteroendocrine and goblet cells were predicted to share the crypt base (Barker, 2014) although these, like Paneth cells, are post-mitotic. Non-Paneth cell lineages were suggested to migrate upwards on the villus in the small intestine, maturing into a functional epithelial cell. This model failed to gain widespread recognition until 2007 when Barker and colleagues identified CBC-specific markers. *In vivo* lineage tracing experiments and *ex vivo* assays provided the first direct evidence that CBCs did function as a stem cell during gut epithelial homeostasis in the mouse colon and small intestine. The Wnt target *Lgr5* was identified as a gene that was selectively expressed at the base of the adult intestinal crypts. Expression of Lgr5 in CBC cells was proven by using *Lgr5-LacZ* and *Lgr5-EGFP* reporter mouse models. In both models, around 14 Lgr5-expressing ($Lgr5^+$) CBC cells were consistently observed at positions 1-4 of the crypt base and were interspersed with Paneth cells in the small intestine, (Barker, 2014; Barker et al., 2007).

Analysis of the colon yielded essentially identical observations. The $Lgr5^+$ cells yielded blue clones emanating from the crypt bottom. These clones contained

colonocytes as well as goblet cells, and essentially remained unchanged during the 60 days of lineage tracing. One important difference to the situation in the mouse small intestine involved the kinetics of clone formation. Blue staining in most crypts was still restricted to the bottom at 5 days, and entirely blue crypts were only rarely observed, implying that the colon stem cells were more often quiescent than their small-intestinal counter- parts. The relative number of entirely blue crypts increased in later days (Barker et al., 2007). This is an important consideration when attempting to reconcile these models with human intestinal homeostasis in which the rates of cell proliferation in the colon are higher than that of the small intestine (Tomasetti and Vogelstein, 2015) suggesting the opposite is true in mice.

Lineage tracing using an *Lgr5-EGFP-ires-Cre-ERT2/R26R-lacZ* mouse small intestine model confirmed the identity of the $Lgr5^{+}$ CBC cells. After stochastic activation of *lacZ* reporter gene in these cells, clones expressing *lacZ* rapidly expanded to form the epithelium from the crypt base to the villus tip. All major epithelial lineages were observed in these clones and they persisted for long periods of time. This validated $Lgr5^{+}$ CBC cells as the self-renewing, multipotent adult intestinal stem cells (Barker, 2014; Barker et al., 2007).

The +4 model of stem cell identity was formulated when cells immediately above the Paneth cells (mouse small intestine), that divided every 24 hours, consistently retained radiolabels in their DNA during crypt neogenesis. Label retention was indicative of cellular quiescence, but the +4 label retaining cells (LRCs) could be further labelled with BrdU, thus confirming their proliferation status as this was incorporated into their DNA (Barker, 2014). An explanation for these conflicting lines of evidence was that the +4 cells had an ability to retain labelled template DNA strands selectively during mitosis as a strategy to maintain fidelity as the newly synthesised DNA strands that could contain potentially dangerous replication errors were segregated to the short-lived TA daughter cells i.e. an asymmetric cell division (Muñoz et al., 2012). This mechanism that can limit DNA damage in long-lived adult stem cells has been challenged as it makes assumptions on the presence of asymmetric division and exempts stem cells from the process of the exchange of DNA between sister chromatids that typically occurs in somatic cells (Reviewed in Rando, 2007). Several markers have been reported (mainly in the mouse small

intestine) to be expressed selectively in the candidate +4 stem cells (Barker, 2014). Many of these markers mark specific epithelial cell populations with cellular characteristics that differ from those of the LRC population. This would suggest that multiple stem cell populations exist, but the failure of many of these mouse models to faithfully report endogenous expression patterns of candidate stem cell markers has caused some question marks to be raised over their robustness (Barker, 2014). *Bmi1* is one purported +4 ASC marker (Sangiorgi & Capecchi, 2008). A *Bmi1-EGFP* reporter mouse small intestinal model confirmed Bmi1 expression mainly at the +4 position within the proximal small intestine. *In vivo* fate mapping using a *Bmi1-ires-CreERT2Ir26r-lacZ* mouse small intestinal model showed that these Bmi1+ populations include self-renewing, multipotent stem cells that contribute to long-term epithelial homeostasis (Sangiorgi & Capecchi, 2008). Further, *in vivo* ablation of this $Bmi1^+$ population blocked epithelial renewal and sorted $Bmi1^+$ cells were able to generate small intestinal epithelial organoids. This population was quiescent, resistant to radiation and not regulated by Wnt. These prompted the suggestion that these +4 cells were reserve stem cells and were distinct from $Lgr5^+$ CBC cells (Barker, 2014). Doubt was cast over this theory when Bmi1 expression was documented throughout the proliferative zones of the mouse small intestinal crypt including in the $Lgr5^+$ CBC cells (Muñoz et al., 2012). Further, Bmi1-driven lineages were observed at random locations throughout the crypt and most of these lineage tracing events were lost over time which suggests that they are short-lived TA progenitors (Itzkovitz et al., 2012; Muñoz et al., 2012). Other +4 markers that have been proposed include Hopx, Lrig1 and Tert (Barker, 2014). Hopx, Bmi1 and Tert were found to be broadly expressed throughout the mouse small intestinal crypt and did not mark any particular population within the crypt in a recent RNA-fluorescent in situ hybridisation study (Itzkovitz et al., 2012; Muñoz et al., 2012). In the mouse colon, $Lrig1^+$cells that underwent Cre-mediated reporter gene activation failed to generate progeny over a long period which indicated that they are quiescent (Muñoz et al., 2012). $Lrig1^+$ cells did however proliferate and regenerate epithelial progeny following radiation induced damage of the mouse colonic epithelia (Powell et al., 2012). Wong et al (2012) also reported LRig1s control of epithelial homeostasis through negative regulation of ErbB signalling (see EGFR gradient in Figure 1.5). In

the mouse small intestine, *in vivo* lineage tracing generated long-term tracing units characteristic of multipotent stem cells. Reporter gene activation was most commonly observed between positions +2 and +5 (Wong et al., 2012).

This uncertainty over cell identity is further complicated by the emerging phenomenon of intestinal cell plasticity. Recent evidence has highlighted the role of dedicated progenitor populations that can revert to the stem cell state upon damage to aid crypt regeneration in the mouse small intestine (Beumer & Clevers, 2016). In one study $Lgr5^+$ cells were selectively removed leading to CBC cell depletion, for at least one week small intestinal crypt homeostasis was not disturbed (Greten et al., 2016) and in another, $DII1^+$ secretory progenitor cells reverted to stem cells upon damage to the mouse small intestinal crypt (van Es et al., 2012). Furthermore, Tetteh and co-workers (2016) showed that enterocytes in mouse small intestinal crypts marked by $Alpi^+$ that can dedifferentiate to replace lost $Lgr5^+$ CBCs. Other studies revealed that DNA methylation and histone marks in intestinal crypt/villus cells and $Lgr5^+$ CBC cells in the mouse small intestine are near identical and may therefore permit the rapid reprogramming into $Lgr5^+$ stem cells during regeneration (Kim et al., 2014). These dedifferentiating progenitors generate multipotent, self-renewing stem cells and Paneth-like cells in the mouse small intestine, thus stemness is not "hard-wired"(Tetteh et al., 2016).

1.2.3 Stem cell-driven intestinal adenomas

The adenoma-carcinoma sequence described in section 1.1.3 is a progression from normal crypt form and functionality to an adenoma and then invasive cancer. As adult stem cells normally generate normal tissue, the concept that tumours are maintained by dedicated stem cells is attractive yet remains controversial.

Schepers and colleagues (2012) provided direct functional evidence of the presence of stem cell activity within primary intestinal adenomas. As discussed above, adenomas are characterised by a loss of function mutation in the tumour suppressor *APC* that results in constitutive activation of the Wnt signalling pathway. $LGR5^+$ CBC cells are found at the bottom of colonic crypts where there is high Wnt (Barker et al., 2007).

In order to investigate whether the stem cell hierarchy of normal crypts was maintained in adenoma crypts, the multicolour Cre reporter system was used to investigate the cells of origin of intestinal adenomas (Schepers et al., 2012). Adenomas developed from Lgr5^{+} *Apc* $^{-}$stem cells and after 4 weeks, large adenomas were observed that either had a single Cre-induced colour or had anatomically separated Confetti-labelled segments, each originating from a single Lgr5^{+} CBC. This indicates that these adenomas were generated by several independent *Apc*-mutant stem cells (Figure 1.6) (Schepers et al., 2012). Barker et al., (2009) in a similar study showed that deletion of *Apc* in Lgr5^{+} CBC cells led to their transformation in days. These transformed cells remain at the crypt bottom while generating microadenomas.

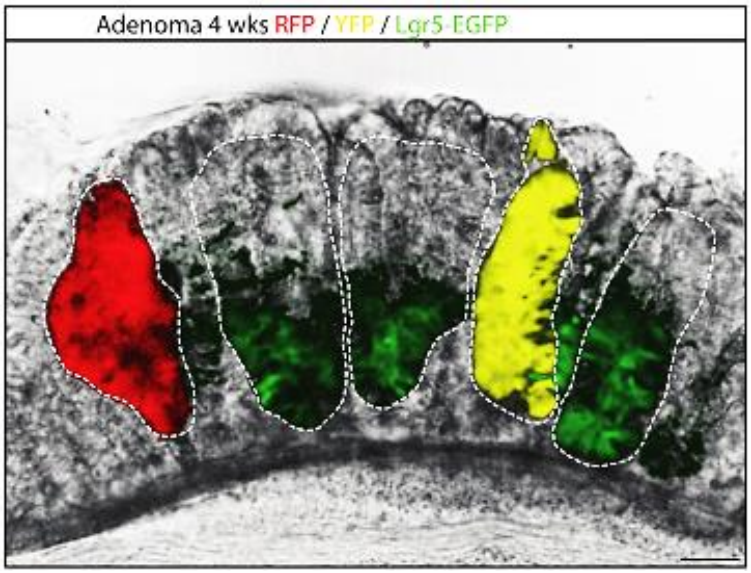

Figure 1.6 Fluorescent imaging of an adenoma in a Lgr5EGFP-Ires-CreERT2/Apcfl/fl/ R26R-Confetti mouse after 4 weeks of tracing. Multiple segments (white dashed lines) can be observed in the adenoma which are marked uniformly by a single colour (or are unmarked). RFP (red fluorescent protein) and YFP (yellow fluorescent protein) tracings are shown in red and yellow, respectively. Lgr5+ stem cells (green) are visible at the base of adenoma segments. Colours are overlaid on a differential interference contrast (DIC) image in grey (ASchepers et al., 2012).

These microadenomas develop into macroscopic adenomas within 3-5 weeks. The distribution of Lgr5^{+} CBC cells at the bottom of the crypt suggests that stem cell-TA progentitor hierarchy is maintained in these precancerous lesions (Barker et al., 2009). Furthermore, Snippert et al., (2014) demonstrated that *Kras* mutated individual Lgr5^{+} stem cells rate is increased compared to wild type and creates a biased drift toward a single *Kras* mutated clonality. This led to increased crypt fission and thus the formation of microadenomas.

As mentioned above, adult stem cell plasticity has become an emerging theme that complicates the simple model of stem cell hierarchy. Schwitalla et al., (2013) suggested that differentiated villus enterocytes could give rise to tumours when β-*catenin*/*Kras* mutations were present in these cells. They observed crypt stem cells in aberrant positions that re-expressed Lgr5 and putative stem cell marker Sox9. These villus cells were also able to form tumours when injected into *Apc*$^{-/-}$ mice (Schwitalla et al., 2013). Another study also suggested that differentiated villus epithelial cells can give rise to tumours upon overexpression of the BMP agonist *Grem1* in these cells. The dedifferentiating Alpi^{+} marked mature enterocytes that dedifferentiated in response to the ablation of Lgr5^{+} cells in the normal mouse intestine do not from tumours when they have mutations in *Apc/Kras in vivo* (Tetteh et al., 2016). *Apc*-mutated enterocyte progenitors failed to generate *ex vivo* tumour organoids but those with both *Apc* and *Kras* mutation were able to. This can be explained by the fact that the short residency of enterocyte progenitors in the crypts might prevent combined *Apc/Kras* mutatated progenitors establishing themselves in the crypts and are therefore unable to progress towards cancer (Tetteh et al., 2016). In contrast to Schwitalla (2016) (using *Xbp1*as a villus enterocyte reporter) and Davis (2015) (using *Xbp1* as a villus enterocyte reporter) studies, Tetteh et al., (2016) did not observe tumours from villus enterocytes upon *Apc/Kras* mutations with either of these reporters. As both of these reporters are in fact expressed in other cell lineages of the epithelium, it is likely the tumours in these models did not originate from enterocytes (Davis et al., 2015; Schwitalla, 2012; Tetteh et al., 2016).

1.2.4 The cell of origin of colorectal cancer

Many recent studies have demonstrated that *Apc* deletion in Lgr5^{+} (A. G. Schepers et al., 2012) , Lrig1^{+} (Powell et al., 2012c), CD133/Prominin1 (Zhu et al., 2009) and other candidate adult intestinal stem cell populations, can provoke rapid adenoma formation (Huels & Sansom, 2015). It seems therefore that the activation of Wnt signalling in stem cells is sufficient for adenoma formation in the mouse intestine. Another recent study, observed that a high fat diet (HFD) promotes stem cell (Lgr5^{+}) and progenitor cells to form adenomas that then go on to form cancer in the mouse

small intestine. Interestingly they report an increased β-catenin activity in intestinal stem cells in HFD mice and observed a 50% increase in the frequency of *Olfm1*$^{+}$ intestinal stem cells (a marker expressed in *Lgr5*$^{+}$ CBCs). At the same time, HFD reduced the number of differentiated niche-generating Paneth cells by 23% and led to a mild reduction in villi length, meaning that HFD increases intestinal stem cell numbers at the expense of differentiation. Surprisingly, a reduction in Paneth cell number does not affect the intestinal stem cell compartments ability to expand (Beyaz et al., 2016).

ISC division operates on a principle of random replacement of ISCs, a process termed 'neutral drift' (Snippert et al., 2014). This principle states that a single ISC in a crypt can be replaced by any other ISC in the crypt. Indeed, Vermeulen and colleagues (2010) showed that if a single ISC acquires a neutral mutation (i.e. non-transforming) it has a high probability of being replaced by a normal stem cell within the crypt and the probability of an *Apc*-mutated ISC becoming fixed and populating the whole crypt is 42%, although a *Kras* mutation in an ISC has a 72% chance of populating a whole crypt. These studies demonstrate the effect these oncogenic mutations have on the fitness of ISCs that harbour them. These probabilities might also explain why it takes CRC years to develop in patients with FAP (germline *APC*$^{mut/+}$*)* (Huels & Sansom, 2015).

Vermeulen and coworkers (2010) highlighted the importance of the tumour microenvironment (stroma) in regulating cancer stem cells through Wnt signalling. They observed that high activity of the Wnt pathway is seen in tumour cells located in close proximity to stromal myofibrobloasts, a similar situation to the normal stem cells of the intestinal crypt. This provided the first direct evidence that external cues may regulate Wnt activity and cancer stemness. Moreover, myofibroblast secreted factors restored the cancer stem cell phenotype both *in vivo* and *in vitro* (Vermeulen et al., 2010).

Considering a non-stem cell origin of intestinal cancer,the two studies mentioned above by Schwitalla (2013) and Davis (2015) suggested that villus cells could be the cells of origin of the precancerous adenoma (although Tetteh (2016) and colleagues failed to replicate their results), furthermore in the context of the cell of origin of colorectal cancer conclusions can't be drawn from cancers that originate from villus

cells. If these studies findings are real however, it seems reasonable to conclude that activating the Wnt pathway in non-stem cells is not sufficient to drive adenoma formation but additional enabling mutations and activation of inflammatory pathways and/or changes in tissue microenvironment enables an increase in the pool of origin to non-stem cells (Huels & Sansom, 2015).

Taken together, these studies (in the mouse small intestine) suggest that Wnt activation in ISCs transfrom these cells into potent intiators of the precancerous adenoma and later carcinoma. In addition, some reports suggest that TA cells can form adenomas upon Wnt activation but with much less potency than ISCs (Tetteh et al., 2016). With present evidence from mouse models it is reasonable to conclude that ISCs are the most potent cells for transformation but there remains the possibility that short-lived progenitors could be the cells of origin although more work is needed to understand their potential to form tumours.

The experiments described above that established our current understanding of stem cell-driven gut homeostasis and tumourigenesis were perfomed in mice. Recent studies have aimed to contextualise these findings in the normal human gut and colorectal lesions. Baker et al (2014) used somatic mtDNA (mitochondrial DNA) mutations to trace clonal lineagesin the normal human colonic crypt. Through analysing clonal mtDNA imprints in the walls of the normal colonic crypts they were able to demonstrate that human intestinal stem cells exhibit one-dimensional neutral drift dynamics. Here, five to six stem cells are present in both normal patients and individuals with FAP (*APC*$^{-/+}$*).* Furthermore, they demonstrate that in *APC*$^{-/-}$ adenomatous crypts, there is an increase in the functional stem cell pool. Interestingly, they show that one colon crypt divides once every 30-40 years and that this rate increases by at least an order of magnitude in adenomas (Baker et al., 2014)

As there is no reliable antibody against LGR5, in order to study its presence in human colorectal tumours, research in humans has been limited to *in situ* hybridisation of *LGR5* mRNA transcripts. Recently Baker et al (2015) used RNA-Fluorescent in situ hybridisation (RNA-FISH) to identify *LGR5*$^{+}$ cells in human adenomas and carcinomas. They found that a small number of *LGR5*$^{+}$ cells were present at the bottom of normal colonic crypts. They went on to show that conventional adenomas express high levels of *LGR5* but they found no evidence of

cellular hierarchy in these lesions. Another sub-class of adenoma (serrated lesions) did however display basal localisation of *LGR5* resembling the cellular hierarchy of the normal crypt and the $Lgr5^{+}$-driven adenomas reported in mice. This study highlights potential differences between the stem cell dynamics of serrated (10-20% of adenomas, BRAF/KRAS, CIMP, MSI) and conventional adenomas (80-90% of adenomas, APC/KRAS/p53) in humans. Furthermore, they observe basal localisation *LGR5* in carcinomas of all stages, interestingly they noted that *LGR5* expression was generally high in invading cells (Baker et al., 2015).

In conclusion, the source of the cells of origin of colorectal cancer in humans and how they can be exploited as therapeutic targets is likely to be complex. However, $Lgr5^{+}$ intestinal stem cell's potency in normal crypt generation, crypt regeneration after damage and intestinal tumourigenesis in mice make them an attractive therapeutic target. Furthermore, their presence in normal human crypts in the hypothesised stem cell zone, basal/extensive expression in adenomas, basal expression in carcinomas, together with stem cell dynamics in humans resembling those in mice, provide proof of their status as the cells of origin in human colorectal cancer.

1.3 Epithelial to mesenchymal transition (EMT)

During development, epithelial cells are highly plastic and are able to switch back and forth between the tightly bound together epithelial and the less tightly bound mesenchymal phenotypes, epithelial to mesenchymal (EMT) and mesenchymal to epithelia (MET) respectively (Kalluri, 2009). Typically, once development is complete, epithelial cells constitute the functional part of the tissue, forming polarised cells that are bound together by tight junctions, while mesenchymal cells are less tightly bound and form connective supportive tissues.

During embryogenesis, EMT underpins a variety of tissue remodelling events including mesoderm formation (Craene & Berx, 2013). The formation of mesoderm represents a key EMT program that generates oligopotent mesenchymal cells that are able to further differentiate into various cell types. Mesoderm forms from the primitive ectoderm, a process that is initiated during gastrulation (Figure 1.7) (van

Roy, 2014; Yang & Weinberg, 2008). During embryonic development, EMT is not an irreversible process that commits cells to one fate or another. Mesenchymal cells can often revert to an epithelial cell state through MET (Yang & Weinberg, 2008). For example, during the formation of the nephron epithelium in the developing kidney, nephric mesenchymal cells begin to aggregate, begin to polarise and eventually develop cell-cell adhesions and differentiate into epithelial cells that form kidney tubules (Cruz-solbes & Youker, 2017; Davies, 1996).

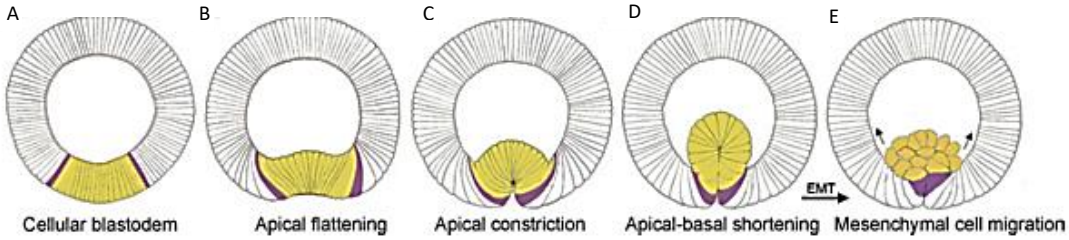

Figure 1.7 Drosophila gastrulation as a model of EMT. (A) The simple Drosophila blastorderm is a simple monolayer epithelium. (B and C) Apical flattening is induced by the transcription factors Twist and Snail in the ventral domain (yellow), leading to the formation of the ventral furrow. (D) Cell shortening in the apica-basal direction accompanies the final 'push', as ventral cells are internalised. (E) Internalised epithelial cells undergo a rapid transformation into a mesenchymal morphology as they lose their adeherens juncions. They then migrate along the extracellular space of the ectoderm to take up new positions the embryo (Baum, Settleman, & Quinlan, 2008).

Once development is complete, epithelial cells exhibit an apical-basal polarity through their association with a lamina layer at their basal surface, often referred to as the basement membrane. If this tissue architecture is disrupted as during wounding, epithelial sheets must migrate in response to injury signals (C. Yan et al., 2010).

The EMT is initiated by a set of well-defined transcription factors; *Snail, Slug, Zeb1/2* that control the expression of E-cadherin (*CDH1*) (Banyard & Bielenberg, 2015). E-cadherin is responsible for the tight junctions formed between epithelial cells. In addition Zeb1, Snail and Slug are capable of repressing the transcription of cell polarity factors also (J. Yang & Weinberg, 2008).

1.3.1 The EMT during cancer

Epithelial cells are the source of colorectal tumours and maintain the attached epithelial phenotype in solid well-differentiated primary tumours (Polyak &

Weinberg, 2009). However, when cancer cells reach a blood vessel, as in an invasive carcinoma, they must undergo a transition into the solitary mesenchymal phenotype and detach from neighbouring cells (Banyard & Bielenberg, 2015; Craene & Berx, 2013). To achieve this, these tumour cells must lose cell-cell adhesions and acquire motility through the EMT process (Figure 1.8) that facilitates tissue remodelling during development and wound healing (J. Yang & Weinberg, 2008).

It has been difficult to identify cancer cells that are undergoing the EMT, most EMT markers are expressed in either epithelial or mesenchymal cells and therefore not specific to cancer cells undergoing the transient EMT process (Thiery & Sleeman, 2006). Although two recent studies have provided evidence that epithelial tumour cells do not need to undergo a transition to a mesenchymal cell state to form metastases (reviewed in Maheswaran & Haber, 2015). Some studies have shed light on EMT in carcinoma cells. For example, TWIST1 is overexpressed in lobular breast carcinoma and diffused-type gastric cancer. Over 90% of these tumour types do not express E-cadherin and exhibit a permanent EMT phenotype (Craene & Berx, 2013).

CRCs that have mutations in *APC* or *CTNB1* (β-catenin), exhibit cells at the invasive front and scattered in the adjacent stroma that express CTNNB1 at a high level in the nucleus (Sánchez-tilló et al., 2011). Also, cells that have undergone EMT have nuclear CTNNB1 accumulation, lose E-cadherin expression and acquire mesenchymal markers like fibronectin (Fodde & Brabletz, 2007).

As well as an extensive network of transcription factors and the membrane proteins that they control, it has become apparent that micro RNAs also play an important role in the regulation of EMT in several cancers. Two independent studies highlighted the importance of the miR-200 family of micro RNAs. The targets of these miRNAs included both ZEB1 and ZEB2 EMT inducing transcription factors Figure 1.8 (Polyak & Weinberg, 2009).

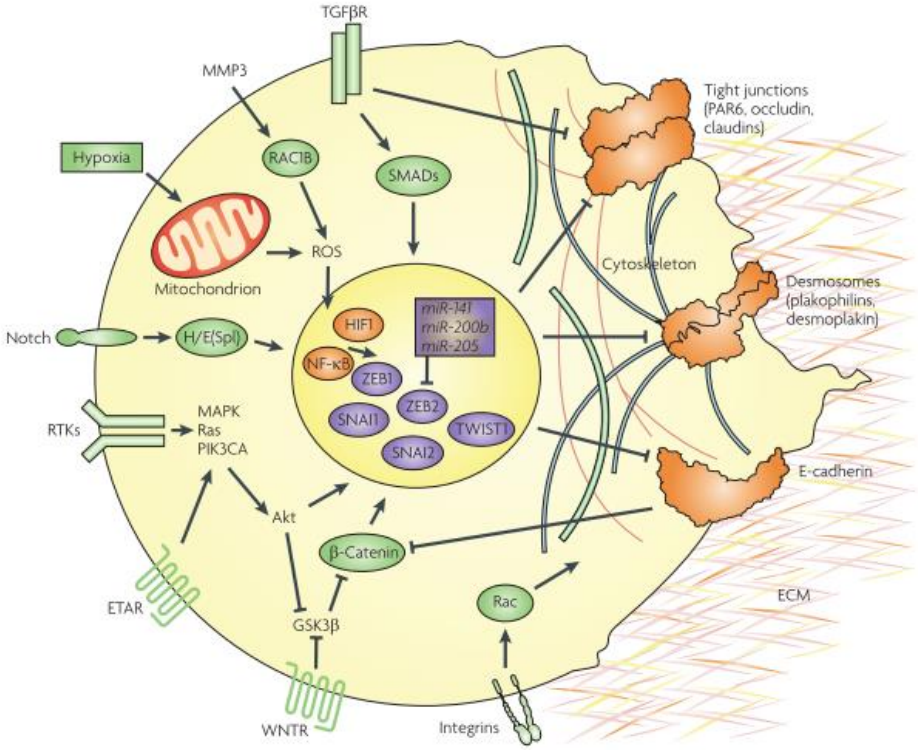

Figure 1.8 A simplified overview of signalling networks regulating epithelial–mesenchymal transitions (EMTs). Selected signalling pathways and some of their downstream effects and interactions are depicted. Transforming growth factor-β (TGFβ), Notch, endothelin A receptor (ETAR), integrins, Wnt, hypoxia and matrix metalloproteinases (MMPs) can induce EMTs through multiple different signalling pathways depending on the cellular context. EMTs and mesenchymal–epithelial transitions (METs) are associated with dramatic changes in the cytoskeleton and extracellular matrix (ECM) composition and attachment that act together to alter cell morphology. Tight junctions and desmosomes can be disrupted in response to EMT signals through protein phosphorylation (for example PAR6A phosphorylation by TGFβ signalling) or by repressing protein levels (for example ZEB1 represses plakophilin 3). Also, the extracellular matrix is completely reorganised as many EMT-inducing factors upregulate the expression of ECM proteins (such as fibronectin and collagens), proteases (such as MMPs) and other remodelling enzymes (such as lysyl oxidase). There is also extensive crosstalk among the EMT-inducing transcription factors (purple circles) and the microRNAs (miRNAs) regulating them. E-cadherin, epithelial cadherin; H/E(Spl), hairy and enhancer of split; WNTR, Wnt receptor (Polyak & Weinberg, 2009).

EMT enables cancer cells with an epithelial phenotype, for example, in the epithelial lining of the human gut to invade through intravasation into adjacent blood vessels (Figure 1.9) (Pattabiraman & Weinberg, 2017). This final progression in the context of the adenoma-carcinoma sequence (defined in 1.1.3) comes after carcinoma *in situ* and is termed invasive carcinoma. It occurs once the basement membrane has been broken down and tumour cells (that have undergone the transcriptionally-induced morphological changes described above i.e. the EMT) are disseminated and migrate into the underlying blood vessels (Mani et al., 2008; Pattabiraman & Weinberg, 2017; J. Yang & Weinberg, 2008). A number of studies have demonstrated that

Epidermal Growth Factor (EGF) family members are activated and stimulate intravasation (Figure 1.9).

They initiate downstream signalling events through PI3K, N-WASP, RhoA, and WASP to induce invadopodia. These are actin-rich protrusions of the plasma membrane that are associated with degradation of the extracellular matrix in cancer invasiveness and metastasis (Chiang, Cabrera, & Segall, 2016).

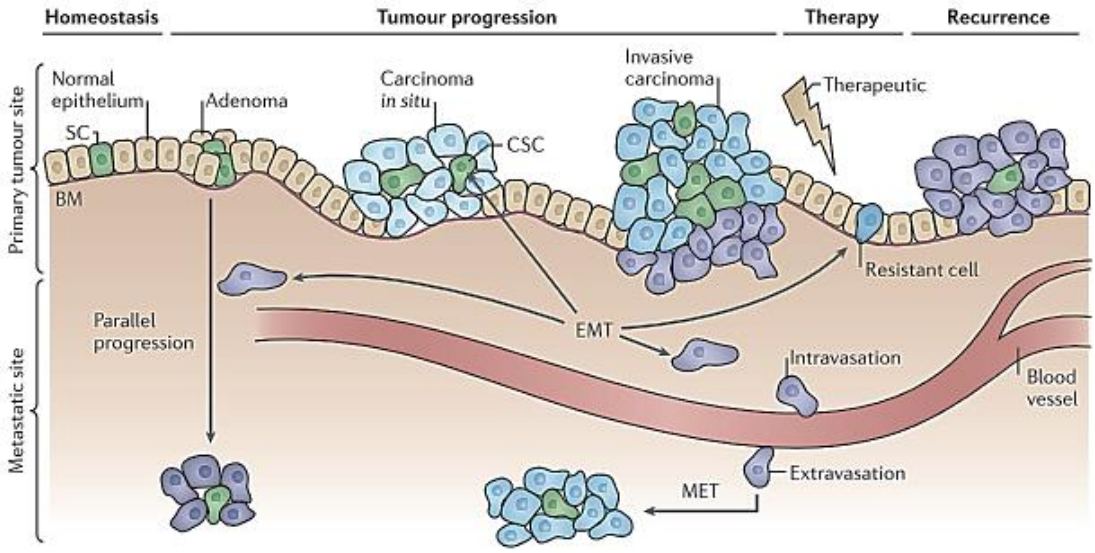

Figure 1.9 EMT in the context of cancer progression. EMT- inducing transcription factors may redefine the phenotypic status of the cancer cell, these changes may happen simultaneously with and be similar to transcriptional changes that occur during dedifferentiation into a stem-like cell (discussed above). After carcinoma *in situ* the basement membrane is broken down through expression of proteases driven by the EMT transcriptional program. Cells must then undergo intravasation through the activation of EGF family members. Cells may then travel through the blood stream, undergo the inverse process of MET through extravasation of the blood vessel. Activation of the transcription factor Twist 1 has been implicated in the process of extravasation during MET (Tsai et al., 2012).

1.4 The Wnt signalling pathway

The Wnt signalling pathway is a highly conserved pathway that plays a critical role in embryonic development of all animal species, the regeneration and homeostasis of adult tissues and is aberrant in cancer, a disease in which normal tissue homeostasis is deregulated (Clevers & Nusse, 2012; Reya & Clevers, 2005).

1.4.1 The Wnt signalling pathway during development

The Wnt, Notch, Hedgehog, TGFβ, (transforming growth factor β) – BMP (Bone Morphogenic Protein) and receptor tyrosine kinases are major molecular mechanisms that control embryonic development (Clevers & Nusse, 2012). They operate through morphogens that are present in a gradient on a spatial axis that extends beyond cell and tissue boundaries. Wnt signalling regulates vertebrate body-axes and provides a tissue-specific context that informs a cell of its differentiation status. These signalling systems are tightly regulated during development (Sokol, 2011).

Biochemical analysis revealed direct interactions between components of the Wnt pathway. In particular, β-catenin, when it is not bound to E-cadherin, translocates to the nucleus, where it binds to the transcription factors LEF1 and TCS (T-cell factor) (Schepers & Clevers, 2012), resulting in LEF1 becoming a transcriptional activator. In the canonical Wnt signalling cascade (Figure 1.10), extracellular Wnt stimulates intracellular signals via Frizzeled receptors, the LRP5/6 (LGR5 in humans and mice) co-receptor and the Dishevelled protein to inactivate the β-catenin destruction complex (Holland et al., 2013) , that includes Axin, APC (adenomatous polyposis coli, mutated in FAP patients) and GSK3β. DKK1 is a Wnt agonist that is secreted (Sokol, 2011).

1.4.2 Wnt signalling in cancer

No link was found between Wnt signalling and cancer until it was discovered that the tumour suppressor, *APC*, that is frequently mutated in FAP patients and sporadic colorectal cancer cases, interacted with the Wnt pathway component β-catenin (Rubinfeld et al., 1993). Subsequently it was discovered that FAP patients had deletions in the 5q21-22 led to truncating mutations in *APC* that were present also in sporadic cases of CRC (Klaus & Birchmeier, 2008). Many mutations in *APC* may accumulate before the region that encodes the region that interacts with Axin1 and Axin2, of the β-catenin destruction complex, is affected. *AXIN1* and *AXIN2* loss-of-function mutations have also been observed in rare cases of colorectal cancer (Krings et al., 2000; Satoh et al., 2000).

Gain-of-function mutations of *CTNNB1* have been observed in a fraction of CRCs. Their effects are to prevent the phosphorylation, subsequent ubiquitylation, and proteosomal degradation of β-catenin with around 1% of sporadic CRCs exhibiting activating mutations in *CTNNB1* (Polakis, 2007).

The oncogenic mutations in *APC*, *AXIN1*, *AXIN2* and *CTNB1* lead to the aberrant stabilisation of β-catenin (Figure 1.10). Excess β-catenin migrates to the nucleus, where it interacts with LEF and TCF transcription factors, and the consequent constitutive activation of downstream genes associated with cell proliferation, like *MYC* and cyclin D1 (Clevers & Nusse, 2012; He et al., 1998; Schepers & Clevers, 2012; Shtutman et al., 1999).

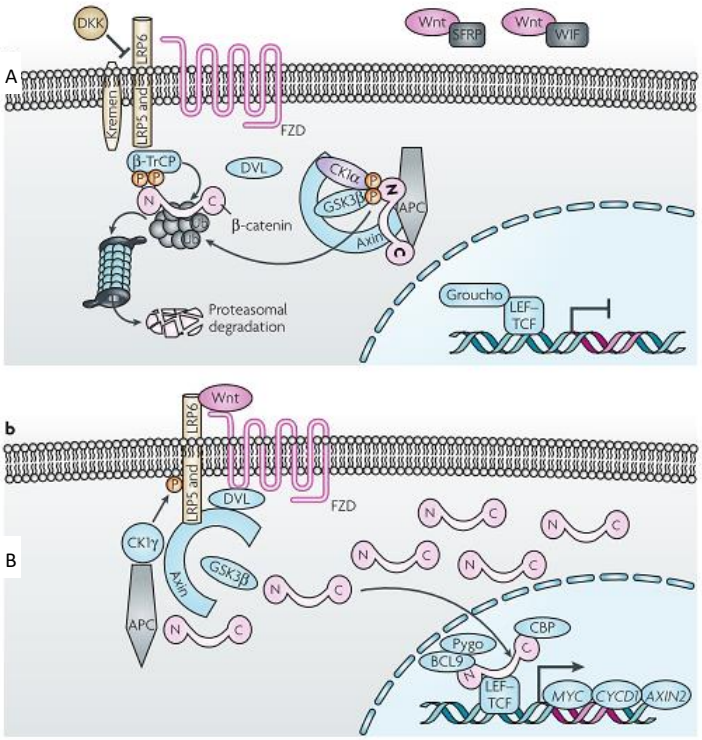

Figure 1.10 The canonical Wnt signalling pathway in vertebrates. (A). When Wnt ligands are absent, β-catenin is recruited into the destruction complex with APC and the axins (β-cateninN---C). Subsequently, β-catenin is phosphorylated on its N-terminus by CK1 α (casein kinase 1α) then ubiquitylation by β-TrCP (β-transducin repeat containing protein, an E3 ubiquitin ligase). After this, β-catenin is proteosomally degraded. Low cytoplasmic levels of β-catenin ensure transcriptional repression of Wnt target genes like Groucho, the transcriptional co-repressor to LEF –TCF. B When external Wnt ligands are present, LPR5 (LDL-related receptor protein 5) and LRP6 re phosphorylated by CK1γ and GSK3 β, and Disheveld (DVL) moleculaes translocate to the plasma membrane to interact with Frizelled (FZD) receptors. When axin interacts with phosphorylated LRP5, LRP6 and Dishevelled, the deactivation complex is destroyed and the subsequent stabilisation of β-catenin which is now free to translocate to the nucleus. Once in the nucleus, β-catenin forms a transcriptionally active complex with TCF and LEF, displacing Grouchos and interacting with BCL9 (B-cell lymphoma 9), Pygopus (Pygo) and CBP (CREB binding protein). CYCD1, cyclin D1; DKK (a secreted inhibitor of Wnt) , Dickkopf; SFRP, secreted Frizzled-related protein; P, phosphorylation; Ub, ubiquitylation; WIF, Wnt inhibitory factor 1 (Klaus & Birchmeier, 2008).

1.5 The T-box transcription factor Brachyury

1.5.1 The evolution of the T-box family of transcription factors

T-box proteins are a group of conserved transcription factors. T-box genes originated in the premetazoan ancestors of metazoans, homologues are present in the extant genomes of early-branching fungi and unicellular holozoans (Sebé-Pedrós & Ruiz-Trillo, 2016).

The T-box genes are essential transcription factors for the embryonic development of metazoa. They have an evolutionary conserved DNA-binding domain of 180-200 amino acids known as the T-box domain. They are expressed in various temporal-spatial axes in the developing embryo (Sebé-Pedrós & Ruiz-Trillo, 2016).

The Brachyury T-box class was the first member of the T-box family. Brachyury (gene symbol: *T*) (short tail in Greek) was discovered in the mouse when mice that were heterozygous for functional Brachyury failed to develop a tail and homozygous mutants were not viable (Kispert & Herrmann, 1993).

The T-box family are united by a conserved binding motif (Figure 1.11) that binds to a consensus sequence that is both species and class specific. For example, the DNA-binding preferential sequence of *Capsaspora* and mouse Brachyury orthologs are almost identical. Different mouse T-box classes also share similar motifs (Sebé-Pedrós & Ruiz-Trillo, 2016).

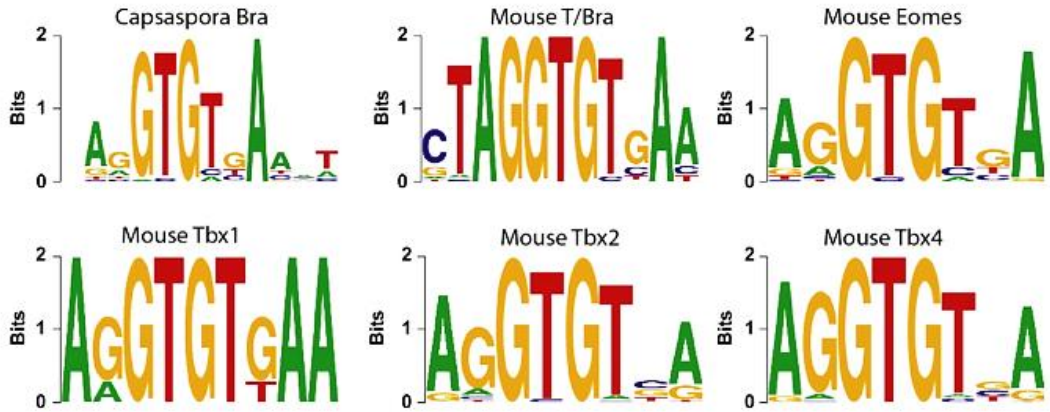

Figure 1.11 Highly conserved T-box binding motifs. Protein-binding microarray experiments reveal the consensus binding sequences of the mouse and Capaspora Brachyury orthologues are almost identical. Different T-box classes in the mouse show similar motifs, indicating conservation of the consensus binding motif across the T-box transcription family (Sebé-Pedrós & Ruiz-Trillo, 2016).

1.5.2 Brachyury during embryonic development

Since the observation of Brachyury expression in invaginating hindgut cells in *Drosophila* (Kispert et al., 1994), further observations in different species led to various researchers proposing a conserved role for Brachyury in morphogenetic movements (Sebé-Pedrós & Ruiz-Trillo, 2016).

Many studies have indicated a conserved function for Brachyury in morphogenetic movements such as those needed during gastrulation (Gross & McClay, 2001; Tagawa, Humphreys, & Satoh, 2000).

In the amphibian embryo, gastrulation is driven by the cells of the mesoderm. Brachyury is a gene that confers mesodermal identity in *Xenopus* (Papin et al., 2002). It is required for the movements that facilitate gastrulation and notochord differentiation. The *Xenopus* Brachyury orthologue *Xbra* is a transcriptional activator. Inhibition of this transcriptional activation leads to an inhibition of morphogenetic movements during gastrulation (see Figure 1.12 for an overview of gastrulation).

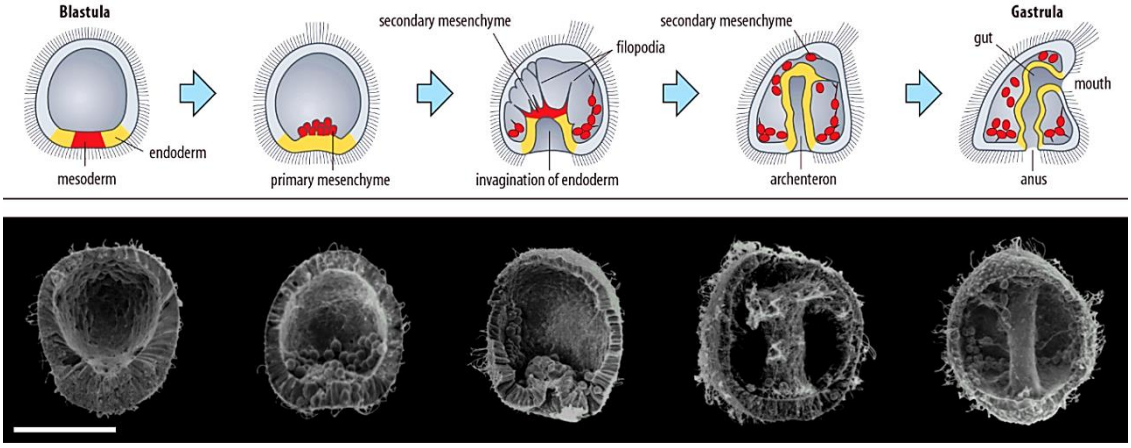

Figure 1.12 Gastrulation in the sea urchin. Gastrulation is the formation of the early gut in metazoans. The vegetal-most cells undergo an epithelial to mesenchymal transition. This transition results in changes to cell shape and cell movement. The primary mesenchyme separate from each other and the surrounding hyaline layer and migrate as single cells into the blastocoel. Then, the endoderm invaginates as a sheet of cells to form the archentereon. Convergent extension of the tip of the gut followed by extension of filopodia from the cylinder until contact with the blastocoel wall follows. Once contact is made, the filopodia pull the elongating gut to fuse with the mouth region (Wei et al., 2009).

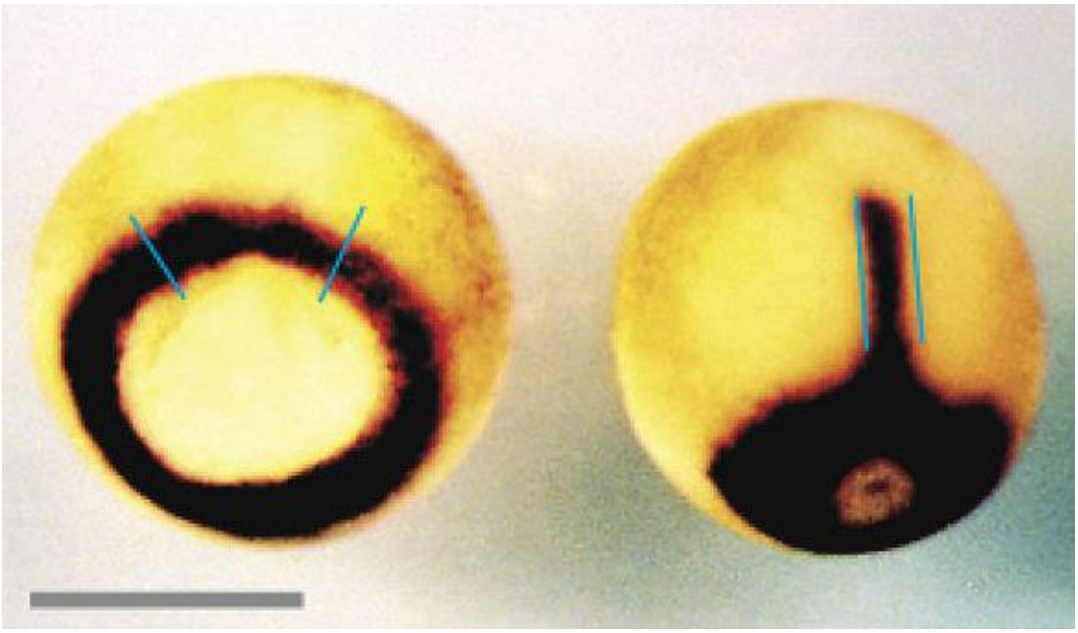

Figure 1.13 RNA FISH of Brachyury during Xenopus gastrulation illustrates convergent extension. Left, Brachyury (dark stain), before gastrulation marks the presumptive mesoderm, present as an equatorial ring when viewed from the vegetal pole. Right, the mesoderm that will form the notochord (identified by the blue lines) converges and extends along the midline (Gentsch et al., 2013).

Brachyury's expression has been observed during gastrulation and the process of convergent extension in both *Xenopus laevis* and ascidian embryos (Yamada et al., 2010, Sebé-Pedrós & Ruiz-Trillo, 2016). Convergent extension is a process that plays a key role during gastrulation and other morphogenetic movements (Yamada et al., 2010). It is a mechanism for elongating a sheet of cells while narrowing its width, and occurs by rearrangement of cells within the sheet, rather than by cell migration or division. Convergent extension happens during the extension of the mesoderm that elongates along the antero-posterior axis during which Brachyury is expressed in gastrulating *Xenopus* and ascidian embryos (Figure 1.13) (Cunliffe & Ingham, 1999). Furthermore, ectopic mesoderm is observed in the animal cap when *XBra* is overexpressed (Cunliffe & Smith, 1992).

Brachyury is essential for normal development in the mouse with embryos that are homozygous for the Brachyury mutation showing severe deficiencies in mesoderm formation (Schulte-Merker et al., 1995). Homozygous mutants have no notochord or allantois (gives rise to umbilical chord) and abnormal somites when they are born, as a result they usually die at E9.5 (Stott et al., 1993). Interestingly, Stott and colleagues demonstrated that increasing dosage of a Brachyury transgene representing the wildtype Brachyury allele is able to rescue the heterozygote *T*-associated tail

phenotype, with an increasing dose of wildtype Brachyury causing a corresponding increased extension of the tail along the anterio-posterior axis (Stott et al., 1993).

1.5.3 Regulation of Brachyury

Brachyury is a target gene of the Wnt/ β-catenin (Arnold et al., 2000), the TGF-β/Activin/Nodal(Larocca et al., 2013) and the FGF signalling pathways (Schulte-Merker & Smith, 1995) . It is downregulated by transcriptional repressors like Snail (Fujiwara et al., 1998) and SIP1 (Papin et al., 2002) in various organisms and biological contexts. Brachyury also exists in feedback loops consisting of *T*, *Foxa2*, and *Sox17* and a positive feedback loop with itself and *Fgf8* (Larocca et al., 2013; Lolas et al., 2014)

Deletion analysis in *X. laevis* by Lerchner et al., (2000) revealed that a 371bp promoter fragment is sufficient to drive reporter gene expression in the *Xenopus* embryo, it is also the smallest promoter fragment that can be activated by activin and FGF.

Mouse primordial germ cells (PGCs), the precursors of spermatozoa and oocytes, are induced in pluripotent epiblast by Bmp4 and Wnt3. Aramaki et al (2013) demonstrated *in vitro* that Wnt3 induces a set of transcription factors (among which one is *T*) associated with mesoderm formation in epiblast-like cells through β-Catenin. In this study, Brachyury was confirmed as downstream effector of Wnt3 signalling, sufficient for the specification of PGCs. Wnt3 activates β-Catenin, which upregulates Brachyury (Aramaki et al., 2013).

Arnold and colleagues (2000) identified Brachyury as a target of the Wnt/β-catenin signalling pathway when trying to identify targets of the latter in early mouse embryonic development. They established a co-culture of fibroblasts expressing different Wnts as feeder cells. These ES cells specifically responding to Wnt were identified by a GFP promoter. In these GFP-positive ES cells, they observed expression of Brachyury. Also, they report two TCF binding sites located 500bp from the Brachyury promoter fragment which are able to bind to the LEF-1/ β-catenin-dependent transactivator (Arnold et al., 2000).

1.5.4 Genes regulated by Brachyury

As outlined above, Brachyury is an essential transcription factor during the morphogenetic movements that facilitate gastrulation. Brachyury targets the core promoters and proximal enhancers of key developmental genes in the mouse, zebrafish and *X. laevis,* with these three diverse species sharing 9 identical high-confidence Brachyury binding sites (Lolas et al., 2014)(Figure 1.14). Lolas and colleagues demonstrated T's function as a genome-wide transcriptional activator in mice (*in vitro* and
in vivo) and observed gene regulatory feedback loop between T, *Foxa2* and *Sox17* directs cells towards stem-cell lineage commitments during streak formation. Brachyury target genes were correlated to mRNA sequencing in the mouse and a crucial

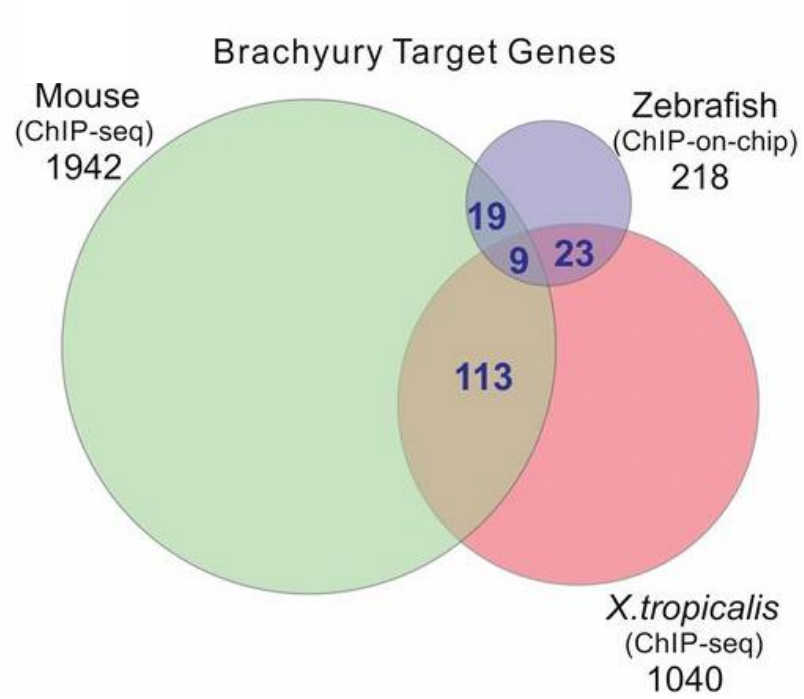

Figure 1.14 Venn diagram showing evolutionary conserved Brachyury target genes. Venn diagram of evolution conservation analyses of common Brachyury target genes in mouse, zebrafish, and X. tropicalis (Lolas et al., 2014).

role for Brachyury in up-regulating multiple key hematopoietic and muscle fate regulators was found (Lolas et al., 2014)

Evans and colleagues (2012) identified 396 potential Brachyury targets in mouse embryonic stem cells, 13.4% of these were transcription factors, with gene ontology analysis identifying a significant enrichment for sequence-specific DNA binding

proteins. Several families of transcription factors were represented, including Ets, paired box, homeobox, winged helix helix forkhead, bZip and zinc finger families. Expression of many of the transcription factors in these families was observed during (*Foxa2, Foxe1*) or just after (*BapX1, EBF2, Erg, Hoxa13, Meis1, Msgn1, Nkx26* and *Slug*) Brachyury expression. Also, several binding sites that differed only 2 to 3 base pairs from the Brachyury consensus binding site were found in close proximity to *Axin2, Ctnb1, Erg, Etv1, Fgf8, Fev, Foxa2, Foxe1, Fyb, Id4, Meis1, and Hoxa3* (Evans et al., 2012).

Several chromatin immunoprecipitation experiments have identified gene ontologies regulated by Brachyury. Evans and colleagues (2012) identified anatomical structural morphogenesis, cell structure morphogenesis and embryonic morphogenesis ontologies as enriched GO terms(Figure 1.15), the bar charts show Gene Ontology (GO) annotations for biological processes. Horizontal bars represent enrichment ratio (observed frequency/expected frequency) and the vertical axis gives the GO term followed by the GO identification number in brackets and hierarchy level. Colour bars indicate statistical significance. GO terms related to the function of Brachyury are highlighted in red boxes. According to this analysis, anatomical structure morphogenesis and embryonic morphogenesis are statistically significant and exhibit a low enrichment ratio. This is consistent with Brachyury's expression as a key transcription factor that orchestrates morphogenetic movements during embryogenesis.

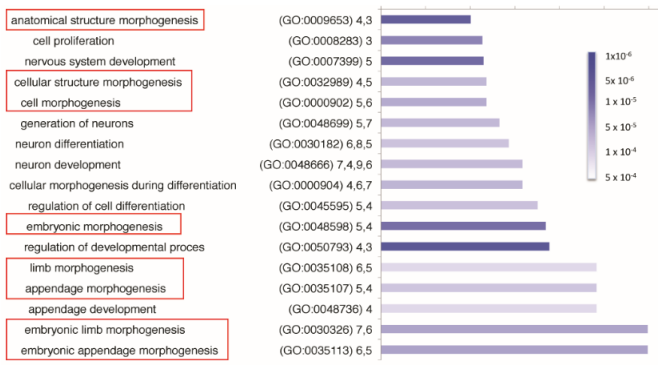

Figure 1.15 Bar charts show gene ontology annotations for biological processes in the mouse embryonic stem cell putatively controlled by Brachyury. Horizontal bars represent enrichment ratio (observed frequency/ expected frequency) and vertical axis gives the GO term followed by the GO identification number in brackets and hierarchy level. Colour bars indicate statistical significance (Evans et al., 2012).

In another study in human embryonic stem cells, biological process ontologies included regulation of gastrulation, negative regulation of Notch signalling pathway,

negative regulation of Wnt receptor signalling pathway, calcium modulating pathway and negative regulation of ectodermal cell fate specification (Faial et al., 2015). In Lolas study (2013), enriched annotation terms associated with seq-regions in mouse, zebrafish and *X. laevis* are shown in (Figure 1.16), these were homologous non-coding regions of the genome of these species that were bound by Brachyury and were associated with nearby genes using the GREAT bioinformatics programme (McClean et al., 2010), the horizontal axis shows the P-value of the annotation terms associated with these nearby genes, interestingly positional regulation of gastrulation and mesoderm morphogenesis are significantly associated with genes nearby Brachyury-bound genomic regions. Furthermore, Kusch and colleagues (2002) previously observed that *Xenopus* and mouse Brachyuyry homologs activate expression of the Brachyury target gene orthopedia in *Drosophila* embryonic cells, providing further evidence of T's conserved function

In another study looking at Brachyury targets in chordoma cells (an osteosarcoma in which the differentiation repertoire of notochordal cells is recapitulated and Brachyury is overexpressed), gene ontology annotation terms enriched compared to control included, developmental maturation, M phase of mitotic cycle, spindle and kinetochore (Nelson et al., 2012) suggesting Brachyury has a role in cell division and differentiation in chordomas. Brachyury has been reported to regulate gene expression through interactions with various other proteins also. Brachyurys interaction with the homeodomain protein Mixl1 in mouse embroyonic stem cells was demonstrated by Pereira et al., (2011). Mixl1 is a transcription factor that is essential for mesoderm and endoderm patterning during mammalian embryogenesis. Mixl1 and Brachyury associated via their DNA-binding domains and

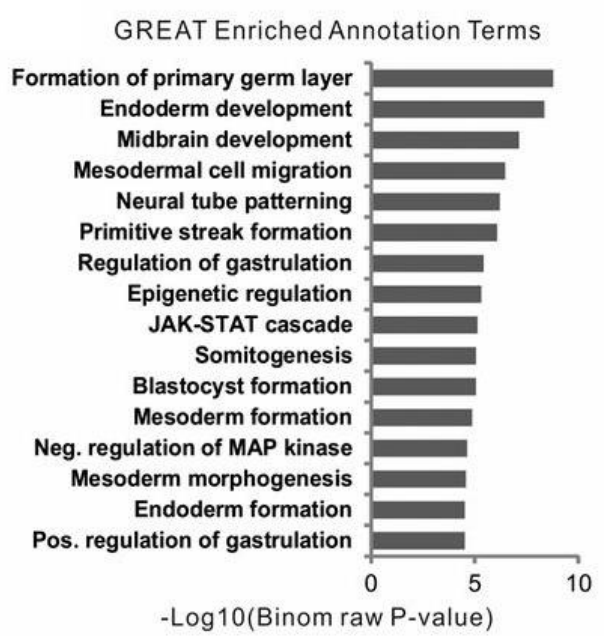

Figure 1.16 Enriched annotation terms associated with Seq-regions. Annotation terms common to all three species using common Brachyury ChIP-seqs (Lolas et al., 2014)

Mixl1 transcriptional activity was negatively regulated through this interaction and similar interactions with other T-box proteins (Pereira et al., 2011). Messenger and colleagues (2005) showed that Xbra in *X. laevis* interacts with Smad1, part of the BMP signalling pathway. When this interaction is interfered with, ectopic activation of the homeobox protein goosecoid is observed and consequent anteriorisation of the embryo. Brachyury-type proteins have also been reported to dimerise via the Brachyury domain (Minguillon et al., 2003). The conservation of this domain suggests that all members of the Brachyury subgroup can dimerise. Kusch et al (2002) demonstrated that a mouse Brachyury-GAL4 fusion construct could interact with the *Drosophila* Brachyury ortholog's Byn interaction domain, indicating that this ability to dimerise between two distant Brachyury homologs is a conserved feature of Brachyury proteins (Kusch et al., 2002).

1.5.5 T function chordomas

Brachyury was first recognised in the context of cancer in chordomas, a rare bone cancer that is thought to originate from notochordal remnants. A duplicated region on 6q27 that contains only *T*, that is important in notochordal development and is expressed in most sporadic chordomas, was reported in four multiplex families with familial chordoma (Yang et al., 2009). Prior to this study Oakley et al., (2008) had identified Brachyury along with SOX9 and podoplanin as new markers in skull base chordomas. Brachyury is a major susceptibility gene in familial chordoma through its increased copy number and hence amplified protein expression. ChIP-seq experiments in chordomas revealed that the Brachyury binding motif is identical to that found in the embryonic development of mouse, *Xenopus* and zebrafish with Brachyury mainly activating transcription of its target genes. Brachyury was reported to regulate several growth factors that regulate proliferation, components of the cell cycle and extracellular components (Nelson et al., 2012).

Nelson and colleagues(2012) also reported that Brachyury knockdown diminishes proliferation and migration in chordoma cells through downregulation of genes that control these processes such as *EGF, TGFA,* and *NUSAP1*. *NUSAP1* is a spindle checkpoint gene that is expressed exclusively in proliferating cells and is required for mitosis. Growth factor receptors including EGFR and FGFR are activated in

chordomas (Siu et al., 2013). Brachyury appears to activate expression of several growth factors, including CTGF, PTN, TGFA, EGF, FGF1, and BMP6 and indirectly regulates the expression of selected integrins (ITGA2 and ITGA10), ECM, and adhesion proteins (LAMA2, HAPLN1, OLFM4, PKP2, COL6A3, CNTNAP2, CHST4, MMP16, and ADAMTS3) (Nelson et al., 2012). Connective tissue growth regulating factor (CTGF), which is required for ECM formation during development and is known to play a pivotal role in regulating the functions of integrins, signalling, and motility in the notochord, was identified as a direct Brachyury target in chordoma (Nelson et al., 2012).

In a study of 181 patients, twelve (7%) tumours displayed amplification of the Brachyury locus and an additional two cases showed focal amplification; 70/181 (39%) tumours were polysomic for chromosome 6, and 8/181 (4.5%) primary tumours showed a minor allelic gain of T. Knockdown of Brachyury in chordoma cell line U-CH1 resulted in a significant decrease in cell proliferation (Presneau et al., 2011). Another study failed to show amplification of Brachyury but did report FGFR expression in 47 out of 50 chordoma patients studied (Shalaby et al., 2009). FGFR is known to regulate Brachyury in ascidian, *Xenopus* and zebrafish embryogenesis, however, no amplification or mutation of the *FGFR* gene was observed and no conclusions could be drawn on the specific signalling pathways active in Brachyury-driven chrodomas (Presneau et al., 2011).

1.5.6 Brachyury expression in cancer

After the discovery of the amplification of Brachyury in chordomas and following the rationale that it is an inducer of mesoderm, Palena et al., (2007) reported Brachyury expression in several tumours of the small intestine, stomach, kidney, bladder, uterus, ovary and testis.

Overexpression of Brachyury in the human pancreatic cell line PANC-1 demonstrated that Brachyury promotes changes in this cell line consistent with the EMT, induction of mesenchymal markers, downregulation of epithelial markers and increases cell migration and invasion (Fernando et al., 2010). Overexpression of Brachyury in a lung cancer cell line (H460) reduced E-cadherin transcript levels, while

inhibition of Brachyury caused downregulation of mesenchymal markers and a reduced capacity to form lung metastases in a mouse xenograft model (Fernando et al., 2010). In a later study, Brachyury expression was identified as a prognostic marker in primary lung carcinomas, with high Brachyury expressing tumours predicting poor prognosis in both 5 year-survival and overall survival rates (Roselli et al., 2012). Brachyury mRNA expression was significantly inversely correlated with *CDH1* expression and positively correlated with *SLUG* expression, the latter has been reported to repress expression of the former (Haro et al., 2013) and was also linked to the EMT and metastasis in oral squamous cell carcinoma (Shimoda et al., 2012) and hepatocellular carcinoma (Du et al., 2014).

Brachyury expression has recently been associated with the ability of breast cancer cells to form mammospheres *in vitro* and to confer docetaxel resistance. Further, high Brachyury expression is associated with a higher risk of recurrence and distant metastases in breast cancer (Palena et al., 2014). In another study, increased Brachyury expression during TGF-β1-induced EMT was observed with a subsequent autocrine loop being sustained between Brachyury and TGF-β1 in mesenchymal-like tumour cells (Larocca et al., 2013). In this cellular context, Brachyury overexpression promoted upregulation of TGF-β1 at the mRNA and protein levels with Brachyury inducing transcriptional activation of TGF-β1 through binding to its promoter. Furthermore, inhibition of TGF-β1 signalling decreased Brachyury levels and reversed the EMT (Larocca et al., 2013).

Brachyury was recently implicated in prostate cancer progression and aggressiveness (Pinto et al., 2014). Increased Brachyury was observed in primary prostate cancers and metastatic tumours when compared with normal tissues both at protein and at mRNA levels. Cytoplasmic localisation of Brachyury was correlated with highly aggressive tumours and nuclear Brachyury was correlated with tumour invasion. *In vitro*, Brachyury-positive cells had a higher viability, proliferation, migration, and invasion rates than Brachyury-negative cells. Furthermore, a majority of genes co-expressed with Brachyury were found to be functionally clustered in the categories of immune response, cell membrane/receptor activity, development, cell motility and chemotaxis, pathways in cancer and response to hormone stimulus. (Pinto et al., 2014)

Sarkar and colleagues from our lab (2011) observed that Brachyury expression in the CRC cell line SW480 is cell density-dependent, with higher Brachyury expression detected in low-density (invasive front-mesenchymal-like cells) *in vitro*. Furthermore, they demonstrated that Brachyury regulates the expression of Nanog, a key regulator of pluripotency. However, both Fernando et al., (2010) and Sarkar et al., (2011) suggest that Brachyury alone is insufficient to induce the EMT.

Immunohistochemistry studies by Roselli (2012) and colleagues reported Brachyury expression in 41% of primary lung carcinomas, including 48% of adenocarcinomas and 25% of squamous cell carcinomas, they also report that Brachyury was negative in all normal tissues studied apart from the testis. The expression of Brachyury mRNA was also shown in primary lung tumour tissues, predominantly in tumours of higher stages (stages 2–4) than among those of stage 1 or histologically normal lung (Roselli et al., 2012) and in a separate study was correlated with a poor prognosis of lung cancer (Haro et al., 2013). Furthermore, when Brachyury was depleted by siRNA-mediated knockdown, H460 lung cancer cells were more sensitive to EGFR inhibitor treatment *in vitro* (Roselli et al., 2012).

When lung carcinoma cells were treated *in vitro* with chemotherapeutic agents, slow growing Brachyury high cells were selected for, furthermore treatment of *in vivo* tumour xenograft models with these same agents selected for the growth of resistant Brachyury high tumours with lower rates of growth (Huang et al., 2013). High levels of Brachyury in these tumours are suggested to inhibit the cell cycle through cyclin D, pRb and p21 downregulation conferring chemotherapeutic a radiation resistance on these cells.

Other studies have suggested that Brachyury is a potential regulator of EMT and CSCs in adenoid cystic carcinoma (Shimoda et al., 2012). When Brachyury is depleted in cells of this oral cancer, cells are sensitised to chemotherapy and radiation *in vitro* (Kobayashi et al., 2014)

To date, only one clinical investigation has been conducted into Brachyury expression in colorectal cancer. In this study, 748 patients were immunohistochemically analysed and correlated with classical and new markers of colorectal cancer. Significantly lower survival rates were seen in patients at the early stages of colorectal cancer while they found no correlation with the presence of

Brachyury in later tumour stages. They concluded that a subset of colorectal cancers showed an ability to metastasise at the early stages of tumour growth, this group can be predicted by IHC analysis of Brachyury expression (Kilic et al., 2011).

In the present study we aimed to characterise Brachyury expression in the normal human small intestine, colon and precancerous adenomas of this tissue. We aimed also to use a robustly validated antibody to determine the relationship between Brachyury expression and survival outcomes of patients with primary colorectal cancer and associated metastases. Finally, we tested several small molecule inhibitors designed to inhibit Brachyury and its effect on cancer cell lines in which it is expressed.

2 Materials and Methods

2.1 Cell Line Culture

SW480 cells [American Type Culture Collection (ATCC) #CCL-228] and NTERA2 (ATCC CRL-1973) were cultured in Dulbecco's modified Eagle's Medium (DMEM; Life Technologies, #61965) supplemented with 10% Foetal Bovine Serum (FBS, Life Technologies, #10270). H460 cells (ATCC #HTB-177) were grown in Roswell Park Memorial Institute medium (RPMI, Life Technologies #61870) supplemented with 10% FBS (Life Technologies #10270). SW620 cells ATCC #227. All cell lines were cultured at 37°C with 5% CO_2 for SW480, SW620 and H460 and 10% CO_2 for NTERA2 in a humidified incubator.

2.2 siRNA Transfection of Cell Lines

Gene-specific mRNA depletion was performed using commercial siRNA molecules from Qiagen and Ambion at a final concentration of 5nM. BRACHYURY and negative control sequences can be found in Table 2.1. HiPerfect was used as a transfection reagent with transfection experiments being carried out using the manufacturers instructions. Briefly, 150 ng of siRNA was added to 6µl HiPerfect together with 100µl of serum-free media and vortexed. The transfection complexes were incubated for 15 min at room temperature on a rotary incubator. Cells were plated on 6-well plates at a seeding density of 1.5 x 10^5 cells per well. The siRNA complexes were added drop-wise and plates gently rotated before being placed back into the incubator. siRNA transfection experiments consisted of 3 siRNA treatments, the first on plating the cells, the second 24 hours after seeding and the third 48 hours after seeding. The cells were harvested 72 hours after seeding. This consisted of RNA and whole cell lysate extract to assay knockdown efficacy through measuring RNA and protein levels by qRT-PCR and Western blot.

Table 2.1 – siRNA molecules used to transfect parental cells

siRNA	Name	Catalogue Number	Target Sequence
Non-interfering	NI	SI03650318 (Qiagen)	Proprietary
Brachyury	Hs_T_5	SI04133521 (Qiagen)	5'-GAGGATGTTTCCGGTGCTGAA-3'
Brachyury	SI3703	43925420 (Ambion)	5'-CGGUGACUGCUUAUCAGAAtt-3'
Brachyury	SI3704	43925420 (Ambion)	5'-GCUCUUAAAAUUAAGUACAtt-3'
Brachyury	Si3705	43925420 (Ambion)	5'-GGAACAAUUCUCCAACCUAtt-3'

2.3 Protein Extraction

2.3.1 Whole-cell lysates

M-PER lysis buffer (ThermoFisher scientific, #78503) was used to prepare whole-cell lysates, a proprietary non-denaturing detergent in 25nM bicine buffer (pH 7.6). To counter protein degradation and dephosphorylation the buffer was supplemented with protease and phosphatase inhibitors (ThermoFisher Scientific, Halt Protease Inhibitor Cocktail, #87785; Phosphatase Inhibitor Cocktail, #78420) following the manufacturer's protocol. BCA (bicinchonionic acid) assay (BCA protein Assay kit, ThermoFisher Scientific, #23227) was used to measure protein concentration of whole cell lysates using the nanodrop spectrophotometer (ThermoFIsher #ND-2000).

2.3.2 Subcellular Protein Fractionation

Subcellular protein fractionation was carried out using the NE-PER nuclear and cytoplasmic extraction reagents (ThermoFIsher #78833). Nuclear and cytoplasmic extracts were prepared from cultured cells after trypsinisation. Reagents CERI and NER were supplemented with protease and phosphatase inhibitors (ThermoFisher, Halt Protease Inhibitor Cocktail, #87785; Phosphotase Inhibitor Cocktail, #78420). BCA (bicinchonionic acid) assay (BCA protein Assay kit, ThermoFisher Scientific, #23227) was used to measure protein concentration of whole cell lysates using the nanodrop spectrophotometer (ThermoFIsher #ND-2000).

2.4 Western Blot

For each sample approximately 30µg of protein was used. These samples were made up to 13µl with water and finally; 2µl reducing agent and (ThermoFisher #B0009 10x Bolt Sample Reducting Agent) and 5µl Sample Buffer (Thermo Fisher #B0007 4X Bolt LDS Sample Buffer). The samples were heated to 70°C for 10 min before electrophoresis. Precision plus protein dual colour standards (Biorad #161037) protein ladder was run alongside the samples in a 4-12% precast polyacrylamide gel (Thermo Fisher #NW04125BOX) immersed in 2-(N-morpholino)ethanesulfonic acid [ThermoFIsher MES DSD Running Buffer (20X) #NP0002] for 30 min at 100V. Proteins were then transferred to a PVDF (Polyvinylidene fluoride) (Immonilon-P, Millipore, #IPV00010) using the wet transfer system from Biorad #170353 at 1 A for 90 min in transfer buffer (25mM Tris, 192mM Glycine). PVDF membranes were soaked in 100% methanol prior to transfer.

Membranes were blocked for one hour at room temperature with Milk Solution [10% dry milk in PBS/ 0.5% Tween 20 (Sigma, #P2287)]. Primary antibodies were diluted in milk solution and incubated with the membrane overnight at 4°C. The membrane was blocked for 10 min in milk solution before adding HRP (Horse Radish Peroxidase)-conjugated secondary antibody that was also diluted in milk solution and incubated for one hour at room temperature. Finally, three 10 min washes were performed using Tween20 PBS (Phosphate-buffered saline)] solution (0.1% Tween in 1x PBS) to remove any non-specifically bound and unbound antibodies. A list detailing all of the antibodies used for Western Blotting are shown in Table 2.2 (please see figure 3.3 for epitope map)

Western lightning Plus-ECL (Perkin Elmer #NEL105001EA) was used as a substrate for the HRP-conjugated secondary antibodies following the manufacturer's protocol. Chemiluminescent signals were detected using digital camera detection with the Biorad ChemiDoc XRS+ imaging system (#17089265).

Table 2.2- Antibodies used for Western Blotting Analysis

Primary Antibody Target	Cat. No.	Company/Source	Lot No.	Host	Clonality	WB Dilution
Brachyury	AB140661	AbCam		Mouse	Monoclonal	1 in 500
Brachyury	AF2085	R & D Systems		Goat	Polyclonal	1 in 500
Brachyury	AB57480	AbCam		Mouse	Polyclonal	1 in 500
LRIG1	12752	Cell Signaling		Rabbit	Polyclonal	1 in 1000
EGFR	4267	Cell Signaling		Rabbit	Monoclonal	1 in 1000
α-tubulin	T6074	Sigma		Mouse	Monoclonal	1 in 8000
Lamin B	Sc-6217	Santa Cruz		Goat	Polyclonal	1 in 500
HA Tag	3274	Cell Signaling		Rabbit	Monoclonal	1 in 1000
Secondary Antibody Cat. No.	**Company/Source**	**Host**	**Species Reactivity**	**Isotype**	**Conjugate**	**Western Blot dilution**
Ab97051	AbCam	Rabbit	Goat	IgG	HRP	1 in 25000
7074	Cell Signalling	Goat	Rabbit	IgG	HRP	1 in 3000
7076	Cell Signalling	Horse	Mouse	IgG	HRP	1 in 3000

2.5 RNA Extraction (Parental cells) and first strand cDNA synthesis

2.5.1 RNA Extraction

Total RNA was extracted from parental cells using RNeasy Plus Mini Kit (Qiagen, #74136) following the manufacturer's instructions. Briefly, cells were washed with PBS in a 6-well plate, to avoid any gene expression changes during trypsinisation, cells were lysed directly with the highly denaturing RLT (RNase Later) Plus buffer

which simultaneously inactivates RNases. The volume of buffer was adjusted depending on cell number (350µl for 5 x 10^6 cells; 600µl for up to 1x 10^7 cells). Lysates were passed through a gDNA eliminator spin column in order to eliminate genomic DNA. The flow-through was then applied to an RNeasy spin column after the addition of an equal volume of 70% ethanol followed by two wash steps before elution. The final RNA purity (260/280nm ratio of ~2) and concentration was determined using the NanoDrop 2000c (ThermoFisher #ND-2000). If RNA quality needed to be assessed more accurately; samples were assayed with the Experion Automated Electrophoresis System (Biorad, #7017000). n.b. whereas the NanoDrop only provides the 260nm/280nm ratio which only provides information about protein or phenol contamination and does not give appropriate and full information according RNA integrity, the experion automated electrophoresis system uses the 28S/18S ratio which represents RNA integrity.

2.5.2 First strand cDNA Synthesis.

First strand cDNA synthesis from an RNA template was carried out using SuperScriptIII First-Strand Synthesis System (ThermoFisher # 1808-051). 1µg of RNA was reverse transcribed following the manufacturer's protocol. Samples were treated with RNase H to eliminate residual RNA. The final cDNA was then diluted 1 in 8 with DNase and RNAse-free water. Non-reverse transcriptase samples were included for all samples with no amplification being observed in subsequent qPCR (quantitative real-time polymerase chain reactions.)

2.6 Quantitative Real-time PCR

qPCR reactions were performed using GoTaq qPCR Master Mix (Promega #A6001), in the CFX96 Real-Time PCR Detection System C100 (Biorad #185-5096). Triplicate reactions were prepared for each sample in a final volume of 20µl. Each reaction consisted of 1xqPCR Master mix, 1.5µl of diluted cDNA (equivalent to 9.375ng/µl of RNA), and primers to a final concentration of 0.1µM. The PCR program consisted of an initial denaturation at 95°C for 10 minutes, followed by 40 cycles of 95°C for 15

seconds and 60 °C for 1 minute. This was followed by melt curve analysis from 60°C to 95°C (in 0.5°C increments). Custom-designed primers and commercial primers (Quantitect Primer Assays, Qiagen) were used in this study with the all primers used listed in below Table 2.3please see figure 3.3 for primer map).

Table 2.3 Brachyury transcript variant specific primers and Quantitect primer assays used for RT-PCR with SYBR green detection.

Gene/Target	**Primer Name**	**Primer Sequence**
Brachyury TV1	Bry Q TV1 Fw Bry Q TV1 Rv	5'-GTGACAGGTACCCAACCCTG-3' 5'-GGTGAGTTGCAGAATAGGTTGGA-3'
Brachyury TV2	Bry Q TV2 Fw Bry Q TV2 Rv	5'-ACCTGGGTACTCCCAATCCTA-3' 5'-GCTGGACCAATTGTCATGGG-3'
Gene/Target	**Quantitect Primer Assay**	**Cat. Number**
Brachyury	Hs_T_1	QT00062314
LRIG1	Hs_LRIG1_1	QT00087430
HSP09A	Hs_HSP90AB1	QT01002624
GAPDH	Hs_GAPDH_1	QT00079247
CROT	Hs_CROT_1	QT00021112
SFRP5	Hs_SFRP5_2	QT01008182
IGFBP3	Hs_IGFBP3_1	QT00072737
CCNE1	Hs_CCNE1_1	QT00041986
MTHFD2	Hs_MTHFD2	QT00081592
MSH2	Hs_MSH2_1	QT00032466
MFN2	Hs_MFN2_1	QT00057589
MEST	Hs_MEST_1	QT00048426

2.7 Immunohistochemistry (IHC)

Human tissue was obtained from patients undergoing tumour resection during treatment for adenomas and carcinomas of the colon and rectum with normal adjacent tissue obtained simultaneously, both with approval from the North Wales Research Ethics committee – West). Tissue Microarrays (TMAs) were prepared by Dr Rui Manuel Reis' laboratory (Portugal). Ethical approval was obtained by him for the

use of these patients TMAs in this study separately. Tissue samples were fixed in formalin and embedded in molten paraffin (FFPE). Tissue blocks were cut to 4µm sections. Haemotoxylin and Eosin (H&E) staining was performed to assess tissue architecture and morphology. This was followed by automated IHC staining using an optimised protocol on the Ventana Benchmark XT(Roche #). Antigen retrieval and antibody dilutions are listed in Table 2.4.

Table 2.4: Primary antibodies for IHC and dilutions, antigen retrieval and sources.

Antibody	Clonality	Host	Dilution	Antigen Retrieval	Source	Cat. No.
Anti-Brachyury	Monoclonal	Mouse	1in750-1in1500	Protease (4 min) 1° Ab 1hr + UltraWash	AbCam	Ab57480
Anti-Brachyury	Monoclonal	Mouse	1in50-1in300	STD CC1, 1° Ab 1hr +UltraWash	AbCam	Ab140661
Anti-$p27^{Kip1}$	Monoclonal	Rabbit	1in2500	STD CC1, 1° Ab 1hr +UltraWash	Cell Signaling	3686
Anti-KI67	Monoclonal	Mouse	1in300	STD CC1, 1° Ab 1hr +UltraWash	LSBio	LS-B64433

*CC1 and protease I are Ventana Products (Roche #)

2.7.1 Kaplan-Meier Survival Analysis

Patients were stratified according to Brachyury staining pattern. Brachyury staining categories were compared with each other for survival and disease-free survival. Univariate analysis of disease-free survival was carried out by the Kaplan–Meyer method using the Log rank test on GraphPad Prism 6.

2.8 Indirect Immunofluorescence (IF)

SW480, SW620 and NTERA2 cells were seeded at $1x10^5$ cells per well in a 24-well plate, cultured on glass coverslips until 70-80% confluent at which point they were

Table 2.5 Primary and Secondary antibodies used for IF with dilutions, sources, excitation and emission wavelengths

Primary Antibody	Clonality	Host	Dilution	Source	Cat. No.
Anti-Brachyury	Monoclonal	Mouse	1in100	AbCam	Ab140661
Anti-HA Tag	Monoclonal	Rabbit	1in100	Cell Signaling	#3724

Secondary Antibody	Host	Species Reactivity	Dye/Label	Excitation/ Emission	Dilution	Cat. No.
Anti-mouse IgG	Goat	Mouse	AlexaFluor® 488	495/519	1:500	A11029
Anti-mouse IgG	Goat	Mouse	AlexaFluor® 568	578/603	1:500	A11031
Anti-rabbit IgG	Goat	Rabbit	AlexaFluor® 488	495/519	1:500	A11034
Anti-rabbit IgG	Goat	Rabbit	AlexaFluor® 568	578/603	1:500	A11011

fixed in 4% paraformaldehyde (PFA) in warm PBS for 15 min then cells were washed three times in PBS for 5 min each. The specimen was blocked for one hour in a blocking buffer that contained 0.3% of the permeabilising agent Triton-X 100,

1X PBS and 5% FBS. Subsequently, the blocking solution was aspirated and primary antibody (Table 2.5) was applied and incubated overnight at 4°C. Cells were rinsed three times in PBS for 5 min each the next day followed by 2 hour incubation with fluorochrome-conjugated secondary antibody (Table 2.5) diluted in antibody dilution buffer (1XPBS/1%BSA/0.3% TritonX-100) at room temperature in the dark. Cells were once again rinsed three times in 1X PBS for 5 min each. Coverslips were subsequently mounted onto glass slides that had a drop of Gold Antifade Reagent with DAPI (Cell Signalling #8961). Mountant was allowed to cure overnight at room temperature covered by tissue paper. In the morning, nail varnish was used to seal the coverslip. The cells were then imaged on the Zeiss 710 confocal microscope.

2.9 Cloning using the pCMV MCS N-HA qIP vector

2.9.1 Cloning

The protein-coding sequence of Brachyury was cloned into the eukaryotic expression vector, pCMV MCS N-HA (Thermo Scientific, #82017) containing an N-Terminal HA tag as described by the manufacturers protocol. Briefly, the ORF of interest was PCR amplified from H460 cDNA using 1 unit of Phusion High-Fidelity DNA Polymerase (NEB, M0530S) in a 50µl reaction consisting of 2µl cDNA, 1x Phusion HF Buffer, 200mM dNTPs and 0.5µM each of Brachyury-specific primers (Transcript Variant 1) that included the restriction site for *NheI* and *AscI* (underlined). Sequence of multiple cloning site (MCS) underlined below:

5'- GATATCGGCGCGCCTATGAGCTCCCCTGGCACC-3'

3'-GTCGCCACCTTCCATGTGAGCTAGCGATTAC-5'

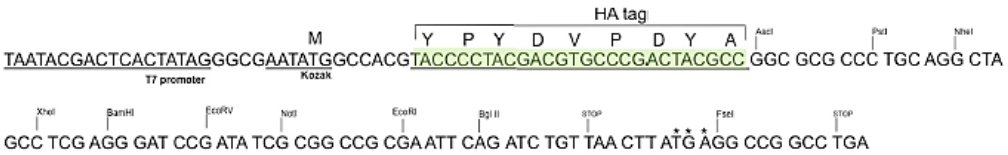

Figure 2.1 Cloning primers and promoter, MCS and HA-Tag sequences of pCMV-HA cloning vector. (A) Brachyury TV1-specific primers for amplification of TV1 with NheI and AscI restriction sites on the 5' ends (underlined). (B) MCS (underlined) of cloning vector pCMV-HA and peptide sequence of N-terminal HA tag (1.1 kDa).

Using the following PCR conditions: Initial denaturation of 98°C for 2 min followed by 35 cycles of 98°C for 5s, 55°C for 30s and 72°C for 30s followed by a final extension of 72°C for 5 min. Primers were designed using primer3 design online tool (primer3.ut.ee/).

The PCR product of Brachyury transcript variant 1 (TV1) was gel purified using the nucleospin gel and PCR clean-up kit from Macherey Nagel (#740609.250). The purified product was digested with *NheI* and *AscI* in a 50µl reaction mixture of 1x cutsmart buffer from NEB (#B2074S) and 10 units of *AscI* (#R013S) and *NheI* (#R0558S) for 1 hour at 37°C. The reaction was incubated for 20 minutes at 65°C to inactivate the enzymes. Subsequently 100ng of the digested PCR product was purified and ligated with 50ng of *AscI* and *NheI*-digested pCMV MCS N-HA vector in a 20µl reaction with 1 µl of T4 DNA Ligase (NEB #M0202) for 1hour at room

temperature followed by heat inactivation at 65°C for 10 minutes. The sequence of pCMVTV1-HA shown in Appendix C.

Five microliters of the ligation reaction was transformed into 50µl of chemically competent JM109 (>10^8cfu/µg Genotype: F´traD36 proA+ B+ lacIq Δ(lacZ)M15/ Δ(lac-proAB) glnV44 e14- gyrA96 recA1 relA1 endA1 thi hsdR17, NEB E4107S). Transformants were selected on LB agarplates containing 100µg/ml of Kanamycin. Resistant colonies were picked and screened using the QIAprep Spin Miniprep Kit (Qiagen, 27106) and restriction digestion. Vectors with inserts of the expected size were verified by nucleotide sequencing using primers from Figure 2.1. Transfection-grade plasmid DNA was isolated using the the HiSpeed Plasmid Midi Kit (Qiagen, 12643).

Glycerol stocks were prepared from liquid cultures, grown in LB liquid broth supplemented with 50µg/ml kanamycin. 0.1 of 10x Hogness buffer (40% glycerol, 36mM K_2HPO_4, 13mMKH_2PO_4, 20mM trisodium citrate, 10mM $MgSO_4$ in deionised water) followed by storage at -80°C.

2.9.2 Purification of PCR products

PCR products were run on a 1.2% agarose gel stained with peqGreen DNA/RNA dye (PeqLab, 37-5000) and the bands were cut out of the gel using a sterile razor blade on a transilluminator. The products were purified using nucleospin gel and PCR clean-up kit from MN (#740609.250) following the manufacturer's instructions. The purified PCR product was eluted from the column using nuclease-free water.

2.9.3 Sequencing of PCR products

DNA samples at a concentration of 5ng/µl, were mixed with 10pmol of the corresponding forward or reverse primer in a total volume of 15µl and sent at room temperature in a 1.5ml Eppendorf tube to Eurofins MWG. Sequencing results were aligned with the expected sequence using NCBI BLAST and for cloning the open reading frame was checked using ORF reader to ensure the sequence was in frame.

Results

3 Brachyury has two splice variants that are expressed in colorectal cancer cells and patient-derived CRC tissue.

3.1 Introduction

Many genes are alternatively spliced and have multiple transcript variants (Nilsen & Graveley, 2010). Many isoforms are cancer-specific and the splicing pattern of transcript variants of many genes is altered during tumourigenesis. The splicing apparatus is itself de-regulated in cancer and in some cases can act as oncogenes that can contribute to positive feedback loops driving oncogenesis (Oltean & Bates, 2013)

Novel splicing alterations in cancers contribute to their altered transcriptome. These alterations can act as both a prognostic signature and new therapeutic targets. Sequencing of the transcriptome of the three most common types of breast cancer TNBC (Triple-negative breast cancer), non-TNBC and HER2-positive breast cancer revealed subtype specific differentially spliced genes and novel splice isoforms, not previously reported in a human transcriptome (Eswaran et al., 2013). The expression of a specific transcript variant of Androgen Receptor (AR) was found to be associated with shorter survival outcomes (Antonarakis et al., 2014) and AR inhibitor resistance (Hickey et al., 2013). Furthermore, He et al.,(2009) identified 15,093 cancer-specific transcript variants of 9,989 genes from 27 types of human cancers. These variations in the expression of transcript variants are likely to produce functionally different proteins from the same gene and generate large variations in the proteome that are specific to cancer cells.

Human Brachyury is not frequently mutated in cancers (Figure 3.1) but it confers major susceptibility to familial chordoma when duplicated (Yang et al., 2009) and so is likely to contribute to oncogenesis through other means which might also include aberrant overexpression as it has been reported to be in several cancers (detected by IHC) (Du et al., 2014; Palena et al., 2007; Pinto et al., 2014)or alternative regulation of splice variants. Brachyury has 4 predicted transcript variants(ENSEMBL accession numbers: ENST00000296946.6, ENST00000366871.7,

ENST00000366876.6, ENST00000461348.2) but only two are listed with the NCBI (NCBI accession numbers AF012130.1 and AF012131.1). According to both databases, the Brachyury gene has 9 exons, with exons 2-9 coding for protein. The primary transcript for transcript variant 1 (TV1) has 9 exons, with exon 2 (2a)

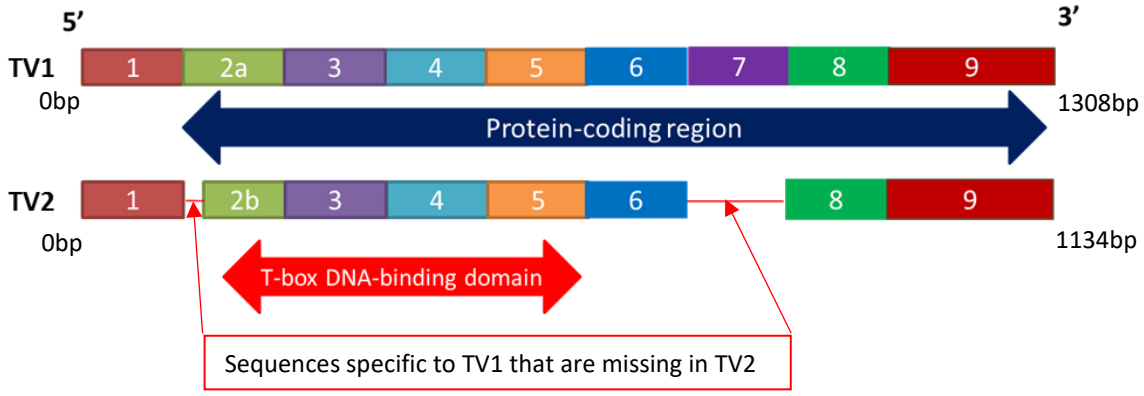

Figure 3.2 Schematic representation of the two splice variants of *T*, TV1 and TV2. Exons 2a/b -9 code for the Brachyury protein. TV1: Exon 7 is TV1-specific and both variants have the evolutionary conserved T-box DNA binding domain. TV2 differs in the 5' UTR and is missing an in-frame exon 7 compared to TV1.

being transcript variant 1-specific. The transcript variant 2 (TV2) primary transcript has all the exons apart from 7. It also has a transcript variant-specific exon 2 (2b) (Figure 3.2).

mRNA expression and/or protein levels of Brachyury (TV1) have been evaluated in many other cancer types but none have distinguished between the two transcript variants (Du et al., 2014; R. I. Fernando et al., 2010; Kilic et al., 2011; Palena et al., 2007; Pinto et al., 2014; Roselli et al., 2012; Vujovic et al., 2006).

We investigated whether expression patterns of the two transcript variants of Brachyury altered in different CRC cells and tried to establish whether or not their expression was restricted to a particular subcellular localisation. It was also important to try to establish the specificity of anti-Brachyury antibodies to the different transcript variants.

3.2 Brachyury splice variants TV1 and TV2 are expressed in SW480 colorectal cancer cells at the mRNA and protein levels.

In order to investigate the levels of Brachyury splice variants at the mRNA and protein levels, we used transcript-specific primers to amplify a small amplicon of each splice variant via qRT-PCR analysis, anti-Brachyury antibodies were used to detect transctipt variant proteins using Western blots. Several anti-Brachyury siRNAs were used to knockdown the expression of the two Brachyury transcripts in SW480 colorectal cancer cells (Figure 3.3).

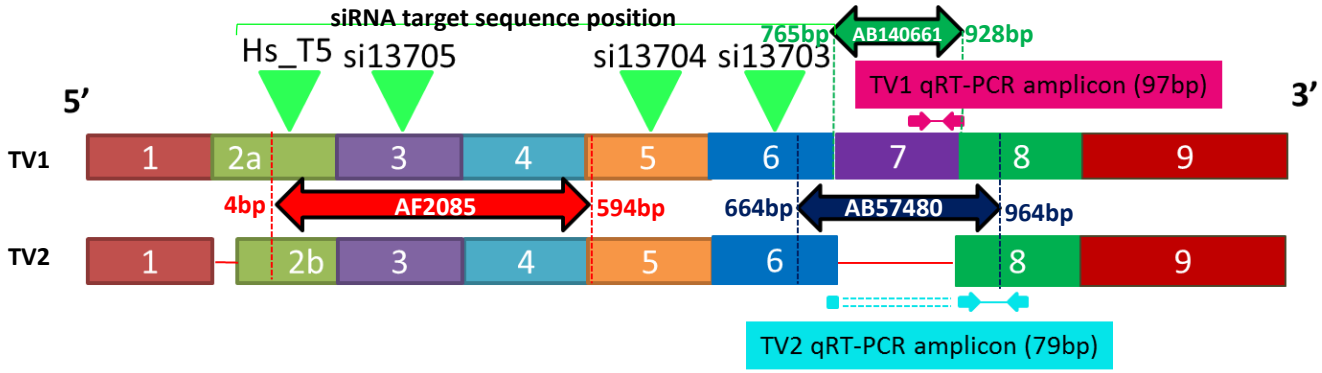

Figure 3.3 Schematic representation of the two Brachyury splice variants, TV1 and TV2, showing the location of qRT-PCR amplicons specific to each transcript variant, anti-Brachyury antibody epitope locations and anti-Brachyury siRNA sequence locations. A and B represent TV1 and TV2 exons respectively. Green arrowheads show the location of respective siRNA sequences in the Brachyury transcripts (Hs_T5, si13705, si13704 and si13703). Double-ended arrows represent the sequence location of small peptide epitopes used to raise anti-Brachyury antibodies; green is the monoclonal mouse anti-Brachyury antibody (AB140661), navy is the monoclonal mouse anti-Brachyury antibody (AB57480) and red is the polyclonal goat anti-Brachyury antibody (AF2085). qRT-PCR amplicons are represented by the small inward facing arrows separated by a line, pink represents the TV1 amplicon that covers the TV1-specific exon 7 and turquoise represents the TV2 amplicon that spans the exon 6/8 boundary specific to TV2.

siRNA-mediated Brachyury depletion was performed in SW480 cells as detailed in section 2.2. Briefly, several Brachyury-specific siRNAs (Figure 3.3) were used to deplete Brachyury expression in SW480 cells, these were harvested with total RNA and protein being extracted with this experiment being performed in triplicate.

qRT-PCR was performed using splice variant-specific primers (Figure 3.3), these were used to quantify the normalised fold expression of each Brachyury transcript variant after siRNA-mediated knockdown (Figure 3.4). TV1 transcript was knocked down using all the siRNAs (listed in Figure 3.3). However, TV2 transcript levels were knocked down by all the siRNAs except for si13705 (05 in Figure 3.4).

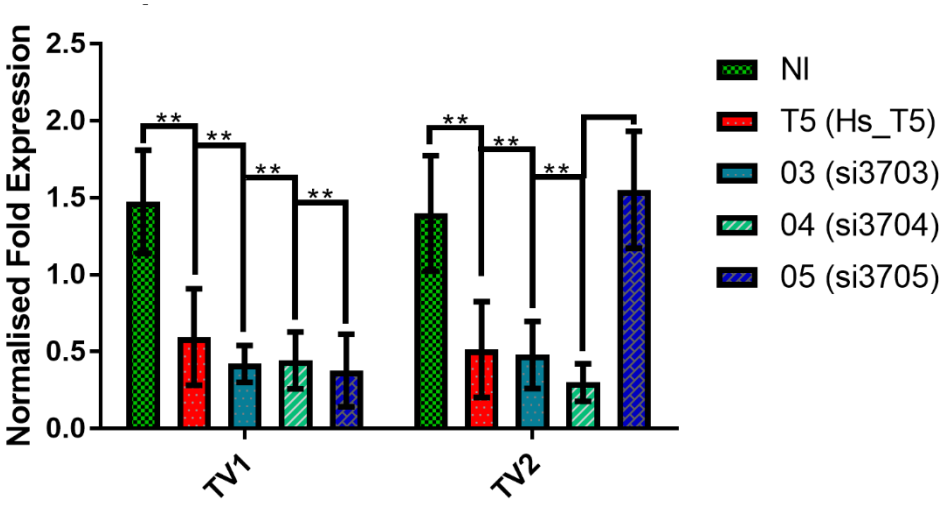

Figure 3.4 qRT-PCR analysis of Brachyury knockdown in SW480 colorectal cancer cells. Levels of both Brachyury splice variants were depleted at the mRNA level across three biological replicates by siRNAs T5, 03 and 04. Transcript variant 2 was not however knocked down by siRNA 05. This siRNA specifically and reproducibly knocked down TV1. Expression levels were normalised with housekeeping genes *HSP09A* and *GAPDH*. P-values of individual pairwise comparisons between controls and siRNA treated samples are represented by *. * <0.05,**,0.01,***<0.001,****<0.0001.

Similarly, the polyclonal goat anti-Brachyury antibody (AF2085) was used to detect Brachyury protein after siRNA-mediated Brachyury knockdown using Western blot analysis using protein prepared from the same samples (Figure 3.5). Western blot analysis using this antibody was used to assess whether or not Brachyury splice variant levels were depleted at the protein level, as we would expect this antibody to detect both splice variants (Figure 3.3). Two bands (approximately 50kDa and 43kDa) were observed in both the untreated control sample of SW480 cells and the scrambled siRNA (NI) control sample (Figure 3.5 A). Levels of both bands were depleted in T5, 03 and 04 samples. The upper band was depleted in the 05 sample and the lower 43kDa band remained. The predicted band size for the two Brachyury isoforms are 47 and 41kDa respectively. The band sizes observed were slightly larger than expected at 50kDa and 43kDa respectively.

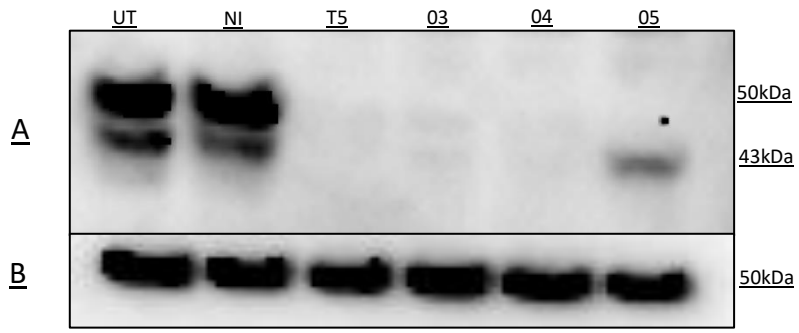

Figure 3.5 Western blot analysis to show knockdown of Brachyury at the protein level and transcript variants. (A) Goat polyclonal anti-Brachyury antibody (AF2085). Three knockdown repeats revealed that T5, 03 and 04 consistently knocked down the upper band and 05 knockdown levels were higher for the upper band. (B) Mouse monoclonal anti-Tubulin antibody to demonstrate equal protein loading. Protein band sizes were determined automatically using the marker and band size function in the ImageLab software package.

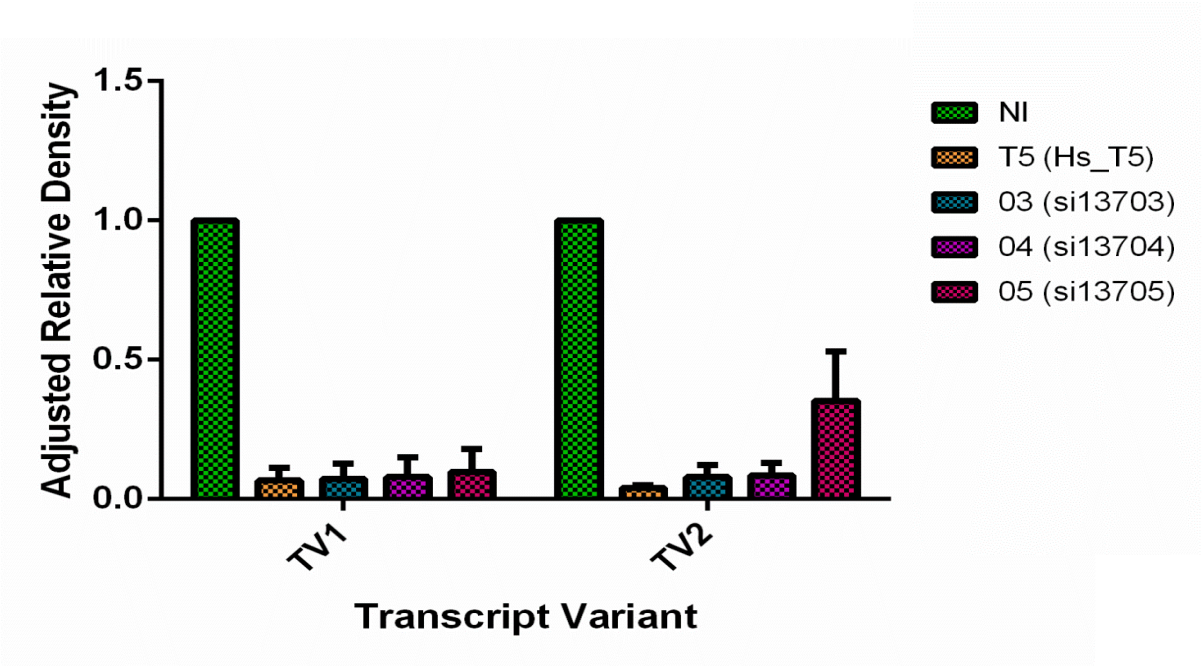

Figure 3.6 Densitometry on three repeats of the Brachyruy knockdown in SW480 to assess the protein levels of the two bands detected by AF2085. Adjusted relative density was used to quantify the protein levels of the two bands corresponding to the two Brachyury splice variants across the 6 samples. The adjusted relative density was calculated using the intensity of the bands of Brachyury and normalising these to the loading control, the ubiquitously expressed cytoplasmic protein α-tubulin.

We confirmed the depletion of the expression of both transcript variants by using densitometry across all three biological replicates (Figure 3.6). This confirmed that the bands corresponding to TV1 and TV2 were knocked down by T5, 03 and 04. The lower band (TV2) was not knocked down completely by 05. qRT-PCR analysis of TV1 and TV2 transcript levels in the 05 sample (Figure 3.4) suggest a 05 knocks down

transcript levels of TV2 specifically, proteins levels in the siRNA sample are depleted however (Figure 3.6). The upper band (corresponding to TV1) being knocked down at both the RNA and protein levels in the sample treated with siRNA 05 and the lower band (corresponding to TV2) remaining.

3.3 TV1 and TV2 are expressed in the nucleus and the cytoplasm of SW480 and SW620 colorectal cancer cells.

We investigated the localisation of the two splice variants in CRC cells. SW480 is a colorectal cancer cell line derived from a primary colorectal tumour and SW620 is a metastatic cell line of colorectal origin derived from the same patient. Each variant was cloned into a pCMV-HATag vector with a constitutive T7 promoter (see section 2.9). The resulting proteins had an N-terminal HA-Tag (1kDa) that would be used to detect them (denoted as TV1-HA and TV2-HA). These plasmids were transfected into SW480 and SW620 CRC cells, 24 hours after seeding, followed by extraction of subcellular fractionation lysates. A Western blot was performed on these samples and the membranes probed with several commercially available antibodies raised against human Brachyury as well as a HATag antibody to detect recombinant proteins (Figure 3.7 and Figure 3.8).

The mouse monoclonal anti-Brachyury antibody (AB140661) detects a band corresponding to TV1 (50kDa) in pCMV-TV1-HA transfected cells (Figure 3.7 and Figure 3.8 A). There is however, a weak band observed in the whole cell lysate for pCMV-HA and pCMV-TV2HA transfected cells (Figure 3.7A) for this antibody that is the same size as TV1, this is likely to be endogenous TV1. AB140661 fails to detect the band corresponding to TV1 in the UT SW480 and SW620 whole lysate samples, this could be due to the antibody's stronger affinity for the much higher quantity of TV1 in the TV1HA transfected samples in which a band is present. AB140661 was raised from a peptide that spans exon 7 (that is specific to TV1) (Figure 3.3). Endogenous levels of Brachyury were not detectable using SC-20109 (Figure 3.7 and Figure 3.8 B) even at high exposures. It is possible that the overexpressed lysate absorbs most of the antibody. The lower band corresponding to TV2 was much stronger for this antibody although in Figure 3.7 B a band corresponding to TV1 is present in the nuclear fraction of SW480 (Figure 3.7B) that is not present in SW620

(Figure 3.8B) that was raised from a peptide from a region spanning from exons 6-9. It is possible that the antibody detects an epitope that is specific to the exon 6-8 boundary that detects TV2 more readily. In Figure 3.7 and Figure 3.8 D, AF2085 (Goat polyclonal anti-Brachyury) detects both transcript variants, which is consistent with Figure 3.5 controls (UT and NI). It must also be noted that the nuclear fractionation control Lamin-B was only detectable in the nuclear fraction when it would be expected in the whole lysate fraction also.

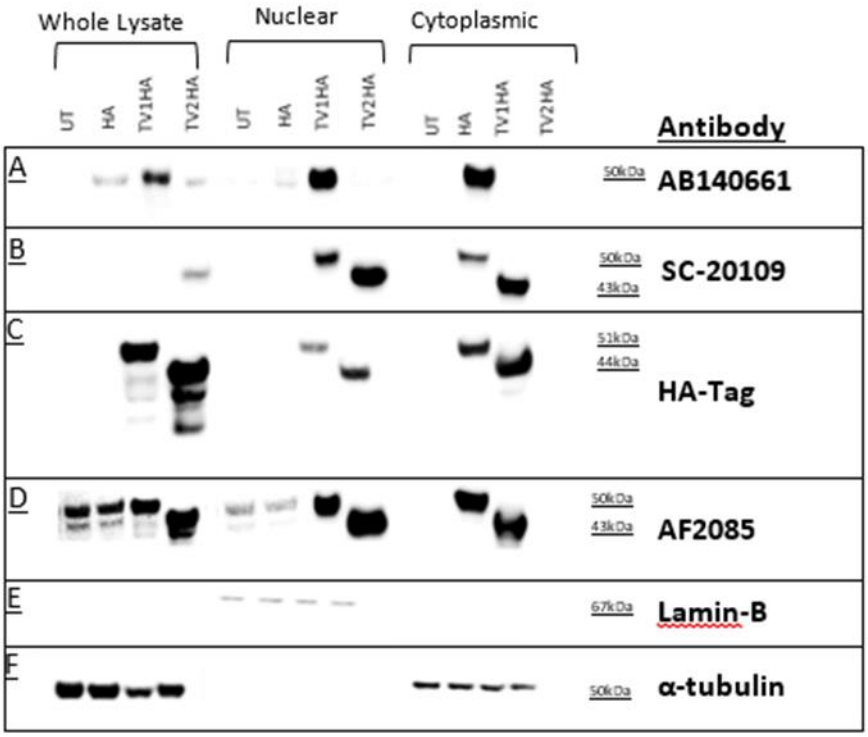

Figure 3.7 Western blot on pCMVTV1-HA and pCMVTV2-HA transfected SW480 cells. SW480 cells were transfected with plasmids containing pCMVTV1-HA and pCMVTV2-HA under a constitutively active T7 promoter. Untreated and vector only samples (UT and HA) were included to show endogenous protein. (A) AB140661 predominantly detects a band corresponding to TV1 (50kDa). (B) SC-20109 detects a lower band corresponding to the same size as TV2 (43kDa) but also detects the upper TV1 band in the nuclear pCMVTV1-HA sample (C) HA Tag antibody detecting the HA tag of pCMVTV1-HA and pCMVTV2-HA corresponding to slightly larger sizes of 51kDa and 44kDa respectively. (D) AF2085 detects both splice variants (E) Lamin-B and (F) α-tubulin were used as nuclear and cytoplasmic markers respectively. **Protein band sizes were determined automatically using the marker and band size function in the ImageLab software package.**

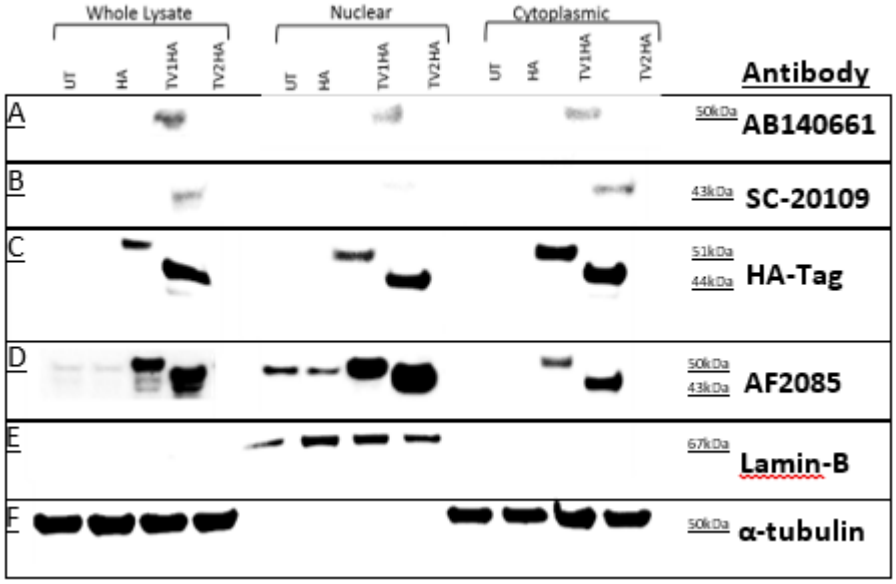

Figure 3.8 Western blot on pCMVTV1-HA and pCMVTV2-HA transfected SW620 cells. SW620 cells were transfected with plasmids containing pCMVTV1-HA and pCMVTV2-HA under a constitutively active T7 promoter. Untreated and vector only samples (UT and HA) were included to show endogenous protein. (A) AB140661 predominantly detects a band corresponding to TV1 (50kDa). (B) SC-20109 detects a lower band corresponding to the same size as TV2 (43kDa) but also detects the upper TV1 band in the nuclear pCMVTV1-HA sample (C) HA Tag antibody detecting the HA tag of pCMVTV1-HA and pCMVTV2-HA corresponding to slightly larger sizes of 51kDa and 44kDa respectively. (D) AF2085 detects both splice variants (E) Lamin-B and (F) α-tubulin were used as nuclear and cytoplasmic markers respectively.

3.4 TV1 and TV2 expression in matched normal and colorectal cancer tissue

To further investigate the two transcript variants expression in colorectal cancer, we used RNA from colorectal cancer tissue (orange bars) and the adjacent normal tissue (green bars) from these patients to assess the expression of the two splice variants (n=21) (Figure 3.9). Heterogeneity was observed in both the normal adjacent and cancer tissue. Expression of both transcripts were observed in the normal and cancer tissue samples with the normal tissue exhibiting a higher expression of Brachyury in many cases. Neither TV1 or TV2 seem cancer specific and restricted to cancer tissue. These results suggest that there is not a simple correlation between the expression of Brachyury splice variants in cancer tissue and normal adjacent tissue. As well as the 21 matched normal adjacent/cancer samples from 21 patients. 1 sample prepared from an adenomatous polyp was prepared (29P). This sample showed a similar level of expression of TV1 as many of the cancer samples.

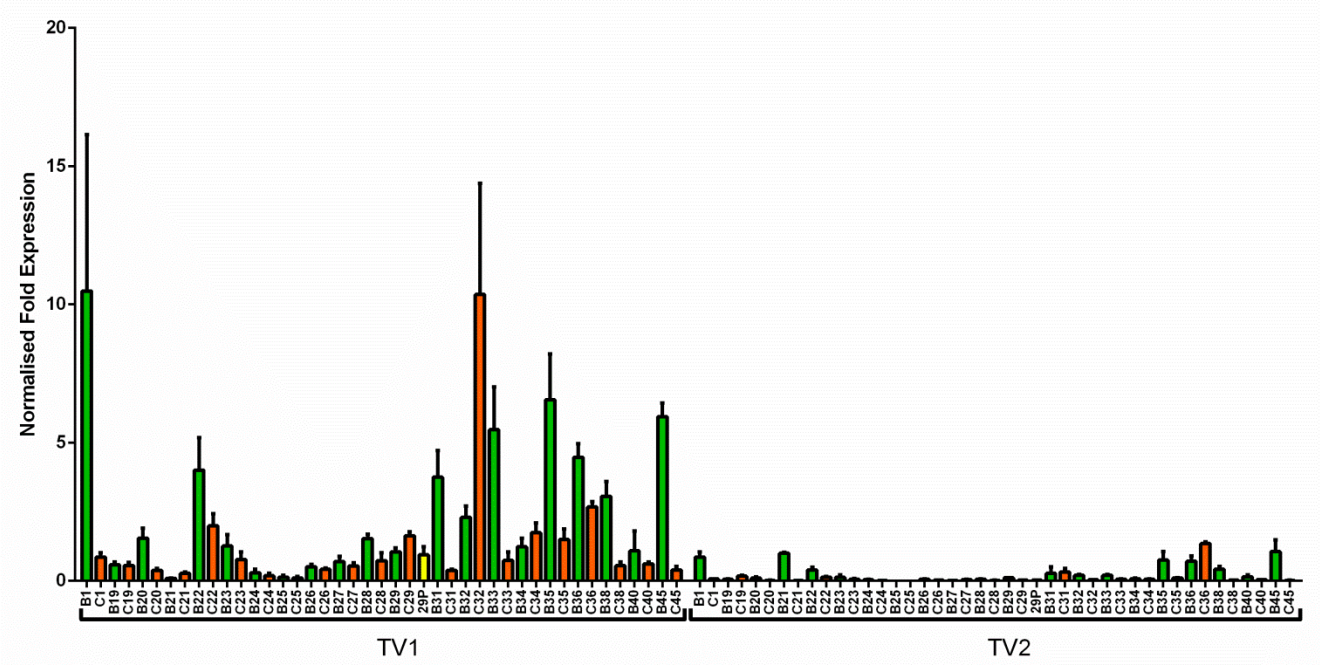

Figure 3.9 Normalised fold expression of TV1 and TV2 in matched cancer and normal samples. qRT-PCR was carried out on matched cancer samples normalised to the ubiquitously expressed transcript *HSP09A*. Variable expression of both transcripts was observed although higher levels of TV1 transcript were evident across normal (green), cancer tissue (orange) and adenomatous polyp (yellow).

3.5 Immunofluorescence analysis of colorectal cancer cells transfected with pCMVTV1-HA and pCMVTV2-HA to assess the specificity of the AB140661 antibody to TV1.

AB140661 is a monoclonal antibody that we found to be suitable for immunohistochemistry analysis, so we wished to validate its specificity in colorectal cancer cell lines using immunofluorescence. Western blotting analysis with this antibody suggests that it detected TV1 (Figure 3.7 AFigure 3.8 A). To further investigate its specificity, we used colorectal cancer cells that had been transfected with pCMVTV1-HA and separately pCMVTV2-HA to see, when co-stained with HA antibody and AB140661, whether similar localisation patterns were observed. SW480 and SW620 cells were used as these were previously used for the subcellular fractionation western blot (Figure 3.7 and Figure 3.8 A). Also, NTERA2, an embryonal carcinoma cell line that is Brachyury negative, (confirmed by qRT-PCR with Ct values of consistently greater than 38, Figure 7.3) was used as a negative endogenous Brachyury control.

siRNA-mediated knockdown of Brachyury was first carried out in H460 cells (a Brachyury-high lung cancer cell line, cT values consistently less than 22, Figure 7.2) to further assess the specificity of AB140661 by Western blotting (Figure 3.10). A band corresponding to the same size as TV1 (50kDa) is depleted by the siRNAs T5 and T6. A faint lower band is also observed in H460 cell lysate (43kDa) corresponding to TV2 and is detected with lower affinity than the upper 50kDa band corresponding to TV1.

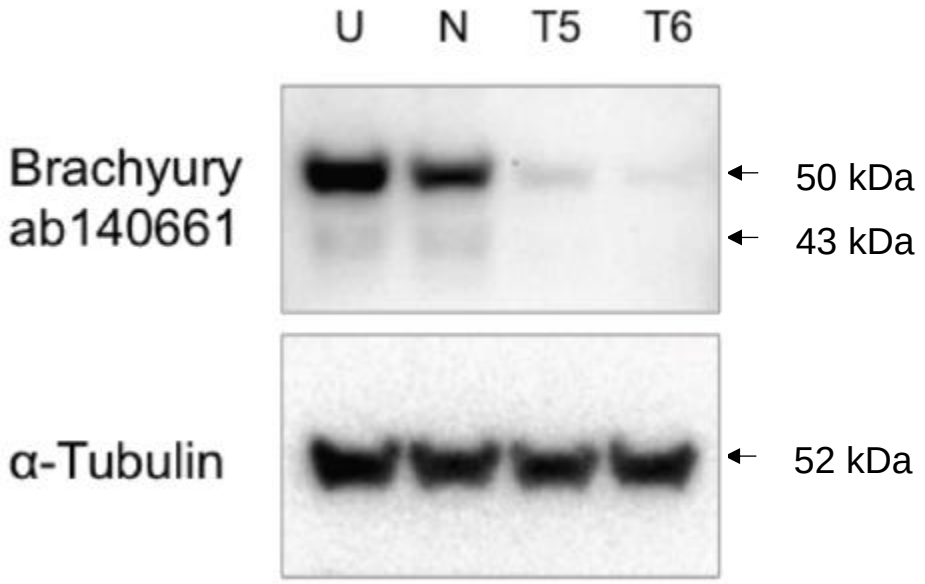

Figure 3.10 anti-Brachyury antibody (AB140661) validation. siRNA depletion of Brachyury and Western blot using the mouse monoclonal antibody from Abcam (AB140661). U–untreated H460 cells, N–negative control siRNA treated H460 cells, T5–Brachyury siRNA (Hs_T_5) treated H460 cells, T6–Brachyury siRNA (Hs_T_6) treated H460 cells, loading was controlled by α-Tubulin.

SW480 cells were transfected with the pCMVHA control vector to show endogenous Brachyury expression which is seen in both the nucleus and cytoplasm (Figure 3.11) The HA-Tag vector is not expressed and this therefore acts as a transfection-only control. The cytoplasmic fraction of the Western blot did not show a band for the untreated or HA vector only sample; it is possible that the antibody gets used up by pCMVTV1HA lysate in the lane immediately to the right of it (Figure 3.7 A and Figure 3.8 A). Overexposures of this membrane did not show a strong band in the cytoplasmic fraction (Figure 3.7 A and Figure 3.8 A).

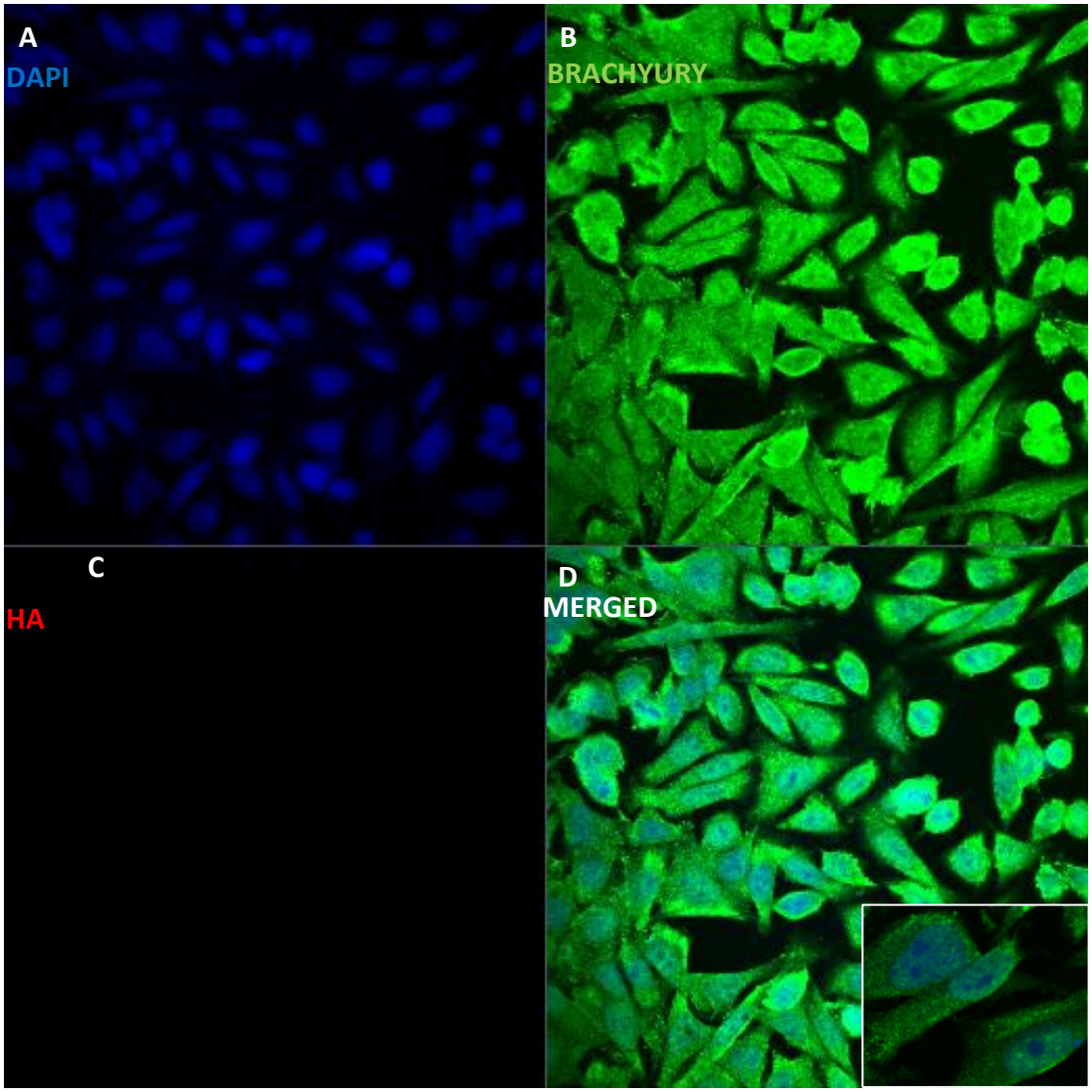

Figure 3.11 Immunofluorescence of SW480 cells transfected with pCMV-HA (40x). SW480 cells were transfected with the pCMV-HA vector only. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) red channel is the HA tag antibody. (D) Merged image(63x). Images acquired on Zeiss 710 confocal microscope. Zoomed inset shows representative cell showing only endogenouscytoplasmic Brachyury labelled with AB140661.

Overlap of the signal of the AB140661 antibody (green) and HA-Tag antibody (red) was consistently observed in SW480 cells transfected with pCMVTV1-HA (Figure 3.12). Two cells in this representative image are transfected with pCMVTV1-HA (Figure 3.12). These cells were easily identifiable as they were high in both the red and green channels (Figure 3.12 D). The signalsof these antibodies in the merged channel image seems to be concentrated around the nucleus of the cell.

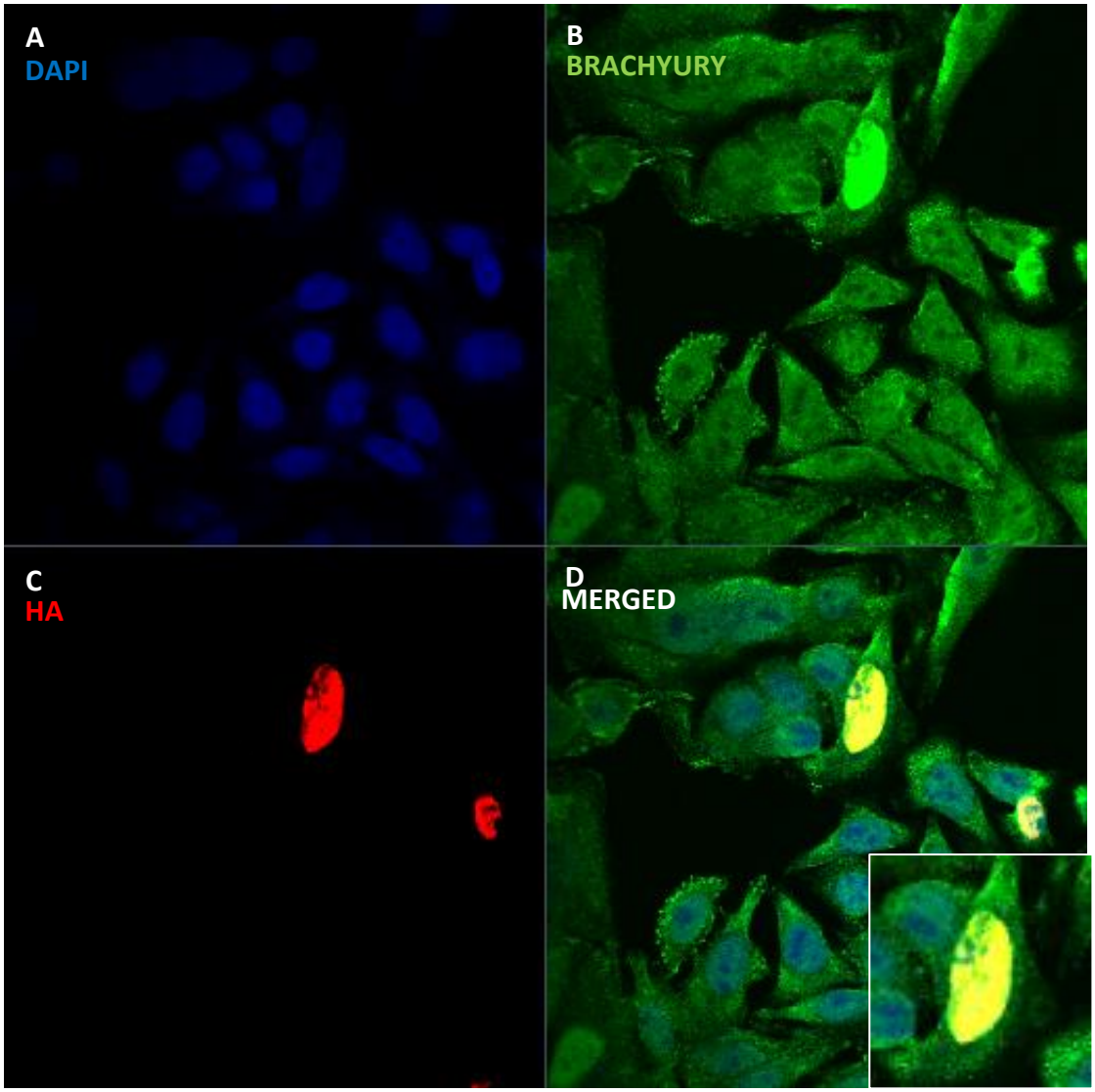

Figure 3.12 Immunofluorescence of SW480 cells transfected with pCMV-TV1HA (63x). SW480 cells were transfected with the pCMV-TV1HA overexpression vector. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) red channel is the HA tag antibody. (D) In the merged channel in the bottom right image overlap of the HAtag antibody and AB140661 (yellow) was observed. Images acquired on Zeiss 710 confocal microscope. Zoomed inset shows representative pCMV-TV1HA transfected cell with significant perinuclear overlap in the green and red channels.

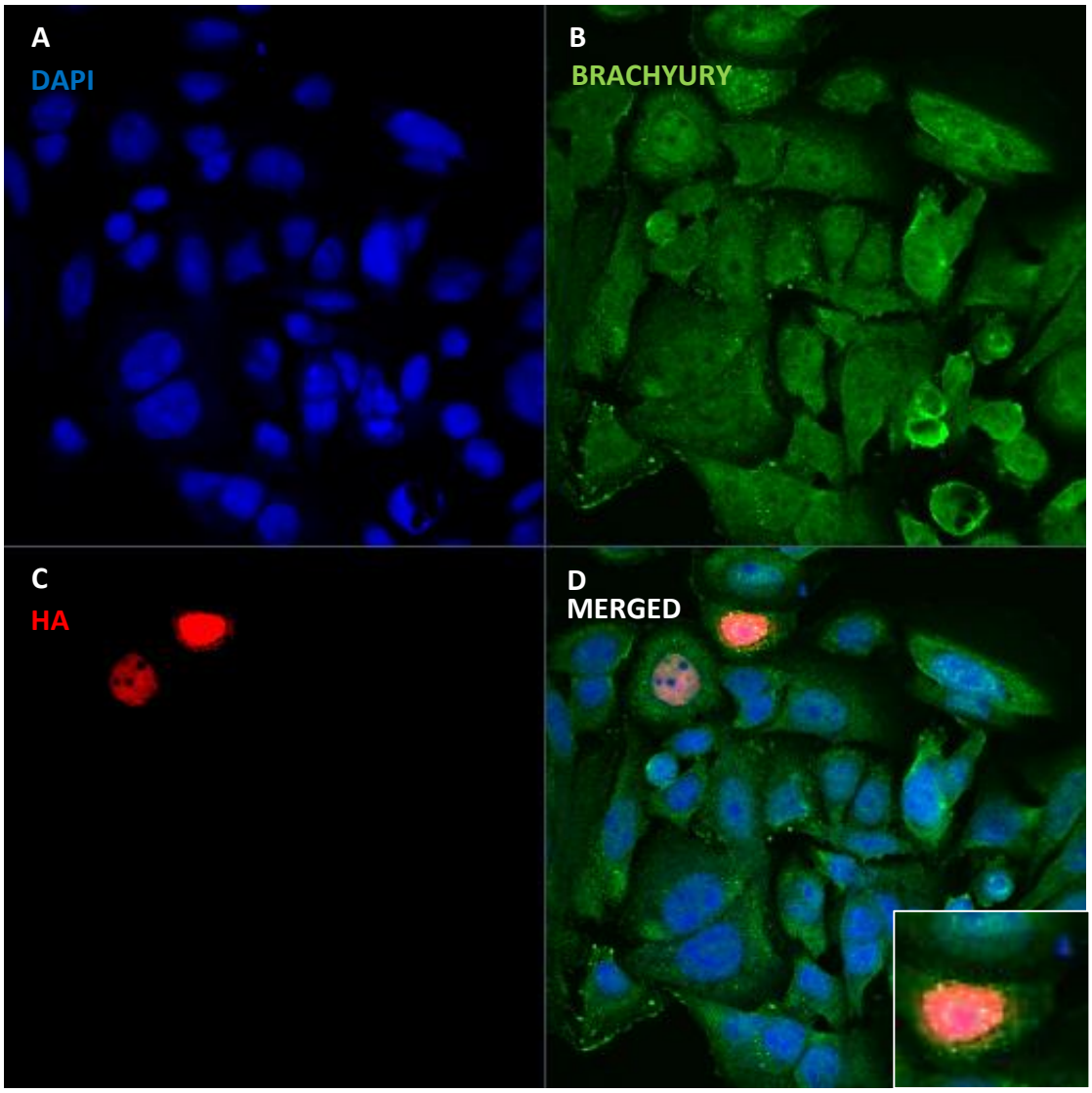

Figure 3.13 Immunofluorescence of SW480 cells transfected with pCMVTV2-HA (63x). SW480 cells were transfected with the pCMVTV2-HA overexpression vector. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) red channel is the HA tag antibody. (D) In the combined image the red colour of the HA tag antibody identifying the overexpressed TV2 HA in two transfected cells is much more visible amongst the green colour of AB140661 than cells transfected with pCMVTV1-HA in Figure 3.12. Inset in image in D highlights a representative pCMVTV2-HA transfected cell. Images acquired on Zeiss 710 confocal microscope.

When TV2 was overexpressed in SW480 cells, less overlap was observed when HA and AB140661 antibodies were co-stained (Figure 3.13 D). Cells transfected with pCMVTV2-HA display yellow punctate areas of overlap predominantly in the nucleus but also show perinuclear overlap. These cells are different to pCMVTV1-HA transfected SW480 cells that appear more diffusely yellow (Figure 3.12 D). This result is consistent with AB140661 detecting TV1 with a much greater affinity than TV2 or, alternatively, TV1 being much more abundant than TV2 at the protein level.

In order to assess difference in the localisation of the transcript variants in a model *in vitro* system for metastatic cancer, SW620 cells were first transfected with the pCMV-HA vector only in order to establish any differences in the subcellular localisation of endogenous Brachyury in this metastatic cell line. Although the SW620 cells have a different morphology and are less confluent, the endogenous

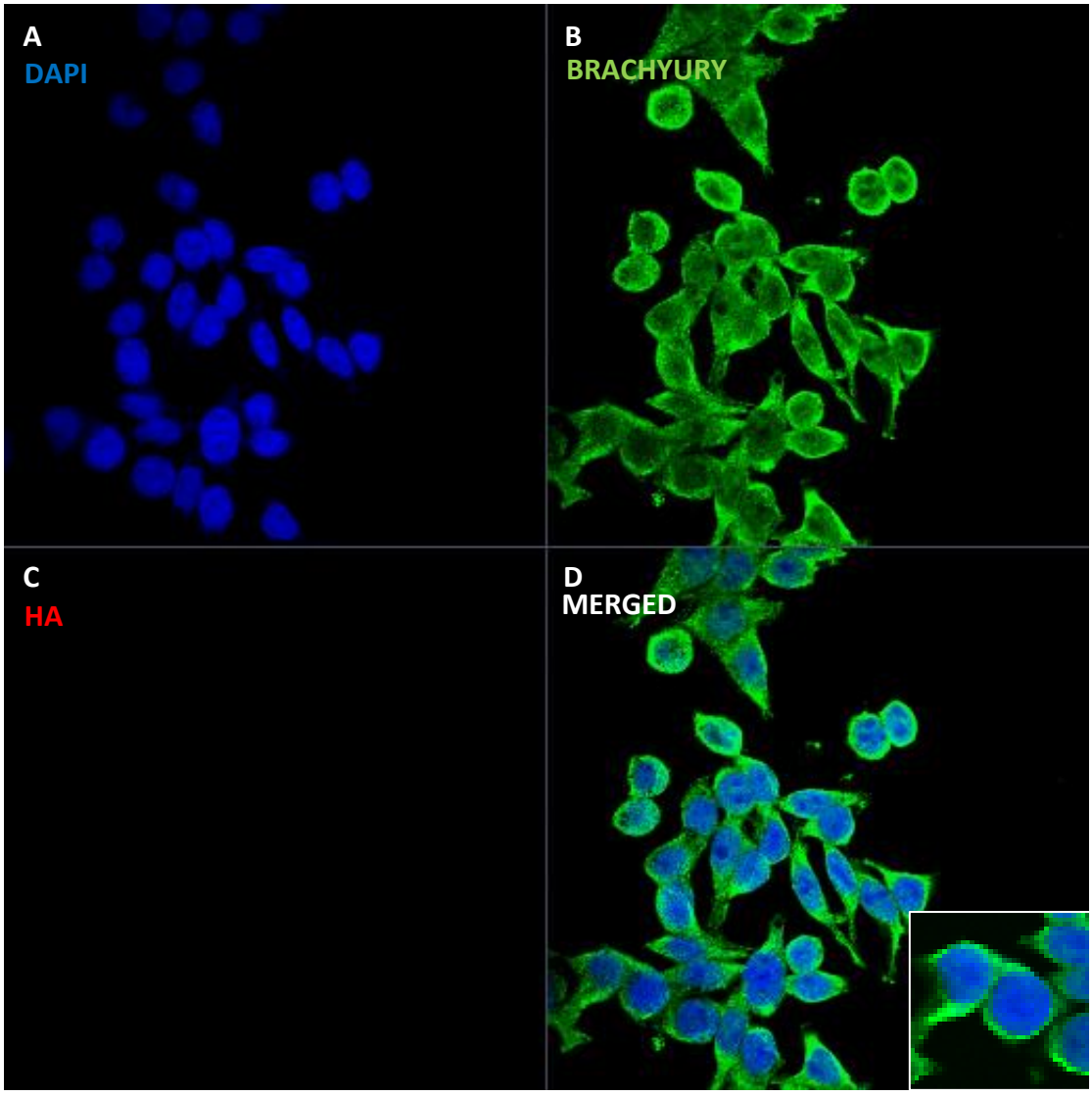

Figure 3.14 Immunofluorescence of SW620 cells transfected with pCMV-HA (63x). SW620 cells were transfected with the pCMV-HA vector only. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 representing endogenous Brachyury (cytoplasmic) and (C) is the red channel of the HA tag antibody. (D) Merged channels image with zoomed inset. Images acquired on Zeiss 710 confocal microscope. **Zoomed inset shows representative cell showing only endogenous cytoplasmic Brachyury labelled with AB140661.**

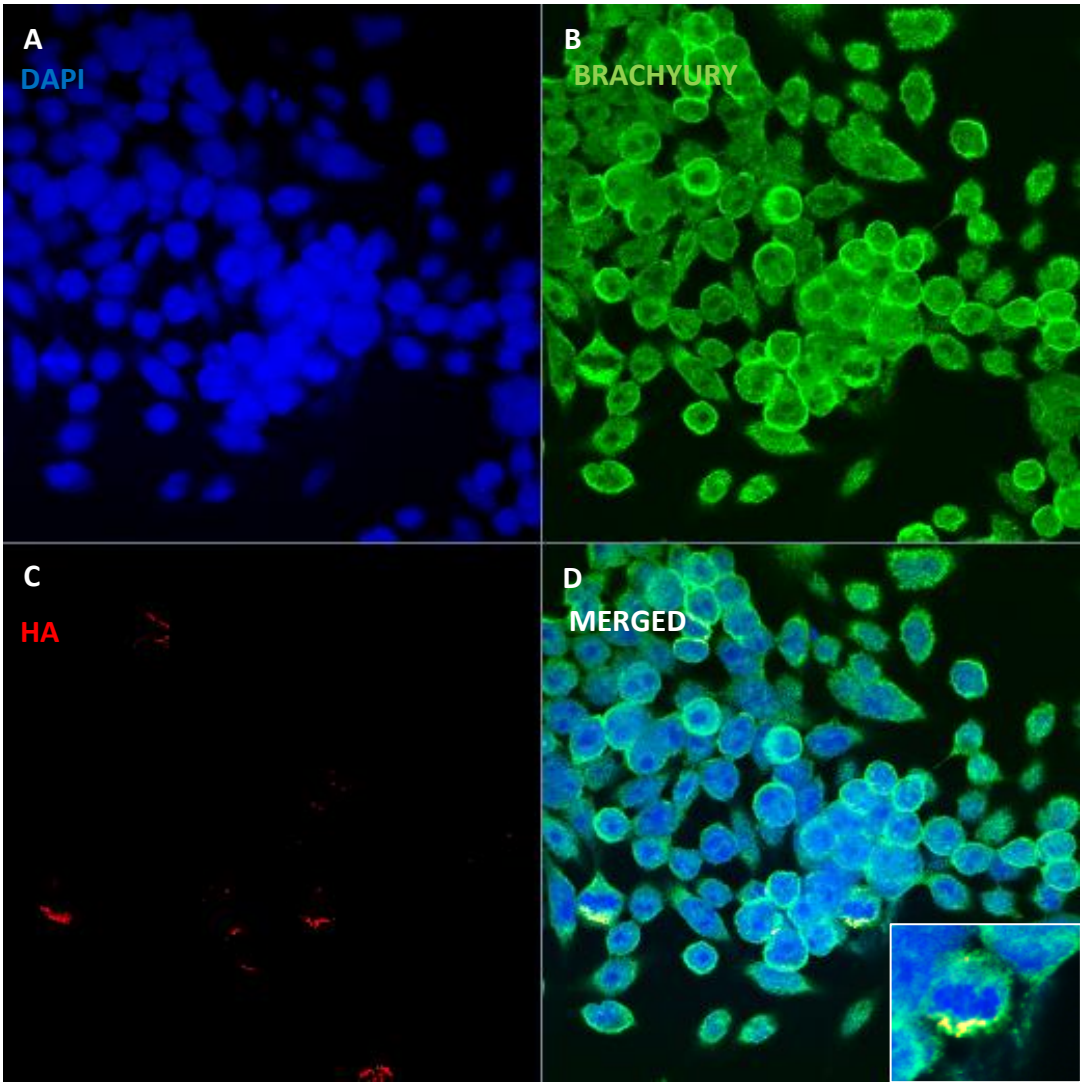

Figure 3.15 Immunofluorescence of SW620 cells transfected with pCMV-TV1HA (63x). SW620 cells were transfected with the pCMV-TV1HA overexpression vector. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 (cytoplasmic) and (C) is the red channel of the HA tag antibody. (D) Merged channels with the zoomed inset of a representative pCMV-TV1HA transfected cell with significant cytoplasmic overlap in the green and red channels.Images acquired on Zeiss 710 confocal microscope.

Brachyury staining pattern in SW620 is quite different from that of SW480 (nuclear and cytoplasmic) (Figure 3.11), with Brachyury being present mainly in the cytoplasm in SW620 (Figure 3.14).

SW620 cells transfected with overexpressing TV1HA showed yellow co-staining in the cytoplasm. The bottom right merged channel (inset) shows one transfected cell that shows this yellow cytoplasmic overlap of the AB140661 (green) and HATag (red) merged signals (Figure 3.15 D inset).

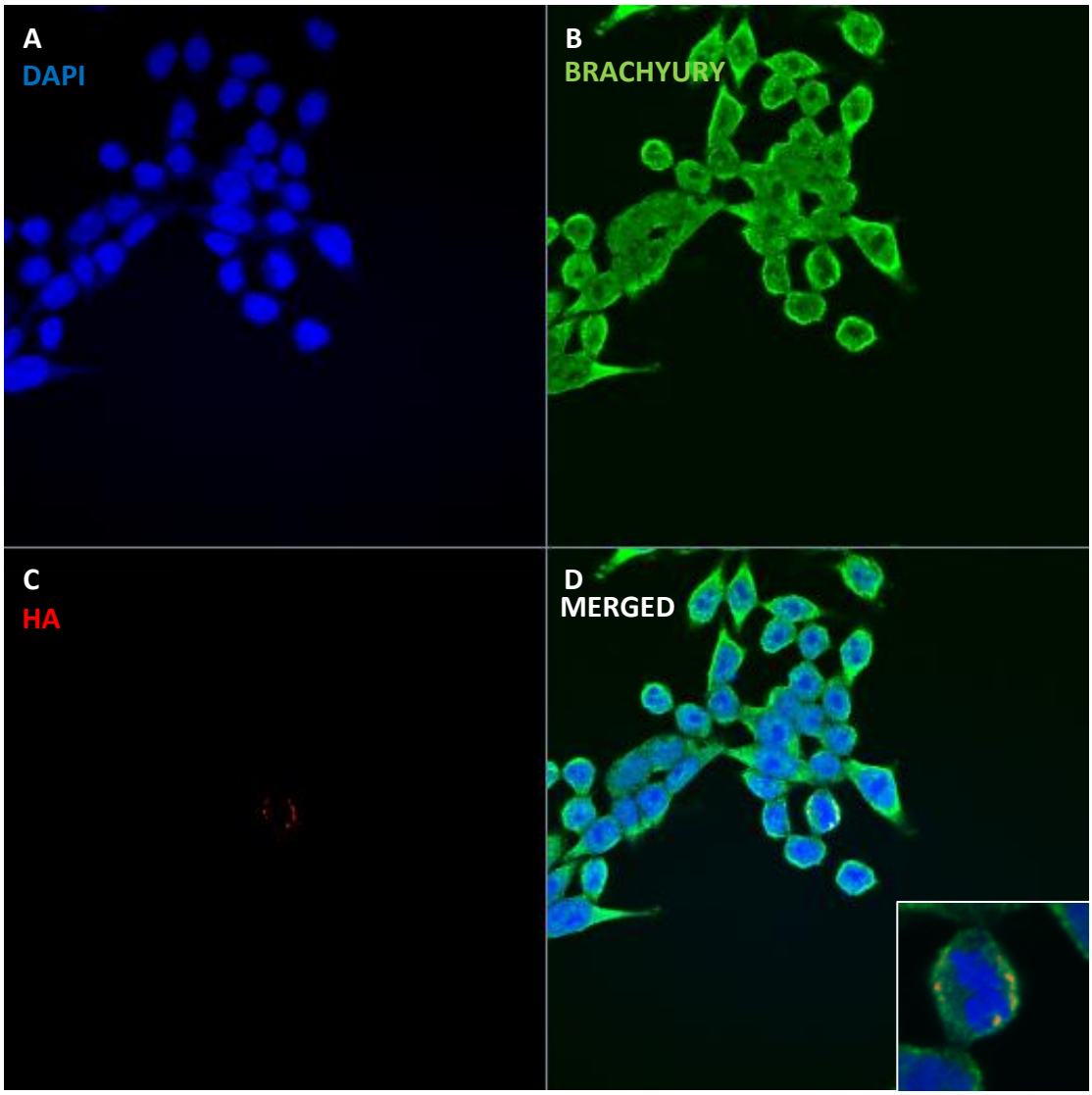

Figure 3.16 Immunofluorescence of SW620 cells transfected with pCMV-TV2HA (63x). SW620 cells were transfected with the pCMV-TV2HA overexpression vector. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) is the red channel of the HA tag antibody. (D) The bottom right panel shows little overlap of the HAtag antibody and AB140661 signals, with three red dots being visible in the pCMVTV2-HA transfected SW620 cell in the bottom right inset combined channel image. Images acquired on Zeiss 710 confocal microscope.

Next, SW620 cells were also transfected with pCMV-TV2HA to establish whether AB140661 detects this transcript variant in SW620 cells. Although the distinction between the pCMVTV1-HA and pCMVTV2-HA transfection of SW620 (Figure 3.15 and Figure 3.16) is less clear than that of SW480 (Figure 3.12 and Figure 3.13), red punctate staining in the pCMVTV2-HA transfected SW620 is more prominent than in the pCMVTV1-HA transfected SW620 cells (bottom right inset Figure 3.15 and Figure 3.16).

In order to remove the background of endogenous Brachyury, we chose a Brachyury negative cell line, NTERA2. They are high in vimentin and we have observed that they exhibit low expression of Brachyury at the RNA level. For this reason, they were chosen as a Brachyury negative cell line so that only the transfected cell would display Brachyury antigens for detection with AB140661.

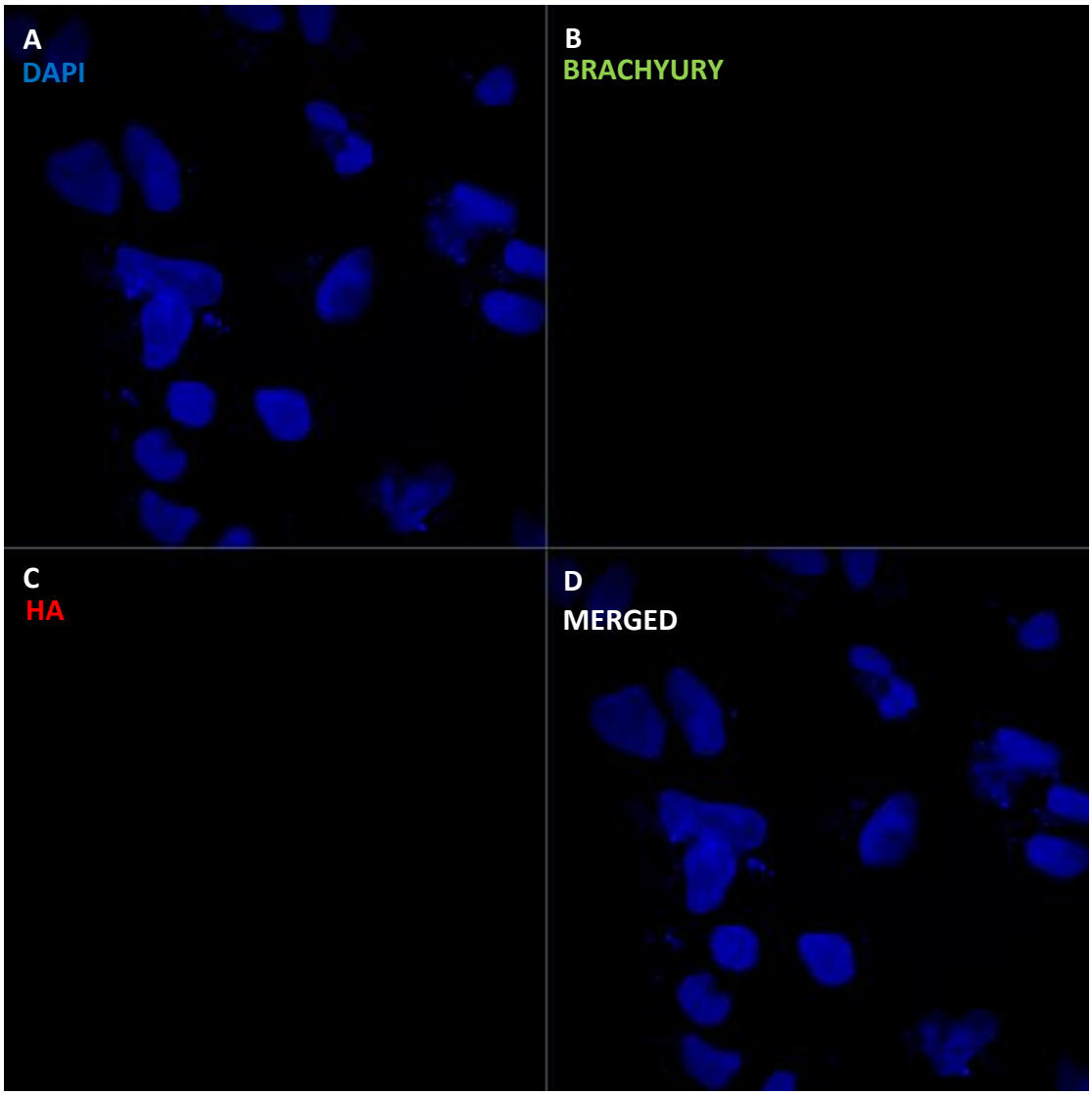

Figure 3.17 Immunofluorescence of NTERA2 cells transfected with pCMV-HA (63x). (A) DAPI (blue), (B) AB140661 (green) and (C) HAtag antibody. (D) Shows the merged channel image. The same gain, offset and laser intensity setting were saved from the previous two cell lines and no endogenous Brachyury was observed. Images acquired on Zeiss 710 confocal microscope.

NTERA2 cells were firstly transfected with the HA vector only and the endogenous levels of Brachyury were assessed. No endogenous Brachyury was observed using the same microscope settings as those used in SW480 and SW620 (Figure 3.17). NTERA2 cells were also transfected with TV1-HA to assess the specificity of AB140661 to this splice variant Figure 3.18. No endogenous Brachyury was observed

and overlap in the transfected cells between the HA tag antibody that identified TV1HA and AB140661 is seen in the combined channel Figure 3.18 D. This result is consistent with AB140661 detecting TV1.

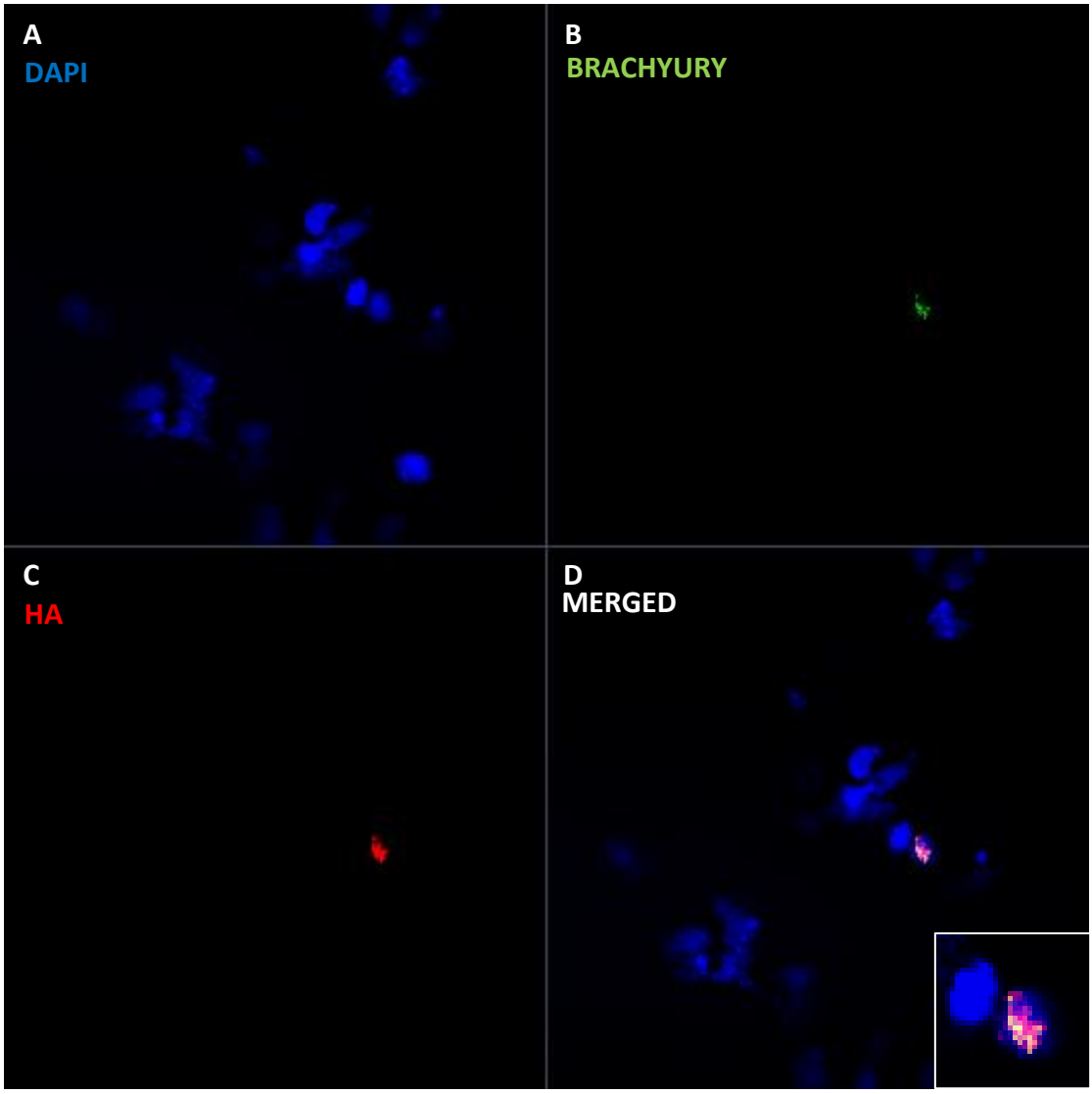

Figure 3.18 Immunofluorescence of NTERA2 cells transfected with pCMV-TV1HA (63x).
(A) DAPI (blue), (B) AB140661 (green) and (C) HAtag antibody. (D) Shows the merged channel image. The same gain, offset and laser intensity setting were saved from the previous two cell lines and no endogenous Brachyury was observed. Images acquired on Zeiss 710 confocal microscope No endogenous Brachyury was observed, but a clear overlap of the HA Tag antibody signal with that of the AB140661 antobody signal is observed in the combined image (D) representative pCMV-TV1HA transfected cell with significant cytoplasmic overlap in the green and red channels is shown in the inset. Images acquired on Zeiss 710 confocal microscope.

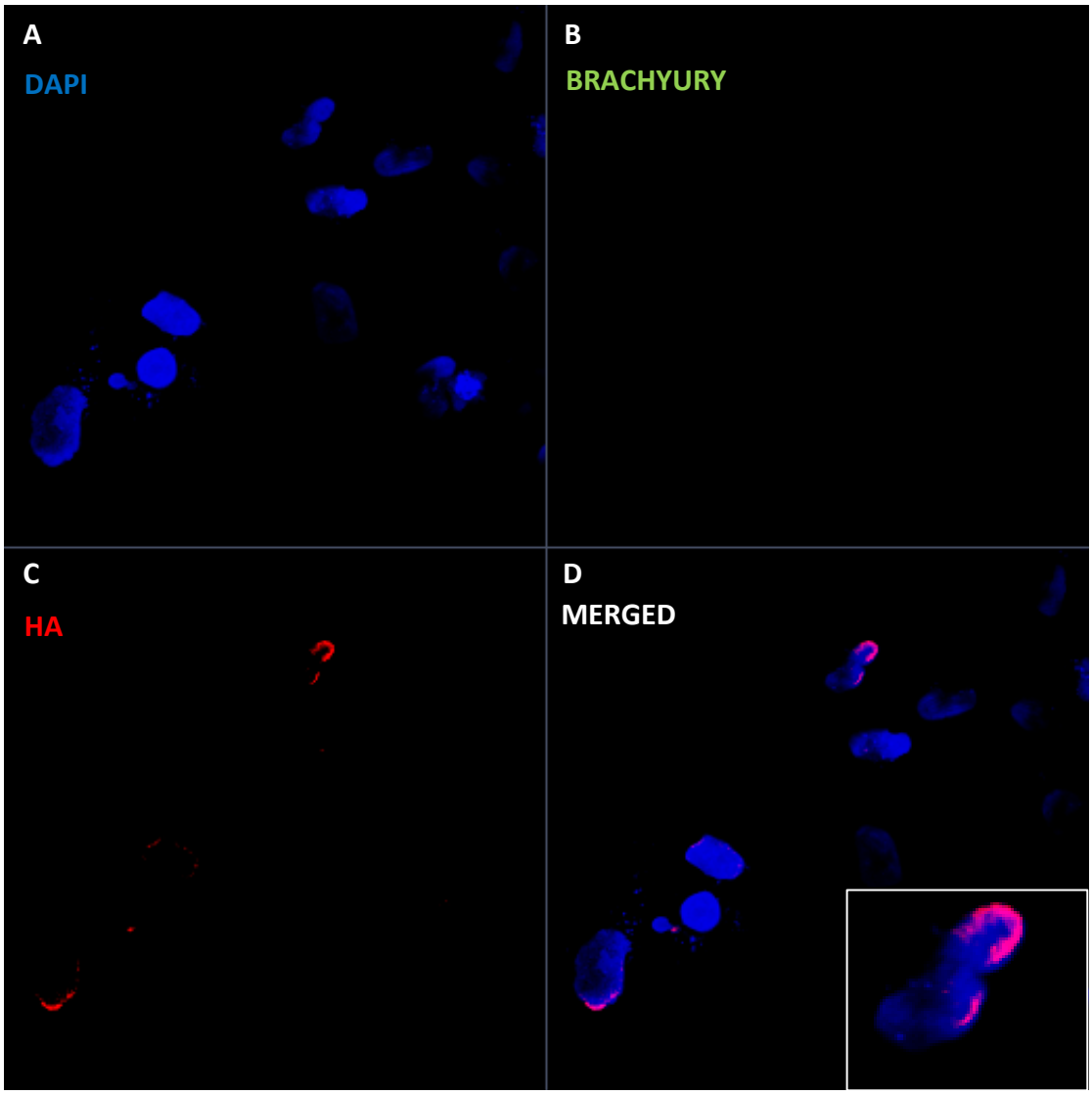

Figure 3.19 Immunofluorescence of NTERA2 cells transfected with pCMV-TV2HA (63x). (A) DAPI (blue), (B) AB140661 (green) and (C) HAtag antibody. (D) shows the merged channel image. The same gain, offset and laser intensity setting were saved from the previous two cell lines and no endogenous Brachyury was observed. Images acquired on Zeiss 710 confocal microscope. Cells transfected with overexpressing TV2HA- pCMV are a clear red colour in the combined image with no green signal from AB140661 observed in the zoomed inset (D (B). Images acquired on Zeiss 710 confocal microscope.

Finally, NTERA2 cells were transfected with TV2HA. No endogenous Brachyury was observed. In these samples, no AB140661 staining and thus no overlap between HAtag and AB140661 antibodies signal was observed in the combined channels. This result suggests that AB140661 is specific to TV1 (Figure 3.19).

Secondary antibody only controls for all three cell lines are shown in Appendix A.

3.6 Discussion

The expression of Brachyury in cancer tissues was reported firstly by Palena and colleagues at the RNA level (Palena, et al., 2007). This group did not however distinguish between the expression of the two transcript variants (TV1 & TV2). The above results demonstrate that two transcript variants are expressed in primary colorectal cancer cells and matched cancer and adjacent normal tissues. In a recent paper from our group, we show that Brachyury expression is restricted to the testis, brain and bone marrow, with Brachyury expression fitting broadly speaking, a cancer testis antigen profile (Jezkova et al., 2016). These are a group of genes that are restricted to the immunoprivileged testis and are also frequently expressed although rarely mutated in cancer tissues, making them ideal therapeutic targets (Caballero & Chen, 2009). Palena and colleagues reported that Brachyury expression was restricted to the testis, however the matched adjacent normal tissue sample qRT-PCR analysis in this chapter and the high expression we observed in the brain and bone marrow reveals that there is sometimes higher expression of Brachyury in the histologically normal adjacent tissue than the cancer tissue. This, together with the expression of two transcript variants that might have distinct functions and expression patterns, makes the picture of Brachyury as a cancer-specific biomarker more complicated. More work is needed to look across all cancers to assess the expression of the two transcript variants and interrogate their function through splice variant-specific knockdown.

Brachyury expression has previously been reported in normal adjacent lung cancer tissue in both the stroma and histologically normal tissue adjacent to the tumour by Western blot and IHC analysis (Roselli et al., 2012). Other normal tissue adjacent to the breast, prostate and oral squamous cell carcinomas were reported as Brachyury positive (Imajyo et al., 2012; Palena et al., 2014; Pinto et al., 2014).

In the matched normal and cancer tissue samples, neither TV1 or TV2 were cancer-specific, although the expression of TV1 was higher compared with TV2. This is consistent with the upper band in Western blots of untreated SW480 and SW620 colorectal cancer cells being more intense than the lower band corresponding to TV1

and TV2 respectively. It is not possible to conclude that either variant is aberrantly switched on in cancer tissue, but certainly in some cancers both are present.

As we wished to assess the expression of Brachyury across many colorectal cancer samples, and knowing that two transcript variants are expressed in primary colorectal cancer cells, it was important that we could be certain what splice variant any antibody that would be suitable to IHC detected. Any conclusions drawn from the IHC analysis could then be attributed to the transcript variant detected by this antibody. AB140661 (mouse monoclonal anti-Brachyury) was validated by siRNA-mediated knockdown and seemed to have a higher affinity to TV1 in both Western blot and IF analyses.

Apparent endogenous Brachyury detected by this antibody revealed that the protein is expressed throughout the cytoplasm of the cell in IF analysis. In order to eliminate the possibility that the endogenous widespread cytoplasmic staining we observed with AB140661 in non-transfected cells was specific, future experiments could use siRNA to see if indeed this cytoplasmic signal is reduced after siRNA transfection with simultaneous qRT-PCR measurement of Brachyury transcript levels. When either transcript variant was overexpressed however, although the endogenous cytoplasmic expression was still visible, more nuclear or perinuclear staining is observed in these pCMVTV1-HA transfected SW480 cells (Figure 3.12), moreover, there is a clear difference between these and pCMVTV1-HA transfected SW620 cells in which the nucleus is clearly negative and TV1 is observed in the cytoplasm. It could be that Brachyury is sequestered in the cytoplasm in untreated SW480 cells, but with an excess (overexpression) there is a limit on the amount of Brachyury that can be sequestered away from the nucleus and therefore it accumulates in and around it. It is quite possible that as well as being sequestered away from the nucleus so that it cannot transcribe its target gene regulatory network, that Brachyury has functions in the mitochondria and/or membrane transport that has been previously reported for other transcription factors (Marinov et al., 2014; Ono et al., 2015; Park & Dolmetsch, 2006). Indeed, RNAseq analysis carried out by our lab revealed that Brachyury influences the expression of the mitochondrial bifunctional enzyme MTHFD2.

As Brachyury is a transcription factor, we would expect to see its expression in and around the nucleus. However, transcription factors have been reported in other subcellular localisations including the cell membrane (Ono et al., 2015) and mitochondria (Park et al., 2006), sometimes being sequestered in the cytoplasm (Rai et al., 2016) or functioning to translocate some target protein to the membrane or regulate the expression of mitochondrial genes respectively (Park et al., 2006).

Although there is no evidence of the cytoplasmic expression of Brachyury during embryogenesis, the work done to determine its expression during development is through RNA-FISH. In cancer biology a lot of evidence suggests that Brachyury, a nuclear transcription factor, is expressed in the cytoplasm.

In a recent study investigating Brachyury's expression in prostate cancer, Brachyury was found to be overexpressed at both the RNA and protein levels (Pinto et al., 2014) and nuclear Brachyury was correlated with invasion. Brachyury was reported as absent in non-neoplastic tissues, displayed a cytoplasmic staining or both nuclear and cytoplasmic staining in primary adenocarcinomas and metastatic lesions, in this study SC-20109 (Figure 3.7 and Figure 3.8 B) was used. Modifications to Brachyury that are subcellular location-specific might mask or enhance their reactivity with certain antibodies also. For example, SC-20109 detects both transcript variants in SW480, a cell line from a primary tumour but only detects TV2 in SW620, a metastatic cell line. This could be as a result of TV1 being modified in such a way in a metastatic context that the epitope usually recognised by this antibody is masked.

In lung cancer Brachyury was reported to be expressed in the nucleus and cytoplasm in several different histopathological classifications (Roselli et al., 2012). In this study ab57480 was used that was raised from an epitope spanning exons 6-8 and could in theory detect both transcript variants.

In chordomas, in which Brachyury is amplified, SC-20109 was also used by Vujovic and colleagues to assay Brachyury expression in 53 chordomas (Vujovic et al., 2006). They report that Brachyury expression is restricted to the nucleus. It would be interesting to fractionate chordoma cells from a primary chordoma and a metastatic lesion to establish whether SC-20109 only detects nuclear TV1 in the primary tumour cell line as is observed above with SW480 and SW620 (Figure 3.7 and Figure 3.8 B). It

is likely that these differences in antibody reactivity due to protein modification, dependent on cellular context are common. Imajyo et al also used SC-20109 and reported nuclear and/or cytoplasmic staining in oral squamous cell carcinoma (Imajyo et al., 2012), they did not differentiate between transcript variants and did not validate the antibody by Western blot siRNA-mediated knockdown like the majority of IHC analyses into Brachyury expression.

Finally Kilic and colleagues used AF2085 in a large scale study to assay the expression of Brachyury in colorectal cancer (Kilic et al., 2011). We have shown that this antibody detects both transcript variants, they do not report this. They report cytoplasmic localisation of Brachyury using this antibody with 90% of colorectal adenocarcinomas being positive for Brachyury.

To summarise, Brachyury has two transcript variants that are expressed throughout colorectal cancer cells, normal tissue and colorectal cancer tissue. Several commercial antibodies detect either both transcript variants or one of them. AB140661 predominantly detects TV1 and is specific to Brachyury with no bands being observed on Brachyury RNA depletion. AB140661 detects overexpressed TV1 by both western blot and IF. This was important to establish before continuing to do a large scale IHC study as cancer-specific transcript variants have been reported that enable drug resistance and are associated with worse prognoses.

4 Brachyury defines a subset of enteroendocrine cells in the normal human gut and adenomas

4.1 Introduction

Normal tissue homeostasis is a finely tuned process that generates the appropriate amount functional tissue in response to environmental cues and insults. It has become increasingly evident, that the epithelial lining of the gut responds to both nutritional cues and damage (Beyaz et al., 2016; Tetteh et al., 2016; van Es et al., 2012). When regulated normally, adult intestinal stem cells generate transient intermediate cells that progressively differentiate into functional daughters that perform a well-defined function, i.e. the absorption of vitamins and water in the colon. These cells have an equally well-defined number of cell cycles before they undergo apoptosis and are then, in the case of the intestine, sloughed off into the lumen. Mutations in the stem cells that generate the functional tissue allow them to expand in number (Schepers et al., 2012; Snippert et al., 2014). The stem cell hierarchy then becomes disrupted with normal adult intestinal stem cells progressively undergoing a transformation by accumulating enabling mutations that allow them to evade detection by the immune system and progress to a hyperproliferative state. As a consequence, the normal crypts characteristic of the normal colon progress to dysmorphic adenoma crypts through to invasive carcinoma and metastases (Barker et al., 2007; Hirsch et al., 2014; Schepers et al., 2012).

Recently it has become evident that the generation and regeneration of the gut epithelium can occur through a process of dedifferentiation when $Lgr5^+$ stem cells are depleted (Tetteh et al., 2016). Committed enterocytes marked by Alpi1 dedifferentiated, populated the $Lgr5^+$ region of the crypt and became proliferative stem cells and Paneth-like cells in this mouse model. In another study, a high-fat diet augmented the numbers and function of $Lgr5^+$ intestinal stem cells of the mammalian intestine and increased their tumourigenicity (Beyaz et al., 2016).

Enteroendocrine cells (EECs) are a differentiated cell type of the secretory lineage present in small numbers at various positions within the intestinal crypts.

Commitment to endocrine differentiation appears to be regulated by the Notch signalling pathway. In precursor and mature endocrine cells, Notch is inactive, this allows the expression of Math1 and neurogenin3. Furthermore, combined EGFR/Wnt/Notch inhibitition in $Lgr5^{+}$ CBCs of the mouse small intestine produced EECs with high purity *in vitro* (Basak et al., 2016). Differentiating precursor cells activate Notch in neighbouring cells to switch off pro-endocrine factors and inhibit endocrine differentiation (Schonhoff et al., 2004). The first factor involved in endocrine specification appears to be Math1, committing cells to one of three secretory lineages: goblet, Paneth and enteroendocrine (Sei et al., 2011). Neurogenin 3 is a consistent downstream target in endocrine cells and appears to be necessary for their differentiation (Sei et al., 2011). Pax4, Pax6,BETA2/ NeuroD, and pancreatic-duodenal homeobox 1 are all transcription factors that have been implicated in enteroendocrine differentiation. (Schonhoff et al., 2004).

The enteroendocrine system dictates how the body responds to the ingestion of nutrients by employing a diverse set of hormones to manifest a wide range of physiological responses inside and outside of the gut. Gut hormones modulate glucose tolerance and food intake (Gribble & Reimann, 2016). The nutritional contents of food modulate the transcriptional profile and enteroendocrine cell differentiation, which modifies the gut hormone response (Moran-Ramos et al., 2012).

Recent evidence in *Drosophila melanogaster* shows that enteroendocrine cells generate a niche that regulates intestinal stem cell division and a high nutrient diet stimulates increased stem cell divisions. Gut hormones such as Tachykinin secreted by the enteroendocrine cells regulate the expression of an insulin-like peptide DILP3 expression in visceral muscle that promotes intestinal growth. In flies that are enteroendocrine cell-less (EEC-less), diet stimulated midgut growth that depends on insulin-like peptide 3 expression is defective, moreover EEC-less flies have reduced ISC division (Amcheslavsky et al., 2014a). In a separate study, Scopelliti and colleagues (2014) report that EECs act as local regulators of ISC proliferation through modulation of the mesenchymal stem cell niche in the *Drosphila* midgut. Paracrine signalling from the EECs constrains ISC proliferation within the epithelial compartment. EECs secrete the neuroendocrine hormone Burscon that binds to its

receptor DLGR2, ortholog of mammalian LGR4-6 and represses the production of visceral muscle-derived EGF-like growth factor Vein.

At least 14 different types of EECs are defined by the gut hormones that they produce and a specific regional distribution. They are identified by markers of the secretion pathways that define the whole population of EECs, i.e. Chromogranin A a large dense core vesicle (LDCV) marker and synaptophysin a synaptic-like microvesicle (SLMV) marker (Rindi et al., 2004). Mice lacking individual enteroendocrine hormones and/or their receptors have shed light on the function of these hormones in regulating energy homeostasis. Mellitzer and co-workers (2010) generated mice that lacked the proendocrine transcription factor *Ngn3*. These *Ngn3* mutant mice were EEC-cell deficient and lacked all gut hormones. These mice had an expanded proliferative crypt compartment and accelerated cell turnover, along with impaired lipid absorption and reduced weight gain resulting in a high frequency of death in the first week of life.

Brachyury has established functions during development that enable morphogenetic movements during gastrulation (Tada et al., 2000; Yamada et al., 2010). Recently Brachyury was identified as a transcriptional activator of essential differentiation pathways and stem cell specification via feedback loop during primitive streak formation (Lolas et al., 2014). Together with its association with several different solid tumour types in human (Kilic et al., 2011; Palena et al., 2014; Palena, et al., 2007; Pinto et al., 2014; Roselli et al., 2012) and its apparent absence in normal adult tissue (Palena, et al., 2007) we aimed to characterise a low level of expression at the RNA level that we observed in the normal human small intestine, colon and rectum through IHC analysis. Here, we report that Brachyury expression is confined to the enteroendocrine cells of the human gut. Expression of Brachyury was observed in normal crypts of the small intestine, colon, rectum and precancerous adenoma lesions.

4.2 Brachyury identifies a class of enteroendocrine cells in normal human intestinal crypts

The anti-brachyury antibody AB140661 was used to immunostain histologically normal colon and small intestine tissue samples (Figure 4.1 A-E, further see section 2.7 for more details of the source of the normal tissue). We demonstrate using this antibody that Brachyury exhibits a strong cytoplasmic subcellular location in the EEC cells in which they are present. Most colonic crypts contain at least one Brachyury positive EEC cell. Most crypts of the colon Figure 4.1 A-C) and small intestine (Figure 4.1 C & D) contain at least one Brachyury positive cell in the previously defined stem cell zone. Some crypts contain Brachyury-positive cells in the transit amplifying zone of the crypt and more rarely in the post-mitotic zone toward the luminal surface. These Brachyury positive cells had a morphology and location typical of EECs, adjacent to the crypt-basement membrane with an abluminal location mainly at the crypt-base. Frequent overlap is observed between CHGA (a marker of differentiated EECs) and Brachyury but not all Brachyury positive cells are CHGA positive (Jezkova et al., 2016). Cells that were positive for CHGA were negative for the nuclear proliferation marker KI67 indicating that these cells are not proliferating although strong Brachyury positive cells cytoplasmic KI67 colocalised with an alternative KI67 antibody (Jezkova et al., 2016).

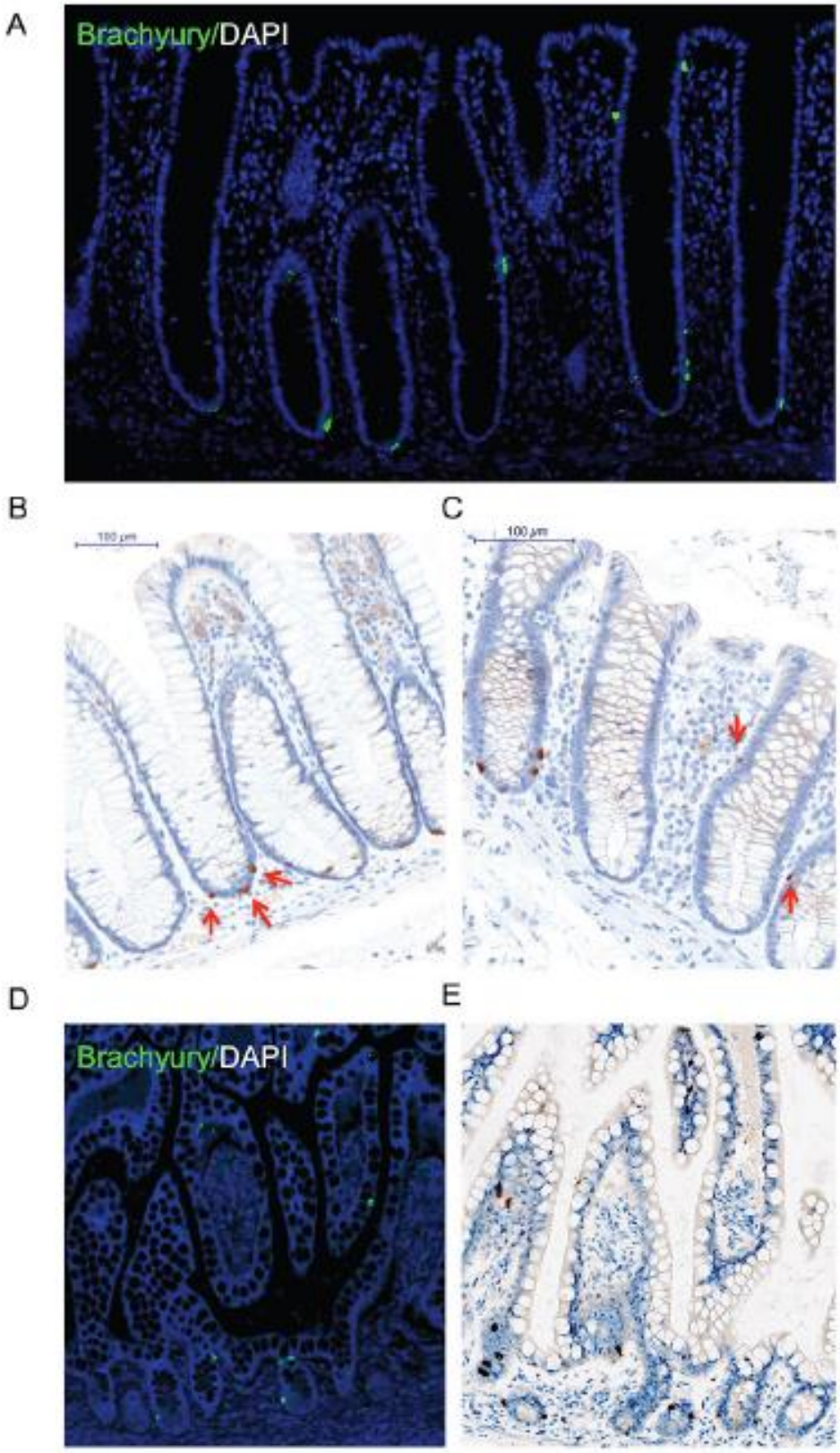

Figure 4.1 Immuno-detection of Brachyury in normal human intestinal crypts. (A) Brachyury-IF on normal right-side colon FFPE sections; DAPI staining, blue; Brachyury staining green/AB140661. (B, C) Brachyury-IHC staining/AB140661 in normal human colon crypts (×20). Red arrows in B illustrate Brachyury-positive cells at crypt-base (consistent with SC zone), some crypts contain single Brachyury-positive cells in the TA zone (indicated by red arrows in (C). (D) Brachyury-IF: normal small intestinal FFPE sections. (E) Brachyury-IHC in normal human small intestine (×20).

4.3 Brachyury is expressed in the crypts of the normal human colon with Brachyury positive cells present in a position consistent with the +4 stem cell.

Expression of Brachyury at the mRNA level in various normal colon tissue samples led us to ask if Brachyury is expressed in the crypts of the normal human gut. Firstly we stained normal human crypts with an antibody against the proliferation marker KI67 to define the region of the crypts that were actively proliferating (Figure 4.2 A).

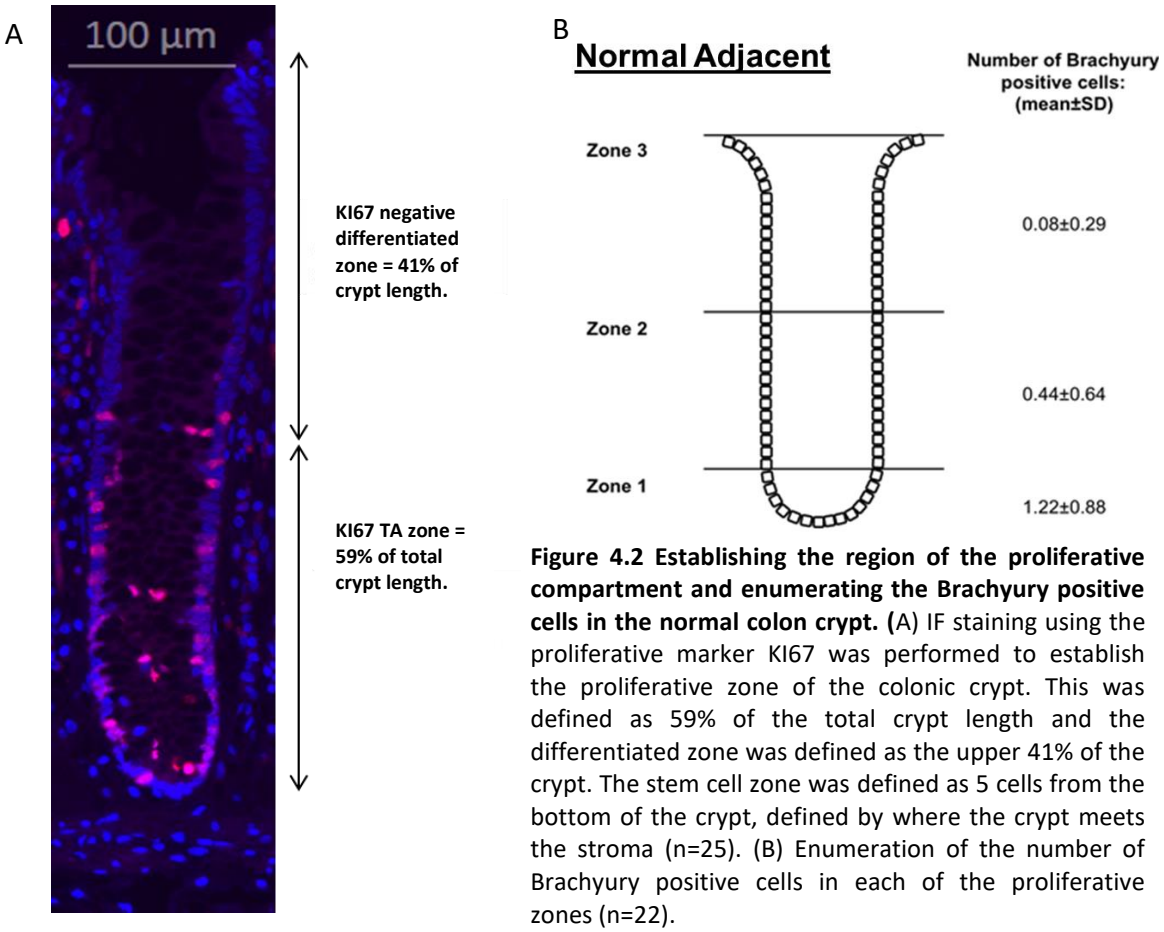

Figure 4.2 Establishing the region of the proliferative compartment and enumerating the Brachyury positive cells in the normal colon crypt. (A) IF staining using the proliferative marker KI67 was performed to establish the proliferative zone of the colonic crypt. This was defined as 59% of the total crypt length and the differentiated zone was defined as the upper 41% of the crypt. The stem cell zone was defined as 5 cells from the bottom of the crypt, defined by where the crypt meets the stroma (n=25). (B) Enumeration of the number of Brachyury positive cells in each of the proliferative zones (n=22).

Significantly higher numbers of Brachyury positive cells were observed in the stem cell zone when compared with either of the other proliferative zones in the adjacent normal tissue of 22 patients and 1930 pathologically normal crypts (Figure 4.2 B).

4.4 An increased frequency of Brachyury positive EECs in rectal crypts.

In order to enumerate the number of Brachyury positive EECs in the colon compared to the rectum (Figure 4.3) the number of positive cells were counted across all available full crypts in histologically normal colon and rectal tissue samples to establish if there were more Brachyury positive EECs in the rectum. This would establish the distribution of Brachyury positive EECs in the human GI tract. Histologically normal full crypts were counted in 22 normal colon samples and 51 rectum samples. A statistically significant higher number of Brachyury positive EEC cells were observed in the rectum (an average of 2.34 cells per crypt n = 49) compared with the colon (an average of 1.22 cells per crypt n = 21) ($P<0.0001$). A no-antibody control was also performed and is shown in Appendix B: Figure 10.4B.

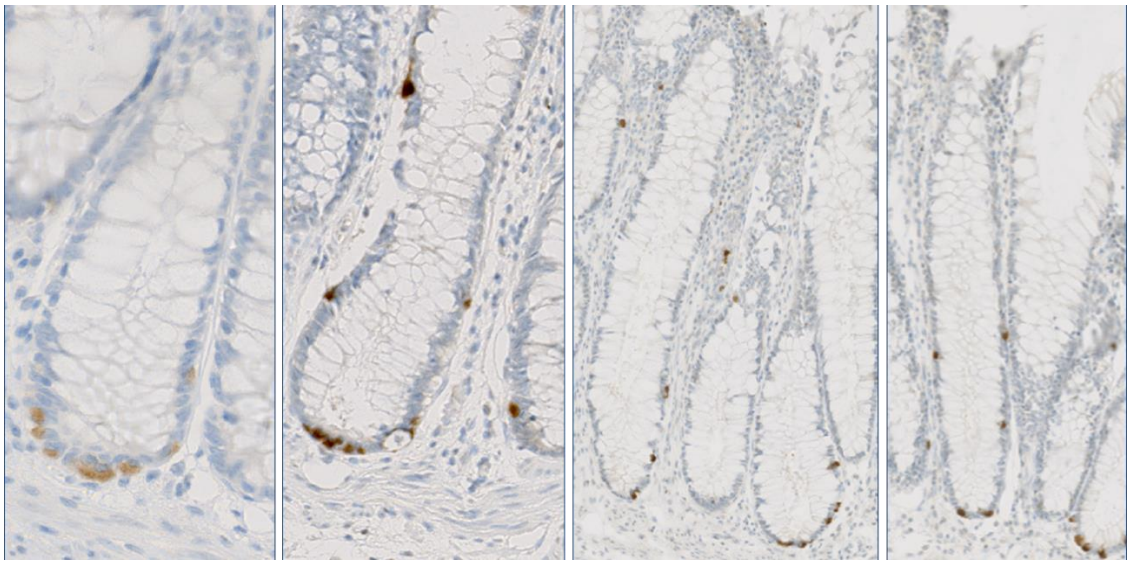

Figure 4.3 Increased frequency of Brachyury positive EECs in histologically normal rectal samples. IHC staining with AB140661 on histologically normal rectal samples. More Brachyury positive EECs were observed in the normal rectal samples studied.

4.5 Brachyury positive EECs are present in colorectal adenomas.

There is a well-established progression from normal colon crypts through a dysmorphic adenoma and finally carcinoma. As we had now characterised Brachyury expression in the normal small intestine, colon and rectum, we stained available adenoma tissue with anti-Brachyury antibody AB140661 to assay the presence of Brachyury in this pre-cancerous lesion.

There are 4 types of histologically distinct colorectal adenomas; serrated, tubular, tubulovillous and hyperplastic polyps. The latter 3 constitute the classical colorectal progression pathway (80-90% of patients) through mutations in *APC/KRAS/p53/CIN* (Figure 4.5 A-C). The former constitutes the serrated pathway of progression that occurs in 10-20% of patients (Figure 4.4) (Baker et al., 2015). We attempted to assay for Brachyury positive EECs in each of these different type of adenoma as a preliminary study before attaining more adenoma tissue samples.

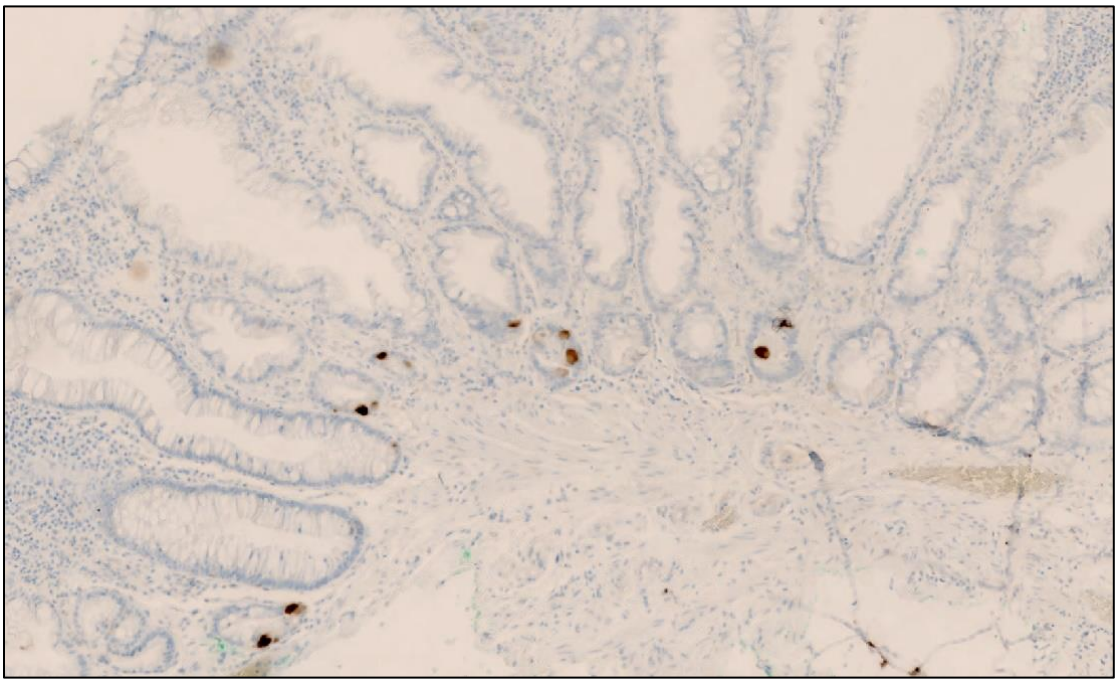

Figure 4.4 Brachyury positive EECs expression in precancerous serrated adenoma. Brachyury expression is observed in the EEC cells of this serrated adenoma. Positive EECs are observed at the abluminally near the epithelial stroma border.

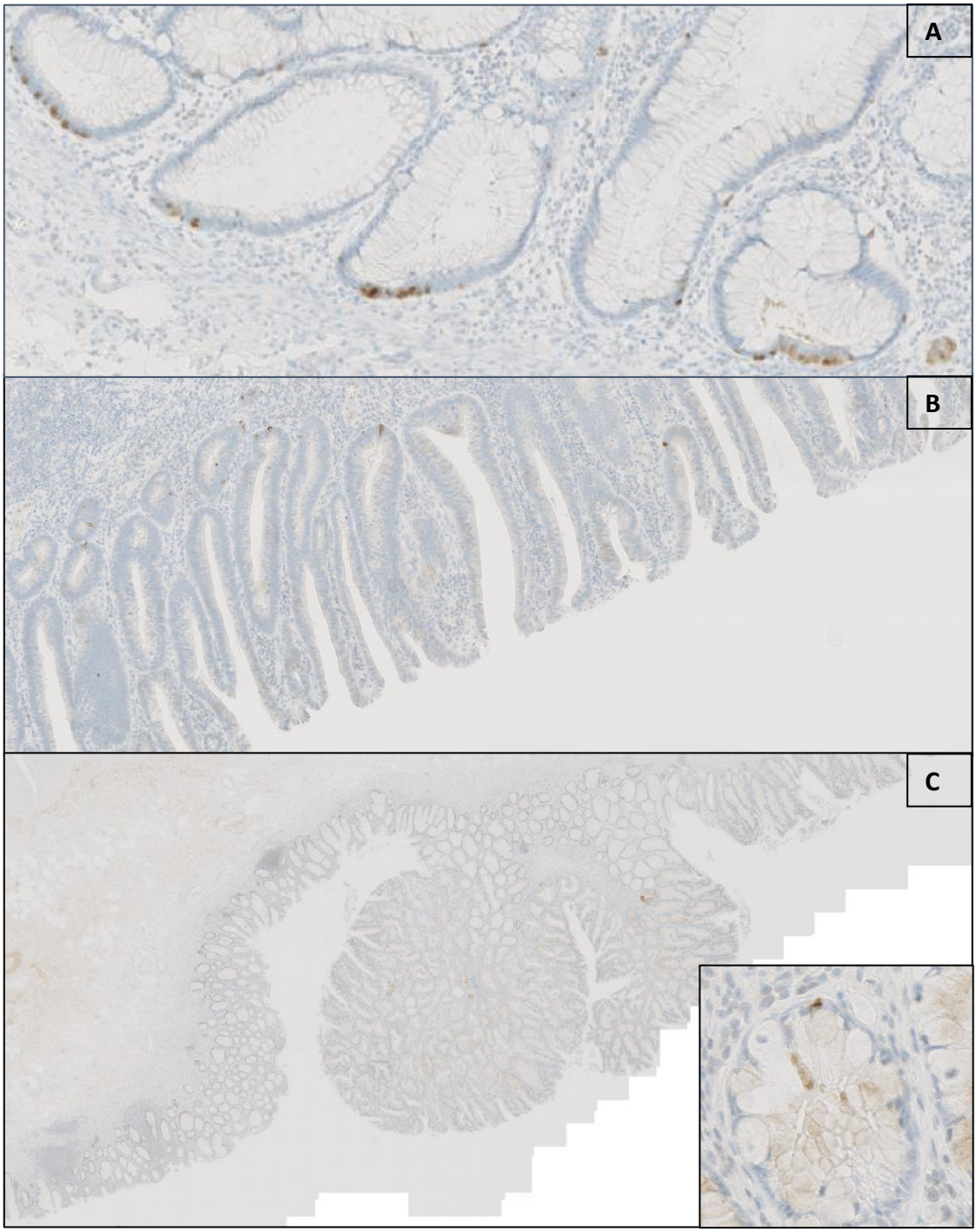

Figure 4.5 Brachyury expression in adenomas of the classical colorectal progression pathway. Brachyury is expressed in tubular adenomas (A), tubulovillous adenomas (B) and a hyperplastic polyp with 40x inset (C).

Brachyury positive EECs resembled those in the normal colon and small intestine epithelium. Cells were observed in all regions of adenoma crypts. Interestingly, in some samples many more EECs were observed at the epithelial-stromal border (Figure 4.5 A) in tubular adenomas than were observed in the normal colon crypts.

Furthermore, Brachyury expression was observed in tubullovillous adenomas (Figure 4.5 B) and hyperplastic polyps (Figure 4.5 C).

4.6 Discussion

Mounting evidence suggests that EECs play an important role in the maintenance of the adult intestinal stem cell niche. Evidence from studies in mice and *Drosophila* suggest that they are required for normal tissue homeostasis, with stem cell numbers significantly depleted in EEC-less Drosophila (Mellitzer et al., 2010 and Amcheslavsky et al., 2014 respectively). RNAseq analysis of a Brachyury knockdown from our group suggests that Brachyury is involved in SC maintenance (Jezkova et al., 2016). Adult intestinal stem cells have been shown to generate normal tissue from their position at the bottom of the intestinal crypts and precancerous adenoma crypts (Barker et al., 2007; Schepers et al., 2012). The intestinal crypt has recently been shown to be highly plastic; with both differentiated cell types dedifferentiating in response to $Lgr5^+$ stem cell depletion (Tetteh et al., 2016) and the augmentation of $Lgr5^+$ adult intestinal stem cell numbers and function in response to a high fat diet (Beyaz et al., 2016).

EECs are one of three cell types of the secretory lineage of epithelial cells that make up the intestinal crypts. Goblet cells and Paneth cells are the other two cell types. KIT^+ goblet cells are located in close proximity to the $Lgr5^+$stem cells are likely niche components (Barker, 2014) as well as Paneth cells that are essential generators of the stem cell niche in the small intestine (Sato et al., 2011). EECs, as the third secretory lineage are therefore highly likely to be generators of the niche that $Lgr5^+$ stem cells require for their maintenance. EECs serve an essential niche function that regulates stem cell division in Drosophila (Amcheslavsky et al., 2014), with EECs providing an important link between high-nutrient diet stimulation of intestinal stem cell division through the production of gut hormones that regulate the generation the gut epithelium. Given the evolutionary conserved mechanisms of tissue homeostasis in Drosophila and Brachyurys expression in a subset of these EECs, it is interesting to speculate that these EECs are also niche generators that are essential

for stem cell maintenance in response to nutritional cues in the human gut. These nutritional cues have been shown to dictate stem cell number but no increase was observed in the number of $ChgA^+$ EEC or Alcian $blue^+$ Goblet cells in the small intestines of mice that were fed a high fat diet (Beyaz et al., 2016). As likely niche generating cells, one would not necessarily expect an increased number when stem cell numbers and function are augmented. Paneth cells and $Reg4^+$ niche generating cells were maintained in similar numbers and positions in adenomas of the small intestine and colon compared to normal mouse crypts with increased ISC number (Schepers et al., 2012).

The loss of EECs in mice alters the absorption of lipids and glucose homeostasis and impairs postnatal survival. Mice were deficient in EECs and hormones, exhibited impaired lipid absorption, reduced weight gain and improved glucose homeostasis (Mellitzer et al., 2010). Moreover, EEC-less mice exhibited an enlarged proliferative crypt compartment and accelerated cell turnover. Similar to Beyaz (2016) findings, no increase in Paneth or Goblet cells was observed. Furthermore, a disorganisation of the crypt compartment in mutant mice was observed with blunt or club-shaped villi and crypts frequently observed (Mellitzer et al., 2010). As it is becoming clearer that the epithelial lining of the gut is plastic in response to dietary fat content and EEC-less mice have impaired lipid absorption, EECs could modulate stem cell division rate in this nutritional context in humans also. Given Brachyury's expression in several cancers and its emergence as a regulator of the cell cycle (Jezkova et al., 2014), it is important that its function in normal EECs and EEC-like cells in adenomas is elucidated. As Brachyury KO mice are not viable, it would not be possible to perform experiments similar to Mellitzers group, nonetheless it would be interesting to establish whether Brachyury positive EECs modulate an increase in stem cell number in response to increased dietary fat through conditional tissue-specific KO. More importantly, EECs function in modulating stem cell function and any increase in their tumourgenicity through the possible nutritional induction of and secretion of gut hormones and signalling peptides is important to our understanding of normal tissue homeostasis and cancer progression.

To confirm the identity of Brachyury positive cells in the colon and small intestine as EECs, we co-stained the cells with markers of EECs Chromogranin A (CHGA) and

Synaptophysin. The strongly Brachyury positive cells frequently overlapped with CHGA-positive cells (Jezkova et al., 2016). However not all CHGA-positive cells were positive for Brachyury, so it seems that Brachyury-positive cells are a distinct sub-population of EECs. In the small intestine there are more CHGA-positive cells that are negative for Brachyury, this is interesting considering the higher rates of proliferation in the human colon compared with the small intestine (Vogelstein & Tomasetti, 2015)

In terms of the general distribution of EECs in the human gut, enterochromaffin cells that are marked by CHGA are the most abundant EEC subtype of the colon and rectum. These make up over 70% of the EEC population in the proximal large bowel, which falls to around 40% in the rectum (Rindi et al., 2004). This fall in proportion in the rectum coincides with an increased number of the other EEC subtypes. The D cell type are uncommon although scattered evenly throughout the GI tract, they are marked by Synaptophysin which we observed are Brachyury negative (Rindi et al., 2004; Schonhoff et al., 2004). The third cell type is the L cell. These are sometimes marked by ChgA and/or Synaptophysin amongst other markers and increase in number from proximal to distal being most concentrated in the rectum (Gunawardene et al.,, 2011). As Brachyury positive cells are occasionally CHGA positive (Jezkova et al., 2016) and increase in frequency from proximal to distal GI tracts, they are likely to be the L subtype of EEC..

Brachyury's expression in the secretory EECs of histologically distinct adenomas extends current studies into the molecular mechanisms that manifest themselves in the morphological transformation from normal to adenoma to cancer tissue. *LGR5* mRNA expression was recently reported in a small number of cells at the base of normal human colonic crypts, resembling those of mouse models. Conventional adenomas widely express high levels of *LGR5* although there is no evidence that the cellular hierarchy of the normal colonic crypt is maintained in conventional human adenomas. Serrated adenomas however did display basal *LGR5* mRNA suggesting this lesion resembles normal stem cell hierarchy (A.-M. Baker et al., 2015), indeed we observed basal Brachyury positive EECs in serrated adenomas resembling their distribution in normal crypts. Clearly there are differences between these histologically and morphologically distinct adenomas in terms of their stem cell

hierarchy and architecture. Brachyury's expression in the EECs of each adenoma and its known expression in cancer, suggests it might play an important function in providing a niche for adult intestinal stem cells and their adenoma counterparts on their way to cancer progression.

Given Brachyury is a member of Wnt signalling pathway (Martin & Kimelman, 2008) like Lgr5 (Barker et al., 2007) and the pathways function in regulating stemness in the intestine and throughout proliferative tissues in the body, it is possible that Brachyury might play a role in regulating the stemness of these cells in response to damage. In *Drosophila*, EECs secrete the neuroendocrine hormone Burscon that binds to its receptor DLGR2, an ortholog mammalian of LGR4-6 and represses the production of visceral muscle-derived EGF-like growth factor Vein through cAMP activation (Scopelliti et al., 2014). As Vein overexpression in the adjacent visceral muscle is sufficient to drive ISC hyperproliferation, its suppression by an EEC derived hormone is essential for the regulation of the stem cell compartment in *Drosophila*. We observed many Brachyury positive EECs at border between the epithelium and the stromal muscle (visceral muscle equivalent) in adenomas. It is possible that Brachyury positive EECs act in a similar way to activate EGFR in humans and regulate ISCs in number and proliferation status although this does not explain the increased number of Brachyury positive EECs in histologically normal rectal crypts.

The highly plastic nature of the gut has left open the possibility that previously assumed terminally differentiated post-mitotic cells have the ability to dedifferentiate in response to classic $Lgr5^{+}$ stem cell depletion (Tetteh et al., 2016). It is important that Brachyury's function be understood in terms of the transcriptional network it might control in these EECs. It is known that BETA2 induces cell cycle arrest through an increased expression of p21 in EECs and that the Math1-NGN-BETA2 cascade is the effector of EEC differentiation. Single cell transcriptional profiling of EECs in *ex vivo* human organoids will shed light on Brachyurys function in these cells and their part in this signalling cascade in the specific subset of EECs in which they are expressed. As work from our group has shown Brachyury regulates cell cycle progression through the downregulation of the cell cycle inhibitor p27 (Jezkova et al., 2014), it is possible that in certain cellular contexts Brachyury regulates the cell cycle in these EECs.

Given that Brachyury is expressed in normal crypt and adenoma crypt EECs, it is important that these are characterised in order to understand their differentiation status, in order to establish whether they fit the transcriptional profile of an adult intestinal stem cell in adenomas or if they retain their function as EECs, likely modulators of nutritional cues through the generation of a dynamic tempera-spatial niche.

5 Brachyury influences the expression LRIG1, a pan-ErbB downregulator, in colorectal cancer cells

5.1 Introduction

The two theoretical models (formed mainly from evidence from mouse models) of adult intestinal stem cell identity that have competed over the past four decades are the stem cell zone model, that suggests that crypt base columnar cells marked by Lgr5 (Barker et al., 2007) are the resident stem cells of the colonic crypt and the +4 model of stem cell identity. The latter proposes that adult intestinal stem cells reside within a ring of 16 cells immediately above the niche generating Paneth cells in the mouse small intestine and colon (Barker, 2014). It is now accepted that these +4 stem cells are a separate population to $Lgr5^+$ CBCs that are insensitive to injury and show increased stem cell activity upon damage again in these mouse models for both small and large intestine. For this reason, these cells are considered reserve stem cells capable of replacing CBCs in the injury setting (Beumer & Clevers, 2016).

Although the +4 population has been defined by several markers, including Bmi1 (Sangiorgi & Capecchi, 2008), Tert (Breault et al., 2008) and Lrig1 (Powell et al., 2012a), the molecular circuitry controlling this quiescent state is incompletely understood. Powell and colleagues performed lineage mapping experiments in a tamoxifen-inducible Cre recombinase *in vivo* system in mice that was under the control of the translational start site of the endogenous *Lrig1* locus (Powell et al., 2012). Lrig1 marked predominantly non-cycling, long-lived stem cells with a basal disposition, that, upon crypt injury, proliferate and divide to regenerate the damaged crypts, suggesting that these are a separate population to the quiescent *Bmi1* +4 population observed by Yan and co-workers (2012). Furthermore, the transcriptional profile of $Lrig1^+$ colonic stem cells differed markedly from that of the highly proliferative $Lgr5^+$CBC cells. Interestingly, loss of *Apc* in $Lrig1^+$ cells led to intestinal adenomas as did genetic ablation of *Lrig1* (Powell et al., 2012).

The function of Lrig1 in the intestinal epithelium was revealed through the generation of *Lrig1* knockout mice (Wong et al., 2012). ErbB signalling is essential for intestinal epithelium maintenance following injury and tumour formation (Roberts et al., 2002). Given that ErbB-family ligands and receptors are highly expressed within

the stem-cell niche (Sato et al., 2011), Wong and colleagues hypothesised that strong endogenous regulators must control the pathway in the stem-cell compartment. In Lrig1-knockout mice, congenital extensive crypt hyperplasia was observed along with an extended $Lgr5^{+}$ stem cell compartment. This was directly linked to deregulated *Erbb* expression in the stem cell compartment as ErbB inhibitors rescued normal crypt homeostasis in *Lrig1*-knockout mice (Wong et al., 2012).

We had previously observed that a sub-set of post-mitotic (non-proliferating) cells that were Ki67-negative and Chromogranin A positive, observed at the crypt-base of the mouse small intestine (Sei et al., 2011) and that these cells demonstrate SC markers such as Lgr5. Furthermore, the transcription factor Sox9 is present in the SC/progenitor zone of the mouse small intestine where it regulates proliferation and differentiation in a dose-dependent manner; Lgr5-SCs are Sox9-low, but a population of differentiated EECs expressing ChgA shows high levels of Sox9. This latter population is Ki67-negative (therefore post-mitotic) but does have regenerative capacity following radiation damage (Formeister et al., 2009 and Van Landeghem et al., 2012). Recent work from our group demonstrated that Brachyury-positive cells in the crypt are Sox9- negative (or sub-low) so in this respect, are not similar to the Sox9-high EECs seen in mouse small intestinal crypts (Jezkova et al., 2016).

Given that Brachyury marked a subset of EEC cells in the normal colon crypt (Jezkova et al., 2016) and that these cells were localised close to the +4 region, we investigated any co-localisation of Brachyury with several purported +4 markers by IF on FFPE normal colon tissue (see section 2.7 for more detail on source of tissue). Brachyury positive cells that were positive for the post-mitotic EEC marker CHGA, were found to be negative for the nuclear localised proliferation marker KI67 (Jezkova et al., 2016), indicating that these cells are unlikely to be proliferating. We observed Brachyury and LRIG1 co-localised in occasional cells at the crypt base in a region consistent with the +4 region (Figure 5.1). To investigate if any relationship between Brachyury and LRIG1 in colorectal cancer cells, we performed knockdown experiments on SW480 and SW620 colorectal cancer cells to see if the RNA and protein levels of *LRIG1* were influenced by a Brachyury knockdown.

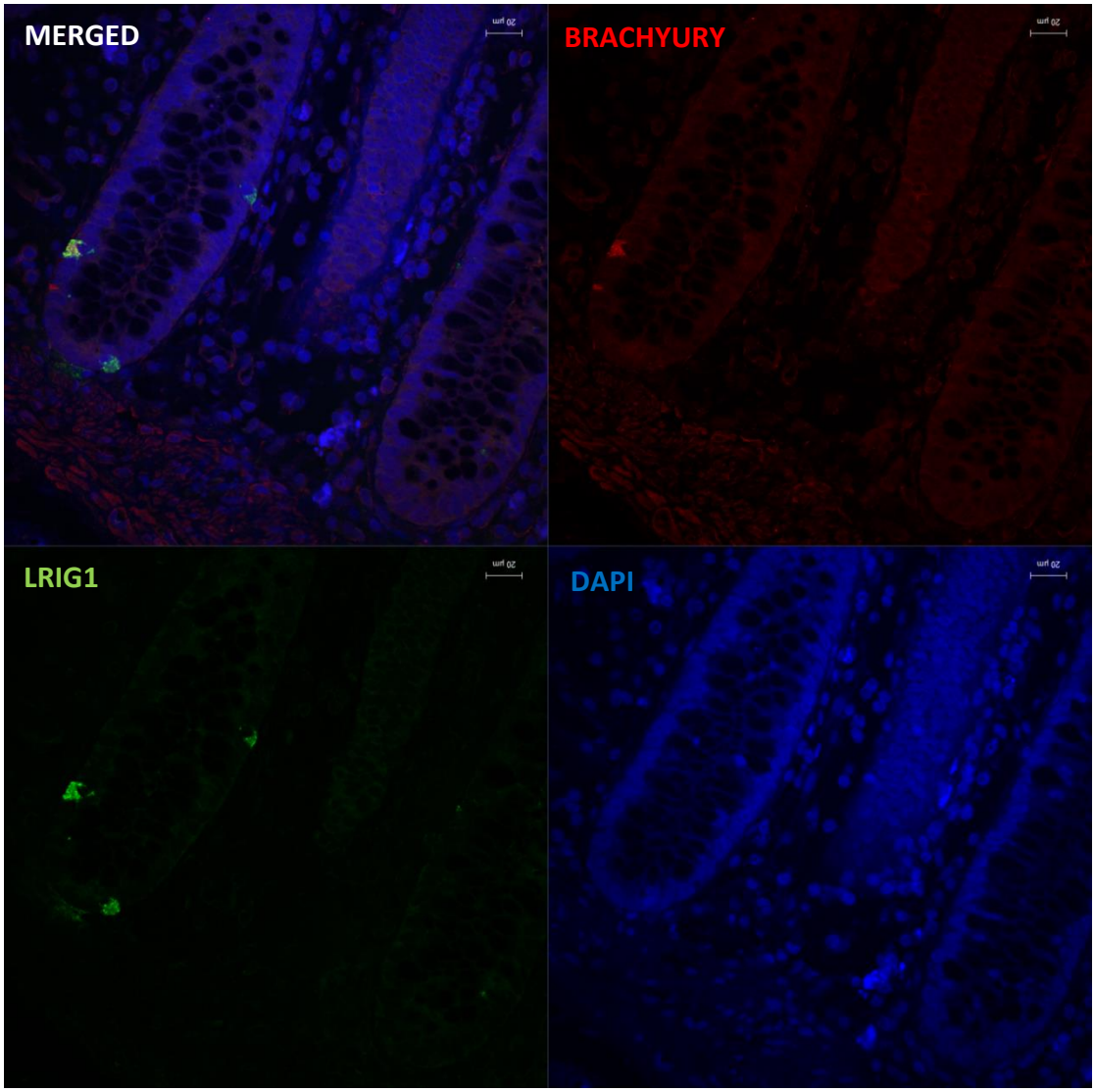

Figure 5.1 Brachyury and LRIG1 co-staining in histologically normal colon tissue (x40). FFPE tissue blocks were immunostained using Brachyury antibody (green) (AB140661) that showed co-localisation with LRIG1 (red). DNA was counterstained with DAPI blue. Brachyury and LRIG1 positive cells are located abluminally in a region consistent with the stem cell or +4 zone. Some crypts contained single Brachyury positive cells in the TA zone of the crypt. Images were acquired on Zeiss LSM710 confocal microscope by J. Jezkova, tissue processed by J.S.W.

5.2 Brachyury knockdown results in upregulation of LRIG1 and downregulation of ERBB family tyrosine kinase EGFR in SW480 colorectal cancer cells.

In order to assay whether Brachyury had an effect on the regulation of *LRIG1* levels we performed a siRNA-mediated knockdown of Brachyury using several siRNAs directed against Brachyury in SW480 (Figure 5.2). We measure the levels of knockdown of Brachyury using qRT-PCR with primers that detect both transcript variants (TV1 and TV2). The levels of *LRIG1* transcript increased significantly with Brachyury knockdown, with lower levels of increase observed in the sample treated with siRNA 05 that knocks down TV1 more than TV2.

To assess if the RNA levels of knockdown were consistent at the protein level in SW480, we extracted total protein from these cells and performed Western blots. One representative blot of three biologically replicated Western blots for each antibody is shown in Figure 5.3 A. The two bands that were detected in the

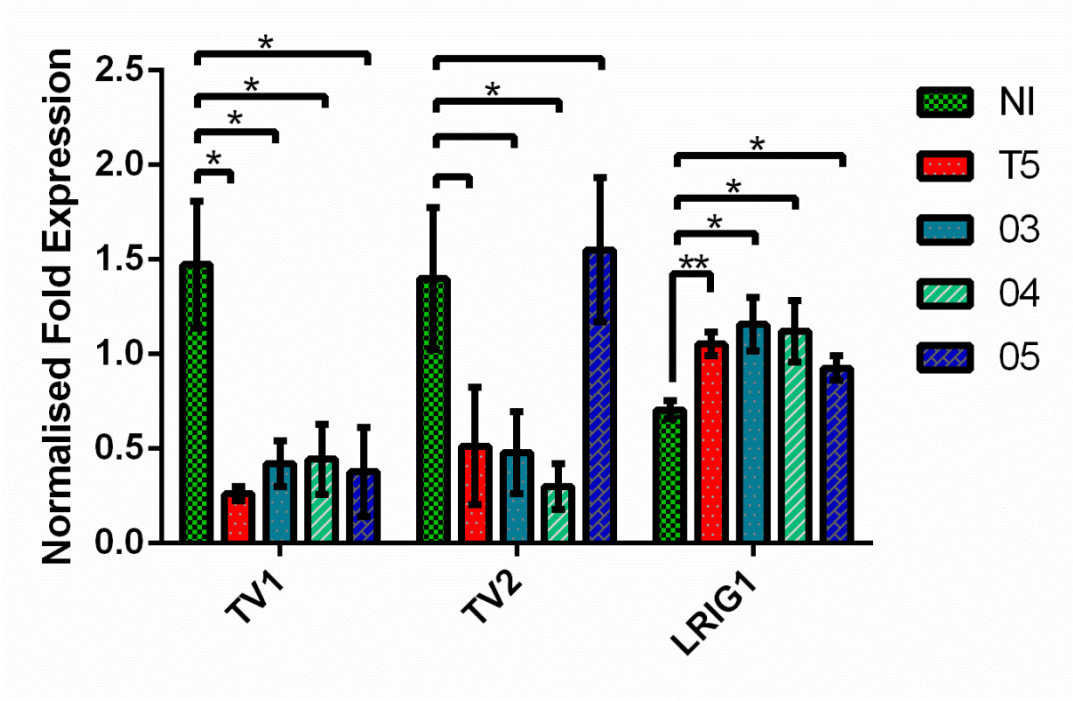

Figure 5.2 qRT-PCR of Brachyury (*T*), (TV1 and TV2) knockdown and *LRIG1* expression levels in SW480 colorectal cancer cells. Samples were treated with NI (non-interfering siRNA), T5 (Brachyury siRNA), 03 (Brachyury siRNA), 04 (Brachyury siRNA), 05 (Brachyury siRNA) and total RNA was extracted from each. qRT-PCR was performed with primers that detect both transcript variants (T), primers that detected transcript variants 1 and 2 (TV1 and TV2) and primers directed against *LRIG1*.The levels of *LRIG1* mRNA increased with decreased Brachyury expression when SW480 cells were treated with 4 different siRNAs directed against Brachyury transcripts in three biologically replicated experiments and with three technical replicates for each qRT-PCR sample. Expression levels were normalised with housekeeping genes *HSP09A* and *GAPDH*. P-values of individual pairwise comparisons are represented by *. * <0.05,**,0.01,***<0.001,****<0.0001.

untreated and non-interfering samples Figure 5.3 A (i) (50kDa and 43kDa, corresponding to TV1 and TV2 respectively) were depleted in the Brachyury siRNA treated samples (apart from the 43 kDa band corresponding to TV2 in the siRNA 05 treated sample, that knocks down TV1 transcript specifically in Figure 5.2).

We observed a significant increase in LRIG1 (Figure 5.3 A (ii) and decrease of EGFR (Figure 5.3 A (iii) compared to control (Figure 5.3B), in Brachyury siRNA treated samples. To quantify the knockdown of Brachyury and increase in LRIG1 expression, we performed densitometry analysis (Figure 5.3B). We normalised the intensity of each Western blot band with the intensity of the loading control band for α-Tubulin, then adjusted these values to the normalised band intensities of the non-interfering siRNA negative control samples.

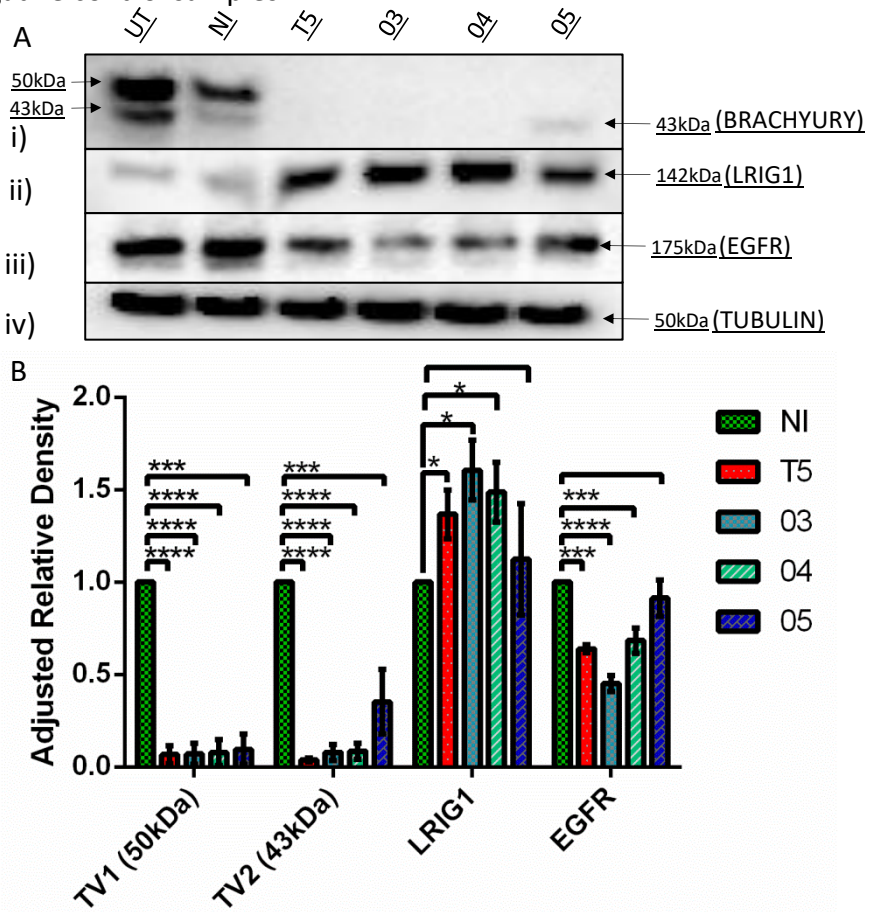

Figure 5.3 Western blot analysis of Brachyury knockdown, LRIG1 and EGFR protein levels in SW480 colorectal cancer cells. Samples were treated with NI (non-interfering siRNA), T5 (Brachyury siRNA), 03 (Brachyury siRNA), 04 (Brachyury siRNA), 05 (Brachyury siRNA). Whole cell protein extracts were performed from three biological replicates, one representative blot is shown. (A) (i) Brachyury (AF2085) levels were assessed to establish the extent of knockdown along with (ii) LRIG1 levels (12752) and (iii) EGFR (4267) with (iv) α-Tubulin (T6704) used to assess equal loading of protein for each sample. (B) Protein densitometry was performed with each sample normalised to α-Tubulin levels and levels compared to the isotype control (NI). All three biological replicates are shown. P-values of individual pairwise comparisons are represented by * <0.05,**,0.01,***<0.001,****<0.0001.

Increased levels of LRIG1 in Brachyury-depleted samples coincided with significantly decreased levels of EGFR Figure 5.3 B. The increase in LRIG1 and decrease in EGFR is greatest in the siRNA 03 sample and is not significant only in the TV1-specific 05-siRNA-treated sample.

5.3 Brachyury knockdown results in upregulation of LRIG1 and downregulation of EGFR in SW620 colorectal cancer cells.

To assess if the correlation between Brachyury and LRIG1 existed in other cell lines, we looked for another cell line in which Brachyury and LRIG1 is expressed. SW620 was the most appropriate cell line as Brachyury and LRIG1 could be detected at the RNA and protein levels. SW620 is a metastatic cell line of colorectal origin, thus also provided a different cellular context to SW480. The same analyses were performed to assay Brachyury knockdown and *LRIG1* transcript and protein levels (Figure 5.4 and Figure 5.5 respectively). Significantly increased levels of *LRIG1* were observed at the mRNA (Figure 5.4) and protein level (Figure 5.5 A (ii) and B) in Brachyury-depleted samples (Figure 5.5 A (i) and B) with decreases in EGFR (Figure 5.5 A (iii) and B) coinciding with Brachyury depletion.

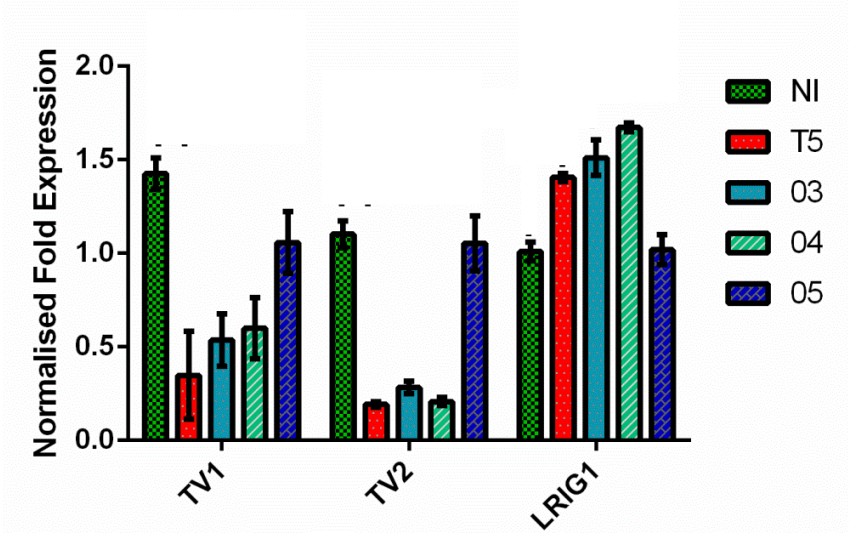

Figure 5.4 qRT-PCR of Brachyury (TV1 and TV2) knockdown and *LRIG1* expression levels in SW620 colorectal cancer cells. Samples were treated with NI (non-interfering siRNA), T5 (Brachyury siRNA), 03 (Brachyury siRNA), 04 (Brachyury siRNA), 05 (Brachyury siRNA) and total RNA was extracted from each. The levels of *LRIG1* mRNA increased with decreased Brachyury expression when SW620 cells were treated with 4 different siRNAs directed against Brachyury transcripts in three technical replicates for each sample. Expression levels were normalised with housekeeping genes *HSP09A* and *GAPDH*. Three technical replicates are shown of one knockdown experiment in SW620.

We observed lower levels of the lower TV2 (43kDa) band in the 05 siRNA treated SW620 protein sample (Figure 5.5B) when compared with the siRNA 05-treated SW480 protein sample (Figure 5.3B). The levels of TV2 RNA in SW480 and SW620 siRNA-05 treated samples were however similar (Figure 5.2 and Figure 5.4 respectively).

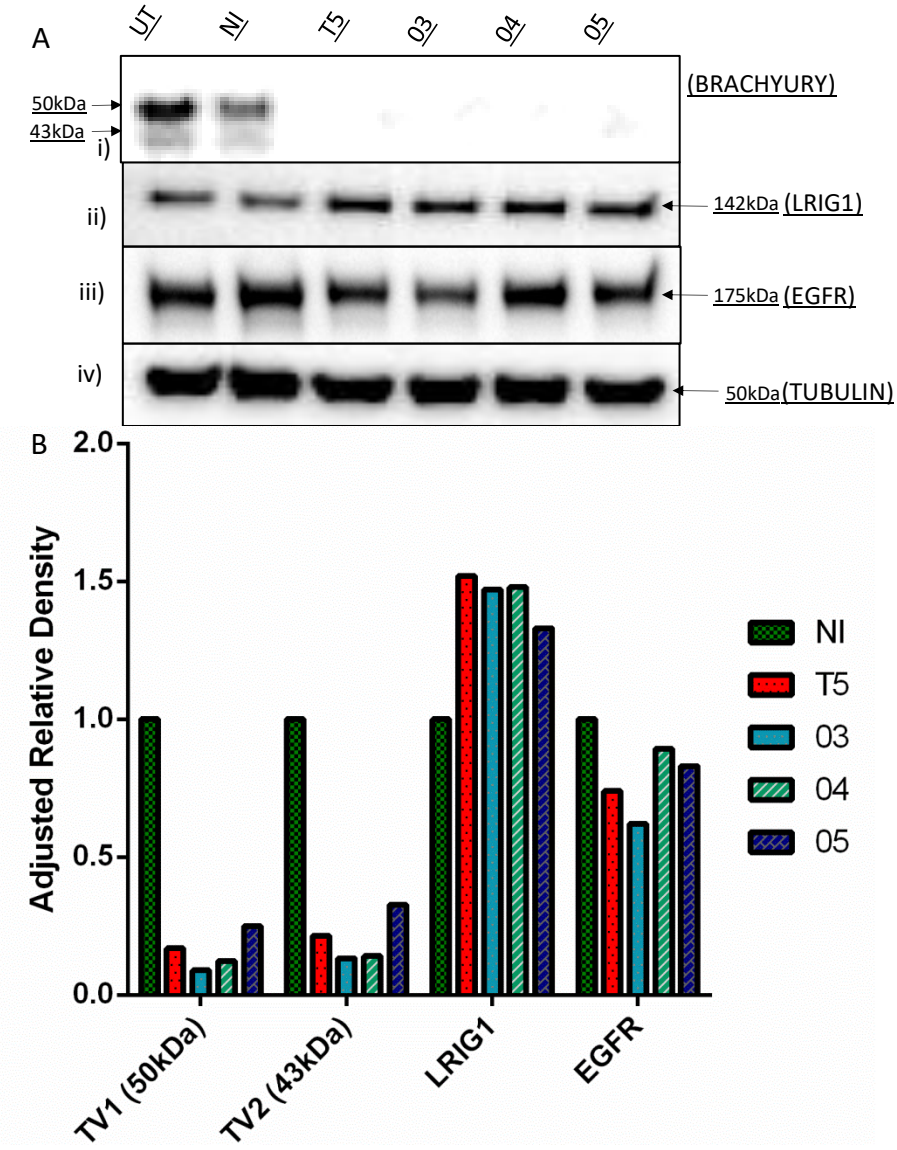

Figure 5.5 Western blot analysis of Brachyury knockdown, LRIG1 and EGFR protein levels in SW620 colorectal cancer cells. Samples were treated with NI (non-interfering siRNA), T5 (Brachyury siRNA), 03 (Brachyury siRNA), 04 (Brachyury siRNA), 05 (Brachyury siRNA). Whole cell protein extracts were performed from three biological replicates, one representative blot is shown. (A) (i) Brachyury (AF2085) levels were assessed to establish the extent of knockdown along with (ii) LRIG1 levels (12752) and (iii) EGFR (4267) with (iv) α-Tubulin (T6704) used to assess equal loading of protein for each sample. (B) Protein densitometry was performed with each sample normalised to α-Tubulin levels and levels compared to the isotype control (NI). One experimentwas performed in SW620. Western blot and associated densitometry of one knockdown experiment in SW620 is shown.

5.4 Discussion

We had previously observed that Brachyury positive cells frequently co-localise with Chromogranin A (CHGA). CHGA is a marker of EEC cells that reside mainly at the bottom of the crypts of the normal colon (Rindi et al., 2004). When EECs were marked with a cholecystokinin-green fluorescent protein in the duodenal crypts of transgenic mice, GFP+ cells were found in the upper crypt and at the crypt base (Sei et al.,2011). Furthermore, many of the GFP+ cells below the +4 position were positive for Lgr5, CD133 and the neuroendocrine transcription factor neurogenin 3. These cells, were not however, actively cycling or transient amplifying cells as they were negative for proliferation markers ki67 and phosphor-Histone H3 and were positive for ChgA (Sei et al., 2011). It seems possible therefore, that Brachyury marks a separate population of non-cycling stem-like cells given that we had observed Brachyury positive cells in a similar basal position to be KI67 negative and CHGA positive (Jezkova et al., 2014), together with the fact that Brachyury co-localises with putative +4 marker Lrig1.

At both the RNA and protein levels we observed less of an increase in the levels of LRIG1 in the siRNA-05 sample. In SW480, we consistently observed higher levels of TV2. Indeed, control levels of TV2 in terms of RNA were observed (Figure 5.2) and significantly higher levels of protein compared to the other Brachyury siRNA treated samples (Figure 5.3). In both the SW480 siRNA-05 samples, the levels of LRIG1 were lower in which transcript variant 1 is knocked down specifically. This suggests that any regulation of LRIG1 by Brachyury might specifically be regulated Transcript Variant 2.

Given that Sei (2011) and colleagues observed ChgA positive EECs were also Lgr5 positive, and $Lgr5^+$ stem cells have previously been distinguished from +4 stem cells (Yan et al., 2012), it is important that further *in situ* hybridisation studies are performed to establish whether or not Brachyury/LRIG1/CHGA positive EEC cells are also positive for LGR5. It remains possible that this population of EECs only dedifferentiates and becomes proliferative when the $Lgr5^+$ CBC compartment is damaged, especially given recent studies into the plasticity of other differentiated cell types of the colon crypt in mice (Tetteh et al., 2016).

Brachyury has previously been reported to be overexpressed in lung cancer (Roselli et al., 2012) and LRIG1 protein expression has been associated with a good prognosis for non-small cell lung cancer (NSCLCs) (An et al., 2015) and (Kvarnbrink et al.,2015). Although we could not identify a lung cancer cell line that was both Brachyury positive and LRIG1 positive to be able to perform a similar knockdown experiment to the one above, both proteins are commonly expressed in NSCLCs and the relationship between the two in this disease should be investigated further. Furthermore, when short-term EGFR inhibition was used to treat EGFR mutant lung cancer cells, this enhanced the cytoxicity of Brachyury-specific T cells, suggesting there exists a synergistic relationship between Brachyury and EGFR combinatorial therapies (Dominguez et al., 2016). This study is consistent with our observation that Brachyury regulates EGFR in cancer and it is possible that this is achieved through downregulation of *LRIG1*.

Little evidence exists of LRIG1 expression in chordomas, a rare sarcoma that frequently exhibits overexpressed Brachyury (Nelson et al., 2012). However, EGFR inhibitor therapy has proven to achieve significant kill effects on chordoma cells and inhibited growth of U-CH1 and MUG-Chor1 patient-derived chordoma cell lines (Scheipl et al., 2016). Furthermore, two small molecule inhibitors of EGFR, erlotinib and gefitinib, inhibit proliferation of the U-CH1 chordoma cell line (Siu et al., 2013), that was previously reported to overexpress Brachyury (Nelson et al., 2012). Erlotinib also significantly inhibits chordoma growth *in vivo* using patient-derived chordoma xenograft mouse models that maintain nuclear expression of Brachyury (Siu et al., 2013). Further studies are required to consider the possibility of using EGFR inhibitor therapy along with Brachyury small molecule inhibitors to treat chordomas that overexpress Brachyury and are responsive to EGFR inhibition.

In summary, our data suggests that there exists a relationship between Brachyury and LRIG1 in the normal colon and colorectal cancer cells. We have shown that Brachyury regulates a key receptor tyrosine kinase in CRC cells (schematic in Figure 5.6) that has been shown previously to mediate growth signals in NSCLCs and chordomas in which Brachyury is overexpressed. It is important that further studies are carried out, including chromatin immunoprecipitation analysis, to establish whether Brachyury directly regulates *LRIG1* in the normal colon and colon cancer

cells *in vitro* in which both are expressed. Furthermore, given that Brachyury has significant effect on the expression of EGFR, it is possible that combinatorial therapies that inhibit both proteins could be beneficial to appropriately stratified patient groups.

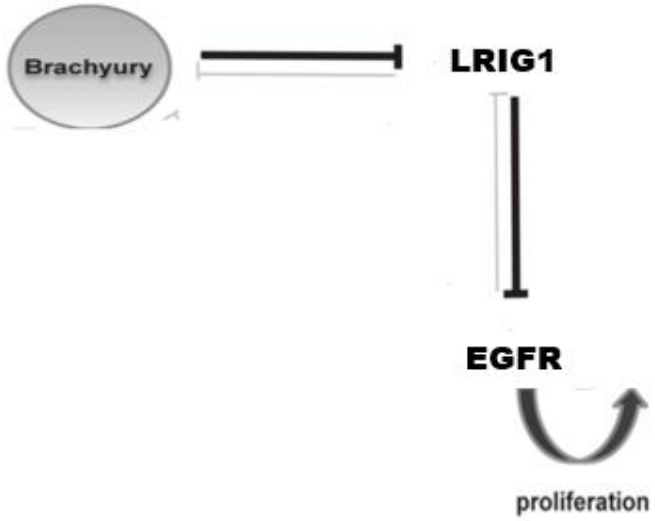

Figure 5.6 Schematic model describing a central role for Brachyury in regulating proliferation through controlling the expression of the negative pan-ErbB regulator EGFR.

6 Brachyury is expressed in colorectal carcinomas and metastases of colorectal origin and confers a poor survival outcome when observed in enteroendocrine-like cells.

6.1 Introduction

Early studies on acute myelogenous leukaemia were consistent with the existence of tumour cells with stem-like properties that resembled the haemopoietic stem cell hierarchy of normal blood homeostasis (Bonnet et al., 1997; Greten et al., 2016.). Following studies extended the evidence that this hierarchy existed in other cancers and led to the concept of a cancer stem cell as the cell of origin of neoplasias (Greten, 2016; Clevers, 2011). A small reservoir of self-sustaining cells that are able to self-renew and maintain tumour tissue (Barker et al., 2009). These cells might be refractory to conventional therapies and like their normal somatic counterparts, are able to regenerate radiation and/or chemotherapy damaged tissue (Greten, 2016; Clevers, 2011).

Little work has been done to establish the role of secretory epithelial cells, these include Paneth cells/Reg4+ secretory cells, Goblet cells and enteroendocrine cells,and their role specifically in generating the cancer stem cell niche. As well as likely contributors of the stem cell niche, it is likely that secretory epithelial cells can themselves be transformed into proliferative cancer cells. Inhibition of Notch in proliferative cells in intestinal crypts and adenomas converts them into goblet cells (van Es et al., 2005) showing that proliferative epithelial cells are plastic and when their niche is manipulated, they may be transformed into a secretory post-mitotic cell type. It is possible that this plasticity converts secretory epithelial cells into proliferative epithelial cells in the dysregulated tumour microenvironment also.

Tumours generated *in vitro* and implanted into mice from SW122 CRCs expressed markers of all three types of differentiated colonic epithelial cells, namely columnar, goblet and enteroendocine cells (Yeung et al., 2010). Chromogranin A, a marker of EECs was present mainly around lumen-like structures of clonogenic megacolonies, suggesting this product is secreted (Yeung et al., 2010). Mucin and CDX1 (marker of

enterocytes) were also present toward the periphery of the colonies, with less differentiated cells towards the centre perhaps representing the CSC population (Yeung et al., 2010).

Little recent work has been done to identify the presence of EECs or EEC-like cells in colorectal cancer. One study associated patients with numerous CHGA positive endocrine tumour cells with a significantly worse prognosis compared with patients without endocrine cells (Hamada et al., 1992). Hamada and colleagues concluded that neuroendocrine differentiation was an independent prognostic factor in CRC. Circulating CHGA has been associated also with prostate cancer patients with hormonerefractory disease (Berruti et al., 2005) and as a serum marker of small cell lung cancer(Lamy et al., 2000).

As we observed Brachyury expression in the EECs of the normal GI tract and distinct histopathological adnenomas, we extended our study to a larger number of primary colorectal adenocarcinomas and metastatic tumours to establish if Brachyury positive EECs are present at every stage of CRC progression and to assess any association their presence has on the prognosis of patients in which they are present.

6.2 Distibution of Brachyury staining in relation to the cell cycle inhibitor $p27^{Kip1}$ and the proliferation marker ki67.

Work from our group highlighted that Brachyury regulates cell cycle progression through down regulating the cell cycle inhibitor $p27^{Kip1}$ in CRC (Jezkova et al., 2014) cells. Having established this association, we aimed to compare the expression of Brachyuyry, $p27^{Kip1}$ and proliferation marker KI67 in consecutive patient derived FFPE sections by IHC (see section 2.6 for details of source of tissue). For this, we used the monoclonal anti-Brachyury antibody AB57480 (Figure 6.1 A-C). Not all cells in the carcinoma were Brachyury positive, but we did observe very prominent nucleolar staining that was confined to the carcinoma, we also observed extensive cytoplasmic staining, with some cells staining very strongly for Brachyury and resembling the EECs observed in normal tissue and adenomas using the other mouse monoclonal antibody (AB140661). This was the first reported nucleolar localisation of Brachyury (Jezkova et al., 2014). Whilst the carcinoma was strongly positive for the proliferation marker KI67 (Figure 6.2 D-F), that defined the tumour margin, $p27^{Kip1}$ staining was only observed at the edge of the tumour (Figure 6.2 A-C) and $p27^{Kip1}$ and Brachyury were largely mutually exclusive (Figure 6.2 A-F).

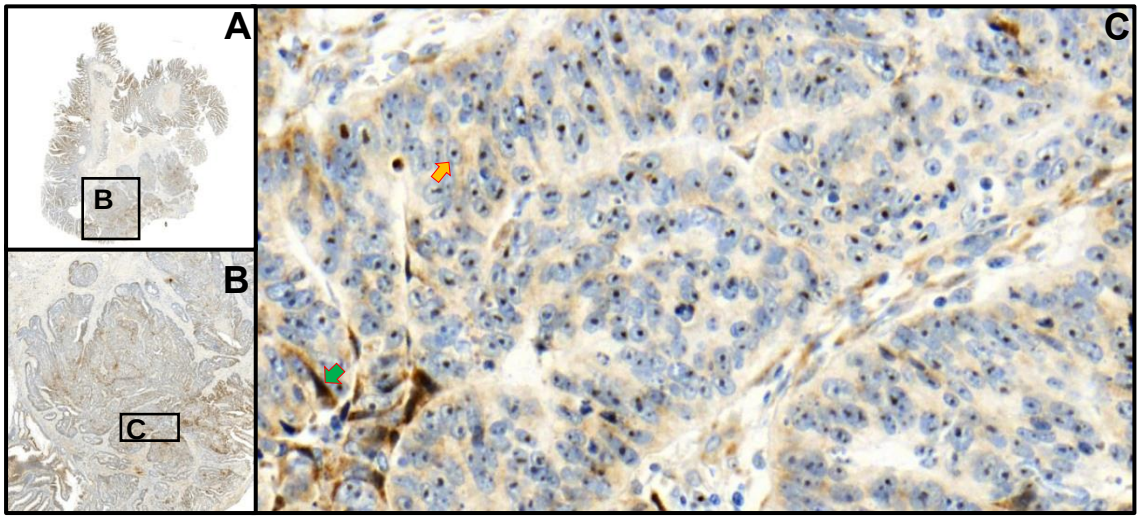

Figure 6.1 Brachyury staining in the cytoplasm and nucleolus of the carcinoma. In this preliminary study, we stained 4 advanced colorectal samples with Brachyury (AB57480) with one representative patient shown here, (A) 2.5x (B) 5x and (C) 20x magnification. Brachyury nucleolar staining was most prominent in the carcinoma of patients and is highlighted by the orange arrow. Some cells with strong cytoplasmic staining resembled EECs/ neuroendocrine cells are highlighted by the green arrow.

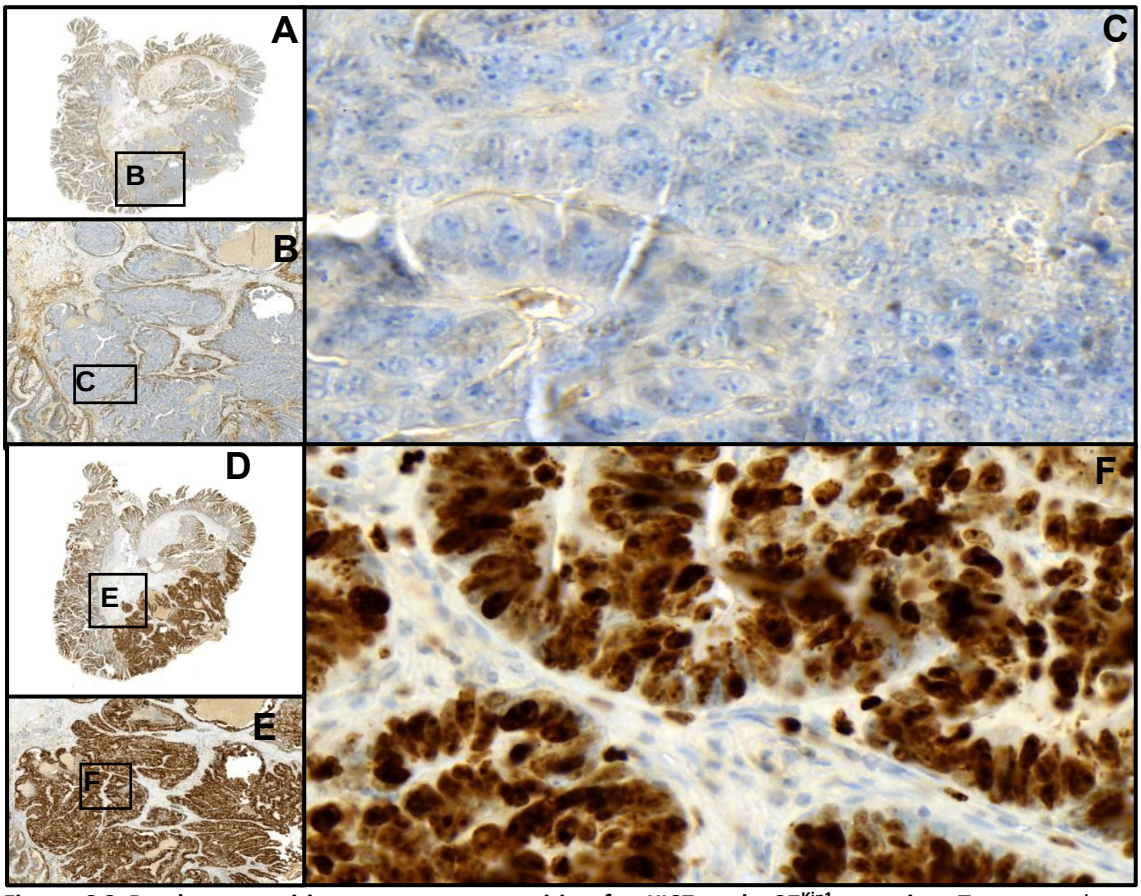

Figure 6.2 Brachyury positive tumours are positive for KI67 and $p27^{Kip1}$ negative. To assess the proliferation status of the Brachyury positive tumour, we stained consecutive sections of the tumour with $p27^{Kip1}$ (A) 2.5x (B) 5x and (C) 20x magnification) and KI67 (A) 2.5x (B) 5x and (C) 20x magnification). C is largely equivalent to region C in Figure 1.1. The equivalent region in F was strongly positive for the proliferation marker KI67.

Four CRC patient samples were studied for IHC analysis with AB57480 all were Brachyury positive (see section 2.6 for the details of the source of the tissue): the results shown are from a single patient and are representative of staining patterns obtained for the four Brachyury positive CRC patients (H&E and negative control shown in Appendix B). We conclude here that the heterogeneity of Brachyury staining within the tumour suggests that it may have region specific functions. This widespread staining is nonetheless consistent with Brachyury controlling cell cycle progression through $p27^{Kip1}$ downregulation. As we could not validate this antibody by Western blot (no signal was detected on western blots of the cell lines SW480 and H460, both known to overexpress Brachyury), we decided to proceed further studies with AB140661 which we were able to validate through western blot and IF.

6.3 Brachyury staining falls into four distinct categories in primary CRCs and metastases.

When primary CRC patient samples were surveyed for Brachyury using monoclonal anti-Brachyury antibody AB140661, four staining categories became clear (Figure 6.3). The first was a weak cytoplasmic staining that the majority of patients fell into in both primary CRCs and metastases (Figure 6.3 A + E). Next, we saw a moderate cytoplasmic staining that was specific to epithelial regions of the carcinoma (Figure 6.3 B + F). Thirdly we saw occasional nuclear staining in regions of the tumour (Figure 6.3 C + G). Finally we saw Brachyury positive EEC-like cells (Figure 6.3 D + H) in a small number of tissue samples that resembled the EECs of normal and adenoma tissue samples.

Occasional nuclear staining (Figure 6.3 C +G) was seen in confined regions of some patient's tumour samples. More of this nuclear staining was observed in metastatic tumour samples although this was still occasional and no histopathological association could be made to these regions.

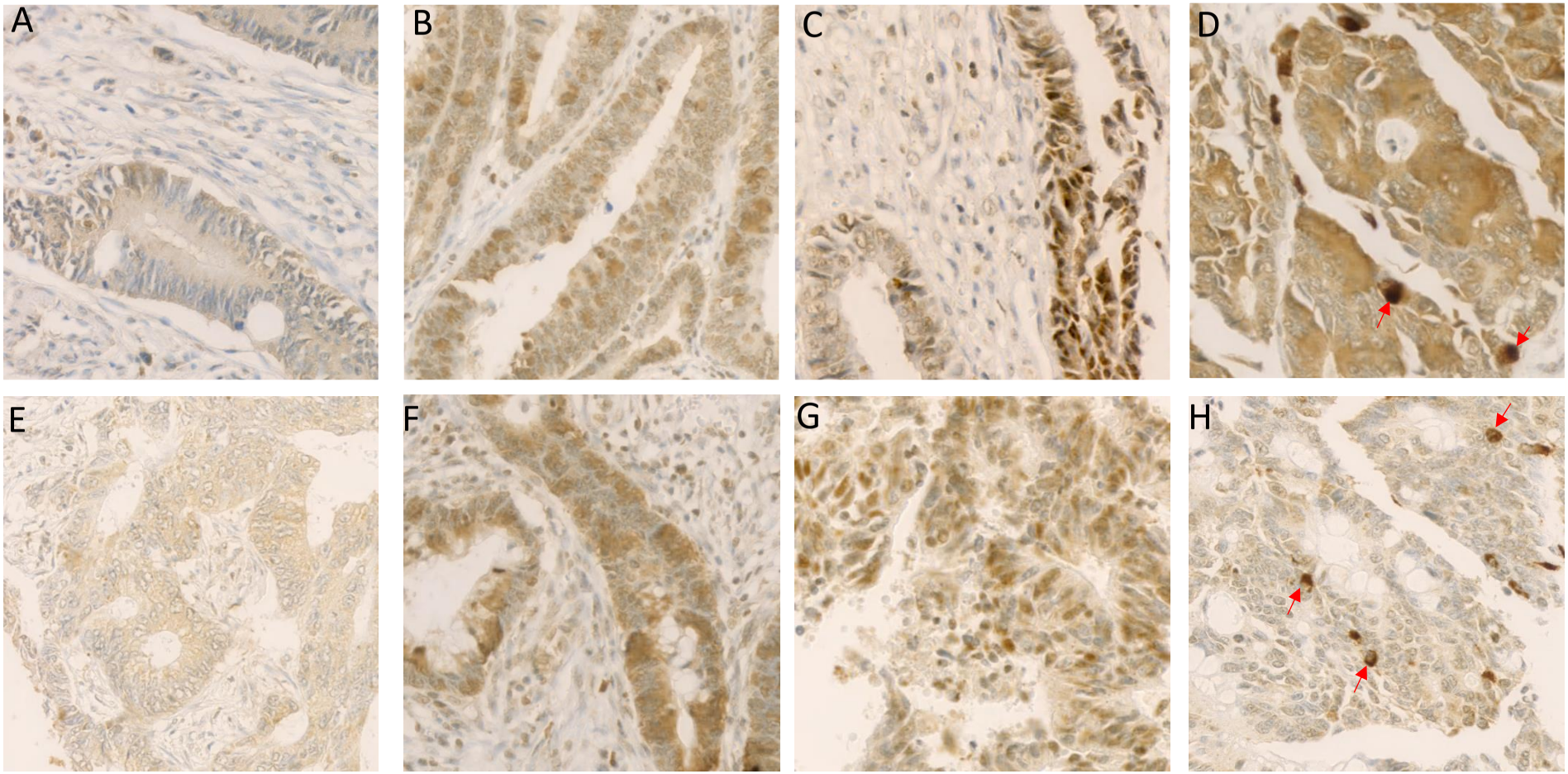

Figure 6.3 Representative TMA of Primary Colorectal Cancers and metastases showing the immunohistochemical expression intensity and pattern of brachyury: (A) Weak Cytoplasmic (10x) (B) Moderate Cytoplasmic (10x) (C) Nuclear (10x) (D) Strong Cytoplasmic with Brachyury positive Enteroendocrine-like cells (10x) in primary CRCs and. (E) Weak Cytoplasmic (10x) (F) Moderate Cytoplasmic (10x) (G) Nuclear (10x) (H) Strong Cytoplasmic with Brachyury positive Enteroendocrine-like cells (10x).

6.4 Brachyury distribution in primary CRCs and metastases

We assessed Brachyury staining distribution in 281 primary colorectal tumour tissue microarray cores (TMAs) and 81 metastatic TMAs of colorectal origin. Four staining patterns were observed; cytoplasmic weak, cytoplasmic moderate, nuclear staining and the strong Brachyury positive EEC-like staining characteristic of the normal colon and colorectal adenomas defined above in Figure 6.3 A-H. Most primary CRC (73%) and metastatic (81%) patient tissue samples fell into the weak staining category (Figure 6.4 A + B respectively). Within the moderately stained group, a small percentage of patients had several strongly cytoplasmic EEC-like cells (2% PCRC and 5% Metastases), so these were categorised as strong Brachyury positive EEC. Within the moderately stained group a small percentage of patients had the presence of nuclear staining (1% PCRC and 8% Metastases), so these were categorised as nuclear staining (Figure 6.4 A + B respectively).

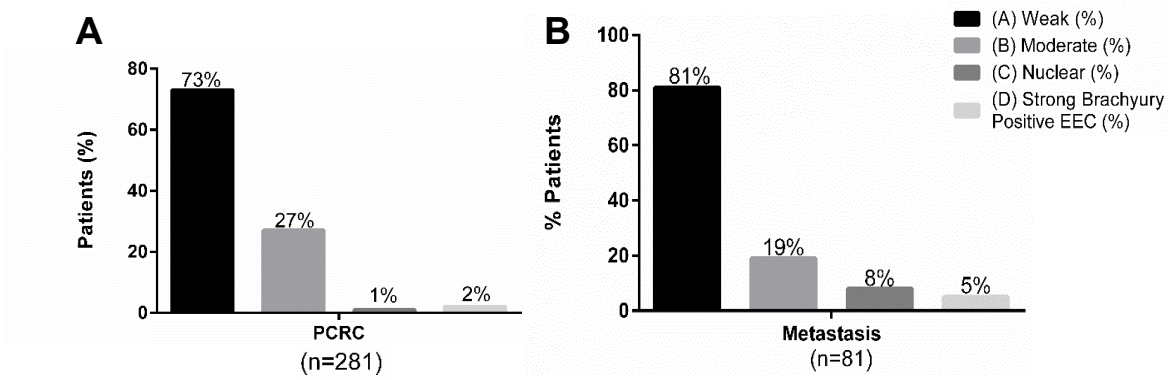

Figure 6.4 Histogram showing the percentage of primary and metastatic patient tissue samples that fall into the brachyury immunohistochemical staining categories. A shows the distribution of Brachyury staining in primary colorectal cancer samples and B shows Brachyury distribution in 81 patients metastatic tumours of colorectal origin.

6.5 Moderate cytoplasmic Brachyury staining confers a poor survival outcome in primary CRCs and a poor progression-free survival rate.

Patients were stratified based on the staining category in which their tissue sample fell into (Figure 6.3 A-H). Patients were scored blindly and three independent observers confirmed the staining pattern for each patient. In the primary CRC patient samples, 204 patients exhibited weak Brachyury staining and in 77 patients moderate Brachyury staining was observed. When these two groups of patients cumulative survival curves were compared using the LogRank test, with the other two groups excluded; the weakly stained Brachyury group had a significantly better survival outcome than the moderately stained Brachyury group ($P<0.0001$) (Figure 6.5 A). As the median survival is the time it takes for 50% for patients to reach 50% survival, and given that 50% of the patients included in this particular comparison are still alive, the median survival cannot be defined.

The moderately stained Brachyury tissue samples also had a significantly worse progression-free survival rate when compared with weak (*P=0.0077*) (Figure 6.5 B).

These qualitative assessments suggest that Brachyury is broadly expressed in primary colorectal tumours with a subset of patients displaying higher expression levels of Brachyury (moderate cytoplasmic (Figure 6.3). In summary, these patients have a significantly worse survival outcome when compared to the 73% of patients in which a weak cytoplasmic staining or low expression was observed. The heterogeneity observed in these tissue samples with AB140661 is consistent with the heterogeneity observed with the AB57480 antibody.

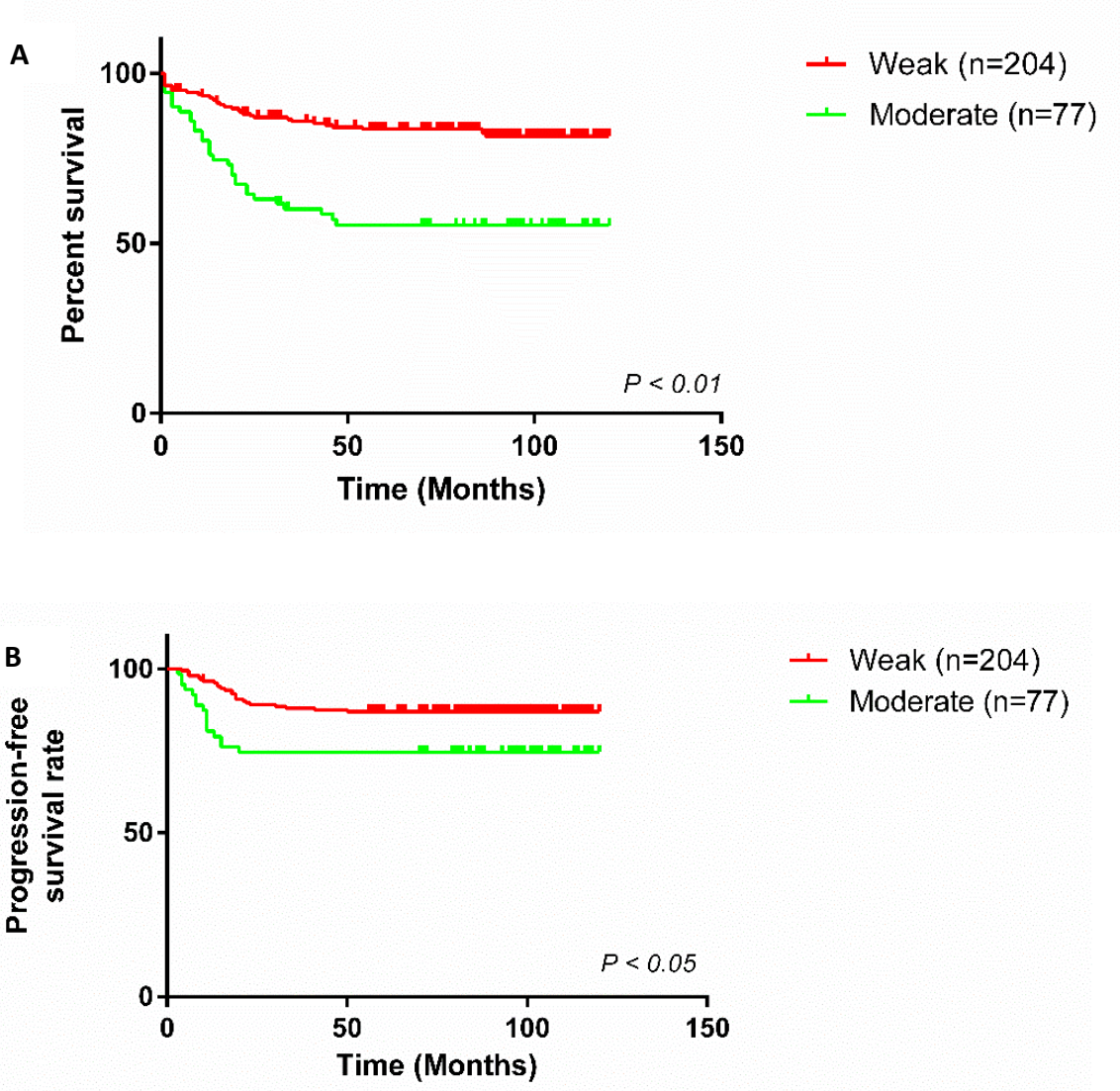

Figure 6.5 Kaplan-Meier overall survival estimates and progression-free survival estimates of primary colorectal cancer patients with weak vs moderate Brachyury expression. PCRC patients were stratified based on staining intensity. (A) The low expressing Brachyury group had a survival advantage over the group that exhibited a moderate expression of Brachyury in primary colorectal cancer tissue samples (*P*<0.0001). (B) The low expressing group also had significantly better progression-free survival rate compared with patients that had tumours with high levels of Brachyury (*P=0.0077)*.

6.6 Brachyury positive EEC-like cells in primary tumours confer a poor survival outcome and progression-free survival rate.

In the primary CRC tissue samples studied, 5 patients had identifiable Brachyury positive EEC-like cells (T^+-$EECs^+$) that were not present in the other 276 patients (T^+EECs^-) (Figure 6.6 A-E). Furthermore, these EECs were within the tumour margin and when T^+-$EECs^+$ patients were compared with the T^+EECs^- patients that did not have Brachyury positive EECs, they had a significantly worse survival outcome (*P=0.0269*) (Figure 6.7 A). These patients also had a significantly worse progression-free survival outcome (*P=0.0269*) (Figure 6.7 B).

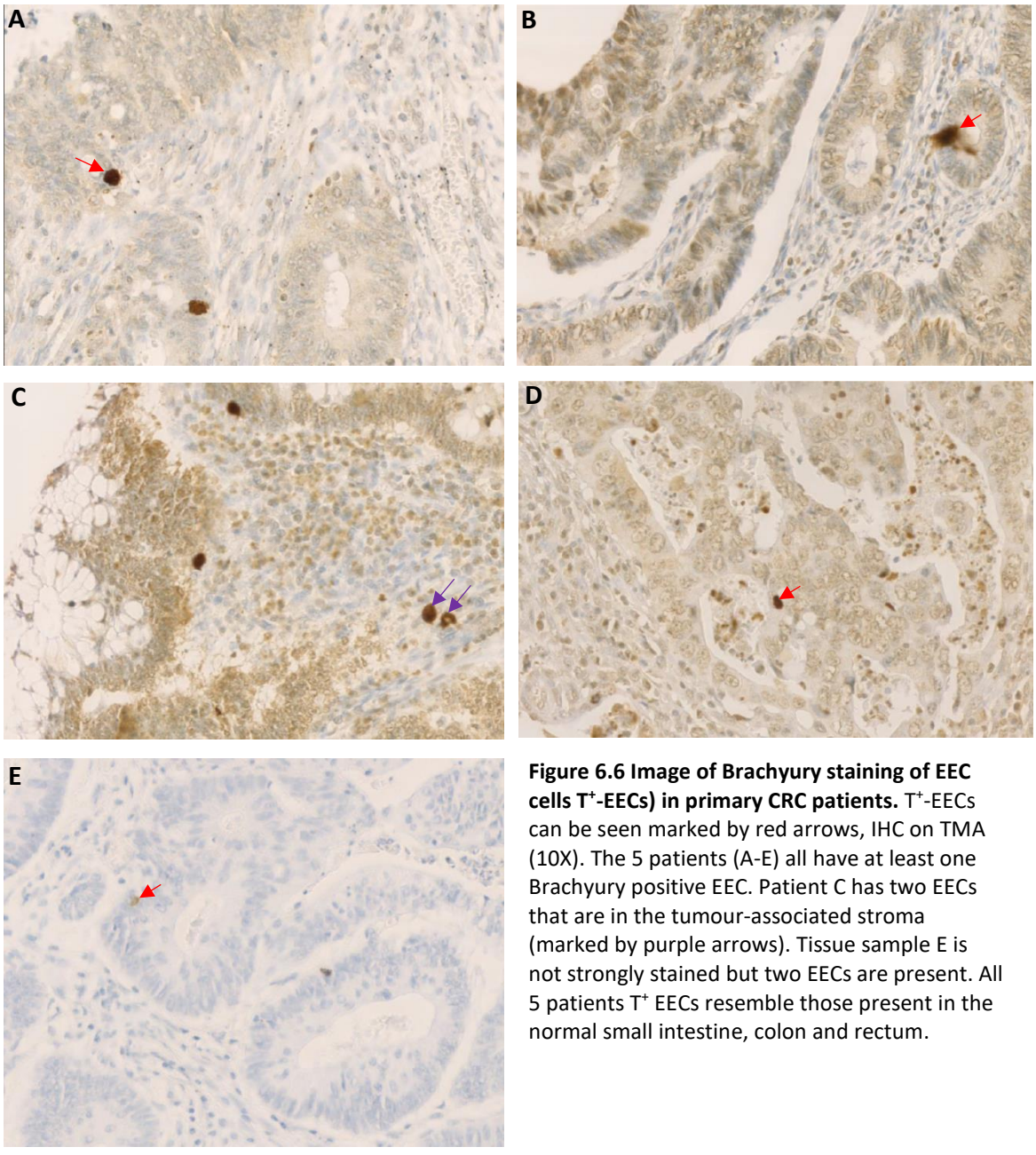

Figure 6.6 Image of Brachyury staining of EEC cells T^+-EECs) in primary CRC patients. T^+-EECs can be seen marked by red arrows, IHC on TMA (10X). The 5 patients (A-E) all have at least one Brachyury positive EEC. Patient C has two EECs that are in the tumour-associated stroma (marked by purple arrows). Tissue sample E is not strongly stained but two EECs are present. All 5 patients T^+ EECs resemble those present in the normal small intestine, colon and rectum.

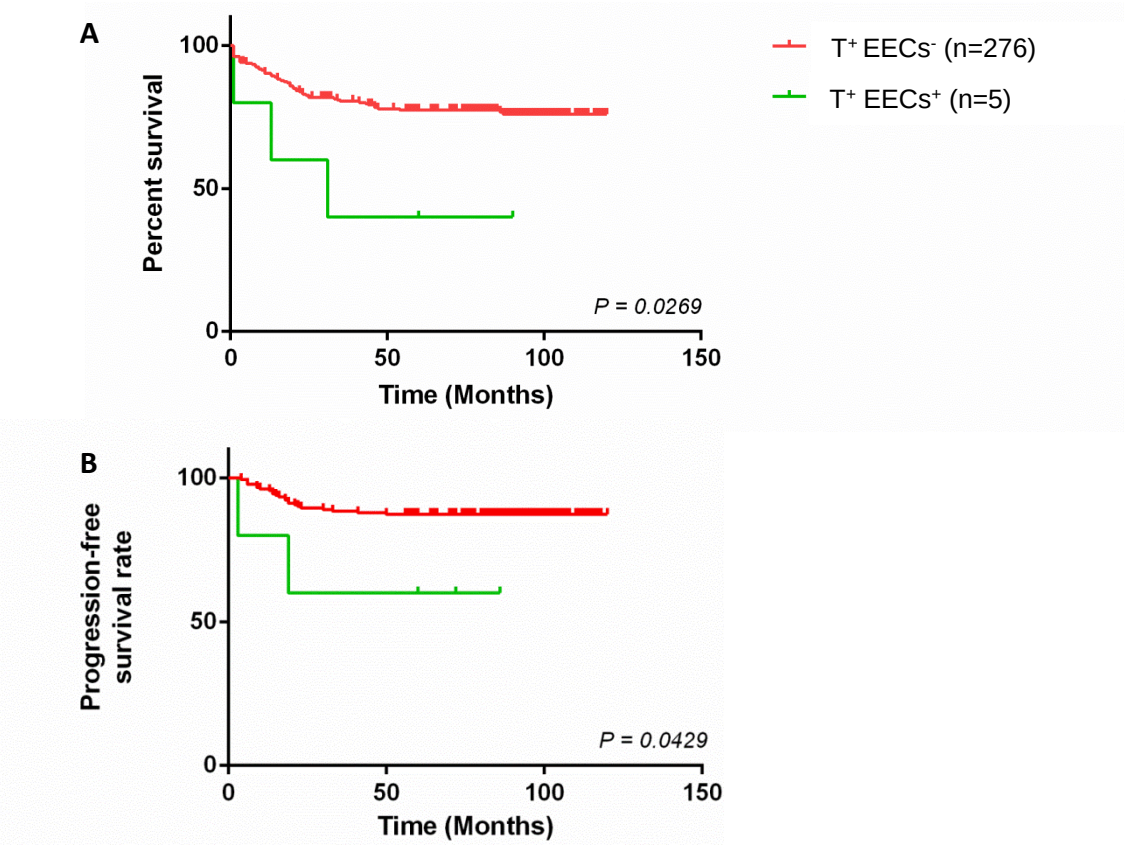

Figure 6.7 : Kaplan-Meier overall survival estimates and progression-free survival rates of colorectal cancer patients stratified by the presence of T^+ EECs and the absence of T^+EECs. (A) When patients were stratified based on the presence of T^+EECs, the survival outcome of patients with T^+EECs was significantly worst when a LogRank comparison was made with the no T^+EECs group's survival curve (*P=0.0269)*. (B) Similarly, patients with T^+EECs also had a worse progression-free survival rate when compared with patients that were T^+EEC negative (*P=0.0429*).

6.7 Brachyury positive EECs in metastatic tissue samples confer a poor disease-free survival rate.

In order to study the expression of Brachyury in colorectal metastases, we stained 81 metastatic patient samples with the AB140661 antibody. Brachyury positive EECs were easily identifiable and again resembled those positive EECs observed in normal, adenomas and primary colorectal cancers (Figure 6.8 A-D).

Brachyury positive EECs were present in a small proportion of patients. These patients had a significantly worse progression-free survival rate when compared with Brachyury EEC negative patients (Figure 6.9). This is the cumulative length of time that patients disease does not relapse and is estimated again using Kaplan-meier survival curves. The curves of the two patient groups were compared using a LogRank test and were found to significantly different (*P<0.01*).

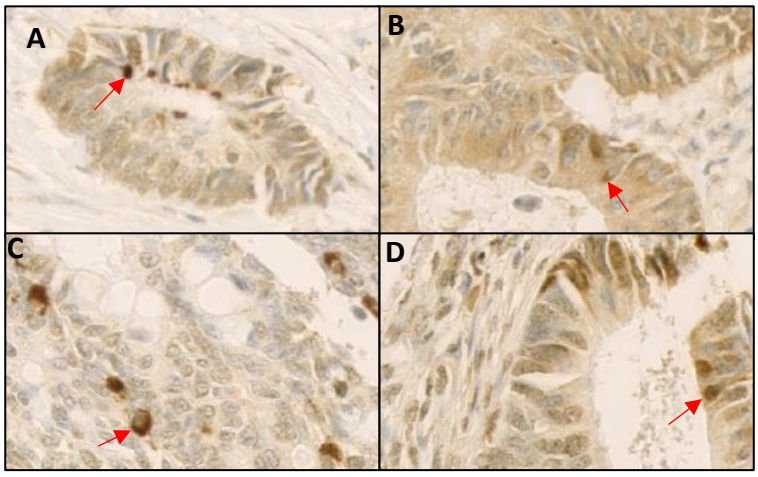

Figure 6.8 Image of Brachyury staining of EEC cells (T^+-EECs) in metastases of CRC patients. T^+-EECs can be seen marked by red arrows IHC on TMA (10X). The 4 patients (A-D) all have at least one Brachyury positive EEC. No EECs were observed in the stromal margin of the tumour, unlike the primary CRC EECs in which occasional EECs were inside the stroma.

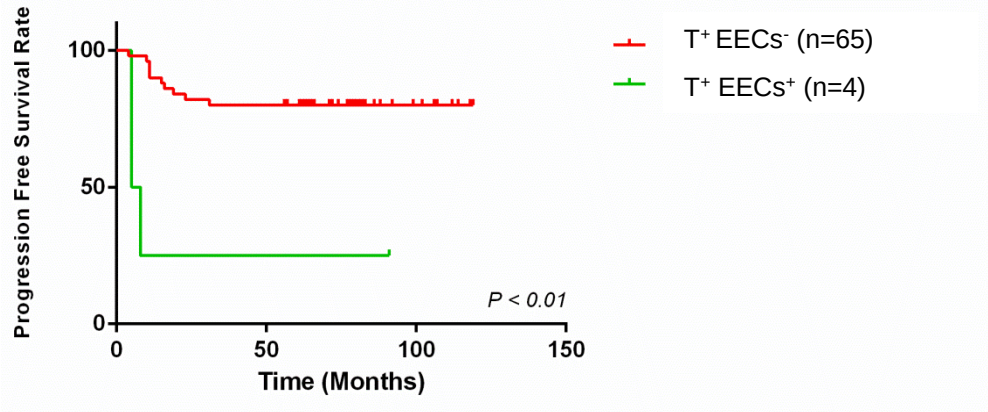

Figure 6.9 Kaplan-Meier metastasis-free overall survival estimates of colorectal cancer patients. Patients were stratified by T^+ EEC-like cells$^+$ vs T^+EEC-like cells$^-$ expression of brachyury in metastatic tissue samples. Patients with Brachyury positive EEC-like cells had a significantly worse disease-free rate than patients who did not display these same cells.

6.8 Cross tabulation analysis of Brachyury positive EECs in CRCs and metastases.

6.8.1 PCRC T^+EECs cross-tabulation analysis.

To investigate correlations between T^+EECs in tissue samples and conventional histological prognostic indices, variables were cross tabulated and χ^2 tests were performed to examine whether they were statistically significant. All variables were cross-tabulated but only those which showed a statistically significant relationship and/or are of interest are tabulated below (Table 6.1).

Interestingly, all 5 T^+EEC patients were males. Four out of the five patients exhibited histologically-defined undifferentiated tumours with one of the five patients having a poorly differentiated tumour (*P<0.0001*). No statistically significant correlation was found between tumour grade and Brachyury EEC positivity although all patients were graded as either T3 or T4. All T^+EEC positive patients had a primary tumour size of less than 4cm, with the mean tumour size being calculated as 4.53 cm. Brachyury positive EECs therefore exist in below average sized primary tumours. Most Brachyury EEC positive patients tumours had invaded into local blood vessels although this relationship was not statistically significant (*P=0.086*).

Prognostic indices	Brachyury +ve EEC	Brachyury -ve EEC	Statistics
Male	5	166	
Female	0	110	*P = 0.07*
Differentiated[x]	0	124	
Poorly differentiated[x]	1	123	
Undifferentiated*	4	29	***P < 0.0001***
Tumour grade 1 and 2	0	52	
Tumour grade 3 and 4	5	224	*P = 0.282*
Tumour < 4 cm	5	142	
Tumour > 5 cm	0	134	***P = 0.031***
Venous vessel invasion (+ve)	4	115	
Venous vessel invasion (-ve)	1	161	*p = 0.086*

Table 6.1 Cross tabulation of prognostic indices and Brachyuyr EEC +/- in primary CRCs.*Undifferentiated/poor differentiation group for Cox regression.[x] Differentiated group for Cox regression.

6.8.2 Metastases T^{+}EECs cross-tabulation analysis.

The same cross tabulation analysis was performed based on Brachyuyry EEC positivity in metastatic tumours. A family history of colorectal cancer was correlated with Brachyury positive EEC status (*P=0.019*) but no other prognostic indices were found to have any correlation to Brachyury EEC status (Table 6.2).

In metastatic tumours, most Brachyury EEC positive patients were male, were positive for venous vessel invasion, lymphatic vessel invasion but were mainly present in patients with primary tumours greater than 5cm in size. Similar to the primary CRC patient samples, all 4 Brachyury metastatic tumours that were Brachyury positive EEC patients had tumours graded 3 and 4. This relationship was again not statistically significant.

Prognostic indices	Brachyury +ve EEC	Brachyury -ve EEC	Statistics
Male	3	51	
Female	1	26	*P = 0 .717*
CRC Family History	2	8	
No CRC Family History	2	69	***P = 0.019***
Venous vessel invasion (+ve)	3	55	
Venous vessel invasion (-ve)	1	22	*P = 0.877*
Lymphatic vessel invasion (+ve)	3	65	
Lymphatic vessel invasion (-ve)	1	12	*P = 0.617*
Tumour < 4 cm	1	34	
Tumour > 5 cm	3	43	*P = 0.451*
Tumour grade 1 and 2	0	6	
Tumour grade 3 and 4	4	71	*P = 0.562*

Table 6.2 Cross tabulation of prognostic indices and Brachyuyr EEC +/- in metastastatic tumours.

6.8.3 Cox regression analysis of Brachyury positive EEC patients in primary colorectal cancers.

Multivariate survival analysis was carried out using Cox regression to estimate hazard ratios for Brachyury EEC status alone and Brachyury EEC status with all other covariates in PCRCs. Brachyury EEC status was found to have a hazard ratio of 4.2 (*P=0.049*). Patients were then grouped into two groups for all the other covariates, with only statistically significant and/or covariates of interest shown in Table 6.3. Differentiation status was grouped into undifferentiated/poorly differentiated (*/xTable 6.1). Finally, tumour stages were grouped into T1/T2 and T3/T4 as with the cross tabulation analysis.

Male patients that were positive for Brachyury EECs had a hazard ratio of 4.5 (*P=0.045*) suggesting that male patients with these T^{+}EEC-like cells had a worse survival outcome than Brachyury EEC-like negative patients. Undifferentiated tumours were associated with Brachyury EECs above (Table 6.1), however when survival was considered with these covariates, the hazard ratio of 3.64 was not significant (*P=0.085*).

Prognostic Indices with Brachyury EEC Status (+/-)	Hazard Ratio	95% CI	P-Value
Male	4.5	1.03-19.7	P = 0.045
Undifferentiated	3.64	.837-15.9	P=0.085
Tumour Stage 3 and 4	4.314	0.971-19.1	P=0.055

Table 6.3 Cox regression analysis including prognostic indices that predicted a poor survival outcome with Brachyury positive EEC status.

6.9 Discussion

Here we show that Brachyury expression throughout the tumour and in Brachyury positive EEC-like cells confers a poor survival outcome and progression-free survival rate in primary colorectal cancers. Overall survival rates in this study are comparable to other large scale studies in which the survival of patients with CRC was evaluated based on stratification of the molecular subtype of their cancer (Phipps et al., 2015) and even other studies that stratified patients based on the number of lymph nodes resected (Storli et al., 2011). Furthermore, Brachyury positive EECs in metastatic tumour samples indicates a poor relapse-free rate in these patients. Interestingly, cross-tabulation revealed a significant association between Brachyury-EEC positive patients and histologically defined undifferentiated tumours (***P < 0.0001***) and is discussed further below in the context of Kilic and colleagues (2011) findings. Other statistics from the cross-tabulation and cox regression analyses had borderline significance and were not deemed interesting enough to further investigate.

In another study, IHC was performed using the goat polyclonal antibody (AF2085) to estimate survival and correlate Brachyury expression with conventional histological indices (Kilic et al., 2011). We observed that AF2085 detected both transcript variants of Brachyury whereas the monoclonal antibody we used detected transcript variant 1. The authors of this study report that patients that were T1-2N0M0 and were Brachyury positive had a significantly worse survival outcome when Brachyury was expressed in their tumour tissue. We found no significant association between moderately stained Brachyury patients of early staged colorectal cancer when overall Brachyury staining was considered as opposed to the EEC-specific staining, with 18 out of these Brachyury moderate patients that were T1/T2 not having any overall survival differences compared with the T3/T4 group. We did however see a statistically significant risk for patients that had moderate (positive) Brachyury staining and were stages T3/T4, with these patients having a hazard ratio of 3.6 (2.2-5.8 and *P<0.0001*). We do therefore report quite different findings to Kilic and coworkers when the survival of positive (moderate) patients at early stages (T1/T2) is considered in comparison to the patients that were Brachyury positive and staged as either T3/T4. This could be due to the specificity of the antibody used with both

antibodies detecting different transcript variants, i.e. AF2085 detects TV2 when AB140661 has a much lower affinity to it if indeed it does detect TV2 at all.

Patients were characterised histologically using Haemotoxylin and Eosin staining to grade the tumours by a pathologist independently. Tumours were graded on a scale (differentiated to undifferentated) to indicate how well differentiated the patients tumours were. Most Brachyury EEC positive EEC tumours fell into the undifferentiated category of tumour (4/5, *P<0.0001*). Interestingly, Kilic and colleagues also reported an association between undifferentiated tumours and Brachyury positivity (*P<0.0001*) (Kilic et al., 2011). This is especially interesting given Brachyurys role in development in which it regulates canonical Wnt signalling (Martin & Kimelman, 2008) and its implication in regulating the expression of Nanog, a gene that is required to maintain pluripotency (Sarkar et al., 2011). Furthermore, it is interesting to speculate that Brachyury positive EECs in undifferentiated tumours contribute to the niche that maintains cancer stem cells that generate this undifferentiated tumour.

In agreement with our groups study that established a correlation between Brachyury expression and p27 –dependent regulation of the cell cycle (Jezkova et al., 2014), Kilic and collegues reported a correlation between Brachyury and low level expression of the cell cycle inhibitors p21 and p27 and a poor survival outcome for these patients (*P=0.006*) (Kilic et al., 2011). In the four patients that we studied, we observed low levels of p27 expression in areas of the tumour that were Brachyury and KI67 positive (see Figure 6.1 and Figure 6.2 for representative image). The association between Brachyury and p27 seems to be consistent therefore. Further investigation is needed however to establish the exact mechanism by which Brachyury might regulate the cell cycle in colorectal cancer.

Only one other study has investigated the prognostic relevance of endocrine cells in colorectal cancer (Hamada et al., 1992). In this study, Chromogranin A (CHGA) IHC staining was used to estimate the survival of primary CRC patients stratified by the number of CHGA positive cells observed in patients tumours. The CHGA-EEC negative group of patients was compared with group II (less than one positive cell/mm^2) and group III (more than one positive cell/mm^2). CRC patients in group III had a significantly worse cumulative survival when compared with patients in group I

(negative for CHGA EECs) (*P<0.01*). This finding is consistent Brachyury positive EECs conferring a worse overall survival outcome on CRC patients in which these EECs are present. We do not know whether these patients are also CHGA positive, but we did observe that not all Brachyury positive EECs are ChgA positive in a small number of patients that we performed colocalisation studies on (Jezkova et al., 2016). In this study, we also observed a direct correlation between Brachyury and *CHGA* mRNA levels in normal tissues (Pearson $r > 0.2$) in all data sets studied. Also, we report that in one out of the four data sets studied, a direct association was observed between Brachyury and *CHGA* mRNA levels (Pearson $r>0.2$). It is likely that the mRNA levels in the small number of cells in the small number of patients that express Brachyury positive EECs are difficult to detect in total RNA whole tissue extracts (Jezkova et al., 2016). Single cell RNA profiling of individual EEC cells would be useful in determining the genes that define this secretory lineage cell and to help determine whether factors that they secrete to indeed contribute to normal adult intestinal stem cell and cancer stem cell niche. Hamada and coworkers failed to find a correlation between tumour differentiation and the number of CHGA positive cells, however, 5 tumours with the most CHGA positive cells were from poorly differentiated tumours. Additionally, morphologic changes such as nuclear hyperchromasia and pleomorphism also indicated that the CG-immunoreactive cells in the cancer tissue were malignant (Hamada et al., 1992).

It is important that further investigation is made into all three gut epithelial secretory lineage cells role in the development of colorectal cancer. Some compelling evidence exists of the role of Paneth cells in the small intestine (Sato et al., 2011) and Reg4 deep crypt secretory cells (Sasaki et al., 2016) in the colon providing the niche for adult intestinal stem cells. It is quite possible that these cells provide the niche for cancer stem cells in rapidly evolving tumours. It is interesting that Brachyuyry, when expressed in a member of the secretory lineage (EECs) in primary colorectal cancers, is present almost exclusively in undifferentiated tumours. More work is needed to clearly define the role of EECs in normal adult intestinal stem cell maintenance and colorectal cancers. Moreover, it would be interesting to establish whether post-mitotic cells of the secretory lineages can dedifferentiate in a

similar manner to committed Alpi^{+} enterocytes dedifferentiate to replenish Lgr5^{+}stem cells when they are ablated (Tetteh et al., 2016).

7 Determination of the effect and specificity of small molecule inhibitors designed to target Brachyury in cancer cells

7.1 Introduction

Given Brachyury's reported expression in several cancers (Du et al., 2014; Haro et al., 2013; Nelson et al., 2012; Palena et al., 2007; Presneau et al., 2011; Sangoi et al., 2011; Shao et al., 1997; Yang et al., 2009), its role in inducing the epithelial to mesenchymal transition (Du et al., 2014; Imajyo et al., 2012; Roselli et al., 2012; Shao et al., 2015) and its function in regulating cell cycle progression (Jezkova et al., 2014; Nelson et al., 2012), and therefore its emergence as a therapeutic cancer target. In collaboration with the drug discovery platform, Cardiff University, small molecule inhibitors (SMIs) were designed that were targeted to the human Brachyury protein. Inhibition of Brachyury has resulted in colorectal cancer cell cycle arrest (Jezkova et al., 2014), increased sensitivity of adenoid cystic carcinoma cells to chemotherapy and radiation (Kobayashi et al., 2014) and increased sensitivity of lung cancer cells to EGFR kinase inhibition (Roselli et al., 2012). Furthermore, in chordomas, in which Brachyury is overexpressed, erlotinib, an EGFR tyrosine kinase inhibitor, inhibits growth of patient-derived chordoma Xenografts that display nuclear Brachyury expression (Scheipl et al., 2016).

Much of the preclinical research into using Brachyury-based therapies has focused on using T-cell mediated cancer immunotherapies (Palena et al., 2007) and cancer vaccine (Heery et al., 2015) approaches. In the first approach, an epitope of human Brachyury was used to expand T lymphocytes from the blood of cancer patients and normal donors and demonstrated an ability to lyse Brachyury-expressing tumour cells (Palena et al., 2007). In a Phase I clinical trial of the second approach, a cancer vaccine targeting Brachyury was used. A Brachyury-specific T cell response in chordoma patients was observed with two patients exhibiting disease control. Furthermore, a patient with colorectal carcinoma who enrolled on the study with a large progressing pelvic mass remained in the study for 1 year or more with disease stability and decreased tumour density (Heery et al., 2015).

Much evidence points to the utility of developing a targeted treatment to inhibit Brachyury in cancers in which it is expressed especially as Brachyury is rarely mutated in cancers, but is overexpressed or duplicated in, for example, NSCLCs and chordomas (Roselli et al., 2012; Yang et al., 2009 respectively). Nonretractable lung cancers and chordomas that are untreatable due to EGFR mutations could be responsive to Brachyury-specific small molecule inhibitors.

Despite promising results from research into a Brachyury-targeted treatment, no *in vitro* or *in vivo* work has been reported on the development of small molecule inhibitors (SMIs) directed against Brachyury. Here we report the effect of several SMIs designed to target Brachyury on cell proliferation, cell viability and cytotoxicity in SW480 CRCs, Brachyury high cell line (H460) and Brachyury low cell line (NTERA2). The Cardiff School of Pharmacy' Computer-aided drug design platform (CADD) generated several compounds designed to inhibit binding motifs in the Brachyury protein that would effect its binding to functional complex and therefore impair its ability to function as a transcriptional regulator of proliferation and stemness in cancer. In a preliminary investigation, we tested these inhibitors for specificity using downstream transcriptional and phenotypic readouts discussed below to select promising candidate inhibitors.

7.2 Materials and methods (Specific to Brachyury SMIs)

7.2.1 Proliferation Analysis

For inhibitor experiments, H460 and NTERA2 cells were plated at $1x10^5$ cells/ml in 6-well plates and allowed to attach and recover. Cells were treated with the desired concentration of inhibitor (10 µM). Cells were harvested at day 1, 3 and 5 with RNA being extracted at each timepoint. Viable cells were counted using a BioRad TC20 Automated Cell Counter with the trypan blue exclusion method. Cell counts were plotted for comparison between control (DMSO) and siRNA/ SMI-treated SW480 cells at days 1, 3 and 5.

7.2.2 Preparation of RNA for transcriptional analysis

SW480, and NTERA2 cells were plated at a density of $1x10^5$ cells/ml and H460 at a density of $0.5x10^5$cells/ml H460 in a 6-well plate, cells were treated with 10µM SMIs on day 0 only and Brachyury siRNAs (T5/05) every day. Cells were harvested after three days by trypsinisation and RNA was extracted (see section 2.5.1). qRT-PCR was performed as detailed in 2.6 *HSP09A* and *GAPDH* used form normalisation.

7.2.3 HeatMap Analysis

Gene expression was normalised to the housekeeping genes *HSP09A* and *GAPDH*. Mean normalised fold expression values for each inhibitor and inhibitor concentration were tabled and pasted into QCanvas tool for data clustering and visualisation of genomic data (Kim et al., 2012). Mean normalised expression values for samples were subtracted from controls (non-interfering siRNA for siRNA treated samples and DMSO for SMI treated samples). Subsequently a HeatMap that corresponded to the differences in expression between samples and control was generated. Negative values were therefore samples that increased in expression compared to control and positive values were those that decreased in expression compared to control.

7.2.4 Cell Viability Assay

To assess cell viability after inhibitor treatment the RealTime-GloTM MT Cell Viability Assay (#G9713) from Promega was used. Cells were plated in a white assay 96 well plate at optimal seeding density in 80 µl media (10000 cells (SW480), 2500 cells (H460) and 10000 cells (NTERA2) (80 ul per well). For Brachyury knockdown; 0.3µl of T5 siRNA (20µM), 0.3µl HiPerfect and 19.4µl serum-free media was added to each well. For treatment with inhibitors; drug was diluted appropriately to achieve desired concentration in well. For example; for a 10µM working solution, 10mM SMI stock solution was diluted 1:1 in DMSO, 1µl of this solution was diluted with 99µl of media and 20µl of this was added to the well containing 80µl media and cells, making a total volume of 100µl in each well. For each treatment three wells were used. Cells were incubated at 37°C (5%CO_2 for H460 and SW480 and 10%CO_2 for NTERA2) for 72 hours.

MT Cell Viability Substrate and NanoLuc® Enzyme and cell culture medium was equilibrated to 37°C. 2X RealTime-GloTM reagent was prepared by diluting the MT Cell Viability Substrate and NanoLuc® enzyme in 37°C cell culture medium to a 2X concentration for each reagent. For example, to prepare 1ml of 2X RealTime-GloTM Reagent, 2µl of MT Cell Viability Substrate 1,000X, and 2µl of NanoLuc® Enzyme, 1,000X, was added to 996µl of cell culture medium to form the 2X RealTime-GloTM Reagent. Mix well with a vortex mixer. An equal volume of 2X RealTime-GloTM Reagent to the cells (100 µl per well). Cells were incubated for 60 minutes in the cell culture incubator with the enzyme and substrate. During this incubation a Perkin Elmer Wallac 1420 Victor2 Microplate Reader - GMI Inc was pre-warmed to 37°C. Relative luminescence was measured after 1 hour. DMSO, equivalent to the maximum concentration was used as a negative control with the detergent NP-40 acting as a negative control to ensure the cell viability assay was working.

7.2.5 IC_{50} determination

IC_{50} values were calculated using GraphPad Prism 6. X values were the logarithms of concentrations (10-100µM in 10µM increments) and Y values were the mean relative luminescence of the samples. Non-linear regression dose-response curves were generated with IC_{50} values.

7.3 Effect of SMIs on the expression of transcripts that are potentially regulated by Brachyury

Previously, an RNASeq experiment was performed by our group on RNA from SW480 CRC cells treated with Brachyury depleting siRNA (T5). Expression levels of 1786 genes changed significantly in T5 treated cells compared to those treated with control (non-interfering siRNA). 8 of these genes (Table 7.1) that have different reported cellular functions and change significantly with Brachyury knockdown were chosen to establish how their expression changed in response to treatment with several small molecule inhibitors that were designed to target Brachyury.

Table 7.1 List of 8 Brachyury responsive genes chosen for qRT-PCR analysis of SW480 cells treated with SMIs designed to target Brachyury. Average transcript counts (Avg. count T5) for T5 siRNA treated and non-interfering control (Avg. count NI) treated SW480 CRCs are shown with pairwise comparisons between each (Ad. Pval) for each gene. The expression levels of these genes changed significantly in SW480 cells treated with T5 siRNA compared with those treated with control siRNA. The expression levels of some of these genes, which have different reported cellular functions, might be expected to change if particular Brachyury functions were inhibited by Brachyury SMIs.

Ensembl ID	Gene. Name	Ad. Pval	Avg. count NI	Avg. count T5	Function
ENSG00000105173	CCNE1	<0.000001	1167.66	474.33	Cell Cycle (Nakayama, et al., 2010)
ENSG00000120057	SFRP5	<0.000001	402	1501	Wnt signaling (Lavergne et al., 2011)
ENSG00000146674	IGFBP3	<0.000001	1534.33	3742.66	Proliferation (Yancu et al., 2017)
ENSG00000095002	MSH2	<0.000001	7795	2564.66	DNA Repair (Moller & Wijnen, 2016)
ENSG00000106484	MEST	<0.000001	5380.66	1448	Metabolism (Takahashi et al., 2005)
ENSG00000005469	CROT	<0.000001	482.66	1434.33	Mitochondrial (Le Borgne et la., 2011)
ENSG00000065911	MTHFD2	<0.000001	7373.66	2253	Mitochondrial (Nilsson et al., 2014)
ENSG00000116688	MFN2	<0.000001	10089	3699	Mitchondrial (Chen et al., 2010)

We measured the expression of Brachyury siRNA (T5) responsive genes (Table 7.1), by qRT-PCR in SW480 cells treated with several SMIs designed to target Brachyury. qRT-PCR analysis would provide some indication of the specificity of the SMIs by looking for inhibitors that showed some similar downstream gene expression changes to Brachyury siRNA-mediated depletion. In order to visualise the transcript levels across all 8 genes from Table 7.1, a heat map was generated to show the normalised fold change expression of each gene, when SW480 cells were treated with 10 µM of each inhibitor (SMIs represented by 1-19 in Figure 7.1). siRNA-mediated knockdown of Brachyury using siRNA (T5) (Figure 7.1) (used for RNAseq analysis) was consistent with the gene expression changes we observed for all of the selected genes from the RNAseq (Table 7.1). Although we did not expect Brachyury transcript depletion, we did observe a reduction in Brachyury transcript in cells treated with SMI 3 and its analogue 3.3 along with SMI 18 (Figure 7.1).

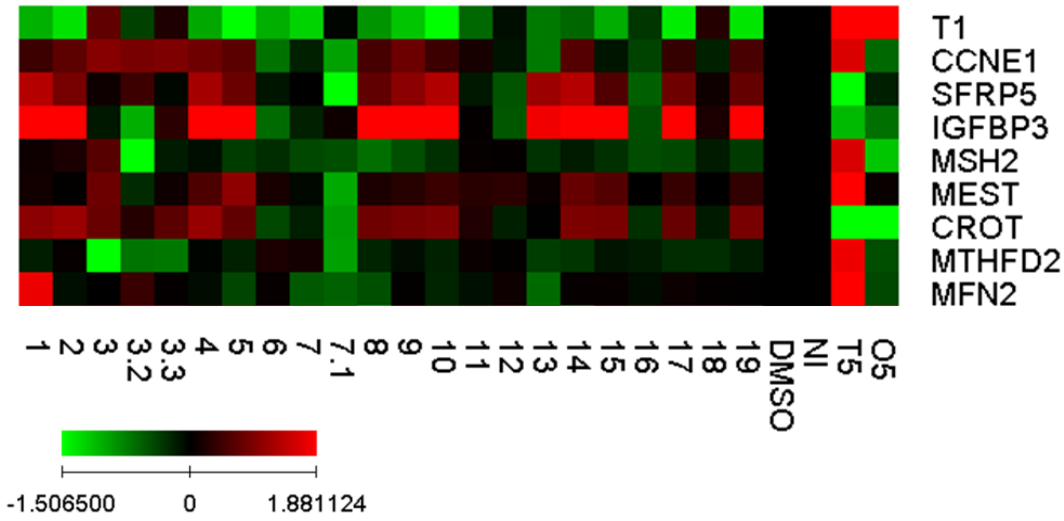

Figure 7.1 qRT-PCR Heat Map of Brachyury responsive genes following treatment with SMIs. On the x axis, each SMI is labelled 1-19, including structural analogues of inhibitors 3 and 7 (3.2, 3.3 and 7.1). The vehicle control (DMSO), followed by the scrambled siRNA (NI) and Brachyury siRNAs (T5 and 05) are also shown for comparison. On the y-axis, Brachyury (T1) is followed by each gene from Table 7.1 with their expression, measured by qRT-PCR, in each sample, represented by a heat map colour. Red (uppermost normalised expression 1.881) represents a reduction in normalised expression (normalised to housekeeping genes *HSP09A* and *GAPDH*) relative to control (DMSO for inhibitors and NI for siRNAs) and green represents an increased normalised expression (uppermost 1.506) relative to control (see key bottom left).

7.4 The effect of Brachyury-targeted SMIs on cell proliferation

Given that SMIs 3 and 3.3 had the effects of changing gene expression patterns of *CCNE1*, *MSH2* and *MEST* in a similar manner to SW480 cells treated with siRNA T5, and also that inhibitor 18 reduced expression of Brachyury (*T1*) (Figure 7.1), we chose these inhibitors to perform further cell proliferation and cell viability experiments. Inhibitors 3 and 3.3 were of particular interest as they are structural analogues and also exhibited some similar gene expression changes (Figure 7.1) to the some genes listed in Table 7.1.

Cell proliferation experiments were performed separately in a Brachyury high expressing cell line (H460) (untreated average cT values of 22) (Figure 7.2) and a Brachyury negative cell line (NTERA2) (Figure 7.3) (see Figure 7.4 B for cT values, average cT for Brachyury of 37). This would allow us to determine Brachyury-specific effects of the SMIs. Cells were plated at the densities described in section 7.2.1 and harvested after 1, 3 and 5 days, after which RNA was extracted. RNA was used for qRT-PCR analysis of Brachyury levels (Figure 7.4). Normalised fold Brachyury expression values corresponding to each of the samples for H460 (Figure 7.4 A) and cT values are shown for NTERA2 (Figure 7.4 B). Significant differences in the number of cells were observed between H460 cells treated with SMIs 3, 3.3,18 and siRNA T5 at day 5 compared with controls (DMSO and non-interfering siRNA respectively) (Figure 7.2). Low cT values of, on average 22, were observed in H460 samples confirming this cell lines high expression of Brachyury. H460 expression was normalised to housekeeping gene *HSP09A* to show relative expression between treated samples and controls (Figure 7.4 A). NTERA2 cell counts at day 5 of cells treated with SMIs/siRNA T5 did not differ significantly from DMSO (Figure 7.3). High cT values for Brachyury of on average 37 (Figure 7.4 B) (a similar cT value to no-template controls) were consistently observed in NTERA2, this being consistent with low/negative expression of Brachyury in this cell line. Furthermore, earlier IF experiments had confirmed no Brachyury protein was observed in NTERA2 cells.

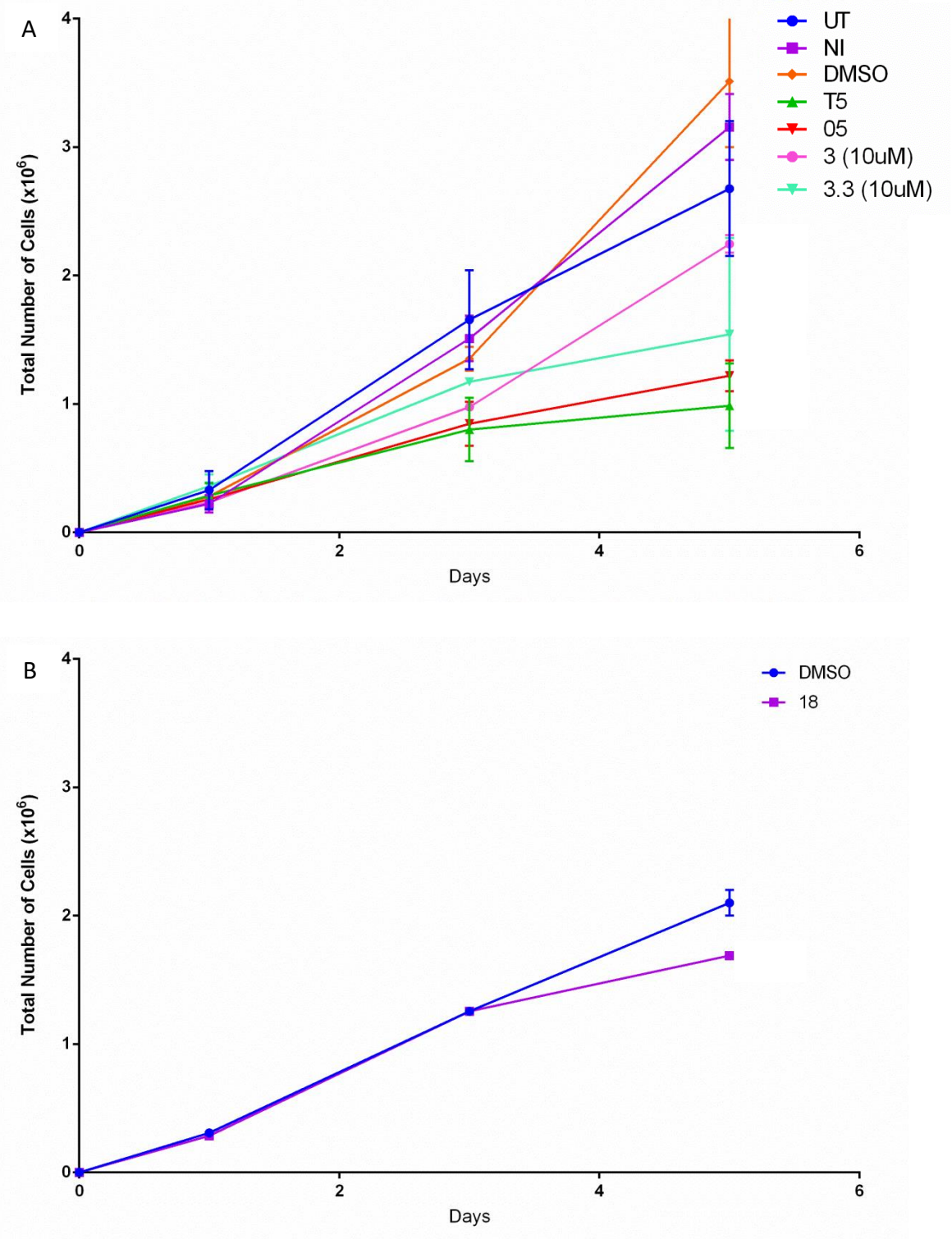

Figure 7.2 H460 cell proliferation assay of cells treated with Brachyury siRNAs and Brachyury SMIs. (A) H460 cells treated with Brachyury siRNAs T5 and 05/Brachyury SMIs 3 and 3.3 were counted at days 1, 3 and 5 after treatment and compared with controls non-interfering siRNA (NI) and DMSO respectively, untreated H460 (UT) is included as a reference for H460 proliferation over 5 days. Pairwise comparisons were made between siRNA/Brachyury SMI treatments their respective controls for day 5 cell counts only.. (B) H460 cells were treated with Brachyury SMI 18 in a separate cell proliferation assay and were compared to control (DMSO treated cells).

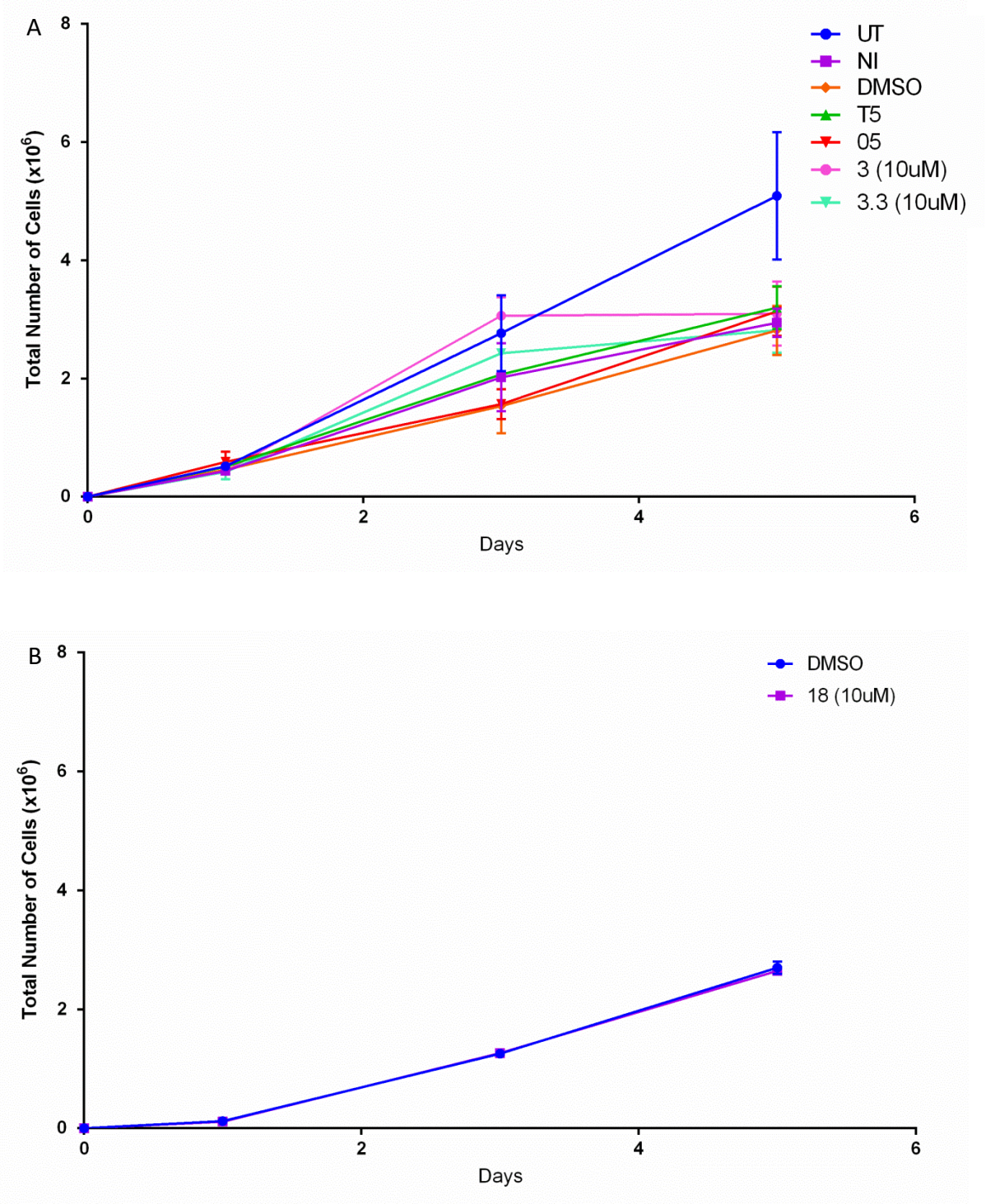

Figure 7.3 NTERA2 cell proliferation assay of cells treated with Brachyury siRNAs and Brachyury SMIs. (A) NTERA2 cells treated with Brachyury siRNAs T5 and 05/Brachyury SMIs 3 and 3.3 were counted at days 1, 3 and 5 after treatment and compared with controls non-interfering siRNA (NI) and DMSO respectively, untreated NTERA2 (UT) is included as a reference for NTERA2 proliferation over 5 days. Pairwise comparisons were made between siRNA/Brachyury SMI treatments their respective controls for day 5 cell counts only. (B) NTERA2 cells were treated with Brachyury SMI 18 in a separate cell proliferation assay and were compared to control (DMSO treated cells).

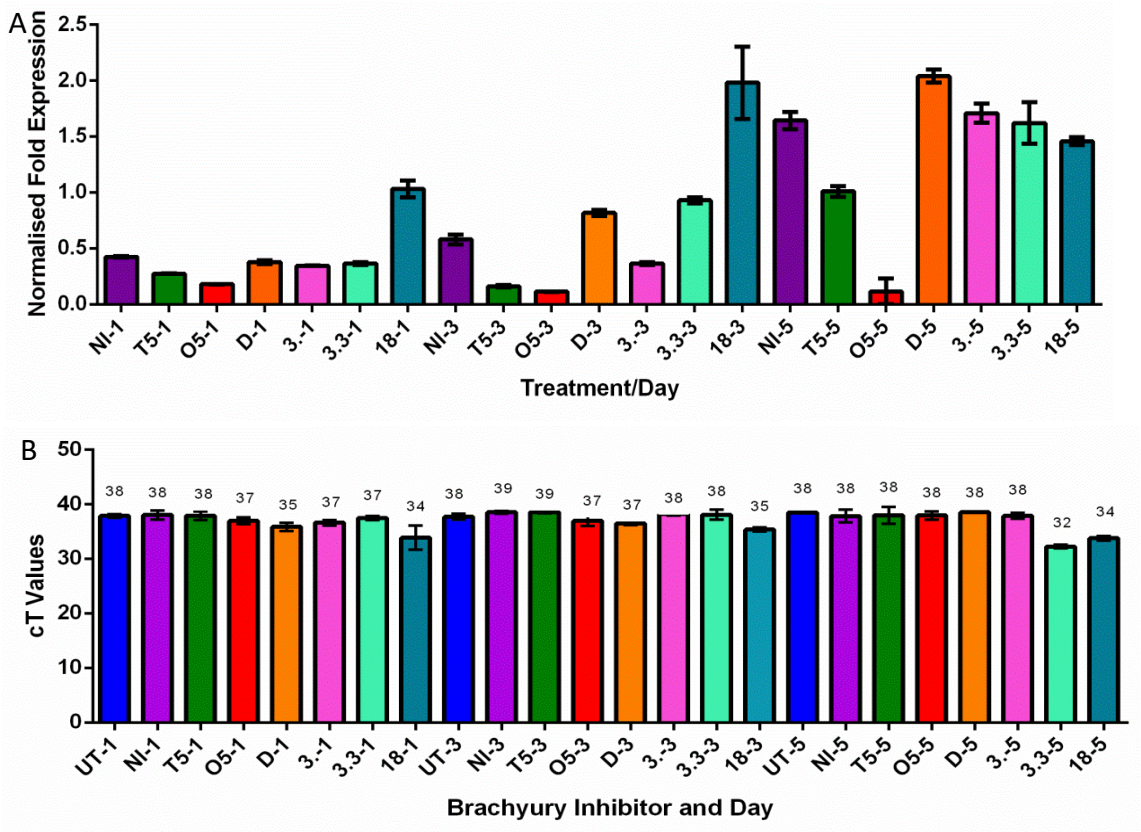

Figure 7.4 qRT-PCR analysis of cells used for cell proliferation assay RNA. qRT-PCR was performed on the reverse transcribed RNA from each day and treatment of the cell proliferation assay (Figure 7.3). (A) qRT-PCR was performed on this reverse transcribed RNA from H460 cells with Brachyury expression normalised to the ubiquitously expressed *HSP09A* and *GAPDH*. (B) NTERA 2 Brachyury expression is given as cT values as expression was similar to that of no template control (38) in NTERA2. (UT = Untreated cells NI = Non-interfering siRNA, T5 = Brachyury siRNA, 05 = Brachyury siRNA, D = DMSO, 3, 3.3 and 18 = SMIs followed by day RNA was extracted, e.g. D-1 = DMSO –Day 1).

7.5 The effect of SMIs on cell viability in Brachyury positive versus Brachyury negative cell lines.

To further confirm the effect of the small molecule inhibitors on cell viability we used a luciferase-based cell viability assay to determine the viability of cells treated with siRNAs/SMIs compared with non-interfering siRNA and DMSO controls. Initially a concentration of 10µM of inhibitors was used to treat SW480 colorectal cancer cells. siRNA T5, SMIs 3, 3.3 and 18 showed a significant reduction in cell viability compared to DMSO treated cells (Figure 7.5).

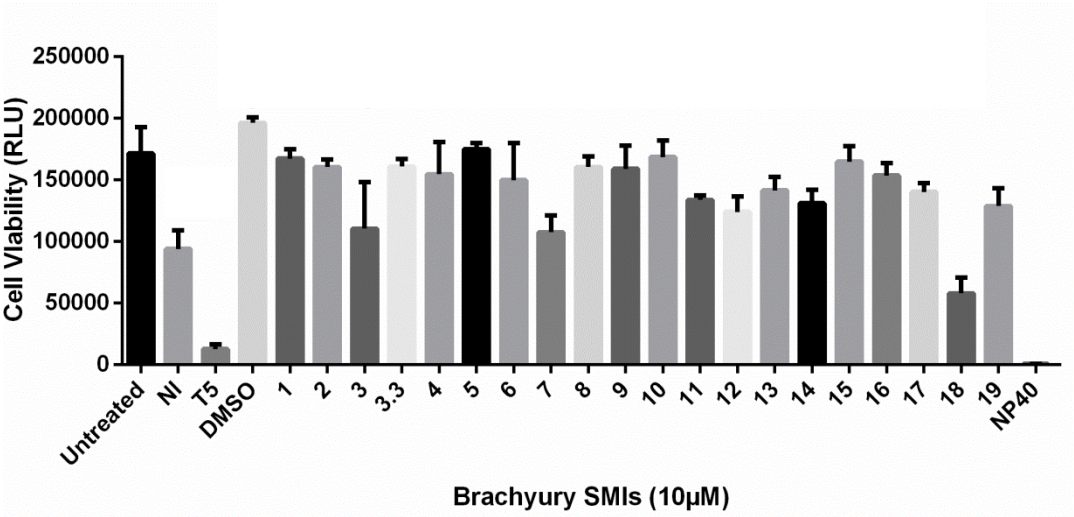

Figure 7.5 Effect of SMIs on cell viability in SW480 colorectal cancer cells. Small inhibitors 1-19 and Brachyury siRNA T5 were used to treat SW480 cells at a concentration of 10µM, their relative luminescence unit (RLU) readings were compared with DMSO (vehicle) and non-interfering (NI) siRNA controls respectively. NP40 is the negative control for the cell viability assay .

After the cell proliferation assay (Figure 7.3) confirmed inhibitor 3 and its analogue 3.3 along with inhibitor 18 had a significant effect on cell proliferation in Brachyury high H460, we performed cell viability analyses on this and the Brachyury negative NTERA2 cell line (Figure 7.7). Titrations of each drug were performed to establish the IC_{50} value of each. The relative luminescence (RLU) of each well in triplicate is a proxy measure of cell viability, these values were plotted against the Log_{10} concentration (µM) of SMI (Figure 7.7). IC_{50} values were determined as the concentration required to observe a 50% decrease in relative luminescence compared with the maximum relative luminescence. Dose-response curves and IC_{50} values were calculated for each individual experiment via sigmoidal dose-response analysis using the Hill fitting equation. IC_{50} values are shown for 3, 3.3 and 18 for H460 and NTERA2 along with

the sigmoidal curves ((log conc.µM vs relative luminescence (cell viability)) in Appendix F. A noticeable increase in IC_{50} concentrations (plotted in Figure 7.6) are observed in the Brachyury-low NTERA2 compared with Brachyury-high H460.

Appendix F

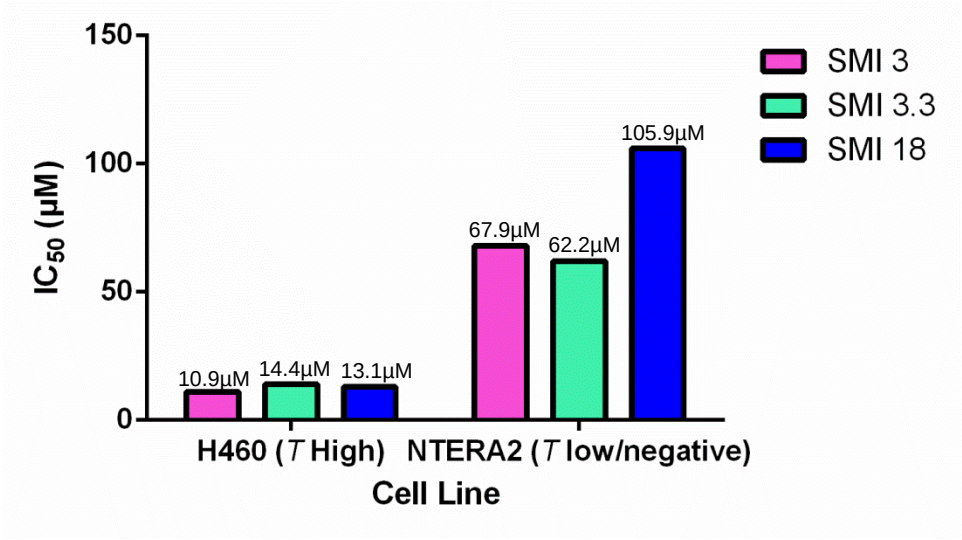

Figure 7.6 Histogram showing IC_{50} values for SMIs in H460 and NTERA2. IC_{50} values are plotted for comparison between Brachyury high H460 and Brachyury low NTERA2.

7.6 Discussion

In this preliminary *in vitro* study we assayed the specificity of several inhibitors designed to inhibit Brachyury function in cancer cells. The expression of several genes that are associated with different functions that changed significantly in response to a Brachyury knockdown was measured. In this preliminary study, we observed some similar changes in gene expression when SW480 cells were treated with small molecule inhibitors in as we observed when they were treated with Brachyury siRNA T5. Although the downstream transcriptional profile upon treatment of SW480 of no one inhibitor matched exactly that of the T5 siRNA treatment some genes were regulated similarly. Below we discuss some genes of interest that were similarly regulated, although further investigation is needed to establish the mechanism and specificity of inhibition.

Transcriptional changes similar to Brachyury siRNA (T5) depletion were observed in the cell cycle related gene CyclinE1 *CCNE1*. *CCNE1* is an activating subunit of cyclin dependent kinase 2 (*CDK2*) (Nakayama et al., 2010). Constitutive expression of this gene results in accelerated entry into S phase of the cell cycle (Ohtsubo et al., 1995; Resnitzky et al., 1994). Furthermore it has been implicated in cisplatin resistance and is amplified, contributing to poor survival in ovarian cancer (Etemadmoghadam et al., 2010; Nakayama et al., 2010). Similar levels of reduction in *CCNE1* were observed for inhibitors 3 and 3.3 in SW480 cells to siRNA-mediated Brachyury (T5) depletion (Figure 7.1) with a subsequent reduction in cell viability (Figure 7.5) which is dependent on efficient cell cycling. This suggests that Brachyury's effect on the viability of cancer cells is a result of, in part, its downregulation of *CCNE1*.

Reduction in the expression of DNA mismatch repair gene *MSH2* was observed also when SW480 cells were treated with inhibitors 3 and 3.3 (Figure 7.1). These changes were similar to the reduction observed for Brachyury siRNA-treated cells.
Inactivation of *MSH2* in mice results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer (de Wind et al., 1995). Furthermore, *MSH2* has been associated with hereditary nonpolyposis colon cancer (HNPCC) and with patients harbouring loss of function mutations at a higher risk of

developing HNPCC (Fishel et al., 1993; Moller et al., 2016). This reduction in *MSH2* expression and impaired DNA repair also likely contributes to the reduction in cell viability observed in cells treated with inhibitors 3 and 3.3.

Another transcript that changed significantly in response to analogues 3, 3.3 and Brachyury siRNA T5 was *MEST* (Mesoderm-specific transcript). MEST is imprinted with the paternal allele being expressed (Takahashi et al., 2005). Loss of *MEST* in mESCs results in growth retardation of a developing mouse foetus (Lefebvre et al., 1998). It has been implicated in metabolic functions and disease, with enhanced expression being observed in the adipose tissue of mice with diet-induced and genetically caused obesity, furthermore, transgenic mice overexpressing *MEST* in adipose tissue, exhibit enhanced expression of the adipose genes. Moreover, adipocytes were markedly enlarged in these transgenic mice (M. Takahashi et al., 2005). Brachyury-dependent reduction in the transcript levels of *MEST* in SW480 cells likely contributes to impaired metabolic function and cell viability.

An inverse expression pattern is observed in the mitochondrial genes *CROT*, *MTHFD2* and *MFN2* compared with siRNA-mediated depletion (Figure 7.1). It is possible that these inhibitors target Brachyury in such a way that they alter Brachyury regulation in a gene/function-specific manner, i.e. occupying a site that would inhibit Brachyury forming a complex with one co-regulating protein but leaving others unaffected. More work is therefore needed to establish where these SMIs bind exactly and how this might affect Brachyury function with potential cancer-specific interacting partners.

Small molecule inhibitors with low IC_{50} values (inhibitor concentration for half-maximum response) are more specific, with cytotoxicity values in the nM scale for some kinase inhibitors being observed (Zhang et al., 2009). For example one compound FMK (Fluoro methyl ketone inhibitor) inhibits RSK1 (ribosomal S6 kinase) with an IC_{50} of 15nM. In contrast, the non-specific tyrosine kinase inhibitor PP1 has an IC50 of 1.2µM against RSK (Zhang et al., 2009). The former is an irreversible inhibitor and these experiments examined specifically kinase inhibition and did not investigate cytotoxicity (Cohen et al., 2005). When the cytotoxicity of 5-Fluorouracil (a small molecule inhibitor of tymidylate synthase and common cancer treatment) was examined by performing a dose-response experiment using a similar cell viability

experiment as seen in Figure 7.5 as the response, IC_{50} values of 28.7µM were observed in HCT116 colorectal cancer cells (Friedrich et al., 2007). In another study testing the inhibition of growth observed when B cell lymphoma HL-60 cells were treated with putative Bcl-2 inhibitors, the most potent inhibitor exhibited an IC50 value of 4µM (Enyedy et al., 2001). We observed IC50 values of 10.89µM, 14.35µM and 13.1µM for Brachyury-high expressing cell line H460 treated with inhibitors 3, 3.3 and 18 respectively. This, together with significantly higher IC50 values observed for the Brachyury negative cell line NTERA2, suggests that these compounds have a Brachyury-specific effect on cell growth, further studies should include samples which have been treated with both Brachyury siRNA and small molecule inhibitors simultaneously, this will ensure Brachyury expression is minimised especially in cell lines which have a heterogeneous differentiation status *in vitro*. Furthermore the similar expression of some Brachyury responsive genes observed in 3 and 3.3 SMI treated SW480 cells to siRNA-mediated (T5) Brachyury knockdown suggests these inhibitors inhibit Brachyury's transcriptional function.

To address the effect of these inhibitors on Brachyury function and their specificity, future experiments may include liquid chromatography mass spectroscopy (LC-MS) to determine binding sites on Brachyury (Mcmanus et al., 2016) along with co-immunoprecipitation experiments to identify Brachyury interacting partners that are cancer specific that might be inhibited from complexing with Brachyury by these SMIs at specific sites.

8 Summary and Final Discussion

8.1 Conclusion and General Discussion

Here, we show for the first time that Brachyury has two transcript variants that are expressed throughout colorectal cancer cells, normal tissue and colorectal cancer tissue.

We detect Brachyury's expression in the secretory EECs of histologically distinct adenomas. This extends current studies into the molecular mechanisms that manifest themselves in the morphological transformation from normal to adenoma to cancer tissue. *LGR5* mRNA expression was recently reported in a small number of cells at the base of normal human colonic crypts, resembling those of mouse models. Conventional adenomas widely express high levels of *LGR5* although there is no evidence that the cellular hierarchy of the normal colonic crypt is maintained in conventional human adenomas. Serrated adenomas however did display basal *LGR5* mRNA suggesting this lesion resembles normal stem cell hierarchy (A.-M. Baker et al., 2015), indeed we observed basal Brachyury positive EECs in serrated adenomas resembling their distribution in normal crypts. Clearly there are differences between these histologically and morphologically distinct adenomas in terms of their stem cell hierarchy and architecture. Brachyury's expression in the EECs of each adenoma and its known expression in cancer, suggests it plays an important function in providing a niche for adult intestinal stem cells and their adenoma counterparts on their way to cancer progression.

Separately we observed a regulatory relationship between Brachyury and the purported +4 stem cell marker LRIG1 in the normal colon and colorectal cancer cells. We observed overlap in the signals of antibodies directed against Brachyury and LRIG1 in cells of the normal human crypt consistent with the +4 position. Our results also suggest that this relationship is transcript variant specific (i.e. LRIG1 is responsive to TV2). We show that Brachyury regulates a key receptor tyrosine kinase in CRC cells that has been shown previously to mediate growth signals in NSCLCs and chordomas in which Brachyury is overexpressed. In the context of the development

of a Brachyury-targeted therapy, inhibition of Brachyury has resulted in colorectal cancer cell cycle arrest (Jezkova et al., 2014), increased sensitivity of adenoid cystic carcinoma cells to chemotherapy and radiation (Kobayashi et al., 2014) and more importantly increased sensitivity of lung cancer cells to EGFR kinase inhibition (Roselli et al., 2012). Furthermore, in chordomas, in which Brachyury is overexpressed, erlotinib, an EGFR tyrosine kinase inhibitor, inhibits growth of patient-derived chordoma Xenografts that display nuclear Brachyury expression (Scheipl et al., 2016). Given the relationship between Brachyury and LRIG1 (known to downregulate EGFR family members), this provides a potential druggable mechanism of interest to cancer therapy, especially in EGFR intractable cancers. It is possible that combinatorial therapies that inhibit both proteins could be beneficial to appropriately stratified patient groups. It is important that further studies are carried out, including chromatin immunoprecipitation analysis, to establish whether Brachyury directly regulates *LRIG1* in the normal colon and colon cancer cells *in vitro* in which both are expressed. Nonretractable lung cancers and chordomas that are untreatable due to EGFR mutations could be responsive to Brachyury-specific small molecule inhibitors.

Only one other study has investigated the prognostic relevance of endocrine cells in colorectal cancer (Hamada et al., 1992). In this study, Chromogranin A (CHGA) IHC staining was used to estimate the survival of primary CRC patients stratified by the number of CHGA positive cells observed in patient's tumours. The CHGA-EEC negative group of patients was compared with group II (less than one positive cell/mm^2) and group III (more than one positive cell/mm^2). CRC patients in group III had a significantly worse cumulative survival when compared with patients in group I (negative for CHGA EECs) ($P<0.01$). Given Brachyurys reported role in the aetiology of several cancers and the expression we observed in EECs, we have uncovered a new relevance to the study of these cells in cancer. Interestingly, Brachyury positive EECs and EEC-like cells are present at every stage of the well-defined sequence from the normal colonic crypts to adenomas and finally carcinomas. Furthermore, when these EEC-like cells are present in primary colorectal carcinomas, they confer poor survival and disease-free survival outcomes in the patients in which they are present. To further investigate these findings a large-scale study should be performed

staining this time with both Brachyury, CHGA, LRIG1 and EGFR to establish whether $CHGA^-$ $Brachyury^+$ EEC cells in colorectal tumours affect the survival outcomes of patients in which they are observed to establish whether patients with these Brachyury positive EECs are suitable for a trial with Brachyury and EGFR inhibitors. Here we extend the investigation of the transcription factor Brachyury's role in the aetiology of CRC. Previous work by our group had suggested that Brachyury confers cancer stem cell characteristics of colorectal cancer cells (Sarkar, Shields, Davies, Müller, & Wakeman, 2011) and enables proliferation in CRC cells (Jezkova et al., 2014). These studies confirm previous studies of Brachyury's influence on proliferation in other cancer cells (Pinto et al., 2014; Presneau et al., 2011; Shimoda et al., 2012). Furthermore, evidence for Brachyurys role in inducing the EMT has been reported by several groups also (Du et al., 2014; Fernando et al., 2010; Imajyo et al., 2012; Pinto et al., 2014).

Recent studies have reported the capacity of committed differentiated cells in the intestinal crypts to revert back to the Lgr5 adult intestinal stem cells when the latter are lost (Tetteh et al., 2016). Furthermore, a population of secretory precursors, Paneth or EEC, in the small intestine, act as multipotent stem cells when crypts were damaged (Buczacki et al., 2013). With another study highlighting the ability of a high fat diet to modulate the function and number of ISCs and their capacity to initiate tumours (Beyaz et al., 2016). The evidence we present here, suggests that Brachyury is expressed in a post-mitotic cell type in the normal colon crypt that might have reserve stem cell characteristics and could contribute to their plasticity. Furthermore, evidence from an RNAseq performed by our group suggested Brachyury has function in somatic stem cell maintenance.

Brachyury is an emerging biomarker and anti-cancer therapeutic target (Hamilton et al., 2013; Heery et al., 2015; Palena, et al., 2007). Our observations in this study suggest that Brachyury also has a function in normal tissue homeostasis and that this function might be maintained in tumourigenesis at every stage of the well-defined adenoma to carcinoma sequence. Due to the plasticity of the normal gut epithelium, and the ability of other post-mitotic cell types to repopulate the stem cell niche, Brachyury remains a viable therapeutic target worthy of further investigation

especially given the association established between Brachyury and EGFR through LRIG1 in this study.

8.2 Future directions

The results described in this thesis have laid a foundation for continued research and therefore we propose suggestions for further investigations.

In vivo analysis using an inducible Cre-*loxP* system, in order to perform intestinal specific deletion of Brachyury and look at what effect this has on intestinal homeostasis. Furthermore, given that Lgr5+ ISCs produce enteroendocrine cells when EGFR/Wnt/Notch are inhibited *in vitro* (Basak et al., 2016), one could see if these cells express Brachyury at the RNA and protein level and also deplete/knockout Brachyury expression in them and assay their ability to form tumour organoids. Furthermore, if these cells did generate tumour organoids, these could be implanted into mice to compare both Brachyury and Brachyury KO organoids.

A lineage tracing system, similar to that used by Barker and colleagues (2007) could be used to trace Brachyury expression in the intestinal crypt. Ablation of Lgr5 CBCs could be used to see if post-mitotic Brachyury and Chromogranin A positive EECs can repopulate the stem cell niche and establish whether or not they do act as plastic reserve stem cells. Furthermore, *APC* deletions in Brachyury positive EECs ability to form adenomas could also be assayed using a similar experiments to those performed by Schepers et al., 2012.

We have demonstrated the effect of siRNA-mediated depletion of Brachyury on cancer cell proliferation. If therapeutic RNAi delivery methods become viable as a treatment method, Brachyury will be a suitable target in several cancers in which it is expressed. Further *in vitro* work is needed on the small molecule inhibitors that we have developed. Although we have demonstrated that these SMIs have an effect on the proliferation and viability of cancer cells, future work should focus on biochemical analyses to establish the effect these SMIs have on the ability of Brachyury to form complexes with interacting proteins that may influence its ability to regulate downstream target genes in cancer cells.

9 References

Basak, O., Beumer, J., Wiebrands, K., Seno, H., van Oudenaarden, A., & Clevers, H. (2016). Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells. *Cell Stem Cell*, 1–14. https://doi.org/10.1016/j.stem.2016.11.001

Baum, B., Settleman, J., & Quinlan, M. P. (2008). Transitions between epithelial and mesenchymal states in development and disease. *Seminars in Cell & Developmental Biology*, *19*(3), 294–308. https://doi.org/10.1016/j.semcdb.2008.02.001

Beyaz, S., Mana, M. D., Roper, J., Kedrin, D., Saadatpour, A., Hong, S.-J., … Yilmaz, Ö. H. (2016). High-fat diet enhances stemness and tumorigenicity of intestinal progenitors. *Nature*, *531*(7592), 53–58. https://doi.org/10.1038/nature17173

Buczacki, S. J. a, Zecchini, H. I., Nicholson, A. M., Russell, R., Vermeulen, L., Kemp, R., & Winton, D. J. (2013). Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature*, *495*(7439), 65–9. https://doi.org/10.1038/nature11965

Chen, H., Vermulst, M., Wang, Y. E., Chomyn, A., Prolla, T. A., McCaffery, J. M., & Chan, D. C. (2010). Mitochondrial fusion is required for mtdna stability in skeletal muscle and tolerance of mtDNA mutations. *Cell*, *141*(2), 280–289. https://doi.org/10.1016/j.cell.2010.02.026

Du, R., Wu, S., Lv, X., Fang, H., Wu, S., & Kang, J. (2014a). Overexpression of brachyury contributes to tumor metastasis by inducing epithelial-mesenchymal transition in hepatocellular carcinoma, 1–11.

Du, R., Wu, S., Lv, X., Fang, H., Wu, S., & Kang, J. (2014b). Overexpression of brachyury contributes to tumor metastasis by inducing epithelial-mesenchymal transition in hepatocellular carcinoma. *Journal of Experimental & Clinical Cancer Research*, *33*(1), 105. https://doi.org/10.1186/s13046-014-0105-6

Fernando, R. I., Litzinger, M., Trono, P., Hamilton, D. H., Schlom, J., & Palena, C. (2010). The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells, *120*(2), 533–544. https://doi.org/10.1172/JCI38379.(10).

Hamilton, D. H., Litzinger, M. T., Jales, A., Huang, B., Fernando, R. I., Hodge, J. W., …

Palena, C. (2013). Immunological targeting of tumor cells undergoing an epithelial- mesenchymal transition via a recombinant brachyury-yeast vaccine ABSTRACT :, *4*(10).

Heery, C. R., Singh, B. H., Rauckhorst, M., Marté, J. L., Donahue, R. N., Grenga, I., … Gulley, J. L. (2015). Phase I Trial of a Yeast-Based Therapeutic Cancer Vaccine (GI-6301) Targeting the Transcription Factor Brachyury. *Cancer Immunology Research*, *3*(11), 1248–56. https://doi.org/10.1158/2326-6066.CIR-15-0119

Imajyo, I., Sugiura, T., Kobayashi, Y., Shimoda, M., Ishii, K., Akimoto, N., … Mori, Y. (2012). T-box transcription factor Brachyury expression is correlated with epithelial-mesenchymal transition and lymph node metastasis in oral squamous cell carcinoma. *International Journal of Oncology*, *41*(6), 1985–1995. https://doi.org/10.3892/ijo.2012.1673

Jezkova, J., Williams, J. S., Jones-hutchins, F., Sammut, S. J., Gollins, S., Cree, I., … Jane, A. (2014). Brachyury regulates proliferation of cancer cells via a p27 Kip1-dependent pathway, *1*.

Journal, I., & Disease, C. (2011). Overall survival after resection for colon cancer in a national cohort study was adversely affected by TNM stage , lymph node ... Overall survival after resection for colon cancer in a national cohort study was adversely affected by TNM stage , lymph node ratio , gender , and old age, (May). https://doi.org/10.1007/s00384-011-1244-2

Kilic, N., Feldhaus, S., Kilic, E., Tennstedt, P., Wicklein, D., Wasielewski, R. Von, … Schumacher, U. (2011). Brachyury expression predicts poor prognosis at early stages of colorectal cancer. *European Journal of Cancer (Oxford, England : 1990)*, *47*(7), 1080–5. https://doi.org/10.1016/j.ejca.2010.11.015

Klaus, A., & Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. *Nature Reviews. Cancer*, *8*(5), 387–398. https://doi.org/10.1038/nrc2389

Le Borgne, F., Ben Mohamed, A., Logerot, M., Garnier, E., & Demarquoy, J. (2011). Changes in carnitine octanoyltransferase activity induce alteration in fatty acid metabolism. *Biochemical and Biophysical Research Communications*, *409*(4), 699–704. https://doi.org/10.1016/j.bbrc.2011.05.068

Moller, P., & Wijnen, J. T. (2016). MSH2 Mutation Carriers Are at Higher Risk of

Cancer Than MLH1 Mutation Carriers : A Study of Hereditary Nonpolyposis Colorectal Cancer Families, *19*(20), 4074–4080.

Nilsson, R., Jain, M., Madhusudhan, N., Sheppard, N. G., Strittmatter, L., Kampf, C., … Mootha, V. K. (2014). Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nature Communications*, *5*, 3128. https://doi.org/10.1038/ncomms4128

Palena, C., Polev, D. E., Tsang, K. Y., Fernando, R. I., Litzinger, M., Krukovskaya, L. L., … Schlom, J. (2007). The human T-box mesodermal transcription factor Brachyury is a candidate target for T-cell-mediated cancer immunotherapy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *13*(8), 2471–8. https://doi.org/10.1158/1078-0432.CCR-06-2353

Palena, C., Polev, D. E., Tsang, K. Y., Palena, C., Polev, D. E., Tsang, K. Y., … Schlom, J. (2007). The Human T-Box Mesodermal Transcription Factor Brachyury Is a Candidate Target for T-Cell − Mediated Cancer Immunotherapy The HumanT-Box Mesodermal T ranscription Factor Brachyury Is a arget for T-Cell ^ Mediated Cancer Immunotherapy, 2471–2478. https://doi.org/10.1158/1078-0432.CCR-06-2353

Phipps, A. I., Limburg, P. J., Baron, J. A., Burnett-hartman, A. N., Weisenberger, D. J., Laird, P. W., … Newcomb, P. A. (2015). Association Between Molecular Subtypes of Colorectal Cancer and Patient Survival. *Gastroenterology*, *148*(1), 77–87.e2. https://doi.org/10.1053/j.gastro.2014.09.038

Pinto, F., Pertega-Gomes, N., Pereira, M. S., Vizcaíno, J. R., Monteiro, P., Henrique, R. M., … Reis, R. M. (2014). T-box transcription factor Brachyury is associated with prostate cancer progression and aggressiveness. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*. https://doi.org/10.1158/1078-0432.CCR-14-0421

Powell, A. E., Wang, Y., Li, Y., Poulin, E. J., Means, A. L., Washington, M. K., … Coffey, R. J. (2012). The pan-ErbB negative regulator lrig1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell*, *149*(1), 146–158. https://doi.org/10.1016/j.cell.2012.02.042

Presneau, N., Shalaby, A., Ye, H., Pillay, N., Halai, D., Idowu, B., … Flanagan, A. M.

(2011). Role of the transcription factor T (brachyury) in the pathogenesis of sporadic chordoma: A genetic and functional-based study. *Journal of Pathology*, *223*(3), 327–335. https://doi.org/10.1002/path.2816

Roselli, M., Fernando, R. I., Guadagni, F., Spila, A., Alessandroni, J., Palmirotta, R., … Palena, C. (2012a). Brachyury, a driver of the epithelial-mesenchymal transition, is overexpressed in human lung tumors: An opportunity for novel interventions against lung cancer. *Clinical Cancer Research*, *18*(14), 3868–3879. https://doi.org/10.1158/1078-0432.CCR-11-3211

Roselli, M., Fernando, R. I., Guadagni, F., Spila, A., Alessandroni, J., Palmirotta, R., … Palena, C. (2012b). Brachyury, a driver of the epithelial-mesenchymal transition, is overexpressed in human lung tumors: an opportunity for novel interventions against lung cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *18*(14), 3868–79. https://doi.org/10.1158/1078-0432.CCR-11-3211

Sarkar, D., Shields, B., Davies, M. L., Müller, J., & Wakeman, J. a. (2011). BRACHYURY confers cancer stem cell characteristics on colorectal cancer cells. *International Journal of Cancer. Journal International Du Cancer*, *130*(2), 328–37. https://doi.org/10.1002/ijc.26029

Sei, Y., Lu, X., Liou, A., Zhao, X., & Wank, S. A. (2011). A stem cell marker-expressing subset of enteroendocrine cells resides at the crypt base in the small intestine. *Am J Physiol Gastrointest Liver Physiol*, *300*(7), 345–356. https://doi.org/10.1152/ajpgi.00278.2010

Shimoda, M., Sugiura, T., Imajyo, I., Ishii, K., Chigita, S., Seki, K., … Shirasuna, K. (2012). The T-box transcription factor Brachyury regulates epithelial-mesenchymal transition in association with cancer stem-like cells in adenoid cystic carcinoma cells. *BMC Cancer*, *12*, 377. https://doi.org/10.1186/1471-2407-12-377

Takahashi, M., Kamei, Y., & Ezaki, O. (2005). Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size. *American Journal of Physiology. Endocrinology and Metabolism*, *288*(1), E117–E124. https://doi.org/10.1152/ajpendo.00244.2004

Tetteh, P. W., Basak, O., Farin, H. F., Wiebrands, K., Kretzschmar, K., Begthel, H., …

Clevers, H. (2016). Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. *Cell Stem Cell*, *18*(2), 203–213. https://doi.org/10.1016/j.stem.2016.01.001

Wang, Y., Poulin, E. J., & Coffey, R. J. (2013). LRIG1 is a triple threat: ERBB negative regulator, intestinal stem cell marker and tumour suppressor. *British Journal of Cancer*, *108*(9), 1765–70. https://doi.org/10.1038/bjc.2013.138

Yancu, D., Blouin, M.-J., Birman, E., Florianova, L., Aleynikova, O., Zakikhani, M., … Pollak, M. (2017). A phenotype of IGFBP-3 knockout mice revealed by dextran sulfate-induced colitis. *Journal of Gastroenterology and Hepatology*, *32*(1), 146–153. https://doi.org/10.1111/jgh.13461

10 Appendix

10.1 Appendix A: IF Negative controls.

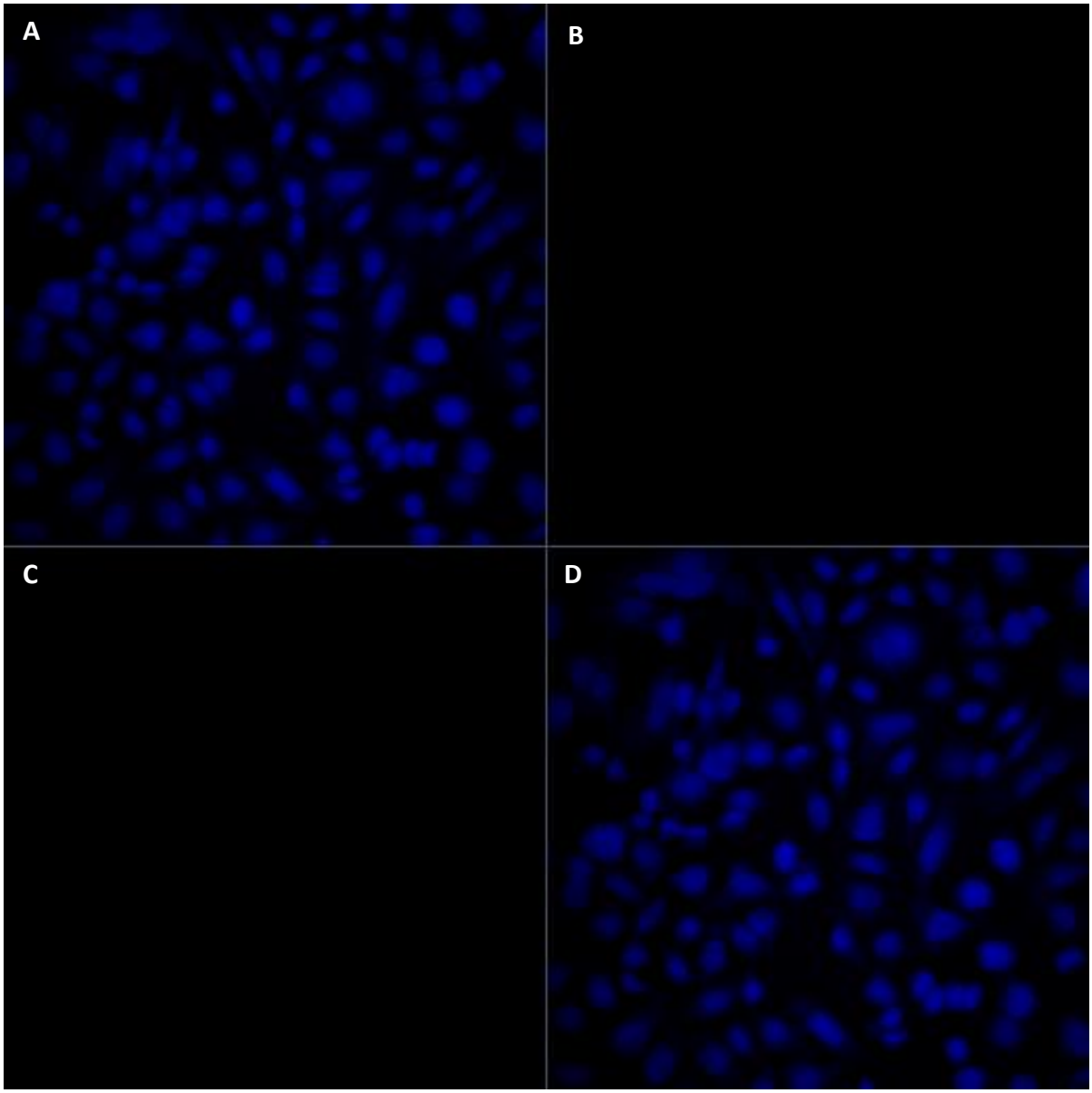

Figure 10.1 Immunofluorescence of SW480 cells transfected with pCMV-HA (63x) Secondary only control. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) red channel is the HA tag antibody. (D) The channels are merged in the bottom right. Images acquired on Zeiss 710 confocal microscope.

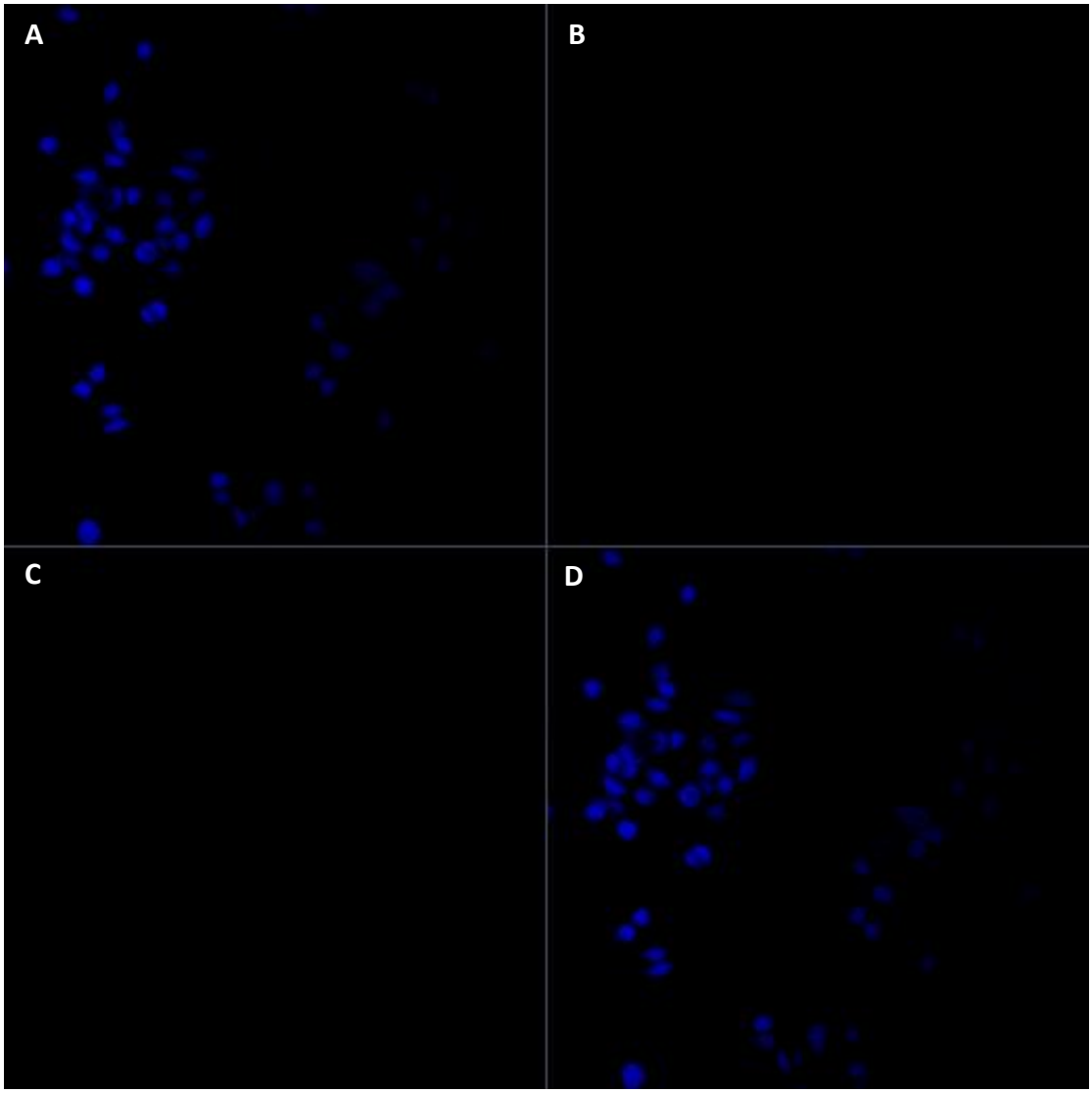

Figure 10.2 Immunofluorescence of SW620 cells transfected with pCMV-HA (63x) Secondary only control. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) red channel is the HA tag antibody. (D) The channels are merged in the bottom right. Images acquired on Zeiss 710 confocal microscope.

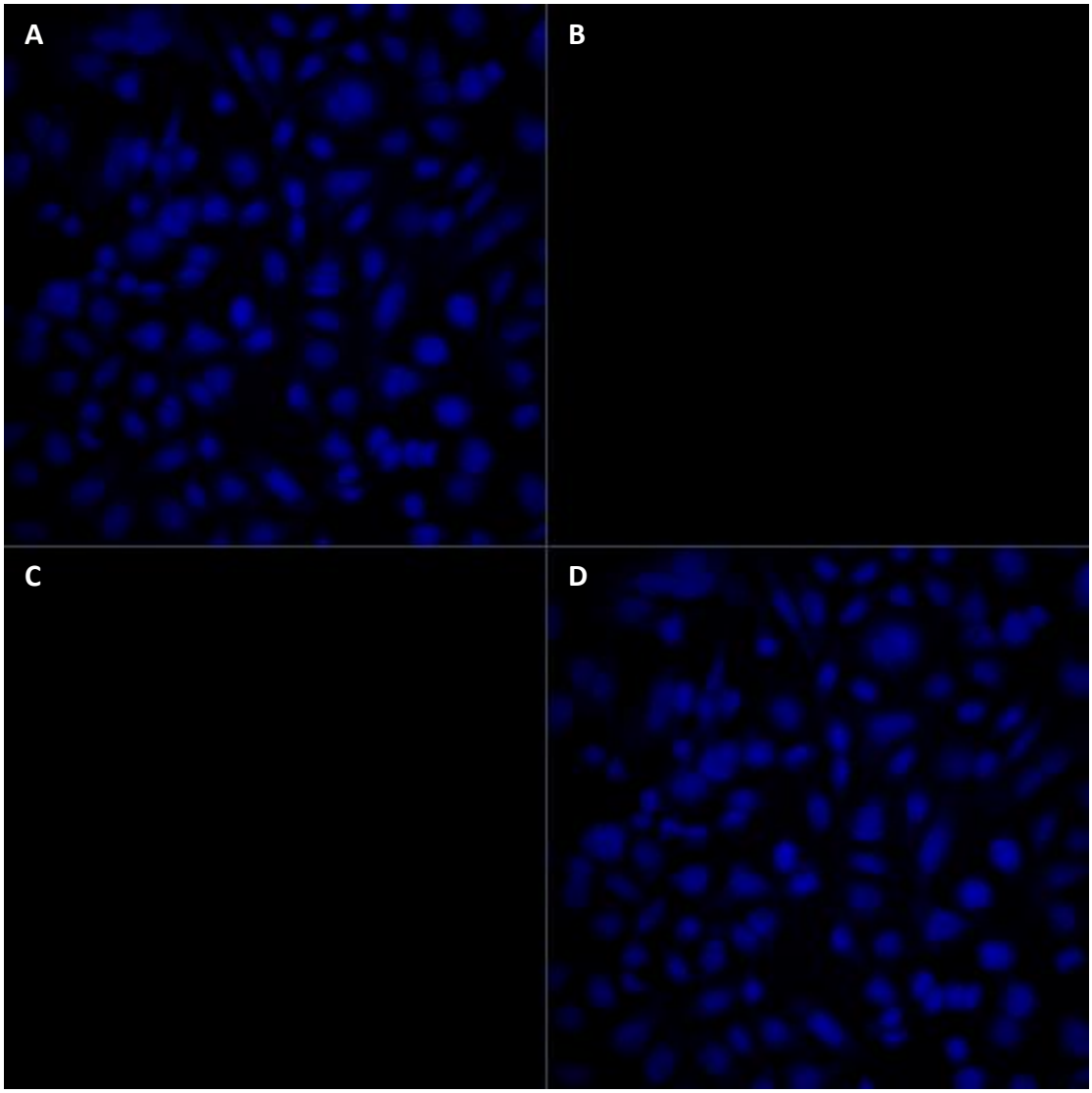

Figure 10.3 Immunofluorescence of NTERA2 cells transfected with pCMV-HA (63x) Secondary only control. (A) The blue channel shows DAPI staining of the nucleus, (B) green is AB140661 and (C) red channel is the HA tag antibody. (D) The channels are merged in the bottom right. Images acquired on Zeiss 710 confocal microscope.

10.2 Appendix B: Supplementary IHC images.

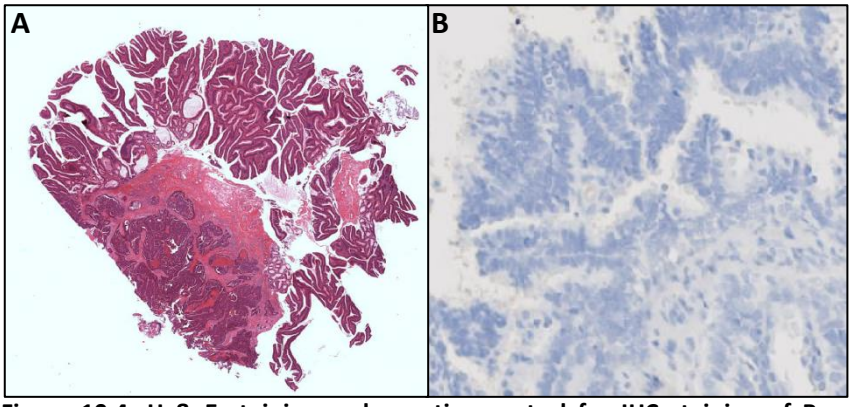

Figure 10.4: H & E staining and negative control for IHC staining of Brachyury and $p27^{Kip1}$ in a colorectal carcinoma. (A) H & E staining (x2) and (B) Negative (x10) (secondary mouse antibody only) control.

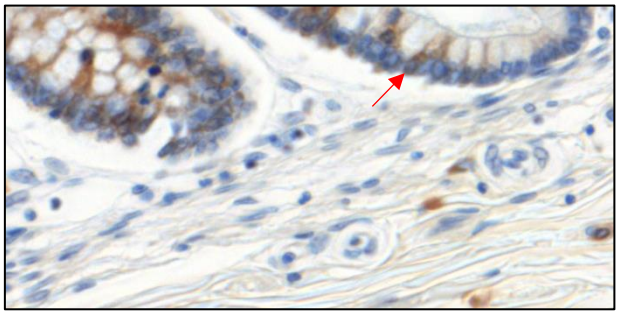

Figure 10.5: IHC staining of normal crypt with anti-Brachyury antibody ab57480 (20x). Red arrow points out cell that was identified as EEC-like and similar to those EECs that were identified by AB140661 in Figure 4.1.

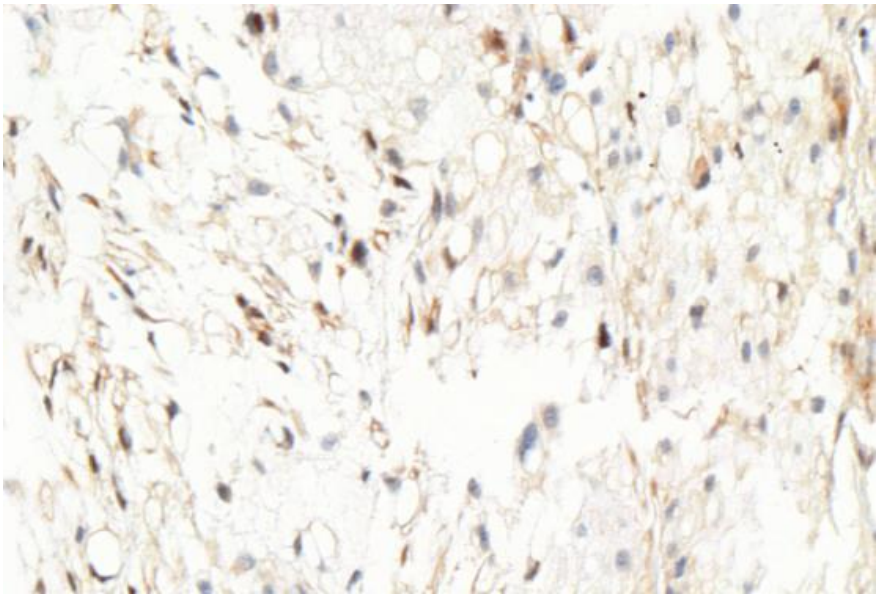

Figure 10.6: Positive control of IHC Chordoma section with anti-Brachyury antibody AB140661: Chordoma patient tissue sample was stained using AB140661 as a positive IHC control (20x).

10.3 Appendix C: Sequence results from pCMVTV1-HA vector

In order to confirm that positive clones had the correct in frame sequence. The pCMVTV1-HA plasmid was purified and sent for sequencing using the primers from Figure 2.1. BLAST and ORF reader were used to confirm Brachyury had been cloned in frame with HA-Tag, Kozak sequence and T7 promoter upstream. The forward primed sequence is shown below.

Forward NHA Primed Clone A29
TACCCCTACGACGTGCCCGACTACGCC HA Tag
AATAT Kozak Sequence
TAATACGACTCACTATAG T7 Promoter
ATG Start Codon

>pCMVTV1-HA
CTGAGACCACTGCTTACTGGCTTATCGAAATTCTAGATAATACGACTCACTATAGGGCGAATATGGCC
ACGTACCCCTACGACGTGCCCGACTACGCCGGCGCGCCTATGAGCTCCCCTGGCACCGAGAGCGCG
GGAAAGAGCCTGCAGTACCGAGTGGACCACCTGCTGAGCGCCGTGGAGAATGAGCTGCAGGCGGG
CAGCGAGAAGGGCGACCCCACAGAGCGCGAACTGCGCGTGGGCCTGGAGGAGAGCGAGCTGTGG
CTGCGCTTCAAGGAGCTCACCAATGAGATGATCGTGACCAAGAACGGCAGGAGGATGTTTCCGGTG
CTGAAGGTGAACGTGTCTGGCCTGGACCCCAACGCCATGTACTCCTTCCTGCTGGACTTCGTGGCGG
CGGACAACCACCGCTGGAAGTACGTGAACGGGGAATGGGTGCCGGGGGGCAAGCCGGAGCCGCA
GGCGCCCAGCTGCGTCTACATCCACCCCGACTCGCCCAACTTCGGGGCCCACTGGATGAAGGCTCCC
GTCTCCTTCAGCAAAGTCAAGCTCACCAACAAGCTCAACGGAGGGGGCCAGATCATGCTGAACTCCT
TGCATAAGTATGAGCCTCGAATCCACATAGTGAGAGTTGGGGATCCACAGCGCATGATCACCAGCCA
CTGCTTCCCTGAGACCCAGTTCATAGCGGTGACTGCTTATCAGAACGAGGAGATCACAGCTCTTAAA
TTAAGTACAATCCATTTGCAAAAGCTTTCCTTGATGCAAAGGAAAGAAGTGATCACAAAGAGATGAT
GGAGGAACCCGGAGACAGCCAGCAACCTGGGTACTCCCAATGGGGGTGGCTTCTTCCTGGAACCAG
CACCCTGTGTCCACCTGCAAATCCTCATCCTCAGTTTGGAGGTGCCCTCTCCCTCCCCTCCACGCACAG
CTGTGACAGGTACCCAACCCTGAGAGCCACCGGGTCCTCACCCTACCCCAGCCCCTATG

10.4 Appendix D: Brachyury regulates proliferation of cancer cells via a p27Kip1-dependent pathway Oncotarget. 2014 Jun; 5(11): 3813–3822.

10.5 Appendix E: Brachyury identifies a class of enteroendocrine cells in normal human intestinal crypts and colorectal cancer. Oncotarget. 2016 Mar 8;7(10):11478-86.

10.6 Appendix F: SMI titrations and IC50¬ values for H460 and NTERA2 cells. (A) SMI 3 titration in Brachyury high cell line H460, followed by SMIs 3.3 and 18 in H460 (B and C respectively). (D) SMI 3 titration in Brachyury low cell line NTERA2, followed by SMIs 3.3 and 18 in H460 (E and F respectively).