

Novel Epigenetic Biomarkers for the Barrett's –
Adenocarcinoma Sequence in Oesophageal
Cancer

By

Mark Paul Dilworth
MChB MRCS

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF MEDICINE

School of Cancer Sciences
College of Medical and Dental Sciences
University of Birmingham
September 2015

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Introduction

Oesophageal adenocarcinoma presents an ever increasing challenge to the NHS, with rising incidence and poor overall survival. There are few robust biomarkers available for this disease – either for detection or stratification. Barrett's Oesophagus, a precursor to adenocarcinoma, is common. In the non-dysplastic setting, few patients will progress to cancer. There are no current biomarkers to aid in this stratification process.

Aims

To assess the role of methylation biomarkers in the context of diagnosis of adenocarcinoma and their role in tumour biology. To provide a methylation stratification tool in the identification of high risk non-dysplastic Barrett's Oesophagus.

Methods

Genome wide methylation assessment was performed with validation using bisulphite pyrosequencing on carefully selected tissues to reveal novel methylation biomarkers

Results

Methylation of *TRIM15*, has been shown to be a robust biomarker in disease identification and has a role in tumour biology.

OR3A4, a long none coding RNA, has been identified as a way to reliably risk stratify the non-dysplastic Barrett's patient and forms the first biomarker of this kind.

Conclusions

Methylation biomarkers play a key role in disease identification and risk of cancer development. They also appear to play a role in tumour biology.

Acknowledgements

I would firstly like to take this opportunity to thank the funders of this research. The University Hospital Birmingham Charities, supported by the Upper GI support group, paid for the majority of the consumables in this study and we are hugely grateful to them for this.

I am thankful also to the team at Sandwell and West Birmingham Hospital group, for providing my clinical employment throughout this period of research.

Clearly huge thanks go to both of my supervisors and their patience as a surgeon invades the foreign territory of a scientific laboratory. In addition to my formal supervisors, I should like to thank Professor Alderson for his ongoing support with this study and Dr Rahul Hejmadi (consultant histopathologist) for taking the time to review all tissues used in this work.

The contribution to this project in terms of academic experience, day-to-day help and support, surgical banter and generally being a good friend by Andrew Beggs will never be forgotten. The success of this work, and its ongoing endeavours, would not have been possible without his input.

I should like to express thanks to all the members of the team, including Jonathan, Dan and Jo for their enthusiasm for positive results and their ability to teach me the workings of a genetic laboratory.

Dr Nigel Trudgill provided huge help in the acquisition of precious tissue samples from a neighbouring NHS trust. I am thankful to him and the pathology team at SWBH for their help in this work.

Contents

| | |
|--|----------|
| 1. Introduction | 1 |
| 1.1 Epidemiology of Oesophageal Adenocarcinoma | 1 |
| 1.2: Oesophageal Squamous Cell Carcinoma | 3 |
| 1.3: Junctional Tumours | 4 |
| 1.4: Survival of Oesophageal Cancer | 5 |
| 1.5: Clinical Features of Oesophageal Cancer | 5 |
| 1.6: Oesophageal Cancer Staging | 6 |
| 1.6.1: Interpretation of Staging Investigations | 7 |
| 1.7: Management of Oesophageal Carcinoma | 8 |
| 1.7.1: Complications of Oesophagectomy | 10 |
| 1.7.2: Alternative Therapeutic Strategies | 11 |
| 1.8: The multidisciplinary Management of the Cancer Patient | 11 |
| 1.9: Barrett's Oesophagus | 12 |
| 1.10: The Development of Oesophageal Adenocarcinoma from Barrett's Oesophagus | 15 |
| 1.11: Epigenetics of Oesophageal Adenocarcinoma and Barrett's Oesophagus | 19 |
| 1.11.1 Epigenetics | 19 |

| | |
|---|----|
| 1.11.2: DNA Methylation | 20 |
| 1.11.3: Methylation in OADC | 21 |
| 1.11.4: Inflammation in the Oesophagus | 23 |
| 1.12: Methylation markers in Oesophageal Adenocarcinoma and Barrett's Oesophagus | 23 |
| 1.12.1: <i>P16/CDKN2A</i> | 24 |
| 1.12.2: <i>Vimentin</i> | 24 |
| 1.12.3: <i>P14ARF</i> | 25 |
| 1.12.4: <i>CDX2</i> | 26 |
| 1.12.5: <i>JAK/STAT/SOCS</i> | 27 |
| 1.12.6: Wnt Signalling and Secreted frizzled-related proteins | 27 |
| 1.12.7: Methylation of Cell Free Circulating DNA as Biomarkers | 31 |
| 1.12.8: Other Markers | 31 |
| 1.13: Methylation in Progression of BO to OADC | 32 |
| 1.14: Summary of Methylation Markers | 35 |
| 1.15: Methylation Assessment Technologies | 36 |
| 1.15.1: Pyrosequencing | 36 |
| 1.15.2: Golden Gate Methylation Array | 37 |

| | |
|--|----|
| 1.15.3: Illumina Methylation 450 Array | 38 |
| 1.14: Thesis Aims | 39 |
| 2. Research and Thesis Hypotheses | 40 |
| 3. Material and Methods | 41 |
| 3.1: Materials | 41 |
| 3.1.1 Tissues | 41 |
| 3.1.2: Chapter 4 Samples | 41 |
| 3.1.3: Chapter 5 Samples | 43 |
| 3.1.4: Chapter 6 Samples | 43 |
| 3.1.4.1: Non-Dysplastic BO samples | 43 |
| 3.1.4.2: Lymph Node Positivity Samples | 45 |
| 3.2: DNA Extraction | 46 |
| 3.3: DNA Quantification | 47 |
| 3.4: Nanodrop | 48 |
| 3.5: Bisulphite Conversion of DNA | 48 |
| 3.5.1: Qiagen Protocol for Bisulphite Conversion | 49 |
| 3.5.2: Qiagen Bisulphite Clean Up | 49 |
| 3.5.3: Zymo Bisulphite Conversion Protocol (for compatibility with the Illumina 450 array sample) | 50 |
| 3.6: Pyrosequencing PCR Reaction | 51 |

| | |
|---|----|
| 3.7: PCR Validation | 53 |
| 3.8: Pyrosequencing | 55 |
| 3.9: Cell Culture | 57 |
| 3.10: DNA Extractions from Cell Lines | 58 |
| 3.11: RNA Extraction for Expression Work | 59 |
| 3.12: Real time qPCR for the Assessment of RNA Quantity within a Cell | 61 |
| 3.13: Protein Extraction from Cells | 62 |
| 3.13.1: Protein Analysis | 65 |
| 3.13.2: Demethylation of OE33 | 68 |
| 3.14: Methylation Arrays | 73 |
| 3.14.1: Sample Preparation | 73 |
| 3.14.2: Array DNA Quantification | 73 |
| 3.14.3: Array DNA Quality Control | 74 |
| 3.14.4: Array DNA Bisulphite Conversion | 75 |
| 3.14.5: DNA Repair Process | 76 |
| 3.14.6: Illumina Methylation 450 Array | 76 |
| 3.14.7: Bioinformatic Analysis of Array Data | 77 |

| | |
|---|-----------|
| 4 – Assessment of Published Methylation Arrays in Oesophageal Adenocarcinoma | 78 |
| 4.1: Introduction | 78 |
| 4.2: Aims | 81 |
| 4.3: Methods | 81 |
| 4.3.1: Samples | 81 |
| 4.3.2: Power Calculation | 82 |
| 4.3.3: Bisulphite Pyrosequencing | 83 |
| 4.3.4: Statistical Analysis | 84 |
| 4.4: Results | 84 |
| 4.5: Conclusions | 92 |
| 4.6: Discussion | 95 |
| 5. Data Mining of TCGA Database to Search for Differentially Methylated Biomarkers of Oesophageal Adenocarcinoma, Independent Validation and Functional Assessment | 97 |
| 5.1: Introduction | 97 |
| 5.2: Aims | 99 |
| 5.3: Materials and Methods | 99 |
| 5.3.1: TCGA Methylation Array Datasets | 99 |

| | |
|---|-----|
| 5.3.2: Bisulphite Pyrosequencing | 100 |
| 5.4: Results | 101 |
| 5.4.1: TCGA Analysis | 101 |
| 5.4.2: TCGA Array Validation | 106 |
| 5.5: Oesophageal Methylation Field Effect Change | 115 |
| 5.5.1: Introduction | 115 |
| 5.5.2: Aims | 116 |
| 5.5.3: Materials and Methods | 116 |
| 5.5.4: Results | 117 |
| 5.5.5: Conclusions | 118 |
| 5.6: Crossover of <i>TRIM15</i> Methylation with the Lower Gastrointestinal Tract | 118 |
| 5.6.1: Introduction | 118 |
| 5.6.2: Aims | 119 |
| 5.6.3: Materials and Methods | 119 |
| 5.6.4: Results | 120 |
| 5.6.5: Conclusions | 121 |
| 5.7: <i>TRIM15</i> as a Functional Biomarker | 122 |

| | |
|--|------------|
| 5.7.1: Introduction | 122 |
| 5.7.2: Aims | 126 |
| 5.7.3: Demethylation of OE33 and its correlation with expression of <i>TRIM15</i> | 126 |
| 5.8: Correlation of RNA Expression with Protein Expression | 127 |
| 5.8.1: Aims | 127 |
| 5.8.2: Methods | 127 |
| 5.8.3: Results | 134 |
| 5.8.4: Conclusions | 135 |
| 6. Novel Approaches in Oesophageal Biomarker Discovery | 138 |
| 6.1: Barrett's Oesophagus as a Clinical Problem | 138 |
| 6.2: Aims | 140 |
| 6.3: Patients and Samples | 140 |
| 6.3.1: Samples for assessment of Barrett's progression | 140 |
| 6.3.2: Tissues for Assessment of Lymph Node Disease | 143 |
| 6.4: Methods | 144 |
| 6.5: Methylation 450 arrays | 144 |
| 6.5: Validation Experiments | 145 |

| | |
|--|------------|
| 6.6: Results | 145 |
| 6.7.1: High Risk Barrett's Biomarker | 145 |
| 6.7.2: Analysis of Pyrosequencing Data | 149 |
| 6.7.3: <i>OR3A4</i> data | 150 |
| 6.7.4: Conclusions for Barrett's Progression | 153 |
| 6.8: Biomarker Discovery of Nodal Disease | 154 |
| 6.8.1: Results | 155 |
| 6.9: Conclusions | 156 |
| 7. Discussion | 161 |
| 7.1: Thesis discussion | 161 |
| 7.2: Follow-up Work | 163 |
| 7.3: <i>TRIM15</i> | 163 |
| 7.4: <i>OR3A4</i> | 165 |
| 7.5: Thesis summary | 167 |
| 8. References | 168 |
| 9. Appendix | |
| 9.1: Gene/primer information | 177 |

List of Tables

| | |
|--|----|
| Table 3.1: Methylation PCR reagents and volumes | 52 |
| Table 3.2: Methylation PCR conditions | 53 |
| Table 3.3: cDNA Production Reagents | 60 |
| Table 3.4: cDNA Production Reaction Conditions | 61 |
| Table 3.5: qPCR Reagents | 62 |
| Table 3.6: Reagents used to attempt to dissolve protein plug | 64 |
| Table 3.7: Ingredient list for Western Blot stacking and running gels | 66 |
| Table 3.8: Thermal conditions for array quality control | 75 |
| Table 4.1: Previously published methylation sites with the ability to differentiate oesophageal tissue types, as determined by a Golden Gate array | 80 |
| Table 4.2: Pyrosequencing data and statistical analysis for attempted validation of <i>ALPL</i> | 85 |
| Table 4.3: Pyrosequencing data and statistical analysis for attempted validation of <i>POMC</i> | 86 |
| Table 4.4: Pyrosequencing data and statistical analysis for attempted validation of <i>RBP1</i> | 87 |
| Table 4.5: Pyrosequencing data and statistical analysis for attempted validation of <i>TIAM1</i> | 88 |

| | |
|--|-----|
| Table 4.6: Pyrosequencing data and statistical analysis for attempted validation of <i>VCAN</i> | 89 |
| Table 5.1: Methylation 450 array data revealing top 20 differentially methylated sites between normal mucosa and OADC | 105 |
| Table 5.2: Analysis of <i>TRIM15</i> methylation results for each of the 4 primers on 8 tumour/normal tissue pairs | 111 |
| Table 5.3: Difference in methylation between normal squamous mucosa and adenocarcinoma of <i>TRIM15</i> cg19115272 | 114 |
| Table 5.4: Logistical regression analysis of <i>TRIM15</i> methylation | 115 |
| Table 5.5: Methylation of <i>TRIM15</i> cg19115272 in native OE33 cells | 126 |
| Table 5.6: Corrected protein content of OE33 cells with and without 5AZA treatment | 134 |
| Table 6.1: Top 20 array identified CpG sites with methylation variation between non- dysplastic samples which progress to OADC vs those which remain static | 146 |
| Table 6.2: Pyrosequencing validation data for <i>OR3A4</i> | 150 |
| Table 6.3: Logistical regression showing the use of <i>OR3A4</i> methylation as a clinical test | 153 |
| Table 6.4: Illumina methylation 450 array data comparing T3N0 tumours with T3N1 tumours | 156 |

List of Figures

| | |
|---|-----|
| Figure 1.1: Risk factors for oesophageal adenocarcinoma | 3 |
| Figure 1.2: Risk factors for oesophageal squamous cell carcinoma | 4 |
| Figure 1.3: TMN staging system for Oesophageal Cancer | 8 |
| Figure 1.4: Short-Term Outcomes Following Open Versus Minimally Invasive Oesophagectomy | 10 |
| Figure 1.5: Oesophageal microscopy | 14 |
| Figure 1.6: The Cytosponge | 17 |
| Figure 1.7: The Wnt Signalling pathway | 29 |
| Figure 1.8: Pyrosequencing Diagram | 37 |
| Figure 3.1: Agarose gel electrophoresis | 54 |
| Figure 3.2: Pyrosequencing Trace | 57 |
| Figure 4.1: Variation in methylation between tissue types of <i>TIAM1</i> | 91 |
| Figure 5.1: TCGA methylation 450 array data comparing methylation between normal oesophageal squamous mucosa and adenocarcinoma | 103 |
| Figure 5.2: Pyrosequencing validation data of <i>TRIM15</i> | 107 |
| Figure 5.3: <i>TRIM15</i> gene and location of array identified probes | 109 |
| Figure 5.4 Methylation difference between normal squamous mucosa and adenocarcinoma of <i>TRIM15</i> cg19115272 | 113 |

| | |
|---|-----|
| Figure 5.5: Methylation variation between colonic tissue states of TRIM15 cg19115272 | 121 |
| Figure 5.6: Correlation between TRIM15 cg19115272 methylation and expression | 125 |
| Figure 5.7: Standardisation of protein samples | 129 |
| Figure 5.8: Coomassie blue stained gel | 133 |
| Figure 5.9: Western blot film following exposure, revealing the obvious variation in protein content between 5AZA treated and untreated OE33 cells | 133 |
| Figure 5.10: Corrected protein content of OE33 cells, with and without 5AZA treatment | 135 |
| Figure 6.1: Variation in methylation between progressors and none progressors in <i>FGR2</i> | 149 |
| Figure 6.2: ROC curve demonstrating the ability of <i>OR3A4</i> methylation to differentiate high risk BO | 151 |
| Figure 6.3: <i>OR3A4</i> methylation differentiating high risk non- dysplastic BO | 152 |

Abbreviations

| | |
|--------|---|
| ADC | Adenocarcinoma |
| BO | Barrett's Oesophagus |
| CT | Computerised Tomography |
| DNA | Deoxyribonucleic Acid |
| EUS | Endoscopic Ultrasound |
| FFPE | Formalin Fixed Paraffin Embedded |
| GI | Gastrointestinal |
| GORD | Gastroesophageal Reflux Disease |
| HBRC | University of Birmingham Human Tissue Biorepository |
| H+E | Haematoxylin and eosin |
| NDBO | Non-dysplastic Barrett's Oesophagus |
| OADC | Oesophageal Adenocarcinoma |
| OC | Oesophageal Cancer |
| OGD | Oesophagogastroduodenoscope |
| OSCC | Oesophageal Squamous Cell Carcinoma |
| PCR | Polymerase Chain Reaction |
| PET-CT | Positron Emission Computerised Tomography |
| RTPCR | Real Time Polymerase Chain Reaction |
| SGH | Sandwell General Hospital |
| UGI | Upper Gastrointestinal |
| UHB | University Hospital Birmingham |
| UK | United Kingdom |
| 5-AZA | 5-aza-2'-deoxycytidine |

1. Introduction

Oesophageal carcinoma is an increasingly common cancer worldwide affecting approximately 450,000 people. It is associated with very poor overall survival (1). This disease has two variants; Oesophageal Squamous Cell Carcinoma (OSCC) and Oesophageal Adenocarcinoma (OADC). While the global predominant pathology is OSCC, this has not been observed in the UK and developed countries have always reported an OADC preponderance, however there is significant variation between countries. Currently oesophageal cancer forms 5% of the digestive tract cancers in the UK (2).

The incidence of OADC in the UK has significantly increased from 0.9 per 100000 in 1971-75 to 4.5 per 100,000 in 1996-2001 (3). This observed increase is associated with multiple potential risk factors which include lifestyle, environmental and genetic. No true occupational relationship has been shown however there is weak evidence suggesting a higher incidence in those who have previously had exposure to asbestos (4). Disease survival correlates with stage of cancer at presentation and there has been shown to be significant survival advantages with the detection of early stage disease (5, 6).

1.1: Epidemiology of Oesophageal Adenocarcinoma

Oesophageal cancer diagnosis is most common in a male population and its incidence increases with age. Due to the usual late stage of disease at presentation, overall survival remains one of the lowest for all solid cancers. In the UK the overall 5 year survival is only 14% (5).

Risk factors for the two predominant oesophageal cancer types vary. As with the majority of tumours, smoking has been demonstrated to be a significant risk factor, but has also been shown to increase the risk of progression to cancer from Barrett's Oesophagus (7-11). OADC has long been considered a disease related to chronic inflammation of the lower oesophagus caused by prolonged acid and bile salt reflux into the gullet. This goes some of the way to surmising why the incidence of this disease is increasing.

There have been numerous publications supporting the correlation between Gastro-oesophageal Reflux Disease (GORD) and obesity and the corresponding increased cancer risk in this obese group (12-14). National statistics from 2005 showed that 22% of men and 23% of women aged 16–64 were obese (15), with an even greater number being overweight (16). It follows therefore that the incidence of obesity induced reflux disease is rising, and with it OADC (17, 18). There is also evidence that obesity is in itself an independent risk factor for OADC without the presence of GORD (19).

In support of an inflammatory aetiology, Non-Steroidal Anti-Inflammatory Drugs (NSAIDS) and aspirin have been shown to decrease the risk of OADC by up to 40% (20). Medication which as a side effect may decrease the pressure of the lower oesophageal sphincter mechanism, for example anti-cholinergic medications and calcium channel blockade, lead to increased reflux which correlates with an increased risk of cancer. However, in a recent meta-analysis, not all drugs with this side effect increased the cancer risk suggesting there might be influences other than the increase in gastric reflux which lead to the higher cancer risk (21) but which are as yet unknown. Multiple studies show the potential protective effects of gastric

Helicobacter pylori infection, decreasing incidence of OADC by up to 40% (22). Presence of a hiatus hernia has been shown to correlate with increased risk of OADC (23). Achalasia, a primary motility disorder of the oesophagus defined as the inability of the lower oesophageal sphincter to relax combined with abnormal motility in the remaining oesophagus, also correlates with increased risk of OADC (24).

Figure 1.1: Risk factors for oesophageal adenocarcinoma, taken from Pennarthur et al in 2013 (1):

| Oesophageal adenocarcinoma |
|--|
| • Symptomatic gastro-oesophageal reflux disease |
| • Barrett's oesophagus |
| • Obesity |
| • Tobacco use |
| • History of thoracic radiation |
| • Diet low in vegetables and fruits |
| • Increased age |
| • Male sex |
| • Medications that relax the lower oesophageal sphincter |
| • Familial history (rare) |

Causative factors, increasing the likelihood of OADC development

1.2: Oesophageal Squamous Cell Carcinoma

Although not to be investigated in this thesis, it is important to recognise that OADC is not the only frequent tumour type to affect the oesophagus. In many countries, oesophageal squamous cell cancer is the predominant tumour type seen and indeed OSCC remains the most common oesophageal malignancy in the world. The incidence of OSCC seen in the UK has remained largely static throughout reporting, while OADC has increased exponentially and continues to do so. Risk factors for the two diseases vary, reflecting their very different pathogenesis and origins. Risk

factors for OSCC are summarised in the figure 1.2 (1). This disease, while associated with an equally poor overall outcome, behaves very differently from OADC, not least because of their location within the oesophagus, which is usually much more proximal (upper and mid oesophagus) than OADC. OSCC's are also associated with varying developmental risk factors. Symptoms as well as diagnostic tests remain equivocal for both pathologies. Curative treatment is again based on surgery and in those without metastatic disease will include oesophagectomy.

Figure 1.2: Risk factors for oesophageal squamous cell carcinoma, taken from Pennarthur et al in 2013 (1)

| Oesophageal SCC risk factors |
|--|
| • Tobacco use |
| • Alcohol consumption |
| • Mutations of enzymes that metabolise alcohol |
| • Achalasia |
| • Caustic injury |
| • History of thoracic radiation |
| • Low socioeconomic status |
| • Poor oral hygiene |
| • Nutritional deficiencies |
| • Non-epidermolytic palmoplantar keratoderma |

Causative factors, increasing the likelihood of SCC development

1.3: Junctional Tumours

Over recent years, tumour classification has been given increasing significance. It is important to ascertain the anatomical location of a gastro-oesophageal tumour as this has a bearing on therapy and prognosis. Therefore the concept of a 'junctional tumour' was established. In 2005, Siewert developed and published a classification which, despite initial poor acceptance, is now used routinely by clinicians in the

staging and assessment of junctional tumours (25). Siewert classified these tumours according to the following:

1. Type I - adenocarcinoma of distal part of the oesophagus (centre located 1-5cm above the anatomic Gastro-oesophageal Junction.)
2. Type II - adenocarcinoma of the real cardia (within 1cm above and 2cm below the GOJ.)
3. Type III - adenocarcinoma of the subcardial stomach (2-5cm below GOJ.)

Identification of these junctional tumours and their differentiation from true lower OADC and proximal gastric tumours is significant because of their varying survival rates, which are overall much lower (26). Theories exist as to why this is observed including the relatively thinner muscularis layer in this region, combined with the relatively greater propensity for the tumour to have access to regional lymph nodes (retroperitoneal, gastric and mediastinal) making early metastatic spread to these lymph nodes much more likely.

1.4: Survival of Oesophageal Cancer

The interest in the study of oesophagogastric cancer is borne largely from this disease being associated with such a poor overall survival. Diagnosis of oesophageal cancer in the UK can lead to an anticipated 5 year survival of only 14% (5). Overall survival correlates with stage of tumour at presentation, and is significantly worsened by the presence lymph node disease.

1.5: Clinical Features of Oesophageal Cancer

Of the patients who have symptomatic Oesophageal Cancer (OC), the most frequent symptom is dysphagia (difficulty swallowing with the sensation of food becoming

“stuck” in the gullet) (27). Daly et al in 2000 reported the incidence of presenting symptoms in OC as dysphagia (74%), weight loss (57.3%), gastrointestinal reflux symptoms (20.5%), odynophagia (pain on swallowing 16.6%), and dyspnoea (the feeling of being short of breath - 12.1%) (27).

Clinical suspicion of such a lesion should lead to prompt investigation, ideally in the form of endoscopy. This involves the passage of an oesophagogastroduodenoscope (OGD) flexible endoscope via the mouth, through the oesophagus and into the stomach. This allows full visual inspection of the entire length of the oesophagus and stomach. Thorough completion of this examination involves cannulation of the duodenum and examining to the second part (D2). This examination has the advantage that any identified abnormality can be biopsied for histological analysis and provide a definitive diagnosis.

1.6: Oesophageal Cancer Staging

Once cancer of the oesophagus has been identified, it is necessary to accurately stage the disease to determine prognosis and treatment options. This will include a contrast enhanced computerised tomography (CT) scan of the thorax, abdomen and pelvis, endoscopic ultrasound (EUS), Positron Emission Tomography (PET) CT and staging laparoscopy.

EUS assesses the depth of the tumour along with regional lymph nodes. It has demonstrated 71% correlation with overall histological disease stage, but it becomes significantly less accurate with early T1 tumours (28). PET-CT further increases the detection of metastatic disease over conventional CT alone (29). Patients with junctional ADC will also require a staging laparoscopy. This involves, under general

anaesthetic, a camera being inserted through the umbilicus and a thorough visual inspection of the abdominal cavity where any abnormality can be biopsied. This allows the identification and sampling of metastatic lesions which are small and would not be seen on conventional imaging (30). The final aspect of this operative staging process is peritoneal lavage. The abdominal cavity is washed with fluid (saline) which is then aspirated, collected and sent for cytological review by a gastrointestinal histopathologist. Presence of peritoneal disease, either visually (confirmed with biopsy) or on review of the fluid aspirate for cytology, will mean that the patient cannot be cured by resection and should therefore be offered palliative treatments.

1.6.1: Interpretation of Staging Investigations

Information obtained from the above investigations, designed to assess disease severity and spread, will be entered into the TNM staging score which provides an internationally understood nomenclature and is now in its 7th edition (31). In this system, T represents the local tumour characteristics; N represents the presence or absence of nodal disease, which can be further broken down into the anatomical location of those nodes. Finally, M represents the presence or absence of distant metastatic disease. This staging tool, while used in numerous cancers, has specific characteristics for each tumour type. Importantly, this staging system allows the multi-disciplinary team to determine whether the patient should be offered aggressive therapy with the intention of cure, or have their symptoms and disease controlled with palliative (none curative) therapy, in an internationally recognised format understood by all.

Figure 1.3: TMN staging system for Oesophageal Cancer (32)

| Primary Tumour (T) | |
|---------------------------------|--|
| TX | Primary tumour cannot be assessed |
| T0 | No evidence of primary tumour |
| Tis | High-grade dysplasia |
| T1 | Tumour invades lamina propria, muscularis mucosae, or submucosa |
| T1a | Tumour invades lamina propria or muscularis mucosae |
| T1b | Tumour invades submucosa |
| T2 | Tumour invades muscularis propria |
| T3 | Tumour invades adventitia |
| T4 | Tumour invades adjacent structures |
| T4a | Resectable tumour invading pleura, pericardium, or diaphragm |
| T4b | Unresectable tumour invading other adjacent structures, such as the aorta, vertebral body, and trachea |
| Regional lymph nodes (N) | |
| NX | Regional lymph node(s) cannot be assessed |
| N0 | No regional lymph node metastasis |
| N1 | Metastasis in 1-2 regional lymph nodes |
| N2 | Metastasis in 3-6 regional lymph nodes |
| N3 | Metastasis in 7 or more regional lymph nodes |
| Distant metastasis (M) | |
| M0 | No distant metastasis |
| M1 | Distant metastasis |

This TMN system allows tumour characteristics to be described in an internationally understood nomenclature.

1.7: Management of Oesophageal Carcinoma

Having progressed through the above staging process, patients with oesophageal carcinoma will be divided into two basic categories - those for whom treatment should be employed for potentially curative intent, and those for whom only palliative treatments are possible. Curative intent will include patients whose disease is limited to the surgical resection field, in the absence of invasion of local structures and similarly the absence of distant metastatic disease.

Patients for whom the staging process predicts no chance of cure or for those who are deemed unfit for major surgical resection will be offered palliative treatment. This may include palliative chemotherapy, the benefit of which would be to slow the rate of progression. Other strategies, particularly in those patients for whom swallowing has been rendered difficult/impossible, include the insertion of an oesophageal stent to dilate the gullet and allow the passage of liquid and food boluses. Should stent insertion not be possible, patients might be offered a procedure to insert a percutaneous feeding tube to preserve nutrition and prevent dehydration while making full use of their functional gastrointestinal tract, thereby avoiding the need for intravenous nutrition therapies. Should none of these be possible, the patient will be offered best supportive care (33).

Patients who have a favourable disease stage, suggesting cure may be possible, will be offered neoadjuvant chemotherapy to control systemic disease which has been shown to significantly improve three year survival (34, 35), followed by surgery. Having completed this course, or if the patient is unable to tolerate the chemotherapy, surgery will be considered.

Surgery remains the only potentially curative treatment for oesophagogastric cancer. For oesophageal and upper junctional tumours this will include an oesophagectomy (removal of the oesophagus). Historically this has been performed through both an abdominal and right sided thoracic incision. The diseased oesophagus is removed, along with corresponding lymph nodes. Gastrointestinal continuity is restored by creating a gastric conduit. The stomach is mobilised from all of its attachments in the abdomen and moved into the chest. Subsequent removal of the greater curve of the stomach creates a neo-oesophagus tube. More recently, this procedure has been

performed using a keyhole technique – a procedure known as “minimally invasive oesophagectomy” (MIO). This is now a treatment option for selected patients in large upper GI surgical units with the necessary experience and skills (36, 37).

1.7.1: Complications of Oesophagectomy

There can be little doubt that oesophagectomy is a significant surgical procedure and therefore carries risk of both morbidity and mortality. The 30 day mortality rate from published UK figures stands currently at <2% (38) This procedure carries significant risk of morbidity in the large part due to it being one of the few surgical procedures where both the abdominal and chest cavity are explored. Morbidity, or disease brought about as a consequence of this operation, can be serious and life-threatening. These are defined in figure 1.4.

Figure 1.4: Short-Term Outcomes Following Open Versus Minimally Invasive Oesophagectomy

| Complication | % Affected |
|--------------------------------|-------------------|
| Angina | 2.9 |
| Myocardial Infarction | 0.7 |
| Congestive Cardiac Failure | 0.9 |
| Atrial Fibrillation | 9.5 |
| Pneumonia | 18.8 |
| Pleural effusion | 16.5 |
| Respiratory failure | 3.8 |
| Other Respiratory Complication | 3.3 |
| Deep Vein Thrombosis | 0.6 |
| Pulmonary Embolism | 1.5 |
| Stroke | 0.2 |
| Renal Failure | 1.9 |

Table adapted from: Short-Term Outcomes Following Open Versus Minimally Invasive Esophagectomy for Cancer in England: A Population-Based National Study. Mamidanna et al. Annals of Surgery 2012 (39). The above table lists complications seen after oesophagectomy, in addition to the percentage of patients who will experience them in the post-operative period.

1.7.2: Alternative Therapeutic Strategies

With the associated risk, strategies for treating OADC avoiding oesophagectomy are being developed, some of which are already in common clinical use. These include endoscopic resection methods for small, early tumours where there is no lymph node involvement (T1a N0 M0) (40). However, while macroscopic lymph node disease is relatively easy to diagnose, microscopic disease is currently impossible to predict and there is a danger such patients will be undertreated by endoscopic therapy alone, if nodes containing microscopic tumour cells are not removed (41). Such patients should be followed carefully after their procedure, and salvage oesophagectomy should be considered early where there is disease recurrence (42).

1.8: The Multidisciplinary Management of the Cancer Patient

The management of oesophageal tumours should now occur in a tertiary referral Upper GI cancer centre. This allows dedicated and specialist care for upper GI pathology across the country, promoting the development of specialist teams for optimal patient management and outcome. This also facilitates formalised research and training in the management of upper GI cancers. Having dedicated centres also allows the development of the multi-disciplinary cancer team. This should in general consist of dedicated surgeons, oncologists, radiologists, histopathologists, dieticians, physiotherapists and clinical nurse specialists for the optimal patient management and outcome.

1.9: Barrett's Oesophagus

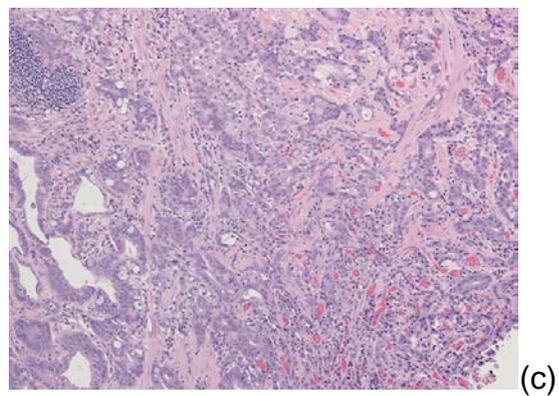
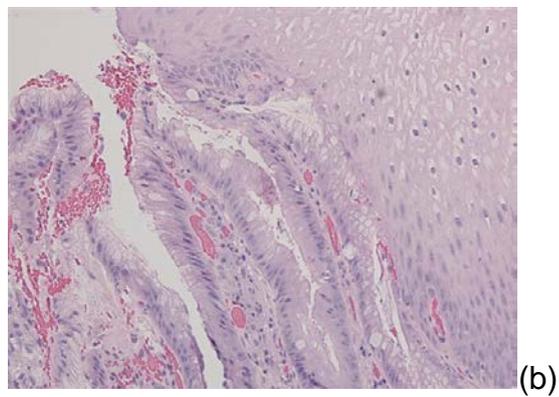
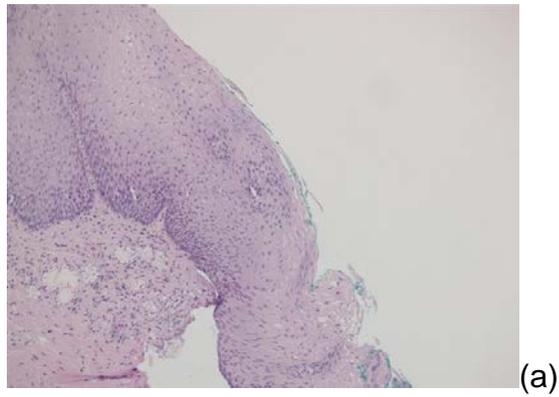
Barrett's Oesophagus is a phenomenon observed in the lower gullet where the normal squamous mucosa changes to a columnar epithelium. The description of Barrett's oesophagus was introduced in 1950 by Dr Norman Barrett. He described a phenomenon of lower oesophageal cell metaplasia; where by normal squamous epithelium is replaced with an intestinal type columnar epithelium (43). Numerous hypotheses exist as to the cause of this event. Initially, this was assumed to be due to a "short oesophagus" syndrome, whereby gastric mucosa was pulled into the thorax as a consequence of the oesophagus not being of full length. There is now consensus that this results from prolonged exposure to stomach acids and bile salts which reflux into the oesophagus causing chronic inflammation and tissue damage (44).

Over time, and driven by chronic inflammation, this metaplastic change is observed (45). For most, this metaplasia may be regarded as protective to the lower oesophagus. However, BO is associated with an increased risk of future development of OADC. For the majority of patients, their Barrett's oesophagus will never develop beyond this benign metaplastic cell change (46, 47). However in a small number of patients, presumably driven by the inflammation process, metaplastic cells will develop dysplasia – a warning sign of malignant potential.

The presence of cellular dysplasia within an area of Barrett's can be challenging to describe, even by experienced and dedicated GI histopathologists (48). Dysplasia can range from mild to moderate and finally severe or high grade. This difficulty in defining grade of dysplasia is usually as a result of the degree of inflammation of the

tissues. Upon reaching high grade dysplasia however, the chance of the patient developing adenocarcinoma is much greater (49) and in situ OADC/early grade OADC is often seen in these specimens should sampling be adequate. These patients therefore require urgent therapy. The overall risk of developing cancer or high grade dysplasia is reported as 0.28% per year in men and 0.13% per year in women (47).

Figure 1.5: Oesophageal Microscopy



Oesophageal microscopy demonstrating Normal oesophageal squamous mucosa (a), Barrett's metaplasia in squamo-columnar junction (b) and Oesophageal Adenocarcinoma (c)

Risk factors for the development of BO are similar to those of OADC and include GORD (50-54), being male (55) and increasing age (51, 53). Other risk factors seem contradictory in the literature, particularly the association with obesity as body mass index seems to have poor correlation (10, 56) whereas abdominal circumference (56) and waist to hip ratio (57) seem superior predictors.

1.10: The Development of Oesophageal Adenocarcinoma from Barrett's Oesophagus

Barrett's adenocarcinoma develops along the well-established metaplasia-dysplasia-carcinoma sequence, a process driven by chronic inflammation (58). It is assumed that OADC originates from Barrett's Oesophagus in the majority of cases. Despite this, most patients will never have had investigations or be known to have BO until their cancer is diagnosed and resected. In many OADC cases, upon histological review, BO will be identified surrounding the tumour. However this is not observed in all cases. It is assumed therefore that where BO cannot be identified, the tumour must have overgrown any remaining BO. It is also assumed that, where BO can be readily identified in these samples, this was the origin of the tumour and that this specimen of BO must be "progressing" BO (BO which has gone on to form a Barrett's Adenocarcinoma (BA)). Recent publication has shown only a small percentage of patients with OADC have a prior diagnosis of BO (7.3%) but where this diagnosis is known, patients are diagnosed with tumours at a significantly earlier stage (59).

The assumption that BO is the origin of the OADC, where there is evidence of metaplasia at the margin of the tumour features repeatedly in genetic research

publications in the field, and is a point which is seemingly accepted with very little argument, despite its obvious pitfalls.

We currently lack the ability to be able to determine, at time of diagnosis of BO, which patients will go on to form OADC. While other premalignant conditions have surveillance programmes in place, specifically designed to detect early neoplastic progression, there has been variation in surveillance programmes employed between individual clinicians and units. While some institutions have employed a programme of endoscopic assessment every 1-2 years, others have discharged patients with BO and only suggest re-investigation should symptoms change. The on-going MRC funded BOSS study will aim to answer which of these is the best course and aims to prospectively monitor BO patients with either frequent endoscopic assessment or a “watch and wait” policy (60), comparing surveillance endoscopy every two years with endoscopy at need.

Due to the lack of clear consensus on the management of the BO patient, guidelines have recently been published by Fitzgerald et al on behalf of the British Society of Gastroenterology (51). These guidelines suggest routine surveillance every three to five years in the BO patient, with treatment suggestions for those with dysplasia. The differentiation between low and high risk BO is currently based on length of BO segment, with the differentiation currently being set at 3cm. There are no routine biomarkers in use to help stratify patients into low or high risk of cancer development.

It is clear that there is a requirement in this condition for a robust method of predicting which patients will progress both to dysplasia and cancer at the time of diagnosis of non-dysplastic BO. It is highly likely if these patients experience

worsening symptoms they will already be advanced along this pathway at the time of re-presentation. This proposed surveillance programme for patients with BO will generate a huge work load given the ever increasing incidence and the relatively low conversion/progression rate. Cost implications do need to be considered and will no doubt be extremely expensive given the current need for OGD as a diagnostic test. As a potential replacement for an OGD, there are on-going studies to assess a novel device known as the “cytosponge.”

Figure 1.6: The Cytosponge (61)



This device is swallowed as a capsule, the coating of which dissolves in the stomach acid to release a small sponge, which is then pulled out through the oesophagus and samples cells in the process. While early in its development, provisional results seem promising in terms of its sensitivity and specificity and this test seems preferable for patients. It also has the advantage of being significantly cheaper (61-63) than standard endoscopy and can be performed in the community.

While surveillance strategies will detect early neoplastic change in the BO patient, there is huge need for a method to be able to risk stratify patients with non-dysplastic BO. Numerous endoscopic treatment modalities now exist for the management of patients with BO including those with dysplasia (64, 65). These include endoscopic mucosal or submucosal resection and ablation therapy, all of which are designed to remove/destroy the areas of abnormality. Having an established treatment proves the need for risk stratification in these patients as the pathology can be arrested prior to the development of a cancer, if high risk patients can be identified more reproducibly. Such a tool would enable a risk stratified surveillance programme allowing a focussed approach, investigating only those deemed to be at high risk of progression to cancer. Those with a predicted low risk would be reinvestigated only should their symptoms change/worsen. This would result in significant advantages to both patients and the health care system.

Avoiding unnecessary endoscopy in a low risk population would have the obvious advantage of not requiring an unpleasant investigation but would have large cost savings over surveying all those with BO. An evidence based surveillance programme for those deemed high risk should see an improved disease detection rate, in addition to diagnosing those who have developed cancer at a much earlier and therefore more treatable stage. This should also therefore lead to improved survival in this group.

A surveillance strategy designed to detect pre-neoplastic dysplasia would correlate with cancer avoidance as we now have acceptable treatment methods for such patients to arrest progression before they develop cancer. This would therefore be a robust and financially viable tool to be implemented in every day clinical practise.

This thesis will aim to develop a robust biomarker for this task, leading to development of a clinically useful test which could be employed to risk stratify this patient group.

1.11: Epigenetics of Oesophageal Adenocarcinoma and Barrett's Oesophagus

1.11.1 Epigenetics

The role of somatic mutation in the initiation of tumourgenesis is well understood, with multiple pathogenic driver mutations being identified across multiple tumours types, most notably *TP53* (66-69). However, the role of epigenetic modification in tumourgenesis is less well understood, but its role in development of many cancers is increasingly being described (70-72). Moreover, epigenetic change may well happen much earlier in the cancer developmental pathway and, as a direct result, provide a more appropriate target for both predicting its development and potentially arresting tumourgenesis should a suitable epigenetic modulator be identified.

While direct genetic effects exert their role by altering the deoxyribonucleic acid (DNA) sequence, epigenetic change elicits the same overall effect while maintaining the original DNA code (73). These effects seem to have greatest impact in the pre-neoplastic phase but also appear to be most prevalent in the inflammatory environment in which many solid tumours flourish (74, 75). Therefore it is likely that epigenetic changes are driven by this inflammatory state.

Epigenetics includes the study of several different mechanisms of transcription control, the most common of which being histone modification and DNA methylation. Histones are crucial in the "packaging" of DNA and they are responsible for the formation of nucleosomes. The major protein components of chromatin are histone

proteins. Together, they may be considered to act as the structure around which DNA will wrap and fold, thus dramatically decreasing the length of the DNA strand. The proteins formed by histones can have their shape altered by a modification in the sequence of these amino acids. Numerous modifications have been described, although very few have a known effect. It is better understood however that modifications in the amino acid sequence of histones can either increase or decrease transcription and a small number of these have now been described (76).

1.11.2: DNA Methylation

This process, in its most simple form, is the addition of a methyl group to the DNA chain and thus represents a form of alkylation. In humans, DNA methylation occurs at cytosine-phosphate-guanine sites (CpG sites). Methylation of a cytosine will result in its conversion to 5-methylcytosine. This process in humans is catalysed by the enzyme DNA methyltransferase.

It is reported that most human CpG sites are methylated (80-90%). However, there are areas within the genome that are “gene rich” and these have become known as CpG islands. These regions are required to be greater than 200 base pairs, and contain a guanine-cytosine percentage of greater than 50%. Interestingly, none of the CpG sites within these areas are methylated in the normal state. These areas are common in the promoter region of most genes, including all those which have ubiquitous expression levels (77, 78).

CpG methylation is now understood on the whole to correlate inversely with expression (79). Therefore a fully methylated gene should be expected to have no expression and visa-versa, although there are some genes where the converse

applies. It follows that if methylation levels are altered in various disease and pre-cancerous tissue states, then expression of genes of interest can be altered. This is of maximal relevance in Tumour Suppressor Genes, where the inactivation of such genes through methylation can lead to uncontrolled cell proliferation and the formation of a tumour.

1.11.3: Methylation in OADC

Aberrant DNA methylation is now known to be a common event in many human tumours and this is seemingly also the case in OADC (80). In the study of the oesophagus, methylation has been reviewed in the context of varying tissue states (BO, low grade dysplasia (LGD), intermediate dysplasia and high grade dysplasia (HGD) as well as OADC) both as an adjunct to diagnosis (specifically in the context of dysplasia where this can be difficult to assess microscopically) but also as a biomarker which can be detected in patients tissue, and perhaps even more usefully in blood.

Numerous implicated genes have been reported in the literature in the assessment of both BO and OADC. The majority have focussed on the discovery of a biomarker for the identification of cancer. Less commonly, researchers have investigated the role of methylation in BO and fewer again have proposed this as a tool for predicting the progression of BO towards OADC (81).

When reviewing the literature in this area, it is apparent that most of this methylation research has been carried out on small numbers of patients. Another concern is that all clinical subsets (defined by tumour stage) of OADC have been treated as one group, rather than distinct entities. There is no data available to gauge whether this

is reasonable, or that all tumours behave in the same way. The variation in clinical behaviour between aggressive and non-aggressive tumours might well suggest an underlying difference in the biology of these tumours, which is likely to be reflected in methylation status.

Several other observations can be drawn specifically in relation to BO. Many of the samples used in the literature fail to discriminate between BO with or without dysplasia and it would be reasonable to assume that dysplastic BO, which by definition already progressed along the metaplasia-dysplasia-carcinoma pathway, will behave in a very different way and therefore have different methylation signatures to pure metaplasia.

Caution must also be used when reviewing the literature in the context of progressing and non-progressing BO. Due to the often long time scale involved before progression to OADC may be seen (20+ years) and the relatively small number of patients with BO who will progress beyond pure metaplasia, it is a challenge to identify suitable patients and therefore tissue for the retrospective analysis of BO progression. Conducting this type of study prospectively, while more rigorous scientifically, would inevitably require a lengthy study period. As a consequence, researchers have routinely taken samples of BO at the edge of tumour resections to be representative of BO which has progressed. This assumes that the genomic changes will be the same as a sample of BO where there was no tumour. This methodology assumes that there is no field effect change in the oesophagus caused by the cancer. In the absence of good quality evidence, this assumption may not be reasonable. It was deemed vital in this study not to make the same assumptions as these remain unproven and this resulted in much time being spent obtaining accurate

tissues to avoid such potential bias and strengthen any conclusions that are generated.

1.11.4: Inflammation in the Oesophagus

It is thought that the primary driver initiating metaplastic change in the oesophagus is chronic acid exposure causing inflammation (82). While in some patients there is an inherent weakness of the lower oesophageal sphincter mechanism, this is in general a disease of the overweight and obese where an increase in intra-abdominal pressure will facilitate and drive gastro-oesophageal reflux in the presence of an otherwise normal sphincter complex (18, 55, 56, 83).

The link to inflammation is supported by the observation that anti-inflammatory drugs can lead to a reduction in the number of cases of OADC (20, 84, 85). A chronic inflammatory state is well known to drive epigenetic change (86) and may activate or suppress numerous pathways due to varying methylation signatures.

1.12: Methylation Markers in Oesophageal Adenocarcinoma and Barrett's Oesophagus

The literature reports numerous methylation markers which have the ability to differentiate an oesophageal tissue sample into normal mucosa and OADC. These reported markers and their relevance are described below, differentiated by the reported gene of interest.

1.12.1: P16/CDKN2A

Many tumour types demonstrate aberrant methylation and/or mutation in *CDKN2A* including colorectal cancer and pancreatic cancer (87). *CDKN2A* is a cyclin-dependent kinase which functions as a tumour suppressor gene by regulating cell cycle progression through the G1/S restriction point. It achieves this function by binding to cyclin-dependant kinases 4 and 6 which in turn prevent phosphorylation of the retinoblastoma protein.

Mutations found in *CDKN2A* in BO and OADC have been shown to be consistent with chronic inflammation and therefore show oxidative DNA damage as a marker (88). Studies have assessed its relevance in both BO and OADC and its ability to be used as a biomarker of either cancer or progression. Several studies have now shown that *CDKN2A* mutations and variation in methylation levels are a very early event in both BO and its surrounding squamous epithelium (89). They also appear to occur and be detectable long before other important mutations for example in *TP53*. *CDKN2A* methylation was assessed in 2002 by Bian et al. They concluded that *CDKN2A* methylation was observed in 0% of normal tissues, 33% of metaplastic BO tissues and 82% of OADC. In addition, loss of *p16* expression correlated with *CDKN2A* methylation ($p < 0.00001$) and that this methylation was the predominant cause of loss of expression (90).

1.12.2: VIM (Vimentin)

VIM hypermethylation has been reported in colorectal carcinoma and is occasionally used as a biomarker to identify cells which have mesenchymal origin, for example identification of certain sarcoma tumours. This protein forms a major component of

the cytoskeleton of mesenchymal cells, which explains its use in the context of sarcoma recognition. In the context of the oesophagus, Moinova et al in 2012 demonstrated that *VIM* is methylated in 91% of BO, 100% of samples showing high grade dysplasia and 81% of OADC specimens (91). However, in normal control oesophageal tissues, *VIM* hypermethylation was not seen.

The use of this marker in serum for non-invasive, non-tissue dependant diagnosis was investigated by Zhai et al. They concluded that *VIM* methylation could be used to distinguish between tissue types, but also that this can be detected in cell free circulating DNA thus providing a potentially useful, none invasive biomarker (64). This study was small and would need validating in a large clinical trial.

Lind et al studied *VIM* hypermethylation in GI cell lines including bile duct, gall bladder, stomach, liver and pancreas. The group showed that overall 75% of these cell lines were methylated. They also report that methylation is seen in other non-GI cancer cell lines including breast and uterus. They conclude therefore as a blood marker of colorectal/Gastro-intestinal malignancy this is insufficiently specific and that stool analysis could be more beneficial to detect GI tumours prior to invasive diagnostic testing (92) but this has not been validated in a large cohort.

1.12.3: *P14ARF*

This tumour suppressor gene is a key regulator of the P53 pathway. Investigations by Huang et al found expression of this gene to be highly significant in relation to both progression of tissue from BO to cancer but it also appears to have a role in prognosis, with those patients demonstrating silenced genes having a far poorer outcome than those with normal expression. The mechanism of this silencing

appears to be both CpG methylation but also Histone-H3 methylation and these appear to be independent in this study.

They also demonstrated that silencing of this gene can be reversed through cell line exposure to 5-aza-2'-deoxycytidine (an inhibitor of the enzyme DNA methyltransferase). The ability to re-express *p14ARF*, when it may be associated with up to 30% of OADC, would be a potentially useful tool in halting the progression of BO to OADC but also improving prognosis in patients in whom the disease is already well established (93).

1.12.4: *CDX2*

Caudal-related homologue 2 (*CDX2*) is a gene involved in the regulation of early embryonic development, specifically axial patterning of the alimentary tract and is expressed in cells from the duodenum to the rectum (94). It is known to be expressed in the colon and small bowel but not in the stomach or the oesophagus. Significant research has looked into the effects of this gene in colorectal cancer with multiple studies demonstrating a loss of expression of this gene (95-97).

In the oesophagus, it has been demonstrated that *CDX2* expression varies depending on the tissue state examined. 37% of oesophagitis specimens showed a greater expression than normal oesophagus tissue, compared to 91% of BO tissues. In OADC, this decreased to 57%. This is compared to 0% being seen in normal oesophageal tissues. However, when examining the methylation patterns for the promoter region of *CDX2* in both normal oesophageal tissue as well as OADC, there was no difference seen despite variation in expression (98). This was also observed by Makita et al where they confirmed expression of *CDX2* irrespective of its promoter

methylation status (99), although the method used (methylation specific polymerase chain reaction - MSP) may have overestimated methylation it is an insensitive technique for ascertainment of methylation status.

1.12.5: JAK/STAT/SOCS

Cytokine signalling depends on the activation of intracellular molecules including JAKs (Janus family kinases) and STATs (signal transducers and activators of transcription). The later discovery of the SOCS (suppressor of cytokine signalling) family of molecules has led to further interest in these pathways and therefore increased investigation. SOCS, which act as negative feedback regulators of cytokine signalling, are novel JAK regulatory proteins 2, 3 and 4.

Tischoff et al in 2007 investigated the role of SOCS-3 and SOCS-1 in the pathogenesis of OADC. They demonstrated that although SOCS-1 has a minor role in this process, SOCS-3 promoter methylation occurred in 74% of OADC. In cell lines, this correlated with a decrease in SOCS-3 transcription and therefore protein expression, however they were able to reverse these levels again using 5-aza-2-deoxycytidine (100).

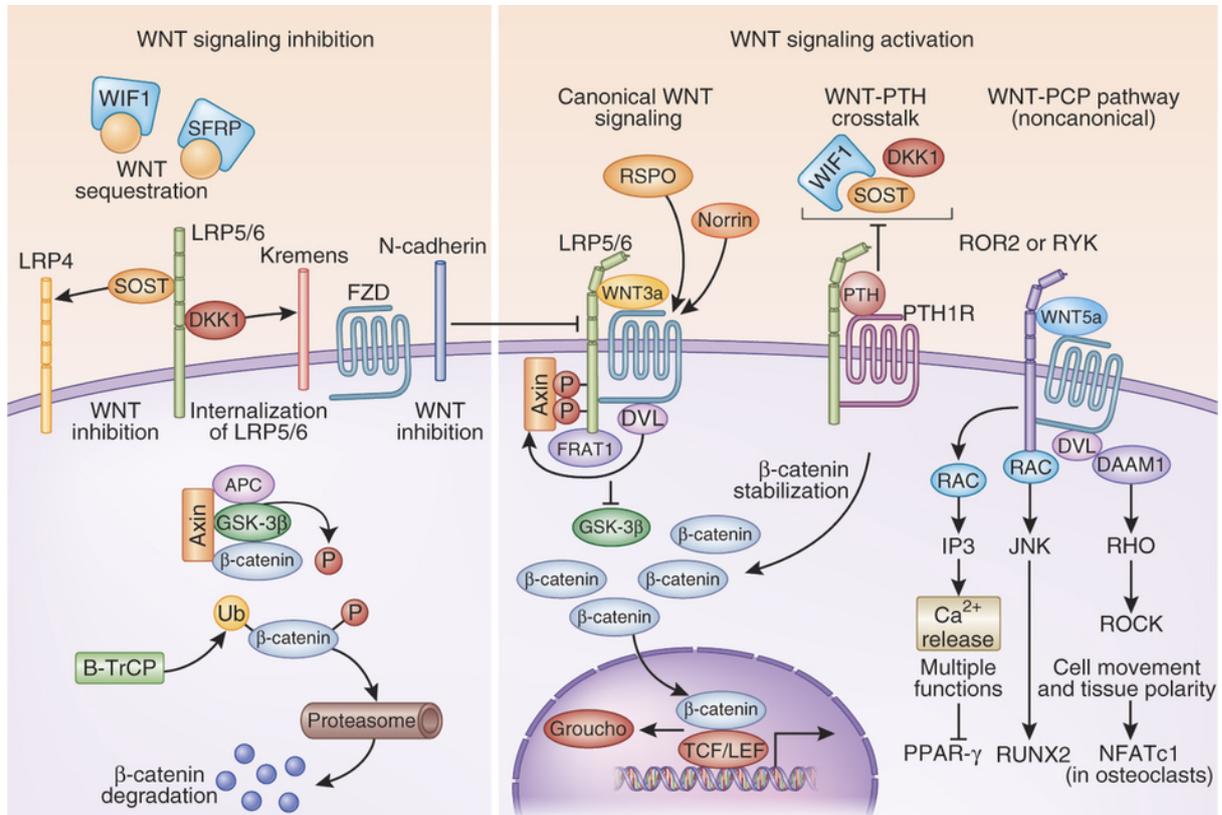
1.12.6: Wnt Signalling and Secreted Frizzled-Related Proteins

The Wnt signalling pathway is an intricate ensemble of proteins that has a major role in gene transcription and cell shape through cytoskeletal regulation and calcium homeostasis. In short, the Wnt signalling pathway begins by the binding of a Wnt ligand onto a Frizzled receptor protein located in the cell membrane. Activation of this frizzled receptor in turn leads to the activation of the Intra-cytoplasmic dishevelled protein. This ultimately results in the build-up of β -catenin within the cell.

β -catenin build up is usually avoided due to it being ubiquitinated which results in its movement to the proteasome for digestion. This is achieved by a complex, formed from protein phosphatase 2A, Axin and APC. Binding disrupts the function of this complex allowing the accumulation of β -catenin which moves to the nucleus and leads to the induction of gene transcription. It follows therefore that any part of this pathway which becomes disrupted would lead to a cell being more tumourigenic due to loss of this regulation.

Figure 1.7: The Wnt Signalling Pathway adapted from Baron R, Kneissel M

(101)



The role of the *sFRP* genes have been evaluated in BO and OADC. The first of these by Zou et al investigated the methylation pattern seen in normal oesophageal epithelium, BO and OADC. The study analysed *sFRP* 1, 2, 4 and 5. It concluded that, as expected, there were low levels of methylation in normal tissues (10, 67, 0 and 14% respectively) which increased dramatically by the time metaplastic tissue was analysed with hypermethylation seen in all genes (81, 89, 78 and 73% respectively). Methylation increased only marginally further between BO and OADC. As anticipated, hypermethylation lead to a decrease in protein expression in cancer

cells and this correlated inversely with grade and stage of tumour (102). This suggests that many of the methylation patterns of OADC are already well established by the time BO is seen. However, this study should be interpreted with caution as many of the BO samples analysed (33/40) showed evidence of dysplasia, with 13 samples being high grade. It is therefore unclear from these findings whether the *sFRP* genes are methylated in BO alone or whether this is seen in the context of dysplasia only. When this group analysed BO vs LGD vs HGD, *sFRP* expression decreased as dysplasia increased. This study proves that the varying levels of dysplasia should be analysed in their own right as individual sub sets and it is not representative to group all degrees of dysplasia into a single BO sample cohort.

Clement analysed *sFRP1* and 2 in addition to *APC* hypermethylation, with subsequent lack of expression being demonstrated in all BO and 95% of OADC. Interestingly, and following the pattern seen above, *WNT2* expression in this study increased as dysplasia worsened (103). This again is in-keeping with the observation that pure BO should not be analysed with dysplastic samples.

The role of Wnt inhibitory factor-1 was analysed in the progression of BO to OADC (104). Findings revealed *WIF-1* methylation and subsequent silencing was greater in OADC than in normal matched tissues. Moreover, *WIF-1* methylation was also seen in BO samples however the level of methylation was greater in those patients with BO who had progressed to cancer and much lower in BO samples where there had been no such progression, again demonstrating a potentially clinically useful biomarker of disease progression.

The presence of free serum hypermethylated DNA of the *APC* gene was analysed in patients with OADC. This was only seen in 25% of OADC patients however its presence was more likely to correlate with poor disease prognosis when matched to survival data (105). This again questions the possibility that biomarkers found in the tumour tissue can be detectable in the blood and this would provide a far more acceptable test to patients should one be identified in this study.

1.12.7: Methylation of Cell Free Circulating DNA as Biomarkers

More recently, genome wide comparisons of methylation patterns seen in cell free circulating DNA have been performed between signatures seen in tissue samples in comparison to those seen in the plasma. Zhai in 2012 found in total 911 loci could discriminate OADC patients from normal, 533 that distinguished BO from OADC and a further 46 that differentiated BO from normal (106). This again suggests a huge area of potential future blood based biomarker development.

1.12.8: Other Markers

Smith *et al* examined the methylation signatures for 9 genes known to be associated with other tumours and reviewed these in OADC and other oesophageal tissue states. The 9 examined genes included *APC*, *CDKN2A*, *ID4*, *MGMT*, *RBP1*, *RUNX3*, *SFRP1*, *TIMP3* and *TMEFF2*. The methylation signatures of these genes were analysed in normal oesophageal tissue, BO (including tissues demonstrating metaplasia only along with high grade dysplastic samples) and OADC. Their comparison demonstrated that BO tissue was closer in methylation pattern to OADC than to normal oesophageal tissue (107).

Numerous other studies have reviewed the methylation patterns of single genes in OADC. The O6-Methylguanine-DNA Methyltransferase (*MGMT*) gene which is involved in DNA repair was investigated and despite hypermethylation being seen in BO and OADC samples, no correlation was seen in terms of disease severity or survival (108).

Kaz *et al*, applied a Illumina GoldenGate methylation array to samples of squamous epithelium, BO, BO with HGD and OADC. This group demonstrated overall methylation to increase between all groups except HGD vs. OADC – again confirming that HGD approaches OADC from a methylation perspective. This group proposed 20 differentially methylated regions and genes between tissues. These genes included *RBP1*, *ALPL*, *POMC*, *VCAN* and *TIAM1* (109).

1.13: Methylation in Progression of BO to OADC

The ability to detect methylation signatures that can predict progression from BO to OADC, the so-called “high risk” BO, would provide a highly useful predictive biomarker. This is particularly the case should there be a marker which can perform this prediction at the stage of non-dysplastic BO. Investigation of this phenomenon is substantially more difficult as it requires sufficient time for patients to progress to OADC. Another weakness is that BO may have been sampled at the edge of a tumour specimen and this taken to represent BO which has progressed to cancer. It is unclear whether this is a valid assumption and whether field effect methylation change will affect any result seen.

Wang *et al* assessed both *APC* and *p16* methylation in normal oesophageal tissue, BO and OADC. A separate subgroup analysis was undertaken between patients

who progressed beyond their baseline pathology (i.e. BO to dysplasia or OADC.) This group demonstrated that samples from patients who progressed had significantly elevated methylation of both *APC* (86% vs. 40%) and *p16* (100% vs. 33%) when compared to those who did not progress. It was also shown that where there was no methylation of *APC* or *p16*, none of the patients progressed to cancer suggesting these to be significant biomarkers (110).

The role of *MGMT* was studied in progression from BO to OADC using methylation-specific PCR on 133 patients. Hypermethylation was demonstrated in 78.9% of OADC and 88.6% of BO samples. In normal tissue, this was seen only in only 21.4% of samples. Assessment of subsequent protein expression by immunohistochemistry correlated hypermethylation with decreased protein levels. This was significant as decreased expression correlated with a worse disease stage as well as lymph node positivity (111).

PKP-1, found in epithelial layers as part of the desmosome complex, was examined by Kaz *et al.* This group showed that while low methylation was seen in normal squamous tissue (9.1%), increasing methylation was observed throughout the states of pathology; BO - 12.8%, HGD/cancer – 33.3%. When this gene was then later knocked down in a BO cell line, it was demonstrated that cells had a greater motility (112), which would be in keeping with the process expected in cancer cell.

A cohort of progressors vs. non-progressors was analysed in a panel of genes including *APC*, *TIMP3* and *TERT*. Samples were collected from patients with BO and then observed for 4-10 years. Over this period, several patients developed OADC and the methylation patterns of these samples were contrasted to patients who over

the same period had experienced no progression. The results demonstrated an obvious difference between those who had and had not progressed. In those samples where disease progression had been seen, 100%, 91% and 92% hypermethylation was observed for the above three genes respectively. When contrasted to patients who had not undergone any progression, the levels observed were 36%, 23% and 17% respectively (113).

A similar study (114) examined several other genes, many of which have been implicated previously. However, they reported significant hypermethylation in progressing samples in *p16*, *RUNX3* and *HPP1*. The finding that *TIMP3* and *APC* were independent risk factors for progression was not replicated in this study.

A much larger study included 50 progressing patients and 145 none progressing patients. Of the panel of genes analysed, significant hypermethylation was observed for *HPP1*, *p16* and *RUNX3* in the progressing patients. Although this study examined a panel of potential methylation candidate genes, only the above three genes showed statistical differences between groups. Results showed *HPP1* ($p=0.0025$), *p16* ($p=0.0066$) and *RUNX3* ($p=0.0002$) were significantly hypermethylated in progressing samples (115).

AKAP12 (A-kinase anchoring protein) is a known tumour suppressor scaffold protein. Jin *et al* examined *AKAP12* methylation in a large study, which examined 259 oesophageal samples in which methylation was 0% for normal tissue, increasing to 38.9% in BO and 52.2% in HDG and OADC. They also demonstrated a field effect of methylation, whereby true normal squamous epithelium has a lower methylation level than normal tissue from a cancer patient (116). This study, proposing a methylation

field effect change in the cancer patient, would bring into question the results of studies which have used BO at the edge of a tumour resection to represent progressing BO.

Agarwal *et al* performed an array based approach to compare the methylomes of progressor vs. non-progressing patients. Progressing patient's original biopsies began either with no dysplasia, indefinite dysplasia or low grade dysplasia, making comparison difficult. Their subsequent analysis of the top 25 differential methylation patterns surprisingly found only 3 gene regions with hypermethylation amongst the progression group (*Pro_MMD2*, *Pro_ZNF358* and *Intra_F10*) with all the others showing hypomethylation (117). This is the first such study that has demonstrated hypomethylation of genes in regard to BO progression.

1.14: Summary of Methylation Markers

There are clearly demonstrable differences in methylation patterns between tissue states. Unfortunately, a lack of uniformity in the samples selected in these studies including peri-tumour BO taken to represent progressing BO, and BO taken in numerous states of dysplasia make it difficult to establish any overall representative findings.

None of the studies use an objective, quantitative genome wide methylation platform to compare tissue types before proceeding to validation exercises and many studies include small sample numbers.

The fact that methylation signatures of varying genes clearly show significant differences between tissues suggests pursuing the study of methylation to be of value in this disease in the hope of biomarker development.

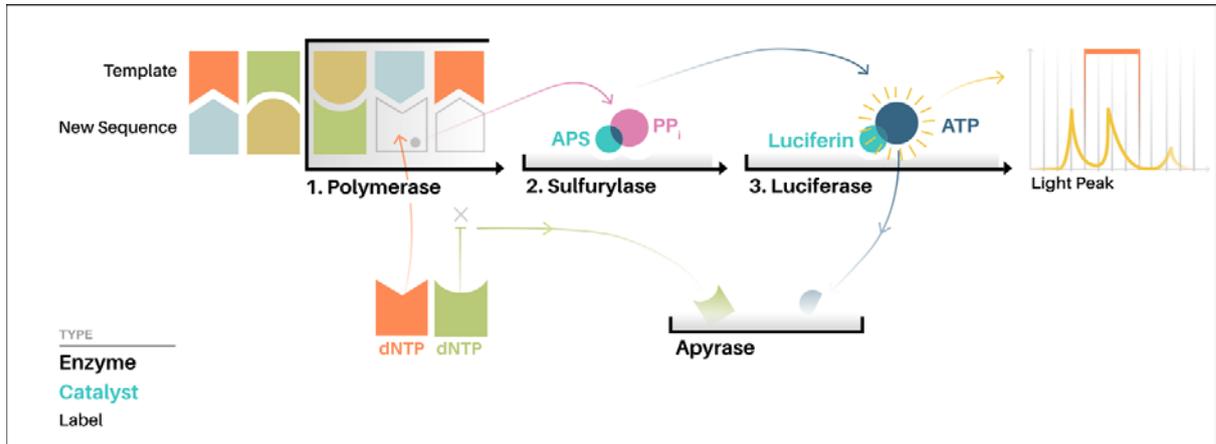
1.15: Methylation Assessment Technologies

1.15.1: Pyrosequencing

This method of methylation assessment provides reliable quantification of methylation at individual CpG sites. The process also incorporates inbuilt quality control mechanisms to ensure complete bisulphite conversion of the starting DNA material.

Pyrosequencing has become known as the “sequencing by synthesis” technique for methylation assessment. Nucleotides are added and incorporated into a DNA chain, which in turn releases pyrophosphate producing a light signal. Methylation pyrosequencing requires the input of bisulphite converted DNA. During the process of bisulphite conversion, un-methylated cytosines are converted to uracil. Those which are methylated remain unchanged. The assessment of methylation therefore comes from the comparison ratio of T and C nucleotides synthesised into the chain (118). This method remains the gold standard for methylation assessment. A diagram demonstrating the above is shown in figure 1.8. The potential problem however, is that bisulphite conversion of DNA is not always complete. Therefore the inbuilt control mechanism within the sequencing platform which makes an assessment of the success of conversion is essential. Any sample in which there is incomplete conversion will be rejected.

Figure 1.8: Pyrosequencing Diagram



This image has been created during "DensityDesign Integrated Course Final Synthesis Studio" at Polytechnic University of Milan, organized by DensityDesign Research Lab. Image is released under CC-BY-SA licence. Attribution goes to "Jacopo Pompili, DensityDesign Research Lab".

1.15.2: Golden Gate Methylation Array

This now out dated methylation assessment platform allows the assessment of 1536 CpG sites. This method again uses a bisulphite converted DNA comparison to assess methylation of a particular CpG site. Four probes are present for each of the CpG sites to be quantified. Hybridisation between cytosine vs thymine (which has almost identical properties to uracil) provides the percentage methylation of a particular CpG site in a given sample. This technology, while still widely reported in the literature, has generally been superseded by the Illumina 450 human methylation system which examines 485,000 probes in a genome wide assessment platform. The reason that this platform is now less commonly used is due to its flaws, which

include a limited panel of genes only, limited CpG' targets (1536) and assay bias meaning methylation values are not reliable.

1.15.3: Illumina Methylation 450 Array

This latest technology for epigenetic-wide association studies, allows the direct comparison of 485000 individual CpG sites per sample. Sites chosen were selected from a panel of experts. As described by Illumina, the array (119):

“covers 99% of RefSeq genes, with an average of 17 CpG sites per gene region distributed across the promoter, 5'UTR, first exon, gene body, and 3'UTR. It covers 96% of CpG islands, with additional coverage in island shores and the regions flanking them. Further content categories requested by the Consortium include:

- CpG sites outside of CpG islands
- Non-CpG methylated sites identified in human stem cells
- Differentially methylated sites identified in tumour versus normal (multiple forms of cancer) and across several tissue types
- CpG islands outside of coding regions
- miRNA promoter regions”

Again the array requires bisulphite converted DNA as a starting material followed by amplification of the entire genome using DNA polymerase. The resulting DNA is then fragmented and purified before being hybridised to the bead chip, which is crucial in this platform. In its most basic format, each CpG has two bead types, one of which

will represent the methylation strand and other the none methylated. Immunohistochemical staining of the comparison samples allows an assessment of methylated to unmethylated DNA strands when the chip is read.

1.14: Thesis Aims

The work presented in this thesis will aim to determine the relevance of methylation in the oesophagus. Specifically, I aim to assess whether methylation patterns play any role as biomarkers of adenocarcinoma utilising pre-existing publications in this area.

Should this be successful, I will employ a novel genome wide approach to seek biomarkers that have yet to be identified but which may be more robust by use of a genome wide interrogation of CpG sites. Any such identified markers might also have implications for tumour biology, and these will be examined.

Finally, I aim to use the same technology but with novel tissue sample selection to determine whether methylation markers can be utilised to predict a cohort of high risk non-dysplastic BO. Additionally, I will assess methylation patterns in tumours with and without lymph node metastasis in an attempt to supplement the current cancer staging process, providing justification for oesophagectomy in early tumours where microscopic nodal disease is currently impossible to predict.

2. Research and Thesis Hypotheses

The literature review shown before demonstrates that there is both a clinical need for biomarker/s to aid in the diagnosis of OADC, and provide inroads to the biology of this tumour.

In addition, a discovery exercise to determine the ability of methylation to stratify the risk of progression in patients with Barrett's Oesophagus towards cancer will be performed. To date, there has been no genome wide methylation study comparing these two groups.

I therefore hypothesise that:

1. Methylation signatures will have the ability to differentiate tissue types to discriminate normal squamous tissue from Barrett's Oesophagus and adenocarcinoma tumour tissue.
2. Tissue discriminating methylation markers will have functional significance within the cell. This will be assessed using an OE33 cell line model and methylation modulation using 5AZA. The effects of demethylation on expression will be assessed
3. Methylation signatures will be able to differentiate between patients with non-dysplastic Barrett's Oesophagus who over time do and do not progress to oesophageal adenocarcinoma.
4. Methylation patterns will be able to supplement the clinical staging system where-by more aggressive tumours with lymph node involvement will reveal a different methylation signature to those in which there is no lymph node involvement.

3: Material and Methods

3.1: Materials

3.1.1 Tissues

Tissues used in this study were a combination of formalin fixed paraffin embedded (FFPE) samples and a smaller collection of fresh tissue samples. FFPE tissue could be used in this context both due to availability but also due to the stability of methylation patterns in historical tissues. In addition, the ability to use FFPE tissue for methylation assessment makes a clinical application easier.

3.1.2: Chapter 4 Samples

All samples were FFPE resection specimens. OADC resection patients were identified from theatre records in the UHB operating log. The log was tracked retrospectively until sufficient samples were identified. Cases were included where the patient had undergone oesophagectomy (open or minimally invasive), and where the operation had been successfully completed. Abandoned procedures or those where the patient had an oesophagectomy for squamous cell cancer or benign diseases were excluded. Other rare tumour types were also excluded from the study.

Tissue was acquired under an ethically approved request from the University of Birmingham Human Biorepository (HBRC). Initially, diagnostic slides were retrieved from the archives and were then reviewed by a consultant histopathologist. The diagnosis was of OADC was confirmed before determining whether there would be sufficient volume of the required tissue type in each sample. Tumour and adjacent normal squamous mucosa was selected from each patient. Where present, and in

sufficient volume to allow sampling, BO was also identified and tissue blocks requested. Where possible, blocks were selected where the entirety of that block represented the tissue of interest. While this was frequently possible, this was not always the case. Some blocks needed cold steel macro dissection and where required the area of interest on each block was identified by the same pathologist on a reference H+E slide, which was later used to cut out the appropriate tissue from the block sections.

The selected block guide was re-submitted to the HBRC, along with instruction as to whether blocks required cutting into scrolls (where the entire block represented the tissue of interest) or trimmed and adhered to microscope slides (where further dissection was needed).

60µm of tissue was requested in total. Previous experience within the research group had determined that a minimum of 60µm of tissue was needed for adequate DNA extraction. Therefore, where there was uniformity of tissue in the blocks and no further dissection was required, 3x20 µm scrolls of tissue were cut. Where further dissection was necessary, 8 µm tissue slices were cut and adhered to a diagnostic glass microscope slide. The tissue was not stained and no cover plate was applied. Eight of these slides per block were cut to allow a minimum of 60 µm of tissue.

Where slides were produced and further dissection was needed, this was performed under microscopy using a scalpel blade. The resulting tissue of interest was placed in a microfuge tube for DNA extraction.

3.1.3: Chapter 5 Samples:

Many of the samples used in Chapter 4 were also utilised in Chapter 5. However, additional samples were also required for revalidation studies. It was desirable that these came from a secondary internal sample cohort, in addition to an external sample cohort. Therefore, the theatre log review was extended to identify a further 9 cases of oesophagectomy for OADC as per the above criteria. In addition a similar cohort of cancer resection specimens were sourced from neighbouring Birmingham Heartlands Hospital utilising an existing tissue transfer agreement in place with HBRC. The cases for inclusion were identified in the local trust. Again all specimens were from OADC patients having undergone oesophagectomy.

All diagnostic H+E slides were again reviewed by the same histopathologist, and blocks cut and prepared according the criteria given above by the HBRC and again supplied as scrolls (in microfuge tubes) or on slides needing further dissection.

3.1.4: Chapter 6 samples:

3.1.4.1: Non-Dysplastic BO samples

In order to determine variation in methylation between progressing NDBO patients vs none progressing NDBO patients, external cohorts had to be sourced as there was a lack of longitudinal data in the local unit to be able to identify these tissues, and therefore extension to surrounding hospitals was required. Specifically, it was important to track patients who had had a diagnosis of BO (non-dysplastic) in whom there were serial biopsies proving the progression or not of the disease process (i.e. the patients were enrolled in BO surveillance). I am very grateful to Dr Nigel Trudgill at Sandwell General Hospital for providing access to his retrospective BO database.

Once this was obtained, all patients' cases were reviewed in turn. Dr Trudgill provided a list containing 37 patients who had progressed beyond base line pathology to OADC. However when this list was scrutinised, many of the patients had dysplastic disease at the original biopsy. Since the hypothesis for this work was to predict progression before dysplasia was seen, each of these cases had to be excluded.

In addition, many of the patients identified had no evidence that they had ever been diagnosed with OADC. This is believed to be a result of the tumour being diagnosed at a different unit. However, these patients were also excluded as definitive evidence of tumour diagnosis was required for tissue validity. While there were several patients who had only one BO biopsy prior to their OADC diagnosis, many had serial biopsies. Where this was the case, each of these temporal biopsies were obtained to allow a separate review of the changes over time as the patient progressed towards cancer. This resulted in 12 patients meeting the inclusion criteria, the majority of which had temporal biopsies. Each of these cases were reviewed again by a local GI histopathologist to ensure original diagnosis of non-dysplastic BO was correct.

None progressing samples were obtained locally. Care was taken in attempting to match these samples in terms of gender and age and year of original diagnosis of BO. To be determined non-progressing, the patient must have been enrolled in a surveillance programme for a minimum of 20 years and have temporal biopsies to within two years of the analysis. If at any point the patient had a dysplastic sample of BO, this patient was excluded. Again, each of the H+E slides were reviewed by a GI histopathologist to ensure no dysplasia was present and that samples represented BO.

Due to the preparation of OGD biopsy samples in blocks to allow microscope examination, multiple biopsies were prepared alongside each other in the same block. Not all of these samples may have been oesophageal biopsies, or have demonstrated NDBO. Therefore, the examination of the slides by a histopathologist was crucial in the identification of the specific biopsy of interest so that the precise tissue was used in downstream analysis. Much of the tissue used in this work, as a result of being from very historical samples and from such small biopsies, yielded very low amounts of DNA which was often of poor quality.

3.1.4.2: Lymph Node Positivity Samples

In order to determine markers of aggressiveness, one of two options was possible. As a marker of lymph node disease was sought, it was felt important to keep all of other aspects of the samples to be the same. Therefore, it was felt there was little point comparing for example a T1N0 tumour with a T3N1 tumour as it might be easily predicted that there would be inherent differences in these samples other than their lymph node status (especially as hypermethylation is likely to be progressive with advancing tumour stage).

This marker would demonstrate maximal clinical use in the early, small T1 tumours where lymph node disease if present is likely to be microscopic and therefore not identifiable by routine imaging techniques. However, developing a cohort of resection patients in sufficient number of both T1N0 and T1N1 proved impossible.

Due to the nature of OADC at presentation, and the need to standardise the tumour samples as much as possible, it was decided to use a T3 sample cohort. While there was an abundance of T3N1 resection samples, difficulties were encountered gaining

enough samples of T3N0 disease locally, therefore some of these samples were again acquired from the local Birmingham Heartlands Trust through the pre used tissue transfer agreement of ethically approved tissues, facilitated by HBRC. Once again, all tissues used had their H+E slides re-examined to determine diagnostic accuracy and were prepared and delivered in the same format described above – either as tissue scrolls where there was sample uniformity, or on microscope slides where further trimming was required.

3.2: DNA Extraction

Tissue (scrolls or dissection sections including wax) were placed into 1.5 ml microfuge tubes. DNA extraction was performed using Qiagen DNeasy Blood and Tissue kits, following the manufacturer's protocol. ATL buffer was added to cover the tissue. This equated to a volume of 360 μ L. 20 μ L of Proteinase K (20mg/mL) was added to each sample and the samples were incubated at 55°C, however they were vortex mixed every 2 hours during the day and left overnight. Due to the large tissue volume in some samples (normally resection specimens), there was occasionally remaining visible tissue seen after 24 hours. When this was the case, a further 20 μ L of Proteinase K was added and the samples left for a further 24 hours with periodic vortexing until no visible tissue remained.

Once digestion was completed, samples were removed from the incubator and immediately centrifuged at 10,000 rpm for 1 minute to allow the wax to sit above the sample. Microfuge tubes were left at room temperature to allow the wax on top to form a solid plug. When solidified, the wax plug was punctured with sterile pipette tip which was subsequently discarded. A new sterile tip was passed through the

opening created in the wax plug and the sample below extracted. This sample was then placed into a new sterile microfuge tube while the wax was discarded. This was repeated for all samples.

Further DNA extraction was performed by following the Qiagen DNeasy blood and Tissue kit protocol for tissue samples. 200µL of Buffer AL was added to each sample and mixed by vortexing. Following this, 200µL of 100% ethanol was added and again the samples were mixed by vortexing. The entire volume was then added to a Qiagen mini spin column and centrifuged at 8000 rpm for one minute. The flow-through was discarded. 500µL of wash buffer AW1 was added to the columns which were centrifuged for 1 minute at 8000 rpm. Subsequently, 500µL of buffer AW2 was added and samples were then centrifuged at 14,000 rpm for three minutes. All flow-through was discarded. Spin columns were placed in new sterile microfuge tubes followed by the addition of 200µL of buffer AE to the centre of the spin columns. Columns were incubated for one minute at room temperature before centrifuging at 8000 rpm for one minute. Where there was predictably a large amount of DNA from each sample (large starting amount of tissue), a second DNA elution was taken. Where starting samples were small (usually BO biopsy samples) this step was not performed.

3.3: DNA Quantification

Each sample of extracted DNA was then analysed on the Nanodrop (ND-2000) to assess DNA yield as well as purity of sample. 260/280 and 260/230 ratios were recorded to assess sample purity and minimise RNA/carryover contamination. Any sample with a 260/280 ratio < 1.7 or 260/230 ratio > 1.8 was not used.

3.4: Nanodrop

This machine uses spectrometry to analyse the DNA sample. UV Light is passed through the sample, and the receiver measures the intensity of the light able to be transmitted through. The resulting graph trace produced will not only provide the quantity of DNA in the sample, but will also define the purity of that sample, or more specifically the amount of contamination present. These purity estimates are provided by the 260/280 and 260/230 measurements. Samples which are outside the usual range were not used in experiments, as is it likely that the contamination would have skewed the results obtained.

3.5: Bisulphite Conversion of DNA

To assess methylation, it is necessary to bisulphite convert the DNA sample. In this conversion, all un-methylated cytosine are converted to uracil, whereas those that are methylated remain unchanged. This ultimately allows a comparison of cytosines to uracils in sequencing experiments. It is the resulting ratio of cytosines and uracil's that allow the pyrosequencer to establish the methylation percentage of any given sample. This bisulphite conversion was performed using a combination of the Qiagen Epiect Plus Bisulphite Conversion Kit and Zymo DNA methylation bisulphite conversion kits. Both achieve the same conversion reactions, but each have varying compatibility with downstream applications which determined the kit to be used and the temperature steps needed.

In order to perform this conversion, 500 ng of DNA was added to a sterile microfuge tube. This will be a variable volume depending on the resulting concentration of DNA achieved after the extraction elution.

3.5.1: Qiagen Protocol for Bisulphite Conversion

RNA free water was added to the 500ng DNA sample to make a total volume of 40µL. 85µL of bisulphite mix was added and finally 15µL of DNA protect buffer. This made a total reaction volume of 140µL and a correctly made mixture turned blue upon the addition of the protect buffer. This reaction volume was then placed in a thermocycler according to the following conditions: 95°C for 5 minutes to denature, 60°C for 25 minutes as incubation, 95°C for 5 minutes as denaturation, 60°C for 85 minutes to incubate, 95°C for 5 minutes to denature and a final 60°C for 175 minutes to incubate.

After completion of this cycle, a hold stage was programmed to last indefinitely. This was avoided as there was suspicion of DNA damage due to the compounds within the conversion mixture. Therefore all samples were cleaned within 30 minutes of the termination of the bisulphite cycle.

3.5.2: Qiagen Bisulphite Clean-up

After the conclusion of the bisulphite conversion reaction it was again necessary to extract the resulting converted DNA for downstream reactions. This process involved the following steps according to the manufacturer's protocol:

310µL of buffer BL was added to each reaction mixture and the sample vortexed before a further 250µL of ethanol (100%) was also added. The sample was again mixed followed by a short centrifuge spin to push the entire sample to the bottom of the microfuge tube. The resulting volume of sample was added to a bisulphite micro-centrifuge tube and spun at maximum speed for 1 minute.

After each spin, the resulting flow through was discarded. 500µL of buffer BW was added to each tube and again spun at maximum speed for 1 minute with the resulting fluid being discarded. The process was repeated with buffer BD, however before spinning the samples were incubated at room temperature for 15 minutes. After discarding the resultant liquid, a further 2 additions of buffer BW were applied and centrifuged. A final 250µL of ethanol was added and spun through the spin column before allowing all alcohol to evaporate.

In order to elute the DNA, 20µL of buffer EB was carefully added directly to the central membrane of the spin column prior to the final centrifugation of 15000g for 1 minute. The resulting DNA was stored at -20°C until needed for downstream PCR reactions. All samples were treated in this way.

Control DNA (100% methylated and 100% unmethylated) was also subjected to bisulphite conversion in this way to be used as both positive and negative controls in the PCR and subsequent pyrosequencing reactions.

3.5.3: Zymo Bisulphite Conversion Protocol (for compatibility with the Illumina 450 Methylation array):

500ng of DNA was used for this conversion reaction. Firstly, as per protocol, CT conversion reagent was made as per instructions. CT conversion reagent (Zymo) is a proprietary formulation of sodium bisulphite used in bisulphite conversion reactions. DNA was added to a sterile microfuge tube diluted with water to make a final volume of 50µL. The resulting mixture was incubated at 37°C for 15 minutes. 100µL of the pre made CT conversion reagent was added to each of the reaction tubes and mixed thoroughly by vortex. The sample was then placed in a thermocycler, and the

samples were all handled according to the modified Illumina array protocol: 95°C for 30 seconds followed by 50°C for 60 minutes. This cycle was repeated 16 times. Therefore this reaction was allowed to run overnight.

On conclusion of the thermal programme, samples were placed in ice for 10 minutes. Meanwhile, 600µL of M binding buffer was added to a Zymo spin column followed by the sample from the conversion reaction which was then vortex mixed. The sample was centrifuged at maximum speed for 30 seconds. 100µL of wash buffer was added and the sample spun again. This was followed by 200µL of desulphonation buffer. On this occasion, samples were incubated for 20 minutes before again spinning at maximum speed. The flow through was discarded after each spin.

200µL of wash buffer was added and centrifuged. This was repeated. Finally, the spin column was placed in a sterile microfuge tube before 10µL of Elution buffer was added to the centre of the spin column membrane. Again the column was centrifuged at maximum speed for 30 seconds to elute the converted DNA. Samples were immediately frozen at -20°C until ready to be used in the PCR reaction.

3.6: Pyrosequencing PCR Reaction

After selection of the appropriate gene/CpG site of interest, methylation insensitive primers were designed and sourced. Primer design was performed using Qiagen PyroMark Primer Design software v2.0. Primers were designed to flank CpG sites of interest. Where the CpG site was an Illumina CG methylation probe, the probe location was retrieved from the UCSC genome browser and FASTA sequence retrieved for -200bp to + 200bp of the target CG dinucleotide. Primer design settings were optimised to design amplicons suitable for FFPE pyrosequencing, with the

optimum amplicon size set to between 80-150bp. All other settings were as per the standard Qiagen design parameters. Primers were ordered from Sigma Aldrich Oligos, with the biotinylated pyrosequencing primer being purified by high performance liquid chromatography and the remainder by desalting.

PCR optimisation was carried out for all reactions, initially via means of a gradient PCR. Subsequent experiments revealed that although the PCR reaction was initially designed for 25µL reaction volumes, these did not routinely produce enough product DNA to obtain downstream results and therefore 50µL reaction volumes were routinely used.

To perform pyrosequencing PCR, a Qiagen PyroMark PCR Gold kit was used. The master mix volumes used in this methylation PCR reaction are shown in the table 3.1, and follow the manufacturer’s guidelines.

Table 3.1: Methylation PCR reagents and volumes

| Ingredient | X1 volume µL | Used volume (x2) µL |
|-------------------|---------------------|----------------------------|
| Master Mix | 12.5 | 25 |
| Coral load Dye | 2.5 | 5 |
| Magnesium (3.5mM) | 1.5 | 3 |
| Q5 reagent | 5 | 10 |
| Primer (20mM) | 2.5 | 5 |
| DNA | 1 | 2 |
| Total | 25 | 50 |

In addition to experimental DNA, each PCR and downstream pyrosequence was performed with 3 control reactions. These included, for all runs, a water control as well as known 100% methylated DNA and also 100% un-methylated DNA.

Table 3.2: Methylation PCR conditions

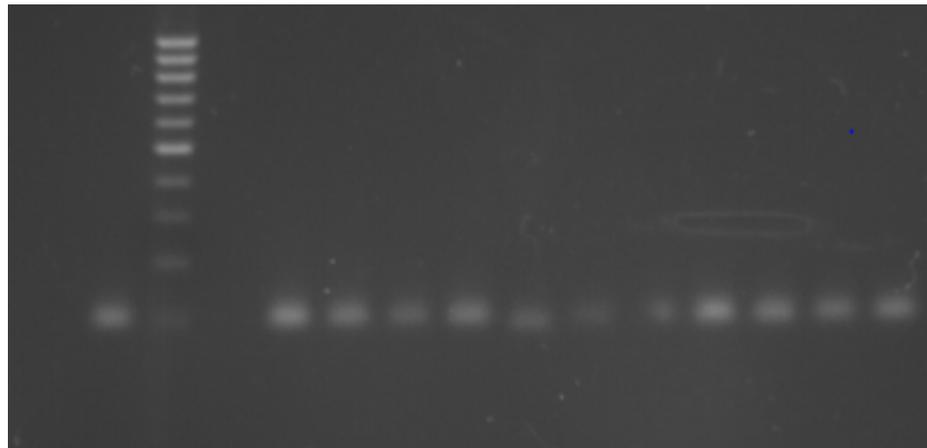
| Step name | Time | Temperature (°C) |
|--------------------------------|-------------|-------------------------|
| Activation | 15 Minutes | 95 |
| Denaturation | 30s | 94 |
| Annealing | 30s | 56 |
| Extension | 30s | 72 |
| Number of cycle repeats | 45 | |
| Final extension | 10 minutes | 72 |

3.7: PCR Validation

Success of the PCR reaction was determined by agarose gel electrophoresis. To perform this, a 1.5% agarose gel was made by boiling 0.75g of agarose powder with 50mls of TE (x1). This was heated in a water bath until all solid powder dissolved. Once cooled slightly, 2µL of ethidium bromide was added and mixed. The resulting solution was then poured into a gel mould before being allowed to cool and solidify. When solid, the resulting gel was placed in an electrophoresis gel tank filled with TE. Samples were loaded into wells. Due to the presence of the Coral Load in the original PCR reaction, no additional gel loading dye was required. Samples were run at 90V for 33 minutes (after trials to show this produced the best quality bands). A DNA ladder (Geneflow PCR Ranger, Product number L3-0001/L3-0024) was also

added to each gel to allow identification of product size. An example of such a gel is shown in figure 3.1

Figure 3.1: Agarose gel electrophoresis



This example of electrophoresis shows post PCR DNA production for all samples. The size of the DNA product can be ascertained from the DNA ladder also shown.

After completion of the run, the gel was exposed to UV light to assess the product band. Only if there was a good product band, but with no bands visible for water were the samples taken forward to downstream pyrosequencing. Multiple issues were encountered with the presence of positive water bands from the PCR. Steps were taken to avoid these which included changing of PCR reagents as well as primers which may have become contaminated. All hardware was sterilised under UV light prior to commencing the PCR reaction. This included all pipettes, tips, and microfuge tubes as well as reaction plates. Reactions were set up in a sterilised PCR hood using filter tips and clean water samples for controls. Intermittently positive

bands were seen in the water control, however when these samples were taken to pyrosequencing, no trace/result was produced. Therefore, as long as the positive and negative controls were giving the expected results and the samples revealed results other than that of the water these results were permitted.

3.8: Pyrosequencing

Pyrosequencing is a semi-quantitative method of determining genetic sequence information that utilises pyrophosphate which releases light by catalysing an ATP reaction. Resulting samples from the PCR were vortex mixed for 20 minutes with binding buffer (37 μ L) and a micro bead mixture (3 μ L). This allows the DNA product from the PCR reaction to bind to the bead in the sample. Only the DNA which was bound to this bead would be carried forward into the pyrosequence. While this was mixing, a solution of annealing buffer (35 μ L) and relevant sequencing primer (5 μ L) was formed with the complimentary primer of the gene of choice being analysed.

This primer solution was added to a pyrosequencing plate with 40 μ L of sequencing solution being added to each well of a 96 well sequencing plate. After being thoroughly mixed, for maximal DNA attachment to the beads, each sample is picked up with the hedgehog device. This device uses negative pressure to lift up the sample, retaining the beads with DNA attached while removing the liquid component. Therefore, as already stated, only DNA attached to the beads will be taken into the pyrosequence, while the remaining DNA in solution will be wasted. The sample is then passed through various wash buffers. These include 70% ethanol (200mls) followed by 0.2M sodium hydroxide solution (200mls) and finally pre supplied wash buffer (Qiagen PyroMark wash buffer (979008)) (200mls).

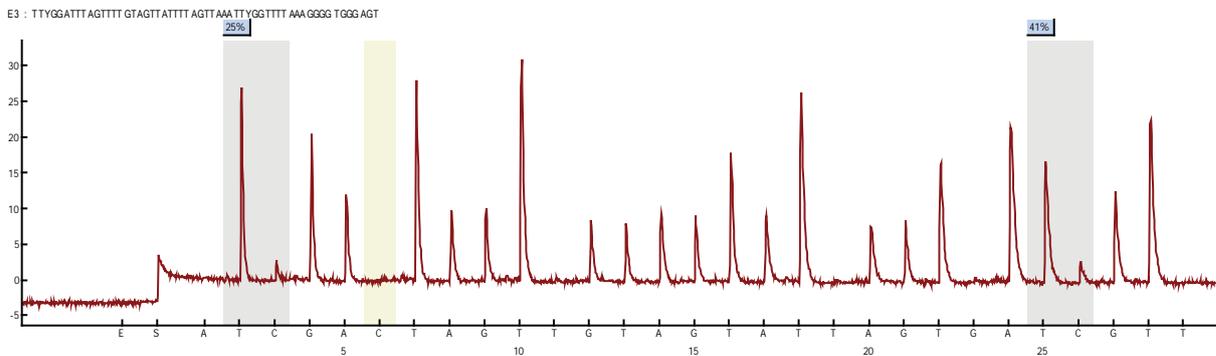
After the washing process, the beads are released into the sequencing primer by withdrawing the suction and allowing the beads to fall away from the suction device into the sequencing plate. Due to the presence of beads in this otherwise colourless liquid, the resulting liquid then assumed a cloudy appearance. Facilitation of the beads to fall into the sequencing solution was often required by agitation. This sequencing plate was then heated to 80°C for 2 minutes on a heat block.

The resulting plate, loaded with bisulphite converted DNA adhered to beads in sequencing primer solution, was loaded into the pyrosequencing machine (PyroMark Q96 ID). The sequence to be examined, based on the sequencing primer, was loaded into the software. Enzyme and substrate were added to the dispensing cartridge as well as each of the 4 nucleotides at a volume specified by the pyrosequencing software (Qiagen PyroMark Gold Q24 Reagents (970802)). The sequence was then allowed to run.

The pyrosequencer adds nucleotides to the sequencing plate in the dispensation order specified (as per the sequencing entered). As it does so, the nucleotides bind to the single-stranded DNA strand resulting from the PCR reaction which is now bound to the bead. This reaction is performed by DNA polymerase enzyme. The reaction also contains a chemoluminescent enzyme, which emits a detectable light each time a base is added, and is proportional to the amount of phosphate released in the polymerase reaction. This allows the quantification between DNA which is methylated and unmethylated by the presence or absence of a methylated cytosine and the nucleotide that is therefore added. This can be analysed as a percentage of the original DNA sequence to produce a value for the percentage methylation at each

position in the sample provided. The pyrosequencing reaction produces a trace, an example of which is shown in figure 3.2

Figure 3.2: Pyrosequencing Trace



Standard pyrosequencing trace - this primer sequence contains two CpG sites. These are shown in blue. Peaks demonstrate the light signal emitted following the addition of nucleotides. The CpG sites highlighted in blue show that these have passed the internal quality control process. Success of sample bisulfite conversion is tested in each sample. The grey bar, in this case without a peak, shows successful sample conversion.

3.9: Cell Culture

Cell line OE 33 was sourced. This was acquired from Sigma Aldrich and is also known as JROECL33. Sigma defines this cell lines characteristic as:

“This cell line was established from the adenocarcinoma of the lower oesophagus (Barrett’s metaplasia) of a 73 year old female patient. The tumour was identified as pathological stage IIA (UICC) and showed poor differentiation. OE33 express HLA-A, -B and -C antigens (MHC class I) and ICAM-1 constitutively. Expression of HLA-DR (MHC class II) can be induced by treatment with interferon-gamma. The cells express epithelial cytokeratins and are tumorigenic in nude mice.”*

*Data taken from ATTC database

This cell line was selected as its origin is from oesophageal adenocarcinoma with evidence of having arisen from BO. Cells were cultured as adhesive cells in tissue culture plates at 37°C in 5% CO₂. Media used was Dulbecco's Modified Eagle's Medium (DMEM) with the addition of 10% bovine calf serum and 1% penicillin/streptomycin mixture. Cells were observed each day while in culture and split before confluence. To split, medium was removed prior to a phosphate buffered saline (PBS) wash. Trypsin was then added (volume would vary depending on plate size used) before returning the plates to the incubator for 10-15 minutes. After this time, the cells were reviewed under the microscope to ensure they had released from the plates – when ready cells would become circular in appearance and move around the plate when tipped from side to side.

Cells were collected in fresh media before being added to a new plate at the desired concentration for pending experiments. To determine cellular concentration, cells were collected in a fixed volume before being stained with trypan blue (which allows the identification of viable vs dead cells) and then counted using a haemocytometer. After re-plating, the cells were then left for growth to establish. Splitting was generally required twice per week.

3.10: DNA Extractions from Cell Lines:

Cells were cultured in 3cm sterile plates. DNA was extracted by again first removing the media and then washing in PBS. 1 ml more PBS was added to the plates before the cells were scraped to detach them from the surface. The resultant solution was added to a microfuge tube and centrifuged at 1300 rpm for 10 minutes. This resulted

in the formation of a cellular pellet. The residual PBS was withdrawn carefully so as not to disturb the cell pellet and then re-suspended in a further 200µL of PBS

Tissue lysis buffer was subsequently added along with 20µL of Proteinase K to digest the cells releasing DNA. DNA was then extracted by the addition of 200µL of buffer AE and ethanol (100%). This resultant volume was transferred to a spin column and centrifuged at 6000g for 1 minute. 500µL of wash buffer 1 was added and the sample spun as before. 500µL of wash buffer 2 was added and the column spun at maximum speed for 3 minutes. The DNA was finally eluted into 200µL of buffer AE before quantifying by nanodrop.

3.11: RNA Extraction for Gene Expression Work:

RNA was extracted from the cells using the TRIzol technique. After removing the media and washing the cells in PBS, 1ml of TRIzol was added to the culture plate and again the cells were scrapped and mixed. The resulting mixture was added to a microfuge tube. The sample was incubated at room temperature for 5 minutes. 0.2 ml of chloroform was added before mixing and again incubating at room temperature for 3 minutes. The resulting microfuge tubes were placed in a refrigerated centrifuge for 15 minutes at 4°C and spun at 12000g. This spin generated a 3 phase separation consisting of phenol-chloroform phase (containing protein), an interphase (containing DNA) and a clear upper aqueous phase in which the RNA is found.

This upper aqueous phase was removed with care and placed in a sterile microfuge tube. To this, 0.5 ml of isopropanol was added and the sample incubated at room temperature for 10 minutes. The sample was again centrifuged at 12000g for 30 minutes at 4°C. This generated an RNA pellet visible in the bottom of the microfuge

tube. The residual liquid was removed before washing the pellet with 1ml of 75% ethanol. This was then vortex mixed and again centrifuged. The ethanol was removed and the sample pellet allowed to dry in room air. The pellet was re-suspended in 20 μ L of RNase free water and finally heated to 60°C for 15 minutes. Samples were then frozen at -80C until ready to make cDNA for expression analysis.

Prior to cDNA production for Real time PCR analysis, RNA samples were nano-dropped for quality analysis and also analysed using the Qubit for RNA quantity.

For the production of cDNA, 2ng of RNA was required in a total volume of 8 μ L with the addition of water. This experiment was performed using the Applied Bioscience High Capacity cDNA kit.

To the resulting RNA solutions, 2 μ L of random Primers (10x hexamers) were added and the reaction heated to 65°C for 5 minutes in the thermocycler before being cooled immediately in ice. After 2 minutes, 10 μ L of pre-prepared reverse transcriptase enzyme master mix volume was added. This master mix was composed as follows (Table 3.3):

Table 3.3: cDNA Production Reagents

| Reagent | Volume (μ L) |
|---|-------------------|
| Water | 6.2 |
| Reverse Transcriptase Buffer | 2.0 |
| Deoxynucleotide Triphosphates (dNTPs) (25x) | 0.8 |
| Reverse Transcriptase Enzyme | 1 |
| Total | 10 |

This reaction volume was then again placed in the thermocycler for the following cycle (Table 3.4):

Table 3.4: cDNA Production Reaction Conditions

| Step | Temperature °C | Time (mins) |
|-------------|-----------------------|--------------------|
| 1 | 25 | 10 |
| 2 | 37 | 120 |
| 3 | 85 | 5 |

The result from the above was the production of the much more stable cDNA from the original RNA which is sensitive to degradation. An assumption was made, as per all RNA expression analyses, that if a greater amount of cDNA was observed in a sample, this correlated to a greater amount of RNA which will in turn correlate to the cell expressing more of the respective protein.

3.12: Real Time qPCR for the Assessment of RNA Quantity within a Cell

This technique was used to determine the relative amount of cDNA from the original RNA extractions. In order to perform qPCR, employing a SybrGreen technique, using a 1:2 ratio the following was added to each reaction. 1.3µL sample was added to 8.7µL of the SybrGreen master mix which contained 1.3µL of primers in addition to 10.5µL of water. The result was a reaction volume of 10µL.

Table 3.5: qPCR Reagents

| Sample | | Volume (μL) |
|-------------------|---------------------------|--|
| RNA sample | | 1.3 |
| Master mix | 1.3 μL Primers | 8.7 |
| | 10.5 μL Water | |
| Total | | 10 |

Samples were run in triplicate in addition to a water sample control. Results produced a Cycle threshold (Ct) value (the value at which increase becomes exponential) from which the delta Ct can be calculated which is the difference between the Ct values. Finally, it was then possible to calculate the delta delta Ct value, which converts the Ct value into the percentage expression change. Given the logarithmic scale of RTPCR results, a change of 1Ct correlates to a 100% increase in expression.

3.13: Protein Extraction from Cells

In order to validate the RNA expression findings, protein was also measured in the cell lines to ensure that there was a correlation between RNA expression and subsequent protein content within the cell. As a result of the TRIzol technique used for RNA extraction, this also allowed protein to be extracted from the same sample. This ensured that methylation, RNA expression and downstream protein analysis was all conducted from the same sample.

As described before, the TRIzol extraction technique requires the cells to be lysed in TRIzol solution before chloroform is added and the sample centrifuged. This chloroform separation forms 3 visible layers – the top clear layer, as described above, contains RNA and was used for RNA extraction. The much smaller interphase layer contained DNA. The residual, much larger volume layer contains protein.

To this, 1.5mls of isopropanol was added. Samples were allowed to incubate for 10 minutes at room temperature before centrifuging again. Samples were spun at 12000g for 10 minutes at 4°C, after which time a visible white protein pellet was formed. This protein pellet was washed as follows:

A solution of 0.3M guanidine hydrochloride in 95% ethanol was made by adding 0.285g of guanidine to a 10 mL solution of 95% ethanol diluted with water. 2mls of this solution was added to the protein pellet. This was left to incubate for 20 minutes before centrifuging at 7500g for 5 minutes at 4°C. This wash was repeated 2 further times. Finally, 2 mls of 100% ethanol was added to the pellet and again incubated for 20 minutes at ambient temperature and spun as before. The ethanol wash was removed before leaving the pellet to air dry.

Re-suspension of the protein pellet was challenging. Initially, the resulting pellet was dissolved in RIPR buffer. However, the yield of this suspension was poor as analysed by protein assay (Pierce BCA protein assay kit 23225). This meant that further dissolving attempts were needed. To do this, untreated native OE33 cells were lysed in TRIzol and the resulting protein pellet was attempted to be dissolved in reagents shown in figure 3.6.

Table 3.6: Reagents used to attempt to dissolve protein plug

| Dissolving agent | Total volume (μL) |
|--|--|
| 1% SDS | 200 |
| 1% SDS at 50°C | 200 |
| RIPR | 200 |
| RIPR + 1% SDS (50:50) | 200 |
| NaOH + RIPR | 200 |
| 0.05% SDS with 4 M urea in TRIS-HCl (with pH 8) | 200 |

The resulting solutions were compared using a protein assay to assess amount of protein obtained. This assay was performed using spectrophotometry in which 10 μL of protein solution was mixed with 50 μL of assay solution and allowed to incubate at 37°C for 30 minutes. On the same 96 well plate, a stock solution of a known protein concentration was added. After analysis, the background noise was subtracted from the sample noise using the known protein stock as a control. This allowed the amount of protein per sample to be calculated. From this, the agent that dissolved the maximum amount of protein was 0.05% SDS with 4 M urea in TRIS-HCl (with pH 8) solution and therefore this was used for the experimental protein dissolving step.

3.13.1: Protein Analysis:

In order to compare the amount of protein in the cells, a western blot technique was employed (Mini-PROTEAN tetra hand cast system (BIO-RAD)). This allows the quantity of protein to be assessed. Western Blots allow protein analysis firstly by using voltage separation of proteins of differing sizes on a gel, alongside a protein ladder which allows the size of the protein fragment to be determined. These invisible proteins are then transferred onto a membrane, allowing proteins of interest to be located using antibody induced “staining” followed by chemoluminescent localisation. Presence or absence of proteins can be discovered using this technique, as well as relative quantity of protein in any given sample (in comparison to another sample on the same blot).

To perform a western blot, an antibody was acquired for the protein to be assessed. Following the protein assay, the total amount of protein per individual sample was calculated. It is important to ensure that each sample to be tested to contains the same overall amount of protein so that any change in amount of singular protein to be analysed would be seen.

20ng protein was mixed with 5 μ L of 5x SDS loading buffer and then heated to 100°C on a heat block. Up to 40 μ L of sample was loaded into a loading gel for the western blot which was so made that it sat on top of the running gel in the electrophoresis tank. Gels were made according to the table 3.7. Samples were allowed to run alongside a protein ladder (10 μ L PageRuler™) for 90 minutes or until the loading dye reached the base of the gel (100V for 15 minutes through the stacking gel,

followed by 180V through the running gel). This process separated each protein in the sample according to size.

Table 3.7: Ingredient list for Western Blot stacking and running gels

| | Running Gel | Loading Gel |
|-----------------------------------|--------------------|--------------------|
| Ammonium Persulfate | 10mg | 10mg |
| diH₂O | 1.9 ml | 3.7 ml |
| Gel Buffer (I) or (II) | 10 ml (I) | 5 ml (II) |
| Polyacrylamide | 8.1 ml | 1.3 μ L |
| Tetramethylethylenediamine | 60 μ L | 60 μ L |

After this, the proteins in the gel were transferred to a membrane (Hybond hydrophobic difluoride (PVDF) membrane (Amersham Hybond-LFP – GE Healthcare)) which would allow the individual protein of interest to be highlighted by the chosen antibody. The membrane was soaked in methanol for 30 seconds prior to use. This transfer process was again performed by electrophoresis. The Running gel was removed from container and placed in a transfer bracket. The transfer was allowed to run for 90 minutes at 90V. At the end of this stage, the gel was discarded if there had been successful transfer of the visible protein ladder.

The membrane was then cut in two at the appropriate place using the ladder as a marker in order to allow one of the halves to contain the protein of interest and the other half to contain a marker protein, which in this example was β -actin. Membranes were incubated in a blocking solution (10% milk in Tris buffered saline) for one hour being agitated on a rocking tray. Antibody containing solutions were

prepared, one containing the β -actin antibody and the other containing the antibody to be examined. The membranes were placed in these solutions and left to incubate at 4°C overnight on a rocking table.

The following day, membranes were washed before being exposed to the secondary antibody. This secondary antibody attaches only to the primary antibody that was specific to the protein being analysed. Membranes were washed with 1 x Tris-Buffered Saline and Tween 20 (TBST) to remove the primary antibody. 4 washes in total were performed. Membranes were then incubated with secondary peroxidase conjugated antibody for 60 minutes at room temperature. A further 4 washes with TBST were performed.

Blots were developed using the Amersham Western Blotting Detection Reagents Kit (GE Healthcare), applying the manufacturer's instructions. Reagents were mixed in a 1:1 ratio. Corresponding membranes were sealed in a plastic bag, in which 5 mL of the detection reagents were placed. The reagent mixture was moved around the bag and therefore over the membrane for 1 minute. Membranes were then placed into a developing cassette (Amersham Hypercassette – GE Healthcare). Amersham Hyperfilm MP (GE Healthcare) was positioned over the membranes and then placed into a developer. Varying time periods were used to ensure a band was produced.

Protein of interest, as well as the β -actin band was analysed to compare the overall protein amounts in each sample. If there was any variation in the β -actin level seen this was adjusted for in the value obtained for the protein of interest. Band intensity was gauged by scanning the resulting films and assessing the band using ImageJ

software. This produced a numerical value for the resulting band and therefore protein concentration could be compared between samples.

3.13.2: Demethylation of OE33

This adherent cell line was cultured in DMEM with the addition of 10% Foetal bovine serum and 1% penicillin and streptomycin antibiotic cover. Cells were cultured in sterile cell culture flasks in this media until reaching sub-confluence at which time they were split. In general this would occur at intervals of between three and four days.

The first stage was to determine the usefulness of this cell for *TRIM15* expression work. To do this, DNA was extracted from OE33 cells, and bisulphite converted before TRIM15-cg19115272 methylation PCR was performed with subsequent pyrosequencing. This allowed the base line methylation of TRIM15-cg19115272 in this cell line to be determined.

In order to demonstrate the correlation between methylation and expression using the OE33 cell line, it was then necessary to be able to reduce methylation in this cell. To do this, the cells were treated with the DNA methyltransferase inhibitor 5-Aza-2'-deoxycytidine.

Initially, 5AZA dilution experiments were performed to assess the necessary dose required to promote demethylation, yet not be toxic to the cell. To perform this, OE33 cells were cultured in 3cm plates with 3 mls DMEM added. Cells were added at an initial concentration to allow sub confluence by day 5. On day two, having allowed the cells time to adhere, the media was changed and 5AZA was added to achieve concentrations of 0.5, 1, 2 and 5 mM. Each day, one plate was used for cell counting

whereas in the remaining cultures the media was changed and more 5AZA added to achieve the desired concentration.

Initially, 5AZA was dissolved in DMSO as per the manufacturers guidelines as solubility was maximal in this solution. However, it should be noted that DMSO is toxic to the cells therefore exposure of culturing cells to DMSO should be limited. On day one, cells were split into sufficient number of plates to allow sacrifice of one plate per AZA dilution per day in order to count cell growth. These experiments confirmed that 1mM 5AZA was the maximal achievable concentration prior to the cells dying.

To conduct the demethylating experiment, OE33 cells were cultured again in DMEM having been split on day one to achieve sub confluent growth by day 5. Again cells were left to adhere overnight prior to any treatment. On days two to four, cells had their media changed with the addition of 5AZA to achieve a concentration of 1mM 5AZA. Experiments were repeated in parallel with control plates receiving DMSO only and two plates each receiving treatment for a final cell harvest allowing for DNA and RNA extractions.

On day 5, cells were harvested initially for DNA extraction. To perform this, media was removed and the cells washed with PBS. This wash PBS was removed before adding an additional 1ml of fresh PBS. Cells were scraped from the plate and collected in the PBS which was transferred to a sterile microfuge tube. This was then centrifuged at 1300 RPM for 10 minutes. This resulted in the formation of a cellular pellet. Excess PBS was removed before re-suspending the pellet in 200 μ L of PBS. DNA was then extracted according to the Qiagen blood and tissue extraction kit. In short, this involves adding 20 μ L of proteinase K prior to the addition of 200 μ L of lysis

buffer and 200µL of 100% ethanol, with a thorough vortex following each addition. The resulting mixture was then added to a micro centrifuge spin column before being spun for 1 minute at 6000g. The resulting fluid was discarded before washing the DNA with wash AW1 followed by AW2. Finally the DNA was eluted with an elution buffer into a clean microfuge tube with a 200µL volume.

DNA was quantified by the Nanodrop before undergoing bisulphite conversion and then methylation specific PCR for TRIM15-cg19115272 (previously described). Finally, the samples were analysed by pyrosequencing to obtain methylation data for the cell lines pre and post treatment with 5AZA.

Given that 5AZA inhibits the enzyme DNA methyltransferase, it would be expected that as the cells passage, observed methylation would half. Therefore, had the treatment been successful, by the end of 3 treatment days and therefore 4 divisions, with an original methylation of 90%, one would expect to achieve a resulting methylation of only 11.25%. It was expected therefore that if methylation were the driving force behind expression of this gene, it would be possible to observe an increase in expression of RNA following this treatment. Were this methylation change not the driving force behind expression of this gene, no difference would be seen.

However, upon analysis of the results obtained from this experiment, no decrease in methylation was observed at this concentration of 5AZA. Numerous reasons why this may be the case exist including all of the below:

- 5AZA dose incorrect.
- Degraded 5AZA due to freeze/thaw cycles.
- OE33 cells resistant to 5AZA molecule.
- *TRIM15* Gene is refractory to demethylation (where a secondary factor is responsible for expression control).

It was necessary to therefore test and modify each of the conditions to determine which of the above was preventing demethylation of this gene in this cell line.

Cell growth curves were plotted for increasing doses of 5AZA until reaching the maximum dose possible while maintaining cell growth.

Other genes were analysed on the same extracted DNA to determine whether *TRIM15* was refractory to methylation change or whether this treatment had failed across the genome. Therefore, the methylation of *sFRP 2* and *4* were also analysed by bisulphite pyrosequencing in native and post 5AZA treated OE33 cell DNA. These genes were selected due to their availability within the laboratory and the reliability of the assay. This also revealed no difference in methylation post treatment.

5AZA was then dissolved in sterile water, as this was also an acceptable solvent as per the manufacturer's guidelines. OE33 cells were again plated on day one at a concentration to allow confluence by day 5. Cells were treated with 5AZA on day 2, 3 and 4. Despite initial dose/growth relationship curves, cells were treated with successively increasing doses with a doubling dose each time (equating to 0.5, 1, 2 and 4mM). The analysis of DNA methylation at the end day 5 revealed that methylation of the candidate loci had decreased from 90% to 59% with the maximum

5AZA dose given. The dose of 5AZA needed to achieve this reduction in methylation was 4mM.

Each dose of 5AZA given was performed in repeats of 4, such that one plate of cells could be analysed for DNA methylation while the remaining plates, if a significant reduction in methylation had been observed, could be used to RNA extraction to quantify expression.

With methylation reduced, it was necessary to ensure that methylation correlated with expression. In order to do this the remaining aliquots of OE33 cells which had been exposed to the same conditions and treated at the same time as the cells used for DNA extraction were then used for RNA extraction according to the TRIZOL protocol (as described in methods). Once extracted, the RNA product was then used to make cDNA as previously described. cDNA for *TRIM15* was then analysed by Real Time QPCR. If *TRIM15* could be detected as cDNA, it must follow that in the cell there was RNA for this gene and therefore by definition the cell was transcribing this gene. The RNA from the 5AZA treated OE33 cells was compared in this reaction to the RNA of native OE33 cells. This would reveal whether the reduction in methylation was able to bring about the anticipated increased expression of *TRIM15*.

RNA expression was calculated from the rtPCR experiments by the machine derived Ct result – the level at which gene replication becomes exponential in the PCR reaction. The calculation of delta Ct – the difference between normal expression and the treated cell line expression shows if there is any expression change. Finally, the calculation of delta delta Ct, which takes into consideration that an increase of 1 delta Ct correlates to a 100% increase in expression due to the logarithmic nature of the

scale applied, shows the real percentage increase in expression as a consequence of 5AZA induced demethylation.

3.14: Methylation Arrays

The most up-to-date methylation array platform is the Illumina methylation 450 array, in which the methylation status of more than 450,000 individual CpG sites are examined (120). Through a bioinformatic analysis of the comparison groups, this then allows identification of variation in the sample cohorts. This is assumed therefore to represent the population variation, but any such identified differences require validating in much larger cohorts in order to prove this difference is not seen by chance alone.

3.14.1: Sample Preparation

Samples as detailed above were sourced as explained. Samples underwent a standard DNA extraction protocol as previously described. Quantification of DNA was again performed using the NanoDrop. In addition, because of the need for accurate quantification, the Qubit fluorimeter was also used to measure DNA concentration as this was believed to provide a more accurate reading.

3.14.2: Array DNA Quantification

In order to perform Qubit DNA quantification, it was first necessary to calibrate the equipment with standard DNA concentrations (provided in the kit). A 1:200 solution was made by adding the specified volumes of dsDNA BR reagent to dsDNA BR buffer (as per the manufacturers guidelines). 190 μ L of this working volume was added to a 0.5ml PCR tube before then adding 10 μ L of sample DNA. This was

incubated for 2 minutes at room temperature to allow binding of the sample DNA. After calibrating, the sample DNA was quantified following the machine instructions. This resulted in a QF value which relates to $\mu\text{g/ml}$. Therefore, in order to calculate the amount of DNA in the original sample the following calculation was used:

$$\text{Concentration of sample} = \text{QF value} \times (200 \div x)$$

Where x = the volume in μL of sample added to the quantification master mix.

3.14.3: Array DNA Quality Control

As a result of the use the FFPE tissue, Illumina recommend a quality control step to ensure the DNA is of sufficient quality to produce an array result. This was undertaken by a qPCR process using the Illumina FFPE QC kit.

Firstly, after quantification, samples were diluted to a $1\text{ ng}/\mu\text{L}$ concentration. $10\mu\text{L}$ of the QCT (the control) solution was added to $990\mu\text{L}$ of water and placed into a 96 well plate. The PCR was performed in a $20\mu\text{L}$ reaction volume which was composed of $10\mu\text{L}$ of PCR master mix, $2\mu\text{L}$ of QCP and $4\mu\text{L}$ of water. $16\mu\text{L}$ of this master stock was added to either $4\mu\text{L}$ of the QCT solution, $4\mu\text{L}$ of the DNA sample or $4\mu\text{L}$ of water to make the overall $20\mu\text{L}$ working volume. Each sample and control at this stage was run in triplicate as per standard QPCR reaction protocols specified by Illumina.

The plate was sealed and centrifuged to ensure all liquid was at the base of the well and then placed in the RTPCR machine (Applied Biosystems 7500). The cycle for this reaction are shown in table 3.8.

Table 3.8: Thermal conditions for array quality control

| Number | Temperature | Time |
|---------------|--------------------|-------------|
| 1 | 50 | 2 mins |
| 1 | 95 | 10 mins |
| 40 cycles | 95 | 30 secs |
| | 57 | 30secs |
| | 72 | 30secs |

Analysis of the data to determine whether the sample was of satisfactory standard to pass the quality control process involved taking the average Ct for each sample and deducting the QCT Ct (i.e. the controls). If the resulting value (Delta Ct) was below 5, the sample had passed the quality control process and was taken forward to the array. Six samples in each category (BO progressors, BO none progressors, T3N1 tumours and T3N0 tumours) were obtained for the array analysis having passed QC.

3.14.4: Array DNA Bisulphite Conversion

After obtaining the required number of samples which passed the QC stage, the Zymo EZ methylation kit was used as per Illumina guidelines with a variation in cycle conditions for array compatible bisulphite conversion. For normal none array conversions, the sample is incubated at 50°C for 12-16 hours. For the array samples 16 cycles were needed of 95°C for 30 seconds followed by 50°C for 60 minutes. Samples were cleaned-up as per the manufacturer's protocol and eluted into 10µL of elution buffer. Samples were stored at -80°C until ready to perform the following array protocol.

3.14.5: DNA Repair Process

The resulting samples, having been obtained from FFPE tissue, required a DNA repair process as per the Illumina 450 methylation array protocol which was followed exactly (121). This process allows the restoration of degraded FFPE DNA into a form which is able to be amplified by the subsequent methylation protocol.

3.14.6: Illumina Methylation 450 Array

Immediately after completion of the DNA repair protocol, the methylation array protocol was commenced to prevent sample degradation. The full Illumina methylation 450 protocol was followed exactly as per the manufacturers instruction (122).

In brief, the protocol can be broken down into the following discreet stages:

1. DNA amplification:
 - Firstly DNA is denatured and neutralised.
 - The DNA is then amplified in an isothermal reaction in a whole genome format.
2. The formed amplified DNA is fragmented.
3. DNA precipitation:
 - The isopropanol precipitated DNA forms a visible blue pellet following centrifugation.
4. The DNA pellet is re-suspended.
5. The samples are then loaded onto the bead chip and allowed to hybridise overnight.

6. Un-hybridised DNA is then washed away from the bead chip.
7. The hybridised DNA strands were then labelled to allow reading by the Illumina IScan.
8. Samples were then read using the IScan. We are very grateful to the Genomics section and Professor Ian Tomlinson at the Wellcome Trust Centre for Human Genetics for scanning the prepared chips.

3.14.7: Bioinformatic Analysis of Array Data

Bioinformatics analysis of the methylation microarrays was carried out by Dr Andrew Beggs, using the ChAMP package (123) via Bioconductor/R. In brief, red/green intensity values were captured from Illumina iDAT files, background corrected and SWAN normalised to produce M values.

M values were analysed using a logistical regression model using Empirical Bayesian shrinkage of moderated t-statistics to correct for small sample size. In short, this allows the direct comparison of CpG sites between groups of similar samples. Each group in this cohort contained 6 samples, which were in turn compared to 6 samples. This small sample size was controlled for by setting stringent FDR Q-values of <0.05. Identification of variable methylated sites allowed the CpG site markers to be highlighted. Methylation primers were then designed and sourced to cover these sites for the validation exercise. Validation of the in house array was performed in the same way as specified previously using bisulphite pyrosequencing. PCR primers were again designed and sourced from Sigma Aldrich.

4 – Assessment of Published Methylation Arrays in Oesophageal Adenocarcinoma

4.1: Introduction

The initial hypothesis of this thesis was to determine whether methylation biomarkers could determine tissue status in the oesophagus. To do this, it was necessary to test both methylation signatures in OADC and the use of recent technological advances in finding these markers. There is no current robust methylation biomarker that has the ability to discriminate tissue type i.e. normal squamous tissue from BO and OADC. To date, the bulk of the literature in this area examines seemingly random markers which are of interest to the researching group, or which have been shown to have an effect elsewhere in the GIT and are therefore investigated in oesophageal tissues.

A comprehensive, genome wide search of differential promoter methylation using a current state of the art platform (HumanMethylation450 arrays) has not been performed, partly because of cost implications, but also because of difficulties with sample selection. One such study (109), while using a much earlier and therefore more basic array platform, did publish markers that may be of interest. If these published markers, which were not validated, did discriminate between tissue types, this would provide the justification either for performing the array again using the latest technology or sourcing this from elsewhere.

In the above mentioned publication, an Illumina GoldenGate methylation bead array was used to highlight variation in methylation between normal oesophageal tissue

samples compared with BO, HGD and also OADC samples. A weakness of this study was that the BO samples, which appear to have been analysed as a collective group, contained non-dysplastic biopsies in addition to those with dysplasia and notably some having high grade dysplasia. It is unlikely that BO samples without dysplasia will have the same methylation profiles as those with dysplasia but this remains unproven. This group published the top 20 of these markers demonstrating differences in methylation patterns between tissue types.

Table 4.1 shows a list of significantly differentially methylated markers identified in this study. 5 of the suggested potentially interesting genes were selected from this publication (highlighted in red) and methylation primers were designed to cover the CpG sites identified from this array and subsequent analysis. These markers were chosen based on availability of commercially produced methylation PCR primers which were all sourced from Qiagen.

Table 4.1: Previously published methylation sites with the ability to differentiate oesophageal tissue types, as determined by the Illumina Golden Gate array (109).

| Gene ID | CpG Location | Avg. β OADC | Avg. β BO |
|-----------------|--------------|-------------------|-----------------|
| <i>SERPINE1</i> | +189 | 0.52 | 0.30 |
| CDP2 | -278 | 0.38 | 0.12 |
| SDP2 | -150 | 0.41 | 0.22 |
| PCMG | -53 | 0.48 | 0.21 |
| <i>CDH11</i> | -203 | 0.49 | 0.25 |
| <i>CDH11</i> | +102 | 0.50 | 0.26 |
| ALPL | -433 | 0.50 | 0.26 |
| SNPGAVCA | -82 | 0.68 | 0.44 |
| <i>TAL1</i> | +122 | 0.35 | 0.16 |
| RBP1 | +158 | 0.66 | 0.40 |
| <i>MMP2</i> | +21 | 0.21 | 0.11 |
| <i>BMP6</i> | -163 | 0.40 | 0.22 |
| TAAT | -117 | 0.54 | 0.28 |
| <i>GPX3</i> | +178 | 0.42 | 0.22 |
| <i>CD40</i> | +58 | 0.51 | 0.34 |
| <i>IGFBP7</i> | -371 | 0.40 | 0.17 |
| <i>RET</i> | -260 | 0.55 | 0.29 |

Gene ID: Gene name as identified from array

CpG location: Hg 19 genome coordinates of identified probe

Avg. β : Mean methylation

Table demonstrating the top 20 methylation differences between OADC and BO from a golden gate array as previously published.

4.2: Aims:

We aimed to validate methylation biomarkers utilising a previously published primitive array by bisulphite pyrosequencing, to see if these markers would successfully validate in our cohort.

4.3: Methods

4.3.1: Samples

All samples used in this validation experiment were FFPE blocks taken from resection specimens of oesophagectomy cases at University Hospital Birmingham. Specimens were identified using operating logs, which were then correlated with histopathology reports to ensure identified patients had oesophageal adenocarcinoma. Cases were selected retrospectively from this log until sufficient cases were identified (as described below in the power calculation).

Identified cases were supplied to the local HBRC department, who extracted the diagnostic slides from the pathology archives. These slides were then all reviewed by a consultant GI histopathologist, to ensure the diagnosis was accurate, and to then further identify blocks of interest within the resection specimen.

A more detailed block guide was relayed back to HBRC, who then kindly trimmed the blocks to provide the tissue. Tissue was either supplied as scrolls, if all the tissue in a specific block was uniform, or cut into thinner sections and adhered to slides if further dissection was needed. Where this was the case, the diagnostic H+E slide was marked by the histopathologist to indicate the area of interest and facilitate further dissection.

Where tissue on slides required further dissection, this was performed under microscope guidance using a sterile scalpel blade. Resulting tissue removed from the slides was placed in a microfuge tube, and further handled as per the scroll tissues.

DNA was extracted from all samples using a Proteinase K digestion along with a buffer for a minimum of 24 hours at 56°C. Samples were vortex mixed regularly to aid digestion. All samples were reviewed after 24 hours. If there was incomplete digestion (visible tissue fragment seen in the sample) a further 20µL of proteinase K was added and the sample incubated again at 56°C.

Samples were cleaned up as described in the methods section, to yield the final DNA sample which was both quantified and qualified using the Nanodrop.

4.3.2: Power Calculation

In order to show statistical variation in samples, a minimum sample cohort was required. It was assumed, that a significantly differentially methylated gene would have a differential methylation of around 45%, therefore using $p_1=0.45$, $p_2=0.90$, $\alpha=0.05$ and $\beta=0.90$, 24 samples would be required in each group to detect this methylation change with a power of 90% (Calculation performed with Stata 11.2, StataCORP TX).

Therefore, 24 resection patients were included in this Chapter. From each of these, normal proximal mucosa (resection margin) was obtained as well as adenocarcinoma. From 8 patients, on histology review, there was BO in sufficient quantity adjacent to the tumour to allow sampling. These 8 samples of BO were also acquired and processed in the same way into order to see whether there was a

methylation variation between these tissue stages, assuming the well described sequence of normal to BO to cancer was followed.

Prepared tissues, either as scrolls or as dissected slides, were placed in microfuge tubes. DNA was extracted as described.

4.3.3: Bisulphite Pyrosequencing

Extracted DNA from FFPE blocks was subjected to quantification using the NanoDrop and then Bisulphite converted using the Qiagen Epitect conversion kit. Pyrosequencing primers were obtained from Qiagen, using the standard design for each gene selected in the validation panel.

The genes selected were *VCAN*, *ALPL*, *TIAM1*, *POMC* and *RPB1*, due to the availability of methylation primers from Qiagen.

Each primer, as supplied by Qiagen, required optimisation in order to produce DNA product of sufficient amount to show a good quality band on agarose gel electrophoresis. Primers were optimised firstly with a gradient PCR reaction to obtain the temperature at which the primer worked best, which was often not that which was specified by the manufacturer. Each primer also needed a variable amount of DNA in order to obtain a good DNA product band in electrophoresis. After this, each gene was analysed by bisulphite pyrosequencing on all 56 samples. Only when the values obtained passed the internal Quality control process of the pyrosequencing software were the results allowed to be entered into the analysis.

4.3.4: Statistical Analysis

In order to analyse the results fully, three comparisons were performed. These include:

1. Normal squamous mucosa vs OADC tissue
2. Normal squamous mucosa vs Barrett's Oesophagus
3. Barrett's oesophagus vs OADC tissue

To do this analysis, the results from each of the individual CpG sites analysed had their methylation compared using a T-test to assess whether there was any notable difference between values, and whether any such difference met the threshold for statistical significance (having a "p" value <0.05).

4.4: Results

Each gene, and individual CpG site, was analysed in turn comparing each of the above tissues.

Therefore, over the next pages statistical logistical regression analysis for each amplicon and individual CpG within the amplicon are presented. Each CpG site was analysed independently from another to highlight significance. Those "P" values shown in red are significant values. Each respective CpG site was analysed in respective tissue type comparison. The following comparisons were therefore performed:

1. Normal squamous mucosa vs OADC tissue
2. Normal squamous mucosa vs Barrett's Oesophagus
3. Barrett's mucosa vs OADC tissue

Tables 4.2: Pyrosequencing data and statistical analysis for attempted validation of ALPL

Raw data (mean)

| Tissue | Mean methylation CpG1 (%) | Mean methylation CpG2 (%) | Mean methylation CpG3 (%) | Mean methylation CpG4 (%) | Mean methylation CpG5 (%) |
|----------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Normal | 49.83 | 44.47 | 41.87 | 46.36 | 51.13 |
| Barretts | 49.94 | 42.92 | 45.69 | 43.93 | 52.98 |
| Tumour | 48.86 | 39.93 | 40 | 43.79 | 47.90 |

Tumour vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.16 | 0.15 | 1.12 | 0.26 | 0.89-1.51 |
| 2 | 0.87 | 0.12 | -1.04 | 0.30 | 0.67-1.13 |
| 3 | 1.07 | 0.13 | 0.56 | 0.57 | 0.84-1.37 |
| 4 | 1.02 | 0.12 | 0.3 | 0.90 | 0.80-1.28 |
| 5 | 0.90 | 0.11 | -0.88 | 0.38 | 0.70-1.15 |

Barrett's Oesophagus vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.13 | 0.19 | 0.72 | 0.47 | 0.81-1.57 |
| 2 | 1.01 | 0.23 | 0.06 | 0.95 | 0.65-1.28 |
| 3 | 1.32 | 0.27 | 1.33 | 0.18 | 0.58-1.97 |
| 4 | 0.77 | 0.16 | -1.28 | 0.20 | 0.51-1.15 |
| 5 | 0.88 | 0.20 | -0.58 | 0.56 | 0.56-1.37 |

Barrett's Oesophagus vs tumour tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 0.66 | 0.20 | -1.39 | 0.16 | 0.37-1.18 |
| 2 | 1.13 | 0.34 | 0.41 | 0.68 | 0.63-2.05 |
| 3 | 1.33 | 0.35 | 1.11 | 0.27 | 0.80-2.22 |
| 4 | 0.76 | 0.18 | -1.19 | 0.23 | 0.48-1.20 |
| 5 | 1.29 | 0.27 | 1.21 | 0.23 | 0.85-1.95 |

Tables 4.3: Pyrosequencing data and statistical analysis for attempted validation of POMC

Raw data (mean)

| Tissue | Mean methylation CpG1 (%) | Mean methylation CpG2 (%) | Mean methylation CpG3 (%) | Mean methylation CpG4 (%) |
|----------|---------------------------|---------------------------|---------------------------|---------------------------|
| Normal | 45.74 | 60.00 | 29.14 | 37.43 |
| Barretts | 41.21 | 51.98 | 30.40 | 36.95 |
| Tumour | 40.00 | 53.85 | 24.45 | 31.45 |

Tumour vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.00 | 0.04 | 0.07 | 0.95 | 0.92-1.09 |
| 2 | 1.04 | 0.05 | 0.85 | 0.39 | 0.95-1.13 |
| 3 | 1.00 | 0.07 | -0.01 | 0.99 | 0.88-1.14 |
| 4 | 0.90 | 0.07 | -1.33 | 0.19 | 0.76-1.05 |

Barrett's Oesophagus vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.01 | 0.07 | 0.09 | 0.93 | 0.88-1.15 |
| 2 | 0.99 | 0.08 | -0.10 | 0.92 | 0.85-1.16 |
| 3 | 1.10 | 0.11 | 0.92 | 0.36 | 0.90-1.33 |
| 4 | 0.92 | 0.12 | -0.63 | 0.53 | 0.71-1.19 |

Barrett's Oesophagus vs tumour tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 0.99 | 0.06 | -0.10 | 0.92 | 0.89-1.11 |
| 2 | 0.94 | 0.06 | -0.93 | 0.53 | 0.83-1.07 |
| 3 | 1.14 | 0.15 | 1.02 | 0.31 | 0.88-1.47 |
| 4 | 1.10 | 0.21 | 0.51 | 0.61 | 0.76-1.60 |

Tables 4.4: Pyrosequencing data and statistical analysis for attempted validation of *RBP1*

Raw data (mean)

| Tissue | Mean methylation CpG1 (%) | Mean methylation CpG2 (%) | Mean methylation CpG3 (%) | Mean methylation CpG4 (%) | Mean methylation CpG5 (%) |
|----------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Normal | 47.69 | 44.71 | 61.16 | 25.77 | 50.25 |
| Barretts | 55.5 | 50.81 | 73.23 | 29.54 | 61.73 |
| Tumour | 49.25 | 45.12 | 60.69 | 24.95 | 50.50 |

Tumour vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.08 | 0.08 | 1.11 | 0.27 | 0.94-1.25 |
| 2 | 1.00 | 0.07 | -0.01 | 0.99 | 0.87-1.15 |
| 3 | 0.96 | 0.07 | -0.65 | 0.52 | 0.84-1.09 |
| 4 | 0.94 | 0.08 | -0.69 | 0.49 | 0.79-1.12 |
| 5 | 1.02 | 0.06 | 0.26 | 0.79 | 0.90-1.15 |

Barrett's Oesophagus vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 0.92 | 0.11 | -0.65 | 0.52 | 0.72-1.18 |
| 2 | 1.00 | 0.09 | -0.05 | 0.96 | 0.83-1.19 |
| 3 | 1.15 | 0.13 | 1.19 | 0.23 | 0.91-1.44 |
| 4 | 0.80 | 0.14 | -1.26 | 0.21 | 0.57-1.13 |
| 5 | 1.08 | 0.13 | 0.66 | 0.51 | 0.86-1.36 |

Barrett's Oesophagus vs tumour tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 0.04 | 0.14 | 0.28 | 0.78 | 0.80-1.36 |
| 2 | 0.97 | 0.16 | -0.19 | 0.85 | 0.71-1.33 |
| 3 | 1.16 | 0.12 | 1.40 | 0.16 | 0.94-1.43 |
| 4 | 0.86 | 0.19 | -0.69 | 0.49 | 0.56-1.32 |
| 5 | 0.98 | 0.11 | -0.22 | 0.83 | 0.78-1.22 |

Tables 4.5: Pyrosequencing data and statistical analysis for attempted validation of *TIAM1*

Raw data (mean)

| Tissue | Mean methylation CpG1 (%) | Mean methylation CpG2 (%) | Mean methylation CpG3 (%) |
|----------|---------------------------|---------------------------|---------------------------|
| Normal | 4.01 | 6.47 | 7.28 |
| Barretts | 11.22 | 16.08 | 12.20 |
| Tumour | 6.18 | 8.45 | 9.77 |

Tumour vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|------|------|--------------------|
| 1 | 1.08 | 0.07 | 1.23 | 0.22 | 0.95-1.23 |
| 2 | 1.01 | 0.05 | 0.14 | 0.89 | 0.91-1.11 |
| 3 | 1.05 | 0.05 | 0.96 | 0.34 | 0.95-1.15 |

Barrett's Oesophagus vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|------|------|--------------------|
| 1 | 1.02 | 0.08 | 0.24 | 0.81 | 0.87-1.20 |
| 2 | 1.14 | 0.07 | 2.11 | 0.04 | 1.01-1.30 |
| 3 | 1.07 | 0.09 | 0.77 | 0.44 | 0.90-1.26 |

Barrett's Oesophagus vs tumour tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|------|------|--------------------|
| 1 | 1.09 | 0.07 | 1.26 | 0.01 | 0.95-1.25 |
| 2 | 1.14 | 0.08 | 1.96 | 0.05 | 1.00-1.30 |
| 3 | 1.01 | 0.06 | 0.11 | 0.92 | 0.90-1.14 |

Tables 4.6: Pyrosequencing data and statistical analysis for attempted validation of VCAN

Raw data (mean)

| Tissue | Mean methylation CpG1 (%) | Mean methylation CpG2 (%) | Mean methylation CpG3 (%) | Mean methylation CpG4 (%) | Mean methylation CpG5 (%) |
|----------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Normal | 7.85 | 27.09 | 7.65 | 27.75 | 5.92 |
| Barretts | 8.85 | 29.90 | 11.38 | 26.69 | 5.49 |
| Tumour | 8.03 | 30.47 | 11.28 | 35.52 | 8.16 |

Tumour vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.00 | 0.03 | -0.09 | 0.93 | 0.93-1.07 |
| 2 | 0.99 | 0.02 | -0.34 | 0.73 | 0.94-1.04 |
| 3 | 1.06 | 0.06 | 1.22 | 0.22 | 0.96-1.18 |

Barrett's Oesophagus vs normal tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.00 | 0.05 | -0.06 | 0.95 | 0.91-1.10 |
| 2 | 0.99 | 0.03 | -0.26 | 0.79 | 0.93-1.06 |
| 3 | 1.07 | 0.07 | 1.09 | 0.28 | 0.95-1.22 |

Barrett's Oesophagus vs tumour tissue:

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|-----|------------|-----------------|-------|------|--------------------|
| 1 | 1.02 | 0.06 | 0.29 | 0.77 | 0.91-1.13 |
| 2 | 0.99 | 0.03 | -0.17 | 0.86 | 0.93-1.06 |
| 3 | 1.00 | 0.05 | 0.06 | 0.96 | 0.90-1.11 |

CpG: Methylation site analysed

Odds ratio: ratio of the probability between the observed probe being associated with event A vs event B e.g tumour/normal tissue

Standard error: Standard deviation of the mean

Z: z-score: this is a score derived from the logistic regression; it is the regression coefficient divided by the standard error

P>|z|: p-value

95% confidence interval: The likelihood of the results being significant

CpG numbers vary depending on number included in each primer – all data presented. See Appendix 9.1 for CpG site information.

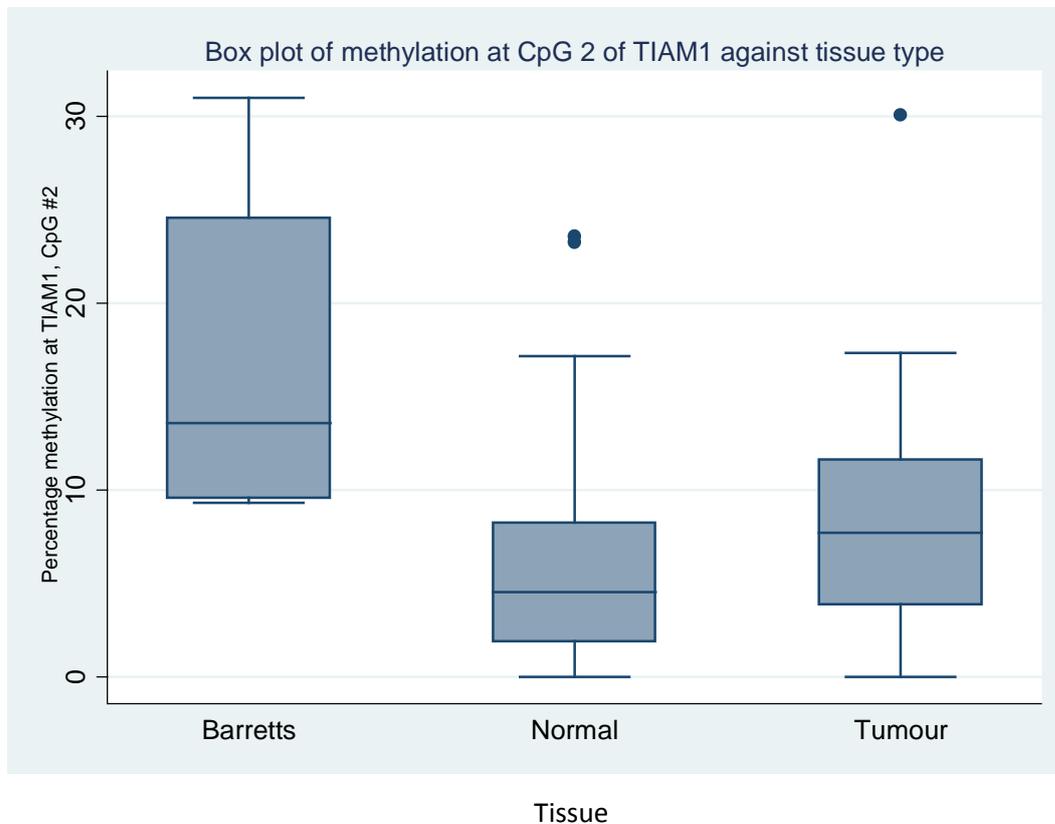
In general, there is little observable difference in methylation between tissues in these data. This is further confirmed in the logistical regression analysis.

The only clear difference seen in this data is the hypermethylation in BO tissue in *TIAM1*. This observation, seen only in CpG site 2, reveals a methylation of 6.47% in normal squamous mucosa. However, methylation in BO mucosa increases to 16.08%. Regression of this difference reveals this to be a statistically different variation, with $p=0.035$ (rounded to 0.04 in the tables shown).

None of the other results obtained show any statistically different results.

This single significant marker can be shown in graphical form to demonstrate the difference in methylation between tissues:

Figure 4.1: Variation in methylation between tissue types of *TIAM1*.



Boxes span the 1st and 3rd quartiles

Band inside the box is the 2nd quartile (the median)

The lines are the “lower adjacent value” which is the first quartile - 1.5 x intra-quartile range and the “upper adjacent value” which is the third quartile + 1.5 x Inter-quartile range

Values above the lines are outlying data points

Figure 4.1 shows the difference in methylation patterns between Normal Squamous epithelium in the oesophagus compared with Barrett’s mucosa and OADC. The difference in methylation between Barrett’s and normal squamous mucosa showing a statistically significant difference ($p=0.04$).

The difference in methylation between tissue types in this analysis is clearly shown in figure 4.1. It is also interesting to note that this same CpG site approaches significance at discriminating between Barrett's and tumour tissue also ($p=0.05$). Interestingly though, the methylation seen in these data show that methylation is almost the same between normal and tumour tissue. It is unclear why BO would have such a hypermethylation of this site.

4.5: Conclusions

Despite the large number of genes studied here, only one marker reached statistical significance in validation despite analysing methylation on a sample size great enough to reveal any difference should this be present. Methylation at the second CpG of the gene *TIAM1* had the ability to differentiate normal squamous epithelium from Barrett's Oesophagus tissue. This same marker comes close to reaching significance ($p=0.05$) at differentiating Barrett's Oesophagus tissue from tumour samples, but fails to meet the agreed statistical threshold of being below $P=0.05$. However, it should be noted that the number of BO samples in this analysis was significantly lower than the 24 samples needed to meet the demand of the power calculation. As a consequence, this observation may be a result of this lower "n". Further analysis of methylation of this gene in a further 16 samples of BO would be required to validate this finding beyond any doubt.

TIAM1, also known as T-cell lymphoma invasion and metastasis protein 1, which in turn activate Rho-like GTPases (124). These proteins are known to play a significant role in cell movement and adhesion (125, 126). The significance of this gene has already been investigated in OADC, specifically invasion and lymph node metastasis.

Expression of *TIAM1* was shown to be higher in patients with more invasive disease as well as those with metastasis (127). Therefore, the demonstrated difference in methylation between normal squamous mucosa and OADC tissue could well be one of the controlling influences over this gene and is worthy of further investigation.

However, this chapter clearly shows that BO samples have a very different methylation pattern from both tumour and normal tissue in this specific CpG site. This statistically significant result shown in the above box plot reveals no overlap with normal squamous mucosa and comes near no overlap with tumour samples. This marker, in addition to the remaining tested markers, does not behave as might be expected. It would be anticipated that methylation would increase in a stepwise progression with normal mucosa having the lowest level of methylation, increasing in BO samples, peaking at high grade dysplasia and reducing slightly from this in tumour samples. This possibly represents the inflammatory status of the tissue examined. Therefore, the pattern demonstrated here, while unexpected, does clearly highlight this as a potentially important marker.

Therefore, returning to the original array publication from which these markers were chosen, there is the suggestion of a potentially important and unique marker. However the majority of markers and CpG sites identified by the performed array failed to validate in this analysis. There are many potential reasons why this may be the case.

While powered sufficiently in terms of the number of normal specimens and tumour samples, there was only a small number of BO patients included. This means that while interesting, no solid conclusions can be drawn from the methylation levels for

BO samples presented in this chapter. While the number of included samples of normal mucosa and tumour met the numbers required, it is still possible that were more specimens included, other markers would have been able to discriminate between tissue types.

It has to be recognised that the array used in the original publication has now been superseded largely by the Illumina 450 methylation array system. However, bisulphite pyrosequencing remains the main stay of methylation array validation and the strict criteria applied to the use of results obtained and their validity provides no element of error.

While unlikely, it is possible that the samples used in this experiment were again not representative of the cancer population. As a result of the chronological nature of the included samples, all tumour stages were included. This translates into T1-T4, N0 and N1 samples being used and it might be possible that there would be some variation in the methylation levels seen between tumour stage (there is no published data to support this). It is unclear which samples were used as “tumour” in the original array publication and therefore no adjustment in the cohort used in the validation set was possible. However, the cohort used in this study was representative of the wide spread of tumours observed in clinical practice, in which any validated marker would have to be robust should it ever be considered for routine clinical use.

Finally, the interpretation of array data published by the original authors may have been inaccurate. Due to the large volume of data produced by an array, and the inherent complexity of that data, interpretation can be a challenge and is liable to

error. No further comment can be made about this as the original data from the arrays was unavailable; therefore the published genes were taken at face value for the validation experiments presented here.

4.6: Discussion

This chapter has shown promise in the search for a methylation biomarker with the ability to discriminate between tumour and normal tissue in the oesophagus, and has revealed one marker from an array which has validated in a powered experiment. It is assumed, in moving forward with this work, that with more modern techniques now available, it may be possible to find further, more robust novel biomarkers.

These initial validation experiments are helpful in that they, along with the reported literature, show that methylation markers can be significantly variable in tissue states relevant to oesophageal adenocarcinoma. However, the markers reported in the Kaz et al (109) paper have largely failed to validate. While this was the first published study using an array platform to assess markers of different disease states (109), validation attempts as performed here have undoubtedly highlighted the potential pitfalls associated with this approach. Despite now being superseded, the fact that only one of the markers highlighted in the array has validated reveals the weaknesses and pitfalls in array experiments. .

While markers from an array may reach the required level of variation between samples to meet statistical significance, this may be a chance finding due to the inevitable small number of tissues analysed on an array. Arguably therefore, these results in isolation are meaningless unless they validate in a powered experiment performed after the initial array experiment on a separate, independent sample

cohort as was performed here. This approach does show that array based discovery has the ability to generate markers of interest, and may identify novel markers other than those which are analysed due to their particular interest to the researching team or because of their highlighted use in other disease processes.

This chapter has therefore provided the justification for using a modern array based approach to investigate whether any more significant methylation markers can be identified in the context of tissue discrimination in oesophageal adenocarcinoma.

In addition, the work presented above demonstrates that an array based approach needs to have the proper experimental design and the correct techniques to analyse these datasets must be chosen.

To continue this work into the next chapter, a more modern array based approach will be utilised to answer the questions posed previously. This chapter has clearly provided the justification for this expensive but promising modern technique.

5: Data Mining of the TCGA Database to Search for Differentially Methylated Biomarkers of Oesophageal Adenocarcinoma, Independent Validation and Functional Assessment

5.1: Introduction:

Methylation arrays have the ability to compare tissue types to reveal signatures specific to that particular tissue. These signatures vary as a consequence of, or are indeed a driving factor of disease. As demonstrated previously, an array based approach is a valid methodology for methylation biomarker discovery but it must be accepted that not all markers identified in an array will validate in larger powered cohorts. Clearly therefore an array must be followed by validation before results can be meaningful.

The array on which Chapter 4 was based (Illumina GoldenGate) used a now outdated array technology, which has been superseded by the Illumina 450 methylation array system.

The Illumina methylation 450 array platform now represents the most up to date technology for biomarker discovery of this nature. Due the work in Chapter 4 only being able to reveal a single validated marker, it was thought necessary to pose the same question using a more modern technique to ascertain whether more robust biomarkers could be found.

Biomarkers have potential use in two regards. Firstly, they have the ability to assist and streamline the diagnostic process. If a methylation profile can differentiate between a tumour specimen and normal mucosal biopsy, this could have far reaching implications for current diagnostics in use in the oesophagus. Not least because of the recent development of strategies to avoid the need for an OGD examination (for example the cytosponge described earlier). As previously discussed, the cytosponge samples cells as it is withdrawn from the oesophagus. Clearly much of the cellular architecture in the resulting sample is lost in comparison to a formal biopsy. However, analysis of the methylation profile of the cells sampled may well increase the sensitivity of this test, and patients would no doubt prefer this to having to undergo an OGD.

Although not essential to biomarker discovery, once validated it is important to explore the functional significance, if any, of such a marker. The link between methylation and expression is clearly described and theoretically therefore it could be anticipated that the variation in methylation seen reflects expression variation within the affected tumour cells (assuming normal mucosal cells have “normal” methylation profiles). The process of exploring functional significance of varying methylation levels opens a huge area of potential work. However, the first stage in any such process is proving the link between methylation and expression of any given gene. It must be accepted that genes will have numerous control mechanisms of which methylation is only one. Therefore, the expected link between methylation and expression is not guaranteed should the overriding cellular control be different.

During the course of this work, the Tumour Cancer Genome Atlas (TCGA) published the results of the genetic and methylation 450 array data for OADC. The TCGA was

established in 2005. Its remit was to categorise the genomic mutations seen in cancer using modern sequencing techniques and subsequent bioinformatic analysis. Methylation data from the TCGA database was acquired using Illumina methylation 450 arrays and so represented a robust modern array technique for OADC biomarker discovery. The methylation array data for OADC was extracted from the database, along with the data for normal oesophageal squamous mucosa and analysed in the hope of discovering potentially informative markers.

5.2: Aims:

To determine whether a modern methylation array based approach can elucidate novel methylation biomarkers that can differentiate tumour from normal squamous mucosa in the oesophagus.

5.3: Materials and Methods

5.3.1: TCGA Methylation Array Datasets:

Level 2 (IDAT files) datasets were downloaded from the oesophageal adenocarcinoma section in the TCGA Data portal for all available (at the time of the study) tumours (n=11) and normal oesophageal tissues (n=10).

Adjusted methylation intensity (red/green) values were analysed using Limma and Bioconductor. Methylation values were background corrected and SWAN normalised and a logistic regression model using empirical Bayesian shrinkage of moderated T-statistics was performed. Differences in methylation between the tissues were validated on the cohort of samples used in Chapter 4 which had already been extracted and bisulphite converted.

5.3.2: Bisulphite Pyrosequencing

Sigma bisulphite pyrosequencing primers were designed and sourced for the desired sequences revealed by the array analysis and run on these samples using an optimised PCR technique previously described.

Once validated on internal samples (24 tumour/normal pairs with 8 BO samples), further samples were sourced from the histopathology department at neighbouring University Hospital Birmingham as an internal validation cohort (9 tumour/normal pairs). As before, each tissue block used in the study underwent dedicated Upper GI histopathological review. Where possible, blocks were chosen to represent the entire tissue type and therefore these sections were cut as scrolls. Where sections required further dissection, tissue was cut onto slides and an H+E slide produced and marked as the tissue guide to allow dissection. Extraction of DNA proceeded as per protocol described previously.

After validation on the second in-house sample set, an external sample cohort was sourced from neighbouring Birmingham Heartlands Hospital covered under the pre-existing ethical approval and tissue transfer arrangement, and facilitated by HBRC (8 tumour/normal pairs). Samples selected for inclusion were identified from patients recruited to the OE03 trial, in which Birmingham Heartlands Hospital participated but whose records were held locally to enable identification of relevant patients. Despite being obtained from a separate unit, all samples had their histology slides reviewed by the same histopathologist to ensure sample control.

Again samples were acquired with the assistance of the HBRC, to which a list of required blocks was provided and these were then sourced under a tissue transfer

agreement which was already in place. Once the blocks were obtained, along with the original diagnostic H+E slides, an independent histopathology review was undertaken to again ensure the samples were of OADC. Specific blocks of interest were identified and this data relayed to HBRC for tissue sectioning. DNA extraction followed the uniform technique.

All samples underwent DNA quantification before proceeding to bisulphite conversion. PCR methodology was standardised once optimisation for the individual primers had been performed.

Pyrosequencing was performed and only results of good quality (which passed the quality control process of the pyrosequencing software) were taken forward for analysis. Samples which failed the PyroMark QC standard were repeated until all results met the required threshold.

5.4: Results:

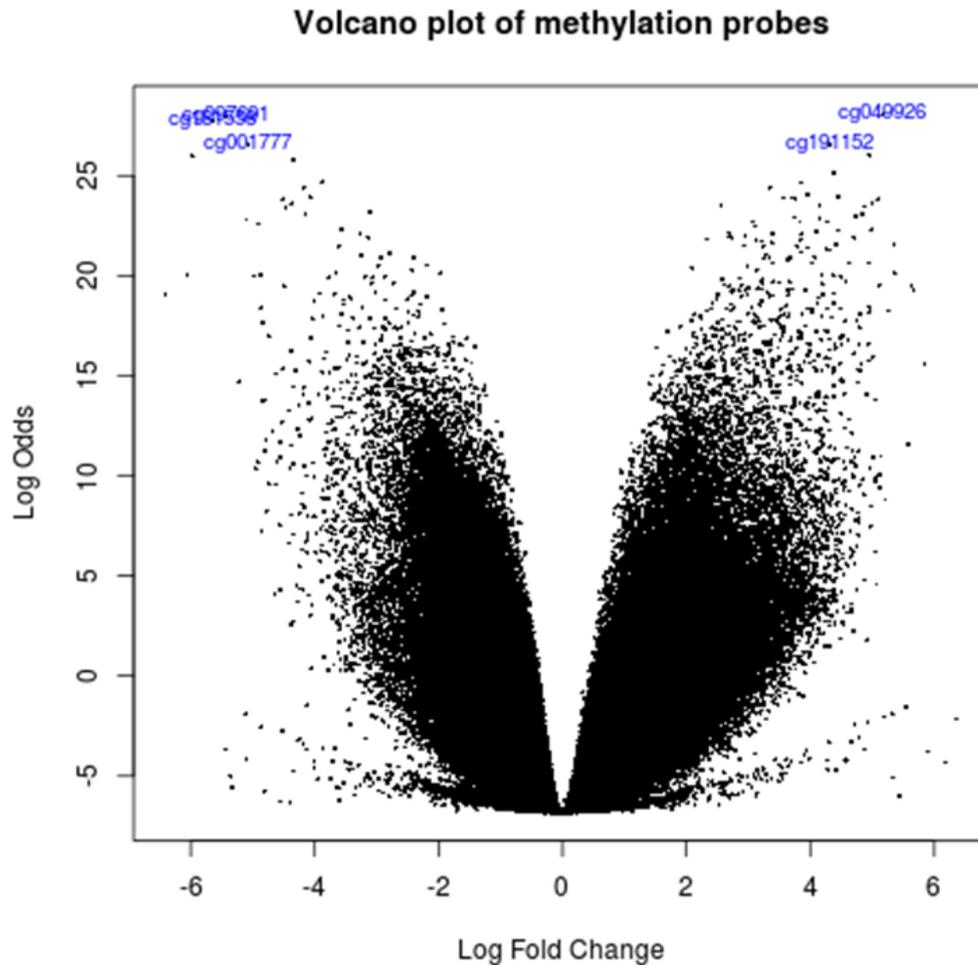
5.4.1: TCGA Analysis:

The TCGA database included 11 OADC samples for analysis. These were identified in the atlas and the data exported. In addition, all of the data for the comparative normal squamous tissue were also exported (n=10). These results were then analysed according to a Bayesian Regression model, to determine methylation variation between the two tissue types and the significance of each highlighted difference.

Figure 5.1 shows an overall representation of the differences in methylation seen in all of the analysed CpG sites (485,000), and also the significance of the variation.

The data shows the change in methylation of a tumour in comparison to normal squamous mucosa. Therefore, data points to the right of 0 show a hypermethylation whereas those to the left show hypomethylation. The position of the data point on the “y” axis reveals the significance of the result, with those highlighted in blue having the maximal variation between tissue type and the greatest statistical significance. These blue sites shown in the graph were therefore selected for validation experiments.

Figure 5.1: TCGA methylation 450 array data comparing methylation between normal oesophageal squamous mucosa and adenocarcinoma



This box plot shows a graphical representation of the variation in methylation between normal squamous epithelium and OADC. Each data point represents one of the 485000 CpG sites examined in the Illumina methylation 450 array. These data assume squamous mucosa have a “normal” methylation, therefore any point to the right of 0 has a hypermethylation change in contrast to data points to the left of 0 having a hypomethylation change. The log odds (y-axis) demonstrates the significance of the methylation difference observed. Data points shown in blue therefore have the maximal differential methylation at the maximal significance.

The above volcano plot is summarised in table 5.1 below. The following table shows the top 20 differentially methylated CpG markers from the array analysis. Each of the

markers shown in the table demonstrates significant ($p < 0.05$) methylation variation between tumour and normal tissue. However, while many sites are revealed as significant differentially methylated in this analysis, it must be acknowledged that these values are Bonferroni adjusted and therefore it is more important to recognise island methylation as a whole rather than isolated individual CpG sites.

Table 5.1: methylation 450 array data revealing top 20 differentially methylated sites between normal mucosa and OADC

| <u>Probe ID</u> | <u>t</u> | <u>P.Value</u> | <u>adj.P.Val</u> | <u>Gene Name (as per PyroMark software)</u> |
|-----------------|-------------|----------------|------------------|---|
| cg04992638 | 24.10550781 | 3.03E-17 | 7.26E-12 | C17orf46;LOC100133991;LOC100133991 |
| cg09769113 | 23.94763281 | 3.48E-17 | 7.26E-12 | TRIM15 |
| cg18155853 | -23.6600785 | 4.49E-17 | 7.26E-12 | LRR1Q4 |
| cg19115272 | 22.04882122 | 1.98E-16 | 1.93E-11 | TRIM15 |
| cg00177787 | 22.04523172 | 1.98E-16 | 1.93E-11 | 0 |
| cg24542751 | 21.37753194 | 3.78E-16 | 2.77E-11 | LOC100133991;LOC100133991;C17orf46 |
| cg05664039 | 21.32108857 | 3.99E-16 | 2.77E-11 | TRIM15 |
| cg02478172 | 21.09300259 | 5.00E-16 | 3.04E-11 | PTPRN2;PTPRN2;PTPRN2 |
| cg23623622 | 20.31337887 | 1.10E-15 | 5.92E-11 | TRIM15 |
| cg11692307 | 19.80733943 | 1.85E-15 | 8.45E-11 | PLCH2 |
| cg11133524 | 19.7769069 | 1.91E-15 | 8.45E-11 | 0 |
| cg02827132 | 19.50514463 | 2.55E-15 | 9.58E-11 | 0 |
| cg08247376 | 19.49946975 | 2.57E-15 | 9.58E-11 | C17orf46;LOC100133991;LOC100133991 |
| cg16601494 | 19.14889075 | 3.74E-15 | 1.19E-10 | C1orf70;C1orf70 |
| cg00330492 | 18.98857777 | 4.45E-15 | 1.19E-10 | 0 |
| cg05649391 | 18.98274573 | 4.47E-15 | 1.19E-10 | MYBPC3 |
| cg26742995 | 18.97517146 | 4.51E-15 | 1.19E-10 | LOC100133991;LOC100133991;C17orf46 |
| cg07922007 | 18.97265613 | 4.52E-15 | 1.19E-10 | 0 |
| cg17892556 | 18.88673894 | 4.97E-15 | 1.19E-10 | ZNF625;ZNF625 |
| cg13768269 | 18.86137023 | 5.11E-15 | 1.19E-10 | 0 |

Probe ID – the Illumina cg probe ID from the Illumina manifest

T = the t value – the size of the difference relative to the variation in the sample

p-value = the raw p-value, not corrected for multiple testing

Adjusted p-value = the p-value corrected for multiple testing

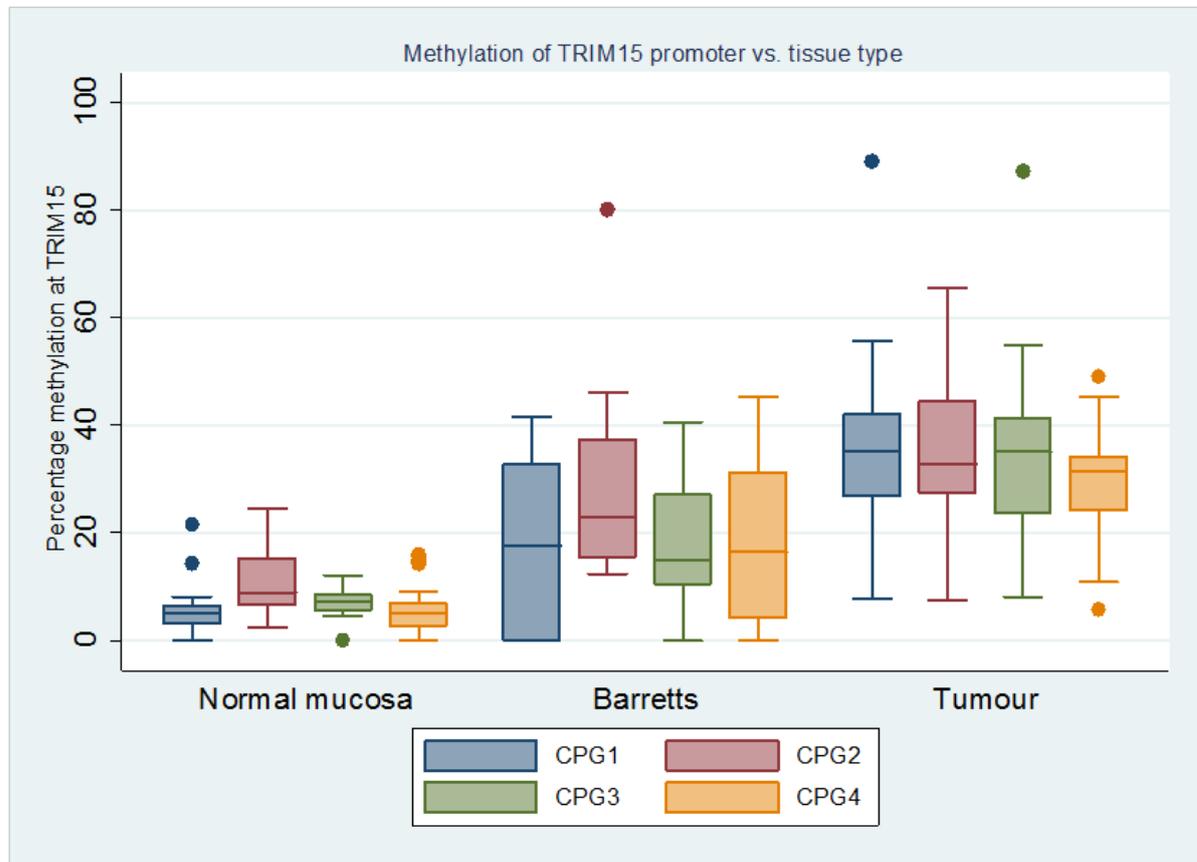
Of note from table 5.1 is that, despite not being the top hit, Tripartite Repeat Motif 15 (*TRIM15*) is seen 4 times in the table of top 20 hit potential markers (highlighted in yellow). As a result of this, *TRIM15* was chosen as the marker with the greatest potential ability to differentiate tissue type, and was therefore taken forward for validation attempts as this seemed a potentially informative marker.

5.4.2: TCGA Array Validation

TRIM15 methylation primer was sourced from Qiagen and using this, bisulphite pyrosequencing was performed on the original 24 tumour/normal pairs. In addition, 8 samples of BO were also analysed in order to determine whether this marker might be able to discriminate out this intermediate tissue type. It was unknown how methylation in BO would be represented by this gene as there was no array data on BO samples. However, the published literature shows a progressive methylation increase from normal tissue through BO and peaking at high grade dysplasia (91). Commonly, methylation is then reduced in the cancer samples (91). This was therefore anticipated here.

The amplicon generated from this primer contained 6 CpG sites, within which were the 4 sites on interest revealed from the array data. The analysis, for 4 significant sites on this amplicon are show in figure 5.2.

Figure 5.2: Pyrosequencing Validation Data of *TRIM15*



The boxes span the 1st and 3rd quartiles

Band inside the box is the 2nd quartile (the median)

The lines are the “lower adjacent value” which is the first quartile - 1.5 x intra-quartile range and the “upper adjacent value” which is the third quartile + 1.5 x Inter-quartile range

Values, represented by dots are outside the expected results range and are therefore shown as outliers.

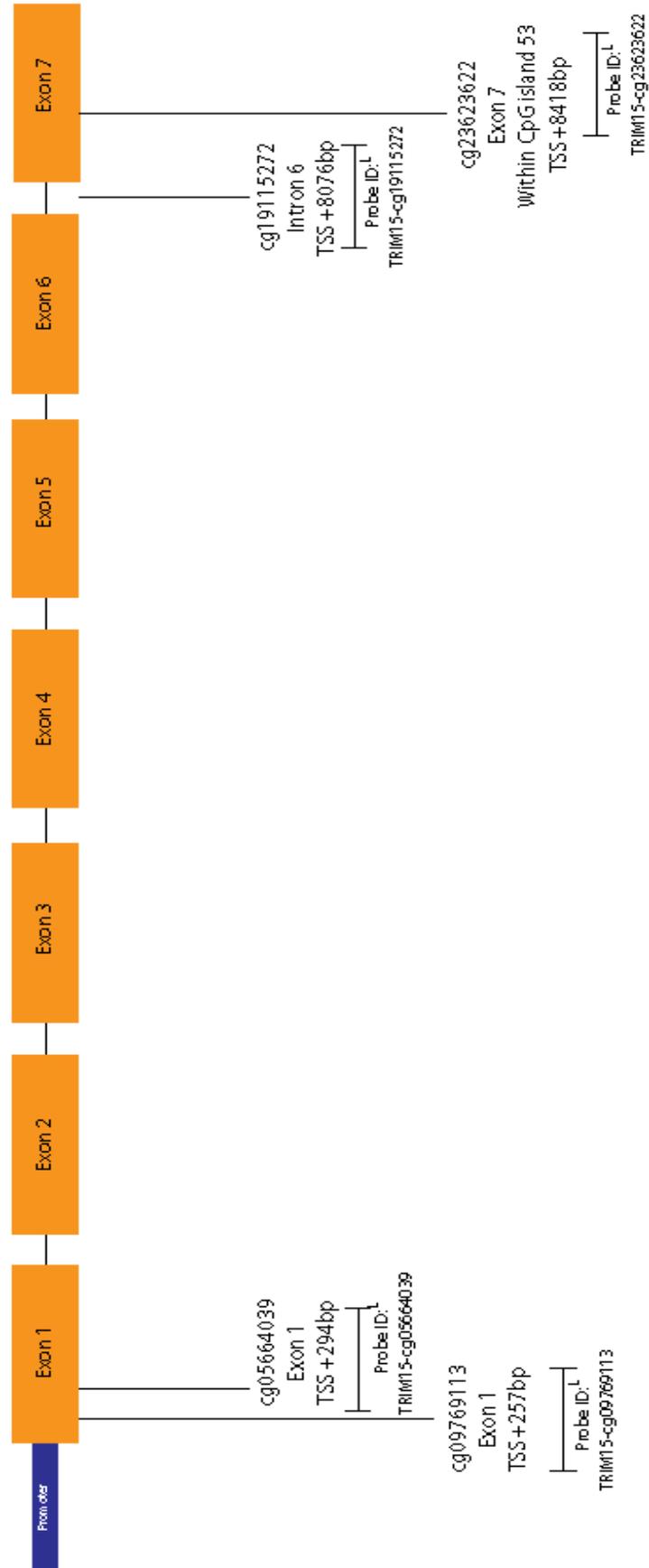
This graph shows the variation in methylation between normal squamous mucosa, Barrett’s mucosa and OADC. The *TRIM15* primer contained 4 CpG sites in the *TRIM15* gene, which are represented above.

It can be clearly seen from figure 5.2 that all 4 sites of *TRIM15* analysed in this experiment were able to differentiate between tissue types, however only the difference between normal mucosa and tumour was statistically significant ($p < 0.05$).

Therefore, to confirm this finding, further samples were acquired (as described) in order to validate this result further using bisulphite pyrosequencing. However, on second round validation samples with the previously used Qiagen *TRIM15* primer, the markers failed to validate.

Re-analysis of the original TCGA database results in fact showed that the primer being used did not exactly cover the 4 CpG sites of interest that were originally identified by the array as discriminatory between tumour and normal oesophageal tissues. As a consequence, 4 new primers were designed, each of which targeted a specific site from the array data. These primers were sourced from Sigma. The position of the respective 4 CpG sites within the *TRIM15* gene are shown in figure 5.3.

Figure 5.3: *TRIM15* gene and location of the array identified probes



Diagrammatic representation of the *TRIM15* gene demonstrating each of the 4 array located CpG sites within this gene which have the ability to discriminate tissue status within the oesophagus.

Initial validation with the new primers and therefore specific target CpG sites was performed on 8 matched samples. This small sample cohort was used to ensure primers were optimised, functional, significant and therefore worth pursuing further. Each of the 4 primer pairs were optimised as previously described to yield maximal DNA amplification, confirmed by the resulting strength of the product band on agarose gel electrophoresis and the quality of the results obtained on the downstream pyrosequencing QC.

Table 5.2 demonstrates the statistical significance of each of the *TRIM15* CpG sites. Of note is that two of the sites failed to reach significance in this small validation experiment, and were therefore excluded from further analysis. However, the first and second sites shown here did have significant “P” values and were therefore taken forward for validation on all three sample cohorts.

Table 5.2: Analysis of TRIM15 methylation results for each of the 4 primers on 8 tumour/normal tissue pairs.

| TRIM15 gene site | p-value |
|-------------------|---------|
| TRIM15-cg09769113 | 0.001 |
| TRIM15-cg19115272 | 0.0027 |
| TRIM15-cg05664039 | 0.280 |
| TRIM15-cg23623622 | 0.0674 |

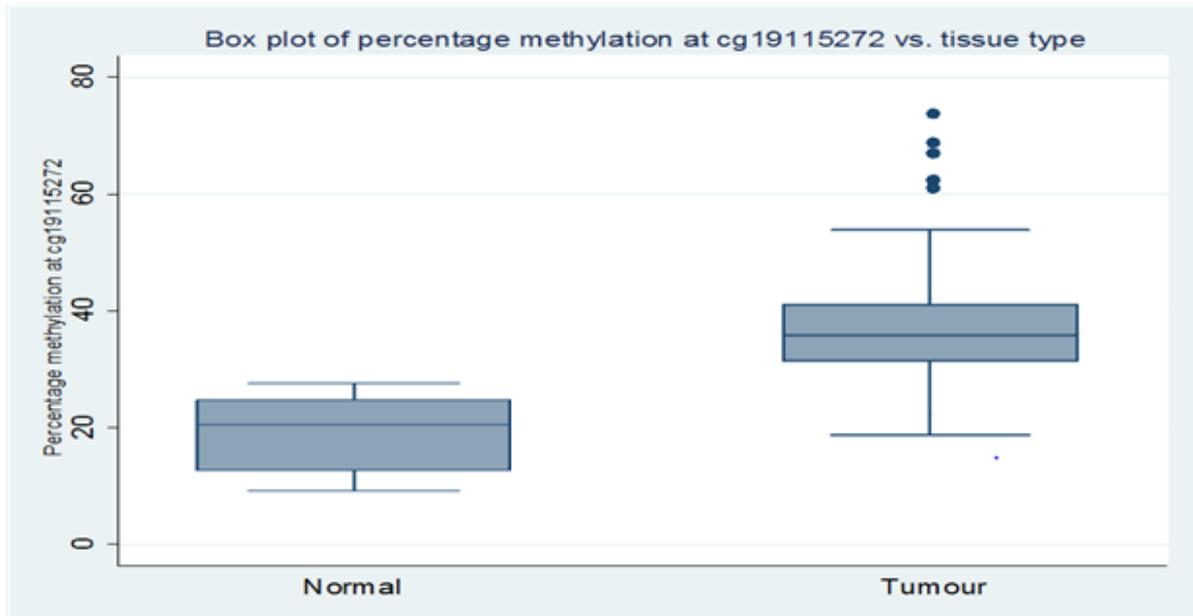
The table above shows the statistical difference in methylation of the *TRIM15* CpG sites identified from the Illumina methylation 450 array. All internal samples, from both the first and secondary cohort were analysed with these two significant *TRIM15* primers (TRIM15-cg09769113 and TRIM15-cg19115272). When they successfully validated and were able to discriminate between tissues, an external sample cohort of the same tissue types (tumour and normal samples) was obtained as described and also analysed. In total, data was obtained for 41 tumour and normal squamous samples.

On completion and analysis of the two CpG's (TRIM15-cg09769113 and TRIM15-cg19115272), only one amplicon validated with statistical significance. While

approaching significance, CpG TRIM15-cg09769113 did not quite meet the $P < 0.05$ threshold and therefore was excluded from further discussion.

To continue with CpG TRIM15-cg19115272 it can be clearly seen (figure 5.3) that there is a stark variation in methylation levels between normal tissue and tumour in this analysis. It can be observed that tumour tissue is hypermethylated in comparison to normal squamous mucosa. While there are outlying points in this data set (those outside of the 95% confidence interval) all of these are seen in the tumour samples and are significantly hypermethylated in comparison to normal tissue. Were these results to have been included in the analysis even greater significance would have been reported.

Figure 5.4: Methylation difference between normal squamous mucosa and adenocarcinoma of TRIM15 cg19115272.



The boxes span the 1st and 3rd quartiles

Band inside the box is the 2nd quartile (the median)

The lines are the “lower adjacent value” which is the first quartile - 1.5 x intra-quartile range and the “upper adjacent value” which is the third quartile + 1.5 x Inter-quartile range

Values, represented by dots are outside the expected results range and are therefore shown as outliers.

Box plot showing the variation in methylation between normal squamous mucosa and OADC in TRIM15 cg19115272. There is statistically different variation in these data ($p < 0.0001$ (5×10^{-8}))

Statistical analysis of the results in the graph (figure 5.3) above, demonstrating their significance are shown in the table 5.3 below.

Table 5.3: Difference in methylation between normal squamous mucosa and adenocarcinoma of TRIM15 cg19115272

| Tissue | Average methylation | 95% CI | p-value |
|--------|---------------------|------------|---------------------------------|
| Normal | 20.6% | 17.3-23.9% | < 0.0001 (5×10^{-8}) |
| Tumour | 36.6% | 32.3-40.9% | |

Therefore from these data, a reliable methylation biomarker, with the ability to differentiate normal from tumour tissue has been presented. Taking this result forward and applying it to a clinical situation, the aim was to develop a methylation marker that might have potential use in the management of patients. Using the above dataset, and applying a logistic regression model, with a pre-determined methylation threshold of 34% and modelling for the known population prevalence of OADC, the following can be determined about CpG TRIM15-cg19115272 methylation in routine clinical practice (table 5.4).

Table 5.4: Logistical regression analysis of *TRIM15* methylation data

| Test | % |
|---------------------------|------|
| Sensitivity | 50 |
| Specificity | 95.3 |
| Positive Predictive Value | 38.9 |
| Negative Predictive Value | 97 |

The overall conclusion from the above results is that, in a population adjusted dataset and setting the methylation threshold at 34%, any tissue sample with a methylation level below this has a 97% chance of not being a tumour. The 34% methylation threshold was determined by modelling using a ROC curve, which allows the threshold to be chosen to maximise sensitivity and specificity. Therefore, this could be used as a reliable method for excluding cancer in an oesophageal tissue sample.

5.5: Oesophageal Methylation Field Effect Change

5.5.1: Introduction:

To date, all of the tissues used in this thesis have been from cancer resection patients. Although the histology review process has meant that there can be certainty about tissue type analysed, with normal tissue having no evidence of tumour cells, it does mean that if there is any potential methylation field effect change

this would influence the results. This is particularly significant if *TRIM15* methylation were to have a future role to play in OADC diagnosis. Because hypermethylation is seen in tumour tissue in comparison to normal squamous epithelium, if there were to be a field effect in the oesophagus then the methylation differential between tissue types would be narrower than would be seen in truly normal tissues. Field effect changes have been observed elsewhere in the GIT, particularly in the colon, and these are the basis under which the ENDCAP study (128) is now investigating methylation field changes to predict ulcerative colitis associated cancer.

The use of oesophageal tissues, as matched samples from a cancer patient, features routinely in methylation research. No previous publications have demonstrated the presence or absence of a field effect in the oesophagus. It was deemed important to demonstrate whether this phenomenon did or did not exist. If it were to be proven in this work, much of the previous literature in this field, and particularly where normal mucosa or BO samples are taken from a cancer patient, would be brought into doubt.

5.5.2: Aims:

To determine whether cancer of the oesophagus creates a field effect methylation change in adjacent normal squamous mucosa.

5.5.3: Materials and Methods:

To test whether any such field change exists in the oesophagus it was necessary to acquire “normal” oesophageal tissue. To do this, endoscopy clinics were targeted to take oesophageal biopsies from patients. Patients who underwent endoscopy demonstrating normal upper GI mucosa were selected for biopsy. Sequential patients demonstrating a macroscopically normal oesophageal appearance

underwent a lower oesophageal biopsy. Full patient consent was obtained prior to OGD examination for inclusion in this research from patients undergoing OGD by an upper GI consultant at UHB. This sample collection was facilitated by HBRC and had full ethical approval.

At the time of research biopsy, a further biopsy was also sent to pathology for confirmation that the tissue used represented normal oesophageal mucosa without any abnormality. This was particularly important as by definition these patients had upper GI symptoms (usually reflux symptoms). Only where the formal histology report revealed normal mucosa were the samples included in this experiment. 9 such patients and therefore biopsy samples were included in this study.

Fresh tissue biopsies were snap frozen in liquid nitrogen and stored at -80°C until all 9 samples had been acquired. When ready for experimentation, the samples were fully thawed and then digested for 24 hours in proteinase K along with digestion buffer. Resulting DNA was acquired according to the previously described protocols. Bisulphite pyrosequencing for TRIM15-cg19115272 was then performed on these samples in order to obtain methylation values for *TRIM15* in normal squamous epithelium from non-cancer patients.

5.5.4: Results:

As presented previously, the mean methylation of TRIM15-cg19115272 in proximal resection margin was 20.6%. However, in the newly analysed “normal” oesophageal samples, mean methylation observed was only 12.4% ($p < 0.05$). This work suggests that there may be a field effect in the oesophagus of a cancer patient, and this is worthy of further investigation extending to other markers of interest.

5.5.5: Conclusions:

This result has far reaching implications for much of the previously published data using proximal resection margins to represent truly normal tissue, including data presented in this thesis. Although using matched tissues for methylation analysis removes many confounding variables in sample acquisition, it clearly adds the confounding factor that it will potentially raise the methylation profile of normal tissue at the proximal resection margin.

It is unclear how this is reflected in other disease states in the oesophagus. For example, it might be that the presence of BO imposes a field effect change in the remaining oesophagus, or in fact that a pre-existing methylation variation from normal pre-exposes the patient to a risk of BO. It might also be that patients who subsequently progress to OADC have a differing methylation signature from those that remain in a static state over time. Unfortunately, no such answers exist to these questions, but this work clearly demonstrates how little is currently understood about methylation in the oesophagus and therefore future work that needs to be done.

5.6: Crossover of *TRIM15* Methylation with the Lower Gastrointestinal Tract

5.6.1: Introduction

As is clear from the introduction to this thesis, there is much overlap between methylation signatures seen in disease states in the upper and lower GI tract. It was therefore hypothesised that this would also be the case with the CpG of interest on the *TRIM15* gene (TRIM15-cg19115272).

This experiment is important to clarify how specific TRIM15-cg19115272 is as a biomarker of OADC. Clearly, if this marker has an increased methylation in other GI associated pathologies then its usefulness as a specific marker of OADC is reduced.

5.6.2: Aims

To demonstrate the anticipated overlap of TRIM15-cg19115272 methylation between the upper and lower GI tract.

5.6.3: Materials and Methods

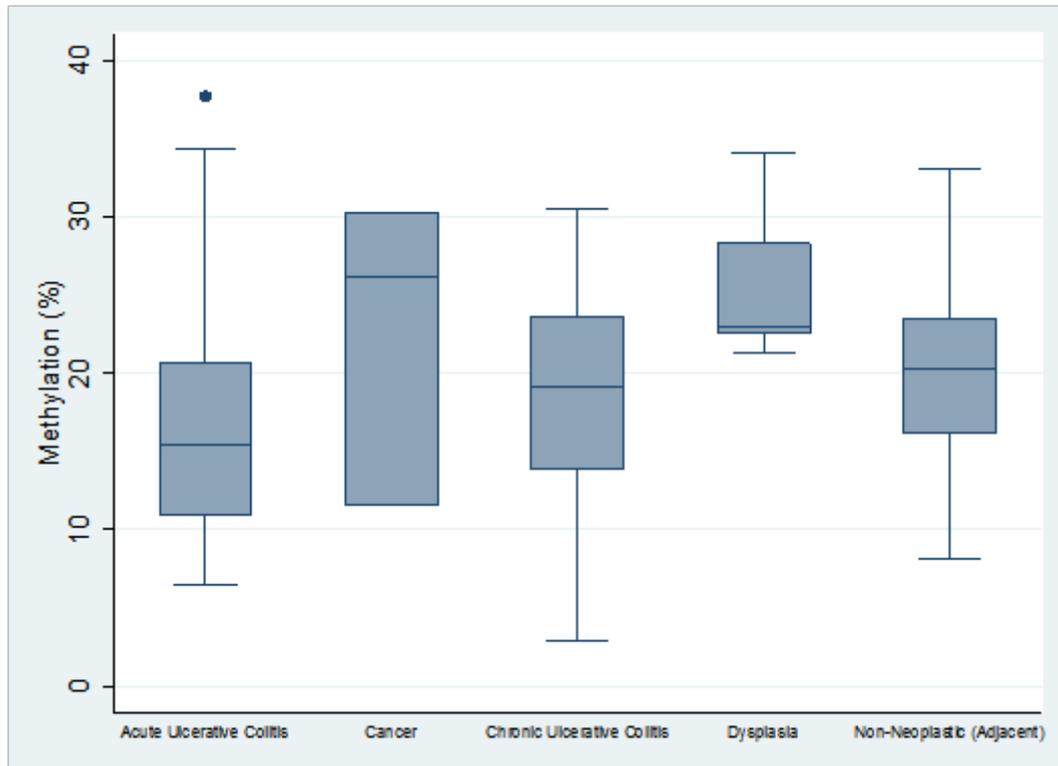
To investigate this anticipated overlap, a pre-existing held cohort of samples were analysed to determine variation in methylation of TRIM15-cg19115272 in progressive disease states in the colon. Tissues examined include those from a similar inflammatory disease process – Ulcerative Colitis (UC). In this disease process, much like that observed in the inflammatory driven metaplasia seen in formation of BO, the chronically inflamed colon due to UC has a much higher than background incidence of colorectal cancer.

Therefore, from a known colitis cohort, samples of normal mucosa were analysed alongside acute colitis, chronic colitis, dysplastic samples and colon cancer. The hypothesis, if there is overlap with the upper GIT, is that there would be a step wise increase in methylation of TRIM15-cg19115272 with normal tissue having the lowest methylation level, increasing through acute colitis, chronic colitis, dysplasia and then a similar or lower level of methylation in cancer. Samples were analysed by bisulphite pyrosequencing using the same TRIM15-cg19115272 primer.

5.6.4: Results:

As can be seen in figure 5.4, the pattern of methylation seen in these colonic tissue types is seemingly random, and there is no statistical difference between any of the tissues analysed. This shows the true significance of the TRIM15-cg19115272 as a methylation biomarker in the upper GI tract.

Figure 5.5: Methylation variation between colonic tissue states of TRIM15 cg19115272



The boxes span the 1st and 3rd quartiles

Band inside the box is the 2nd quartile (the median)

The lines are the “lower adjacent value” which is the first quartile - 1.5 x intra-quartile range and the “upper adjacent value” which is the third quartile + 1.5 x Inter-quartile range

Graphical representation of methylation variation in colonic tissue types, revealing no statistically different results.

5.6.5: Conclusions:

The apparent specificity of TRIM15-cg19115272 methylation at being able to differentiate disease states of oesophageal tissues only increases its potential usefulness in OADC patients. It also opens the possibility that, should this marker be identifiable in samples other than tissue, it should be specific to the upper GI tract,

and results would not be skewed by colorectal pathology. The obvious choices here would be either blood samples or stool, and the sensitivity of this being seen in the upper GI tract alone would be able to tailor investigations accordingly should a positive test in these samples be found.

This finding does raise the possibility of a peripheral, none invasive screening test for OADC. Should DNA be shed by an OADC tumour into either blood (as cell free circulating DNA) or stool and subsequently be detectable in sufficient amounts for methylation analysis, this could be exploited by clinicians to avoid the need to carry out invasive endoscopic examinations. Additionally, there may also be potential use for this methylation marker in the context of post-operative disease monitoring, for which there is no current useful method.

If a recurrent tumour continued to shed DNA into the blood, and do so with increasing quantity as the disease burden increases, this may well represent the first useful disease monitoring test. This would present significant advantages to the patient, particularly if recurrent disease was identifiable early, presenting an earlier indication for imaging to detect tumours. This might present an opportunity to facilitate earlier chemotherapy use in this patient cohort. It might be expected that this would lead to a significant impact on length of survival. While outside the remit of this thesis, this represents important future work to carry out.

5.7: *TRIM15* as a Functional Biomarker

5.7.1: Introduction

Although it is not essential for a biomarker to be of functional significance, it should be established whether this is the case. Those biomarkers that do have a functional

role within the cell clearly have greater biological importance, and may even be driving the disease states in which they predominate.

A literature review of *TRIM15* reveals a paucity of data and knowledge about this gene and its downstream application. Current understanding, based on computational modelling of protein structure, has hypothesised that *TRIM15* is an E3 ligase (129). In short, when a protein within the cell is labelled for degradation, it becomes ubiquitinated by an E3 ligase enzyme (130). The addition of these ubiquitin complexes signals the protein for digestion by the proteasome. However, if this ligase is turned off, for example by hypermethylation of the gene leading to lack of transcription, the protein cannot become ubiquitinated and therefore will not be broken down. It follows therefore that if this protein is somehow involved in normal cell death, and is not digested, a cell otherwise targeted for destruction will survive, and a tumour cell could propagate.

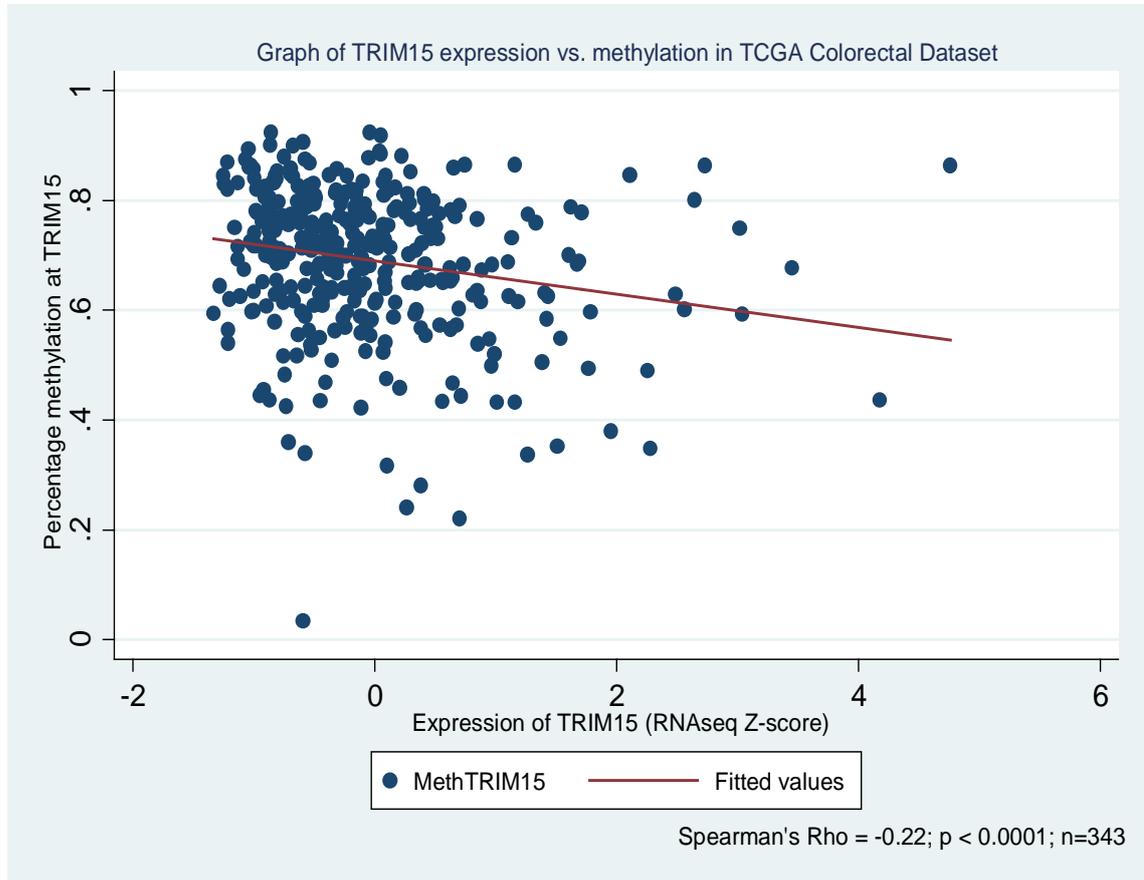
The first method of establishing potential significance of this *TRIM15* biomarker is to prove the correlation, if any, between *TRIM15* methylation and expression. In order for this methylation signature to have functional significance in this context, it is important to prove that the accepted link between hypermethylation and reduced expression applies in this case.

In order to justify functional experiments, it was deemed necessary to firstly assess whether any existing data would reveal a correlation between methylation and expression of *TRIM15*. To do this, methylation and expression data for colorectal tumours was extracted from the TCGA database used before. Colorectal data was used as this was available from the TCGA database, whereas data for Oesophageal

tissue was not available at the time of analysis. The analysis of a different tissue type was acceptable as a proof of principle was all that was required at this stage.

The analysis of 343 data points from the TCGA database showed a clear and significant ($p < 0.001$) correlation between methylation and expression of *TRIM15* (figure 5.5). In short, as methylation of *TRIM15* increases, protein expression is reduced. This is exactly the anticipated trend and should be a finding replicated in the oesophagus. If this is the case, not only would the methylation of TRIM15-cg19115272 be a significant biomarker at disease identification in the oesophagus, but it may also be a component in the biology of OADC. Followed to its natural conclusion, the question of whether reducing methylation in an OADC tumour may in fact be an important drug target in the treatment of this disease is raised.

Figure 5.6: Correlation between TRIM15 cg19115272 methylation and expression



Variation in methylation and expression taken from the TCGA data for colorectal expression analysis as a proof of principle to conduct the same experiments in an oesophageal cancer cell line. The above clearly showing that as methylation increases, expression of *TRIM15* decreases.

5.7.2: Aims

To determine whether there is a correlation between methylation and expression of TRIM15-cg19115272 at the mRNA and protein level.

5.7.3: Demethylation of OE33 and its correlation with expression of *TRIM15*

The methylation of native OE33 cell line for TRIM15-cg19115272 is shown in table 5.5. This value was obtained by pyrosequencing of DNA extracted from the OE33 cell line.

Table 5.5: Methylation of TRIM15 cg19115272 in native OE33 cells

| Assay | Sample | Methylation (%) |
|-------------------|---------------|------------------------|
| TRIM15-cg19115272 | OE33 | 90.36 |

This reveals that native OE33 methylation of TRIM15-cg19115272 is >90%. Repeated validation experiments all produced the same overall result.

There was an assumption that native OE33 cells, where methylation was in the order of 90% would show little to no expression at the RNA level. Therefore, if methylation was reduced, it was hypothesised that the cells would begin to re-express *TRIM15*. QPCR results however did detect expression of *TRIM15* in native OE33 cells, therefore indicating that methylation may not have any controlling impact over this gene. Were this the case, it could be anticipated that there would be no change in expression when the cells were demethylated.

OE33 cells were treated with 5AZA (as described in length in methods section) to decrease methylation. However, with a reduction in methylation of 30% (from 90% to

59% achieved with 5AZA treatment), overall Ct on the QPCR was increased by 1.2. While this does not appear to be a large change in expression, calculation of the Delta delta Ct (the percentage change in expression given that expression doubles for each increment of Ct) equates to an increase in RNA of 225%. It can be concluded therefore that while not the only controlling influence over the *TRIM15* gene, methylation clearly does have a role to play.

5.8: Correlation of RNA Expression with Protein Expression

The final assessment of potential functional significance of *TRIM15* methylation within the cell is to link the observed change in RNA expression with downstream *TRIM15* protein. Clearly, it would be expected that both RNA and protein expression would increase with demethylation. Therefore this was tested.

5.8.1: Aims

To ensure the observed difference in RNA expression of *TRIM15* observed following OE33 demethylation correlates into protein expression.

5.8.2: Methods

To further quantify this difference and to ensure that the observed change in RNA transcription translated into end protein product, a protein assay was necessary. As a result of the method used for RNA extraction (TRIzol protocol), it was necessary to obtain protein from the same sample again using the TRIzol extraction method. This is not the most robust method of protein extraction and there are a number of flaws in addition to the yield being predictably lower than ideal. The protocol is as described in the methods section.

TRIM15 protein was analysed by western blot (previously described) with the purpose of comparing whether the observed doubling of RNA lead to a doubling of protein product. Many problems were encountered with dissolving the protein plug obtained at the end of the TRIzol extraction protocol. However, experiments revealed that a solution of 0.05% SDS with 4M urea in TRIS-HCl (with pH 8) achieved the greatest amount of protein in solution. Even with this, it was still challenging to make the resulting protein plug dissolve in sufficient amounts to allow western blot analysis. Eventually, the plug was subjected to sonification to encourage its physical break down. This resulted in a much greater protein yield and lead to encouraging western blot protein bands.

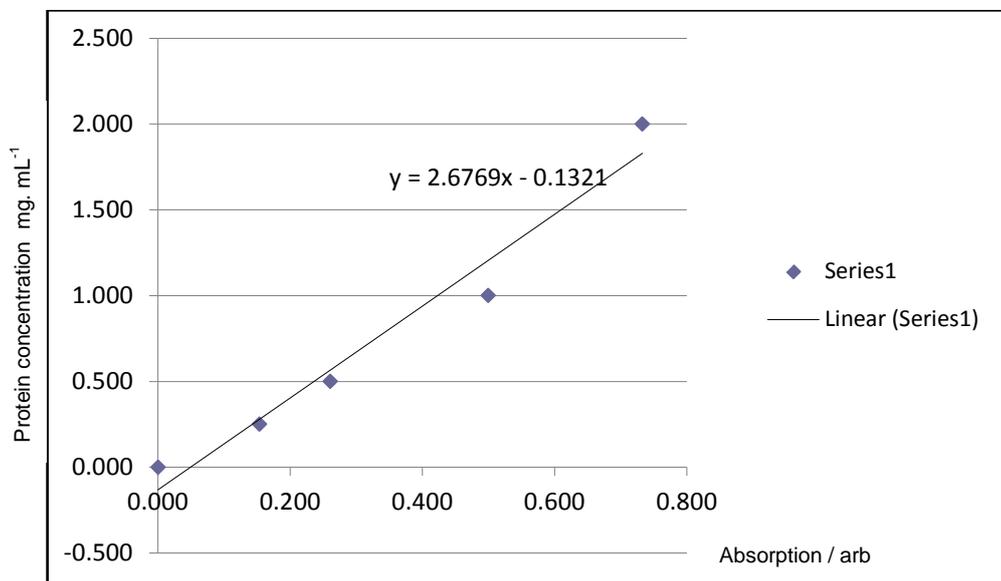
Protein was first quantified by a protein assay. This was an important step as overall total protein must be equal in all loaded samples to ensure that any observed difference in the specific protein (*TRIM15*) under investigation was not seen as a result of overloading one well with protein compared to another.

To perform the protein assay, 10 μ L of protein sample was added to a protein fluorescence buffer. 10 μ L of protein standards at 4 dilutions were also analysed in addition to analysis buffer in the absence of sample. This was allowed to incubate for 30 minutes at 37°C. The 96 well plate with samples was then analysed for protein content. In reality, if the initial green sample solution turned purple during the incubation, protein was present within the sample. True colour change, and therefore protein concentration, was quantified by spectrophotometry.

After analysis on the spectrophotometer, back ground noise (empty well) was subtracted from all samples, followed by subtracting the value given for the assay buffer without sample.

A plot was then performed for the 4 reference proteins as figure 5.6.

Figure 5.7: Standardisation of protein samples



Series 1 = Data points for protein standards.

The spectrophotometer readings for standard protein concentrations (0.25, 0.5, 1 and 2) are plotted on the graph above. This allows the generation of a linear correction to be applied to all measured protein amounts, producing a linear correlation between protein standards and results obtained. "y" shows the correction to be applied to experimental samples.

Figure 5.6 allowed the generation of a linear correlation between the 4 readings from the protein standards and spectrophotometer readings, which allowed the formula shown to be generated. Therefore, by entering the spectrophotometer reading and

adjusting by the given formula, the protein concentration ($\mu\text{g/ml}$) was calculated for the extracted experimental samples.

This allowed the calculation of sample volume required to add 20 μg of protein to each western blot well. Therefore, each well on the blot gel contained the same initial amount of protein, meaning that any observed difference in western blot band intensity was the consequence of a change in expression and not loading bias. A standard western blot was run, with the protein being mixed with loading buffer and heated to 100°C for 5 minutes. The required volume of protein was added to the respective wells so that the total protein in each well was equivalent. Protein was loaded, both from the native cell line and also the treated cell line and these were loaded in triplicate.

Although each well should in theory contain the same initial overall protein amount (from the protein assay calculation), this loading was confirmed with a standard assay which in this case was β -actin.

The protein was allowed to run through the gel alongside a ladder by electrophoresis. After this, the proteins from the gel were transferred to the analysis membrane. Subsequently, the membrane was washed in milk to allow nonspecific protein binding (this stage mops up the protein binding availability of the membrane such that chosen antibodies then only bind to the protein being investigated). The membrane was then agitated overnight in a fresh milk solution containing the *TRIM15* or β -actin antibody at the concentration recommended by the manufacturer.

The following day, membranes were washed, before incubating in a secondary antibody with affinity for the primary antibody. Again the membrane was washed

after this stage before being developed (to show the presence or absence of protein band) onto an x-ray film. Varying exposures were used in order to achieve the optimal protein band picture and allow quantification of protein amount.

The flaw in this experimental plan was discovered only when the x-ray films were developed. Although the 5AZA treated cells demonstrated 200% increase in the amount of *TRIM15* protein, the results also showed a 200% increase in the amount of B-actin protein within the samples from the treated cells in comparison to the native OE33 cells.

The reason for this unexpected finding, which brings the results into question appears to be that β -actin also has a CpG site in its promoter region (131). Given that 5AZA will theoretically reduce methylation in CpG sites across the genome it is possible that the variation in protein expression for β -actin had also been influenced by the demethylation from 5AZA. As a protein standard against which a comparison should be made, this was not a wise selection and a further plan was developed.

To prove the protein loading standardisation in a western blot, any resulting band would usually be compared to a standard protein that would not have been affected by the experiment, which usually involves using β -actin. However, global cell demethylation appeared to have reduced the expression of this protein also. As a result, a total protein stain (coomassie blue) was used. Again, the gel was set up in the standard fashion as previously described. Protein from native OE33 cells was loaded in triplicate next to protein from the 5AZA treated cells.

Protein was also loaded from both samples as a control which would be used for a coomassie blue stain for total protein. No specific protein standard was used

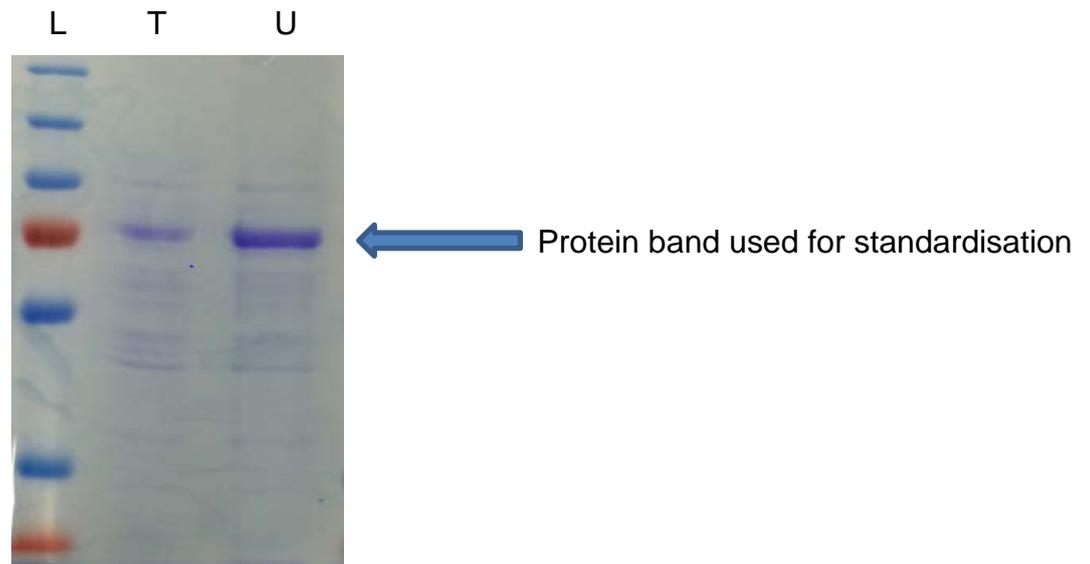
therefore, but since coomassie blue stains all proteins, this would be used as the control to ensure equal loading of wells. After loading, the gel was allowed to run. After which, the experiment protein gel was transferred to a membrane while the control gel was exposed to a coomassie blue stain for 30 minutes. As before, the experimental membrane was washed with milk after transfer, before being exposed overnight to *TRIM15* antibody.

After 30 minutes of coomassie dye exposure, the gel was left to agitate overnight in wash solution, until only the protein bands remained visible and blue, but the remainder of the gel was then clear. Meanwhile, the protein membrane was washed and exposed to the secondary antibody before being exposed to x-rays capture the protein band intensity.

Thereafter, experimental bands were standardised according to the amount of protein represented in the coomassie stain. This allowed comparison of the protein expression levels, standardised for total protein sample content.

The coomassie blue protein standardisation gel is shown in figure 5.7.

Figure 5.8: Coomassie blue stain



The intensity of bands are shown on this gel, with treated (T) cell line protein on the left and untreated (U) on the right. The protein ladder (L) is shown here on the far left. Division of intensity of bands after scanning and assessment by ImageJ software produced the standardisation formula for experimental proteins.

Figure 5.9: Western blot film following exposure, revealing obvious variation in protein content between 5AZA treated (T) and untreated (U) OE33 cells.



All x-ray films, in addition to the coomassie blue protein control gel were. Intensity of all the resulting bands was calculated using ImageJ. Triplicate samples were then grouped to provide the mean intensity of all respective bands. Band intensity was then standardised according to a protein band from the coomassie blue stain such that overall loading of protein difference was removed from generating bias. This

allowed the calculation of average protein amount, the standard deviation and error between triplicate samples and the significance of the difference shown.

5.8.3: Results

The comparison of expression between treated vs untreated cell line protein, with standardisation for the total amount of protein in the sample reveals the following results (table 5.6)

Table 5.6: Corrected protein content of OE33 cells, with and without 5AZA treatment

| Sample | Average | Standard deviation | Standard error | T-test |
|------------------|----------------|---------------------------|-----------------------|---------------|
| Treated | 2.48 | 0.75 | 0.37 | P=0.014 |
| Untreated | 1.00 | 0.00 | 0.028 | |

Average: mean

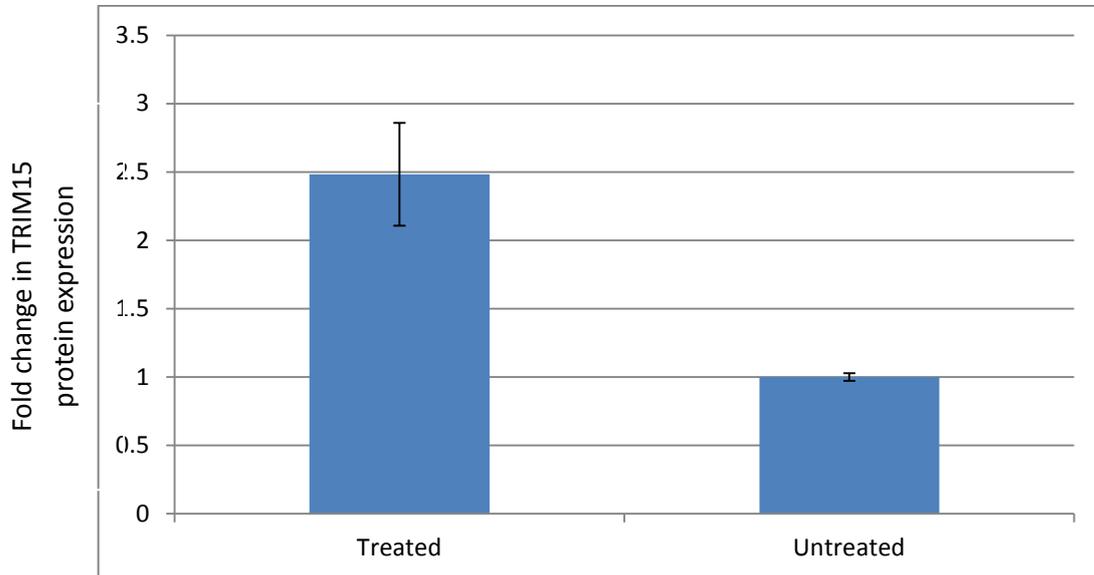
Standard deviation: Variation about the mean

Standard error: Standard deviation of the mean

T-test: Calculation of a statistical p value

This same data can then be shown graphically as figure 5.9.

Figure 5.10: Corrected protein content of OE33 cells, with and without 5AZA treatment



Graphical representation of the protein content comparison between native OE33 cells and those demethylated with 5AZA. This shows a reduction in protein expression of 50% when methylation was reduced. This reduction in protein amount correlates with RNA expression change.

Therefore, these data clearly show that expression of the *TRIM15* protein is doubled in the cell line treated with 5AZA, i.e. where methylation was decreased from 90% to 59%. In addition, the expression of protein level correlates exactly with the expression at the RNA level. This work now provides the preliminary evidence needed for the further exploration of *TRIM15* function.

5.8.4: Conclusions:

The work in this chapter has strengthened the evidence that methylation variation between normal tissues and disease exists in OADC and that these have the ability to be used as clinically useful biomarkers. This data has also proven the usefulness of modern array technology in selecting targets which have a greater ability to

discriminate tissues than randomly selected genes, and appears to supersede previous array technologies as might be anticipated. Caution must be employed however, as there is no guarantee that markers shown in an array will validate in larger sample cohorts and care must be paid to ensuring the validation experiments cover not just the gene in question but the specific CpG sites within that target gene.

This chapter also stressed the importance of accurate bioinformatic analysis of huge data sets to determine not only variation in methylation patterns between tissue states but also their significance. However, it is observed that array data alone is relatively meaningless unless successful validation follows in large powered sample cohorts. This point is very obvious in this data set as 3 of the 4 sites flagged in the *TRIM15* gene failed to subsequently validate.

The discovery of *TRIM15* as a methylation biomarker of OADC is of high scientific interest and clearly warrants further work to discover its function both within a normal cells and its role in OADC. However, the link between methylation and expression at both the RNA and protein level demonstrated in this chapter suggests that *TRIM15* hypermethylation might be an important transcriptional control and may therefore be a driving influence over the development of OADC. If this is the case, and especially knowing that modulation of the gene is possible in a cell line model, this could potentially translate into a drug target in the future.

Although little is known about the function of this gene, the TRIM family of genes and proteins have been previously reported in OADC. *TRIM44* has been reported in the context of OADC on 3 occasions with data suggesting a survival advantage in a

normally functioning gene (132, 133), but also suggests a therapeutic target to control its genomic overexpression (134).

This latest discovery of *TRIM15* in OADC strengthens the case that the TRIM family of proteins might have a driving role in this disease. This needs significant future work in order to determine the precise role of these proteins in the hope of understanding the disease processes of this cancer better. To begin this work, it would be possible to perform a further targeted experiment looking at the downstream effectors of *TRIM15* by carrying out a knockdown experiment with cell lines and looking for a functional pathway using RNAseq and collaborations with other interested groups with the appropriate experience in this field are now established to take this work forward.

6: Novel Approaches in Oesophageal Biomarker

Discovery

6.1: Barrett's Oesophagus as a Clinical Problem

The incidence of BO is increasing year upon year, which is largely thought to be a consequence of obesity driven reflux disease (17, 83, 135, 136). Despite the huge number of patients suffering from BO, it is in its own right a benign non-harmful state. Indeed for many, the development of Barrett's mucosa in a reflux inflamed lower oesophagus is most likely protective. Nevertheless, a small number of patients with this condition will develop OADC, and the incidence of cancer in the BO population is up to 150 times greater than the unaffected population.

The issue currently is that in patients with non-dysplastic BO, there are no robust methods for identifying the small number of patients at high risk of progression to cancer. Surveillance practice of the Barrett's patient varies widely at present, between some who endoscope these patients each year in contrast to others who will never repeat the investigation. Recent guideline publication (137) for the Barrett's patient now formally advises surveillance between 3-5 years, depending on segment length (with low or high risk being determined by a segment <3 or >3cm).

The result is that huge resources will need to be invested in surveying every patient diagnosed with BO. This process takes considerable financial backing, which in the current NHS will not be easy to achieve, and as with any surveillance programme will need to prove it can detect OADC at an earlier stage and lead to improved survival. Therefore, there is a significant requirement for a robust method of risk stratification

in these patients. If a marker could be identified that had the ability to predict development of cancer, this would allow a much more streamlined surveillance programme, targeting only those who were predictably high risk for developing a tumour. This would advantage the patient by resulting in potentially earlier tumour diagnosis and therefore better outcome. In addition, those stratified into the low risk group could avoid an unnecessary investigation which although generally safe, can be associated with complications and is unpleasant.

Indeed, if it were possible to identify these high risk patients and understand the mechanisms involved in tumour formation in this group, it might be possible to develop management strategies designed at preventing the formation of cancer, although current treatments for BO seem to conflict as to whether they reduce the long term cancer incidence (138). Clearly this would provide the necessary rationale for a large scale clinical trial to investigate such patients in more detail.

Stratified surveillance of the high risk Barrett's patient will combine the advantage of an early diagnosis in with ever improving endoscopic mucosal resection techniques. Experienced endoscopists are now able to resect early tumours limited to stage T1a, removing the need for patients to undergo high risk oesophagectomy surgery. However, as stated before, the ability to differentiate a tumour between T1a, where lymphatic involvement would be rare, and T1b where nodal disease is considerably more likely is challenging with current techniques. Clearly, when a patient has lymph node disease, it is crucial that these lymph nodes are removed to guarantee full disease clearance, and this can only be achieved with surgery. Even in those patients with early disease who have had a local resection of their primary tumour,

lymph node recurrence/progression of missed disease is troublesome and necessitates salvage oesophagectomy.

While generally excellent, current staging techniques can struggle to identify the subtle differences between early lesions. Therefore, a marker of a more advanced disease process as evidenced by lymph node disease would be extremely useful in this context, and would provide the justification for high risk surgery in patients with early lesions but in whom there is already local lymphatic involvement present at only a microscopic level.

6.2: Aims

- To determine whether methylation signatures can differentiate between high risk Barrett's which will progress to cancer vs low risk Barrett's who will never progress.
- To determine whether methylation signatures have the ability to determine tumour aggressiveness as evidenced by lymph node metastasis.

6.3: Patients and Samples:

In order to attempt to answer the above question with an array, it was necessary to assemble panels of specific and rare tissues

6.3.1: Samples for Assessment of Barrett's Progression

Several previous groups who have attempted to answer this same question have done so using known cancer patients (110, 111). Many groups have used an overriding assumption that where a patient has visible BO at the edge of a tumour specimen, this tumour must have arisen from this BO and therefore a follow-on

assumption is that this BO represents “progressing Barrett’s.” For this research, it was considered this was an unwise assumption to make. This statement about the tumour definitely having arisen from this BO is likely but not easy to prove. In addition, if visible BO at the tumour edge represents progressing disease, why has it not all been consumed with cancer? The very fact tumour has developed may mean the epigenetic signature of this tissue has already been altered in line with the tumour, as a field change effect is likely and has been demonstrated already in other tissue for example the colon, and further shown in Chapter 5 of this thesis. In an attempt to determine a methylation marker of “progression risk” from non-dysplastic BO to OADC, using tissue of this nature was thought unwise and was therefore excluded as an option.

Therefore, a different strategy to tackle this experiment was taken, although this posed challenges in sample acquisition. Patients and samples were identified at the very origin of the disease process, and tracked throughout time retrospectively. Samples were only included where there was a non-dysplastic BO biopsy and then proof (histopathology) that the same patient developed cancer. Patients who only progressed to dysplasia (of whichever grade) were excluded. Where there was no histopathology proof that the patient had developed cancer, or where this diagnosis had been made in a different hospital, the patient was also excluded. While this removed many potential patients from the experimental cohort, it meant that those who were included were truly representative of the target population and any results should be representative.

It was a particular challenge to find a cohort of BO patients which met these stringent inclusion criteria. To do this, collaborations were established with other centres

within the West Midlands. Often, centres kept no robust database of BO patients, and this made tracking difficult. Clearly a centre could only be included where a surveillance programme was conducted. Often, where a database did exist, it included patients who were dysplastic at first diagnosis and these were excluded. Often, where robust surveillance was in place, patients had multiple biopsies of non-dysplastic Barrett's before developing dysplastic changes prior to their eventual cancer diagnosis. Where there were temporal biopsies, all of these were included in the study where tissue was available. This means that a single patient could contribute several samples to the study. This was also important as it allowed a comparison of markers over time – i.e. if a marker of progression was discovered, it would then have been possible to demonstrate how long prior to cancer development this marker was identifiable in the tissue sample.

An additional challenge included the historical nature of the samples needed for this analysis. It was frequently not possible for the respective pathology department to locate the block of interest. This significantly limited the number of available tissues and of course meant that the sample cohort included was small.

Eventually, 12 patients were included in the progressing cohort and this produced 30 samples. The best 6 (greatest DNA yield in addition to best quality of DNA as per nanodrop data) were put forward for the array analysis. Remaining samples were then used for the follow-up validation exercise of array identified markers.

The non-progressing, never-dysplastic group was easier to obtain, however not without issues. Again, in order to be classed as never dysplastic, the patient had to be enrolled in a surveillance programme. Patients were only included where they

had been in surveillance for more than 15 years, had numerous non-dysplastic biopsies over that time, were still alive and had a non-dysplastic BO biopsy within the last 2 years (showing they were still in surveillance). 37 samples were included in this group. Again the best 6 samples were used in the array experiment with remaining samples used for validation attempts.

6.3.2: Tissues for Assessment of Lymph Node Disease

To discover a marker of lymph node positivity, it was necessary to acquire a group of samples in which as many other variables as possible could be minimised. Therefore, a decision had to be made. The natural pattern of OADC means that the earlier the stage of tumour, the less likely it is for there to be lymph node involvement.

Ideally, a cohort of T1 resected tumours, both with and without lymph node disease was required for this experiment. A review in OADC resection cases revealed that there were insufficient numbers of such specimens available. In order to achieve appropriate numbers of samples, and in an attempt to keep as many variables static as possible, it was necessary to select either a cohort of T2 or T3 tumours. In addition there then had to be enough samples obtainable which had N0 and N1 disease. The review into potential sample numbers showed that using a T2 group would again not yield sufficient numbers for powered experiments to take place. Therefore, T3 tumours were used as these represented the greatest cohort of resected tumours, as would be expected. Even with this being the case, it was challenging to find sufficient number of T3N0 samples, as lymph node involvement would be anticipated by the time this tumour stage was achieved.

6.4: Methods

DNA was extracted from FFPE biopsy samples as previously described. These samples were all qualified and quantified using both Nano-drop and Qubit techniques.

6.5: Methylation 450 arrays

The discovery target gene methylation analysis was performed using an Illumina methylation 450 array, as this represents the latest methylation comparison array platform. The array chip slides were prepared locally however the scanning of the chips occurred in Oxford University, with kind thanks to the Tomlinson group.

The array samples were processed exactly as per the Illumina protocol (reference in Methods section). The subsequent data generated was analysed using a Bayesian regression model in order to compare methylation differences between the two sample groups. Unfortunately, one of the chips containing a BO progressor failed to take the sample all the way across. There was no identifiable reason for this other than failure of the capillary action to draw the sample over the chip. Even though this failure was obvious at the time of sample loading, there was nothing that could be done to rectify the issue. Therefore the results from this chip were reviewed prior to inclusion in the analysis. Although the chip did provide data for 200,000 CpG sites, it was thought best to exclude all of the results from this chip for analysis so as not to skew any other data. The consequence of this was that there was variability in the sample cohort size, but this was not thought to be important as it was the mean values which were taken forward for the comparison analysis, therefore this drop-out could be adjusted for.

6.5: Validation experiments

Validation of any significant result from the above Methylation 450 array was performed by bisulphite pyrosequencing using the remaining samples obtained. This was performed as previously described.

6.6: Results:

6.7.1: High Risk Barrett's Biomarker.

As in Chapter 5, array data was analysed in the same format to allow a comparison of methylomes between the two groups (non-dysplastic BO that progressed to cancer vs. those that remained static). Variation (either hyper or hypomethylation) was ranked in order of statistical significance to allow the identification of differences between the groups.

The comparison between BO samples that progressed to OADC in comparison to those which did not progress is show in table 7.1. There were 44 differentially methylated sites identified.

Table 6.1: Top 20 array identified CpG sites with methylation variation between non-dysplastic samples which progress to OADC vs those which remain static

| <u>Probe ID</u> | <u>t</u> | <u>P.Value</u> | <u>Adj.P.Value</u> | <u>Gene Name</u> |
|-----------------|----------|----------------|--------------------|------------------|
| cg07863524 | -12.26 | 2.02E-08 | 0.0031 | OR3A4 |
| cg24007926 | -11.05 | 6.84E-08 | 0.0032 | NA |
| cg17337672 | -10.95 | 7.54E-08 | 0.0032 | FGFR2 |
| cg02226469 | -10.88 | 8.17E-08 | 0.0032 | NA |
| cg17433294 | -10.51 | 1.22E-07 | 0.0038 | NMUR2 |
| cg18479711 | -10.31 | 1.52E-07 | 0.0039 | HDAC4 |
| cg09011162 | -10.15 | 1.82E-07 | 0.0040 | LMF1 |
| cg19733463 | -9.83 | 2.61E-07 | 0.0045 | NMUR1 |
| cg16150571 | -9.64 | 3.26E-07 | 0.0045 | SNORD116-22 |
| cg24424217 | -9.62 | 3.35E-07 | 0.0045 | ZNF511 |
| cg13164993 | -9.57 | 3.52E-07 | 0.0045 | RBP3 |
| cg14019464 | -9.53 | 3.70E-07 | 0.0045 | TRIB3 |
| cg24581378 | -9.53 | 3.71E-07 | 0.0045 | ZAP70 |
| cg05230642 | -9.36 | 4.54E-07 | 0.0051 | SNORD115-14 |
| cg12297814 | -9.26 | 5.08E-07 | 0.0052 | IGFN1 |
| cg11231240 | -9.23 | 5.33E-07 | 0.0052 | NA |
| cg11864327 | -8.96 | 7.40E-07 | 0.0063 | ZFP2 |
| cg16771467 | -8.92 | 7.78E-07 | 0.0063 | ATP8B1 |
| cg17304276 | -8.88 | 8.14E-07 | 0.0063 | CUX2 |
| cg11443888 | -8.86 | 8.36E-07 | 0.0063 | TMEM151B |

Probe ID – the Illumina cg probe ID from the Illumina manifest

T = the t value – the size of the difference relative to the variation in the sample

p-value = the raw p-value, not corrected for multiple testing

Adjusted p-value = the p-value corrected for multiple testing

Of the top 20 differentially methylated sites as shown, an analysis of the identified genes was conducted to see whether any had been previously implicated in OADC or BO. While most were *de novo* findings, Fibroblast Growth Factor Receptor 2 (*FGFR2*) has been previously reported in the oesophagus. Therefore, *FGFR2* along with the top hit from the array, Olfactory Receptor 3A4 (*OR3A4*) was selected for downstream validation attempts.

Fibroblast Growth Receptor 2 (*FGFR2*), a receptor tyrosine kinase, has already been shown to be amplified in both oesophageal and gastric cancer although, methylation change has not been reported (139). Given the relative hypomethylation of this gene in progressing samples observed in this data and the presumed increase in expression as a consequence of this, this gene probe was selected as a target. Other studies have echoed this expression change. Peterson et al, 2013 showed using immunohistochemistry that at least 51% of OADC overexpressed one of the receptor tyrosine kinases and that the level of over expression is also increased in BO and becomes increasingly so throughout the dysplastic stages (140).

Of further potential interest with this gene is the existence of a known *FGFR* inhibitor dovitinib (141). Although this drug has most affinity for *FGFR2*, there is cross over with other members of the RTK family and this raises the question whether it may have the ability to arrest development of cancer in the Barrett's patient.

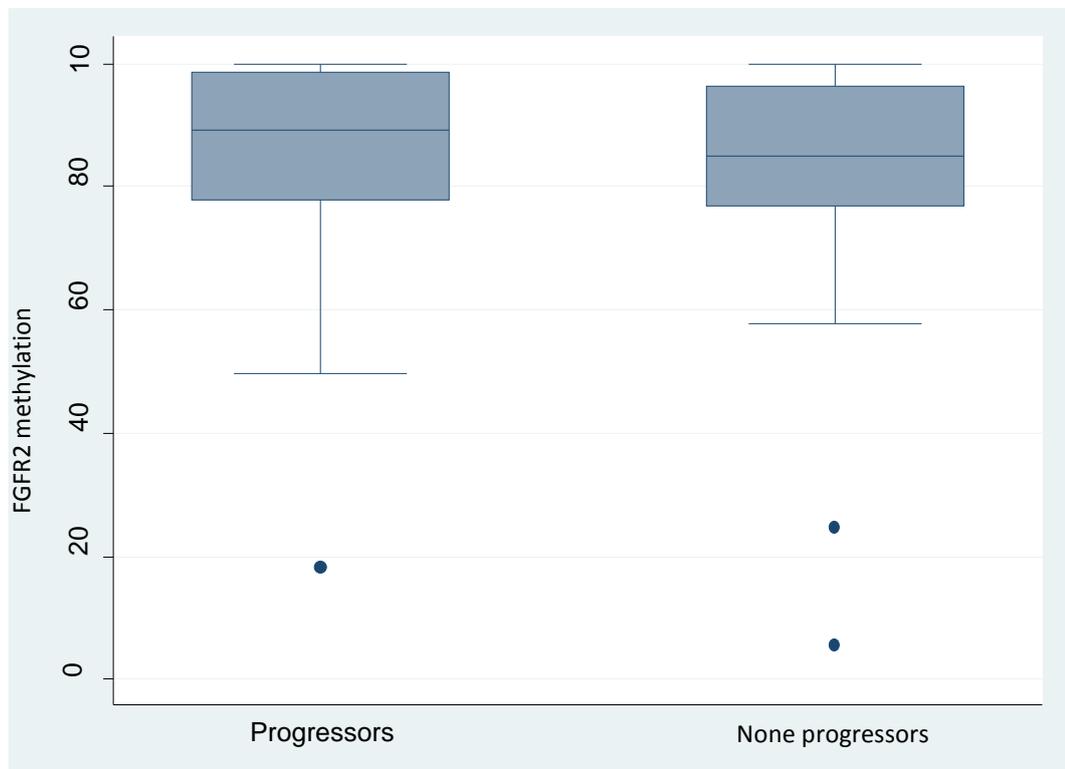
In addition to the potential of *FGFR2*, the top hit marker in this array experiment was *OR3A4*. This marker, showing hypomethylation in progressing patients, has the greatest significance ($p=0.003$) of all the 44 most significant markers shown in this experiment.

As a consequence, methylation insensitive primers were designed around these 2 targets. Once obtained, each of the primers was optimised. This was particularly important as DNA used in this pyrosequencing was sparse and of much poorer quality than that used in previous chapters. Once robust, these markers were analysed using bisulphite pyrosequencing as before.

6.7.2: Analysis of Pyrosequencing Data

Analysis of the pyrosequencing data from *FGFR2* failed to discriminate between samples of BO that progressed to cancer vs those that did not. This result is demonstrated in figure 6.1 below.

Figure 6.1: Variation in methylation between Progressors and none progressors in *FGFR2*



The boxes span the 1st and 3rd quartiles

Band inside the box is the 2nd quartile (the median)

The lines are the “lower adjacent value” which is the first quartile - 1.5 x intra-quartile range and the “upper adjacent value” which is the third quartile + 1.5 x Inter-quartile range

This graph clearly shows no variation in methylation of *FGFR2* between BO progressors and none progressors. This marker therefore failed to validate.

6.7.3: *OR3A4* data

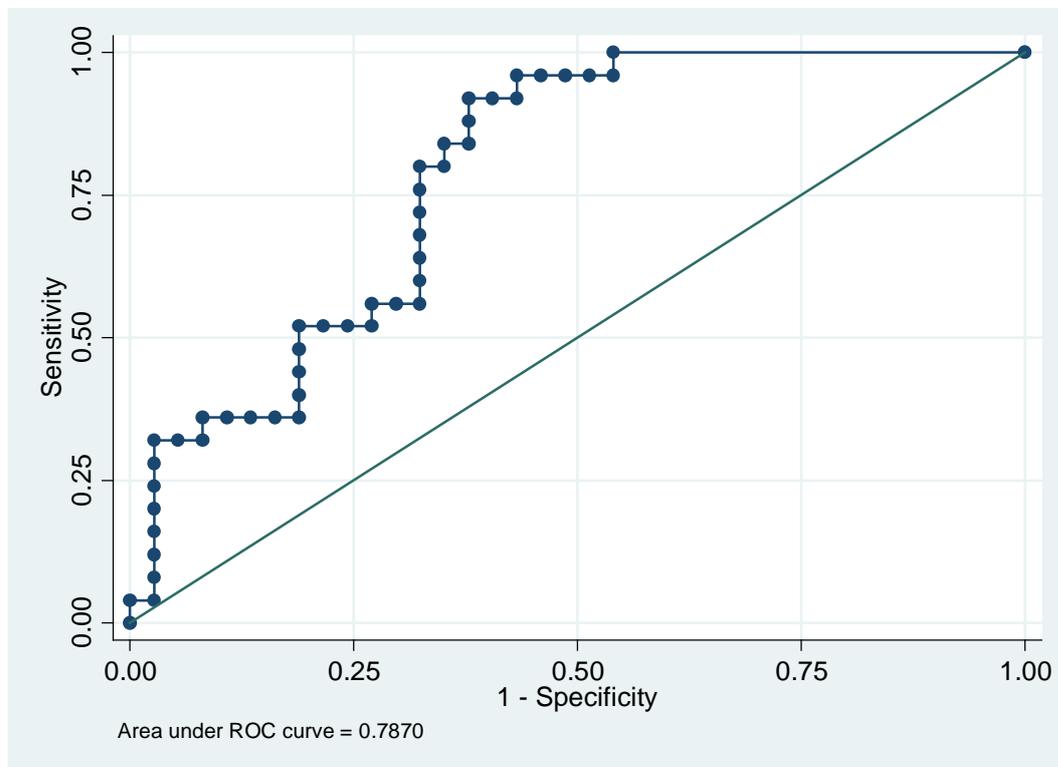
Bisulphite pyrosequencing data for the flagged *OR3A4* CpG has been statistically analysed in table 7.2.

Table 6.2: Pyrosequencing validation data for *OR3A4*.

| CpG | Odds Ratio | Standard. Error | z | P> z | 95% Conf. Interval |
|----------------|------------|-----------------|-------|-------|--------------------|
| 1 <i>OR3A4</i> | 0.94 | 0.02 | -3.47 | 0.001 | 0.91-0.98 |

The significance of the data is presented in the following ROC curve (figure 7.1).

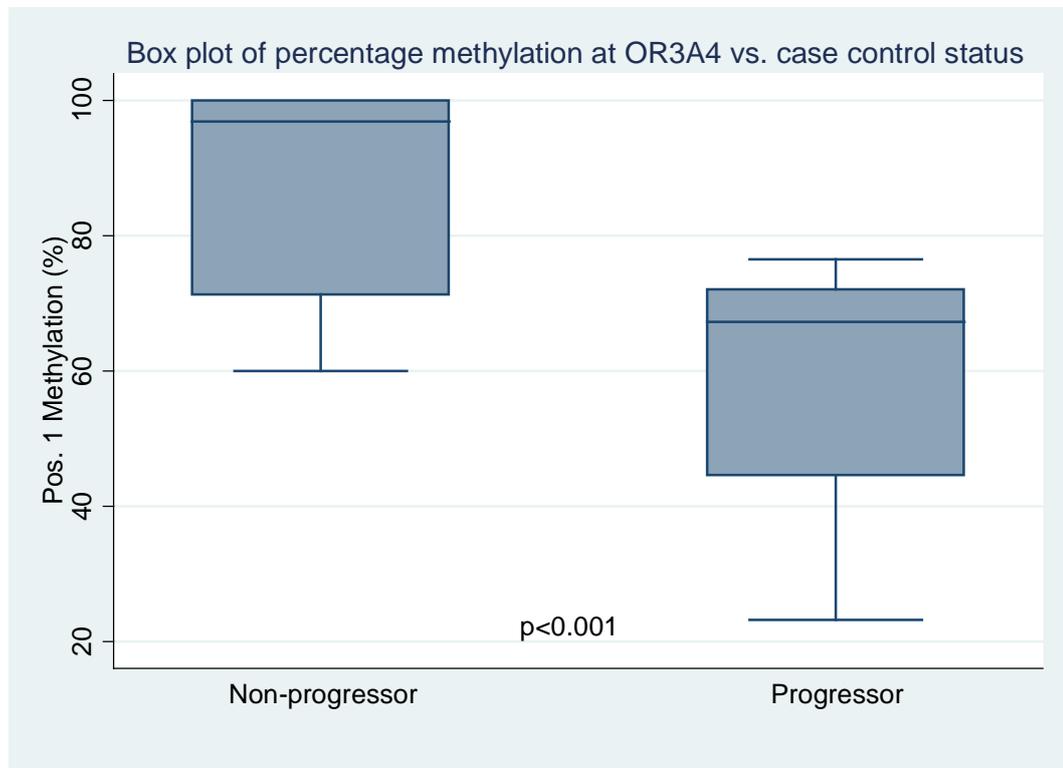
Figure 6.2: ROC curve demonstrating the ability of *OR3A4* methylation to differentiate high risk BO



The above ROC curve demonstrates the performance of *OR3A4* as a clinical test, demonstrating sensitivity and specificity at varying thresholds.

Only one of the 4 amplicons tested in this validation exercise managed to produce a statistically significant result. However, the hypomethylation of *OR3A4* in progressing patients validated with a $p < 0.001$ and is shown in figure 7.2.

Figure 6.3: *OR3A4* methylation differentiating high risk non-dysplastic BO



The boxes span the 1st and 3rd quartiles

Band inside the box is the 2nd quartile (the median)

The lines are the “lower adjacent value” which is the first quartile - 1.5 x intra-quartile range and the “upper adjacent value” which is the third quartile + 1.5 x Inter-quartile range

Logistical regression on this data set, setting a methylation threshold of 80% and modelling for a progression to cancer rate of 0.7% reveals a potential clinical test with the parameters shown in table 7.3.

Table 6.3: Logistical regression showing the use of OR3A4 methylation as a clinical test.

| Test | % |
|-------------|-------|
| Sensitivity | 80 |
| Specificity | 67.6 |
| PPV | 1.71 |
| NPV | 99.80 |

Threshold for methylation level was determined by logistic regression. Sensitivity and specificity tables were formulated based on the above ROC curve and the percentage methylation was chosen to maximise sensitivity and specificity.

6.7.4: Conclusions for Barrett's Progression

While this data does not demonstrate the ability to predict cancer development in the non-dysplastic Barrett's patient, the negative predictive value of 99.8% is extremely useful. Arguably, for the intention of designing a surveillance programme with the ability to risk stratify the non-dysplastic Barretts patient, this is potentially more useful. Of significance is the ability of this test to exclude the development of cancer in this group, and do so 99% of the time when methylation of the NDBO sample is above 80%.

The ability to exclude cancer development in 99.8% of the non-dysplastic Barrett's cohort would facilitate development of a streamlined surveillance programme, and lead not only to huge NHS cost savings but also the avoidance of un-necessary investigations in this group. Although generally safe, it should also be considered that endoscopy does carry an associated morbidity and mortality, which would again be reduced by only offering it to those patients with a high risk of cancer development. While this finding needs further validation in a larger cohort, this data is very promising and needs to be taken forward in the future.

Conceivably, should this marker re-validate in a larger cohort, it is possible that each patient diagnosed with non-dysplastic BO on endoscopy, and proven to be non-dysplastic on histology review should have a sample of their Barretts sent for methylation assessment of *OR3A4*. Any patient with methylation of this site greater than 80% could be reassured and discharged from further review. Conversely, if methylation was shown to be below 80%, the patient should clearly be entered into a surveillance programme. The avoidance of endoscopy in the low risk group would allow a more intensive surveillance of the high risk group, with the intention that should these patients develop a tumour in the future, this would be diagnosed at an earlier more treatable stage. Alternatively, another argument might be to offer these patients treatment for their BO. However, the evidence that such treatments reduce cancer incidence in this group is not currently strong.

6.8: Biomarker Discovery of Nodal Disease

Running in parallel with the experiment reported above was an array investigating whether the presence of nodal disease could be identified using a methylation

marker. To perform this array, 6 samples of T3N1 were compared to 6 samples of T3N0. Ultimately the question was not whether it would be possible to identify nodal disease in a T3 cohort, but whether a marker could be found that would assist in the staging investigation of T1 A/B patients, where knowledge of microscopic disease would alter their management plan significantly.

As these samples were run in parallel with the array already reported, there was no variation in methodology or sample handling. All samples were prepared according to the specific protocol from Illumina as reported in the methods section.

6.8.1: Results

Disappointingly, while every sample in this array produced a result and worked well, upon statistical analysis there were no statistically significant results found. The 20 top hits identified in the array are reported in table 7.4 but none of these reached the required level of statistical significance.

Table 6.4: Illumina methylation 450 array data comparing T3N0 tumours with T3N1 tumours

| Order number | Probe ID | Adjusted p value | Gene name |
|--------------|------------|------------------|-----------|
| 1 | cg22937649 | 0.1175 | NTN1 |
| 2 | cg10561103 | 0.2529 | ADAMTS16 |
| 3 | cg15165154 | 0.2529 | CXXC5 |
| 4 | cg20970354 | 0.3146 | NA |
| 5 | cg10676125 | 0.3146 | ZNF536 |
| 6 | cg23799393 | 0.3146 | BICC1 |
| 7 | cg01526189 | 0.3146 | CDH4 |
| 8 | cg04751761 | 0.3146 | SLC9A3R2 |
| 9 | cg14048780 | 0.3146 | SHC3 |
| 10 | cg26314869 | 0.3146 | NA |
| 11 | cg14643892 | 0.3146 | C12orf56 |
| 12 | cg16524232 | 0.3146 | NA |
| 13 | cg03306486 | 0.3146 | APC2 |
| 14 | cg18181196 | 0.3146 | ESRRG |
| 15 | cg14349131 | 0.3146 | NFIB |
| 16 | cg08296037 | 0.3146 | IFT140 |
| 17 | cg25140795 | 0.3146 | FLNC |
| 18 | cg06300254 | 0.3146 | CHAT |
| 19 | cg07612827 | 0.3146 | FLJ25758 |
| 20 | cg18863595 | 0.3146 | GRM7 |

Probe ID – the Illumina cg probe ID from the Illumina manifest

Adjusted p-value = the p-value corrected for multiple testing

As can be seen, none of the results achieves a p value of <0.05 and are not therefore significant. As a result, no validation exercise was undertaken.

6.9: Conclusions

Despite the disappointment that there were no significant methylation differences between patients with and without nodal disease, the parallel array investigating BO patients with progressing pathology has shown interesting results, and further

strengthens the case for an array based approach to discover methylation biomarkers, followed by validation exercises.

It is unclear at this point why the N1/0 array failed to produce significant results other than to suggest that methylation plays no part in nodal metastasis. It clearly is the case that there are many other pathways not investigated that might explain this mechanism.

However, the fact that some of the results approached statistical significance suggests that this may have been an experimental design issue or perhaps more likely a flaw in patient/tissue selection. It is also possible that there were insufficient samples used in the array to demonstrate the required level of significance. The initial hypothesis was that a methylation marker might assist in the cancer staging process predominantly in T1 disease. In these early tumours, it is known that T1a carries relatively little chance of nodal disease whereas T1b carries a much greater likelihood. It is vital to know whether a patient has nodal disease as T1a tumours can be cured by local mucosal resection, whereas T1b tumours (or any case where there is nodal disease) would need a lymphadenectomy which would necessitate oesophagectomy.

Due to the inability to acquire sufficient numbers of T1 tumour resections to allow array experiments (majority of resected tumours being T3 stage), and requiring both resection samples of T1N0 and T1N1 patients, it was not possible to obtain these tissues in sufficient quantity to allow experiments to proceed in the time frame available. Therefore, recognising that T3 was the predominant resected pathology, and large numbers of these resection samples were available with good volumes of

tissue – these were acquired and used as a surrogate marker of potential methylation signatures of lymph node disease.

However, it is possible that the choice of T3 tumours included in this experiment was a mistake. Appreciating that methylation change seems to occur early in the cancer process, and that T3 disease could be considered a late/advanced tumour, it is possible that any potential methylation signatures had long since gone from this cohort. This supports the evidence that methylation change precedes genomic change.

While outside the remit of this thesis, it should be considered therefore whether this experiment should be repeated using T1 tumours only. However, this is not without potential pitfalls. It should be considered that the majority of T1a tumours are now resected using a mucosal resection technique. While all staging modalities would have shown these tumours to be T1a without nodal metastasis, there is no guarantee that this is the case, and microscopic lymph node disease could not be reliably excluded. Therefore, repetition of this experiment with earlier stage tumours should be treated with caution, and significant follow-up of these patients would be required to ensure there was no local or nodal disease progression. Only doing this would allow patients and therefore tissues to be accurately assigned into their respective groups and the analysis be valid.

The success of the BO progressing vs non-progressing patients is a very interesting finding. Never before has there been any biomarker used routinely in the assessment of non-dysplastic BO samples to risk-stratify their potential for progression to cancer. Due to an ever increasingly obese population, driving reflux

disease in this group, the incidence of BO diagnosis is increasing rapidly. However, until the recent guideline publication in 2014 (137) no formal surveillance programme has been in place and there has been wide variation in practice between never surveying to rigorous surveillance determined largely by unit/clinician preference.

The discovery of this biomarker raises the question as to whether its routine use could change long term clinical practice. This hypomethylation change seen in progressing patients could be applied as a clinical test to all patients with a non-dysplastic BO biopsy in order to determine their long term cancer development risk.

While a second round, external validation cohort is currently being acquired to ensure the validity of this methylation biomarker, should this confirm the findings from the array, a large scale prospective clinical trial should be undertaken to assess the long term use of this potential clinical test.

If successful, the methylation status of *OR3A4* should be investigated in all non-dysplastic BO patients at index biopsy. This, like all other screening tests, would not be without some failure but; it may be an extremely useful tool in risk stratifying the non-dysplastic BO patient to determine the need for long term surveillance. The ability to perform this risk stratification will be invaluable in the NHS. Not only will it potentially prevent many endoscopies in patients at low risk of progression to cancer but it will also save these patients undergoing numerous unpleasant investigations which do of course carry a risk of bleeding and oesophageal perforation, which can lead to death. Moreover, such a marker is likely to be extremely attractive to the NHS as it would lead to reducing the number of surveillance endoscopy examinations which would of course have a huge cost saving, although there is

recognition of the costs involved in introducing this methylation test to clinical laboratories.

In addition to this validated marker, 42 other significantly differentially methylated CpG sites were identified in this array work. Each of these needs validating. It is highly likely that a panel of markers will have wider reaching significance than a solitary signature. Such a panel is likely to include biological, clinical and patient factors, but the results above demonstrate that more focus should spend identifying this panel of characteristics, in the ultimate aim of identifying tumours earlier in a high risk population and preventing cancer associated deaths in this population.

7: Discussion

7.1: Thesis Discussion

This thesis investigated the role of DNA methylation in the oesophagus and specifically Barrett's oesophagus and oesophageal adenocarcinoma. The experimental plan has been logical, firstly building on a pre-existing body of work in Chapter 4 which not only demonstrated that methylation can have a role in disease identification in the oesophagus but also that array based approaches were useful as an investigative technique in identification of novel targets.

Chapter 5 progressed this initial finding and revealed that more modern techniques have a greater information yield, demonstrating a novel methylation biomarker of OADC. Work built on this marker to include a basic functional assessment of *TRIM15* demonstrating the link between methylation and expression at both the RNA and protein level. This reported link suggests a potential biological role that *TRIM15* plays in OADC and future work needs to be done to assess whether this gene, whose current function remains relatively unknown, could be a driving force in the biology of this tumour. If so, this initial finding of a biomarker could in fact lead to the identification of a therapeutic pathway for OADC and significant future work is needed here to investigate this. This marker is made more interesting by the reporting from other groups of another member of the TRIM family of proteins being implicated in OADC (132, 134).

In Chapter 6, using the techniques already identified from previous work, 2 novel questions were proposed. Firstly, whether aggressive tumours could be

differentiated out, as evidenced by lymph node disease. While clinically the question was whether early tumours (T1a/b) could have their nodal status better defined with a methylation biomarker, the paucity of such resection cases meant that T3N0/1 tumours were used in substitute. An array based comparison approach between these groups revealed no overall variation, although it is worth noting that some markers were close to statistical significance. Because no significant markers were revealed from this experiment, the question whether the correct tissue groups were selected for this experiment was raised. Knowing that methylation changes predominate in early stage disease, it has been considered that the selection of T3 tumours was incorrect to allow changes to be seen. Therefore, there may be merit in repeating this experiment with the tissue for which the question was designed, although it will be much more challenging to obtain the correct resection cohort, which was not possible in the timescale allowed for this work.

Of much greater significance from Chapter 6 was the discovery of a biomarker that can discriminate between non-dysplastic BO that either does or does not progress to OADC. This could potentially change clinical practice. The current and increasing incidence of BO, combined with the recently published guidelines on how to manage and survey these patients requires huge man power and expense for a relatively low progression risk to cancer and therefore low tumour diagnostic yield. This marker, with further validation in larger numbers, could provide the first signature with the ability to risk-stratify these patients. This will in turn provide huge NHS cost savings along with the reduction in manpower requirements not to mention removing the need for patients to have an unpleasant investigation which has an attached morbidity and rarely mortality.

Therefore in summary, this body of work has allowed an understanding of the role that methylation biomarkers can have in OADC as well as proving an array based approach to find these potential markers is a useful tool. Perhaps more importantly, it shows that despite very significant and promising markers being identified in an array based approach, unless validation attempts are performed the array data alone should not be relied upon.

Chapter 6 adopted a novel approach into the investigation of biomarkers for BO progression. The bulk of the literature reports methylation changes in specific genes as tissue passes from normal to BO to dysplastic BO and finally OADC. Usually these are not from the same patient, as doing so would require the surveillance of a vast number of patients and the cohorts would be small. This thesis adopted a novel approach of not only performing a whole genome array to assess methylation variation between progressors and none but also rigorous patient selection with exactly the criteria needed to answer the hypothesis. Attention to detail here, along with all the steps to ensure correct tissue handling and processing produced a small patient cohort but none the less strong experimental results.

7.2: Follow-up Work

While significant progress has been made by this thesis, many questions have been raised and therefore future work is needed to explore these. In turn:

7.3: *TRIM15*

TRIM15 has shown to be a reliable methylation biomarker for identification of OADC from normal squamous tissue, which has now validated in a large sample cohort. However, its functional significance has as yet not been elucidated. Now that a link

between methylation and expression has been shown, this raises the question as to whether there is a biologically active role that *TRIM15* plays in the development of OADC. Unfortunately, there are few reports in the literature about *TRIM15*'s function and this needs to be elucidated further.

Interestingly, *TRIM15* is not the only member of the TRIM family of genes to have been shown have a biomarker function in OADC. Another group have already demonstrated a mutational profile of *TRIM44* in OADC, in addition to its potential implication in tumour biology and prognosis (132-134). It is unlikely that these markers have been found by different groups without there being a definitive link to the biology of the TRIM genes in OADC. Clearly this needs extensive further investigation.

Therefore to continue with the work of *TRIM15*, an investigation into expression of this gene in cell lines is ongoing. From this base line experiment, we intend to knock down *TRIM15* in a *TRIM15* expressing cell line using siRNA and then carry out RNAseq to examine differentially expressed pathways and genes. Doing this for all of the TRIM family of genes which have been implicated in OADC will allow similar pathways to be highlighted which should reveal further information about this disease.

Ultimately, if this gene or its family participates in the biology of this tumour, it may lead to the development of this/these as therapeutic targets for disease treatment. While this is currently a distant hope, confidence remains about the potential exploitation of the markers found in this study to date.

7.4: *OR3A4*:

The discovery of this hypomethylated gene in the progression of non-dysplastic BO that develops cancer has potential clinical use in the management of the BO patient. In the initial stage, it is necessary to externally validate this marker on a separate cohort of progressing and none progressing patients. Due to the significance shown in this study to date, but allowing for the small sample size used thus far, testing this signature on an external sample cohort will allow strengthened conclusions to be drawn and build the case for this marker to be used routinely in clinical practice.

Once re-validated in an external secondary retrospective cohort of patients, this marker will then require development in a prospective biomarker study. This will not be an easy or a quick study, as the time scale over which progression to cancer occurs is frequently long (and potentially in excess of 20-30 years). This will also be hampered by the relatively small number of patients who progress to cancer.

In order to increase the potential number of samples that could be included in this study, it may become necessary to expand the inclusion criteria to include those patients who progress to high grade dysplasia. As these patients have historically always been treated in the same way as patients with OADC (by resectional surgery) it is not unreasonable for these to be included also. Of course such patients are now more likely to undergo a mucosal resection than full oesophagectomy, but the methylation signatures of *OR3A4* need to be explored in this context. However, the inclusion of such patients would dramatically increase the “progressing” cohort.

Returning to the biology of OADC, little is known about *OR3A4*. What is known however is that this specific CpG site is within a long non coding RNA. It is

fortunate that a group exists within the University of Birmingham who investigates the role of these long non coding RNA's. We are now collaborating with this group to review what role, if any, this marker plays in cellular function.

There are several additional questions in OADC investigation and management, in which methylation patterns may yet play a role either in biomarker development or biology of the disease.

The most significant of these is the response to neoadjuvant chemotherapy. As started previously, a patient diagnosed with OADC, for whom curative intent treatment is advised would begin with neoadjuvant chemotherapy. However, it is well known that while many patients will have response and shrinking of their tumour, a percentage will have no response and may indeed progress while receiving therapy.

There is lack of any current method to be able to differentiate between patients who will and will not respond to neoadjuvant therapy, and therefore this is given to all patients as a result of proven outcome advantage. However, the effect seen could be even greater if targeted therapy was provided to those who had favourable tumours.

To this end, a panel of tumours has now been assembled which did and did not respond to chemotherapy, as determined by histology Mandard regression score (previously discussed). The intention is to conduct the same methylation comparison using the Illumina human methylation 450 array system as previously described. It is anticipated this experiment will commence in the near future, having now identified suitable tissue cohorts.

7.5: Thesis Summary

In summary, this thesis presents exciting methylation biomarker research in the field of oesophageal adenocarcinoma, its detection, biology and risk stratification using the most modern, robust array technology with subsequent marker validation. It is hopeful this work will provide the basis for significant further research in this field.

8. References:

1. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. *Lancet*. 2013;381(9864):400-12. Epub 2013/02/05.
2. Lepage C, Bouvier AM, Manfredi S, Coatmeur O, Cheynel N, Faivre J. Trends in incidence and management of esophageal adenocarcinoma in a well-defined population. *Gastroenterologie clinique et biologique*. 2005;29(12):1258-63. Epub 2006/03/07.
3. Lepage C, Rachet B, Jooste V, Faivre J, Coleman MP. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. *The American journal of gastroenterology*. 2008;103(11):2694-9. Epub 2008/10/16.
4. Jansson C, Johansson AL, Bergdahl IA, Dickman PW, Plato N, Adami J, et al. Occupational exposures and risk of esophageal and gastric cardia cancers among male Swedish construction workers. *Cancer causes & control : CCC*. 2005;16(6):755-64. Epub 2005/07/29.
5. Enzinger PC, Mayer RJ. Esophageal cancer. *The New England journal of medicine*. 2003;349(23):2241-52. Epub 2003/12/06.
6. Pennathur A, Farkas A, Krasinskas AM, Ferson PF, Gooding WE, Gibson MK, et al. Esophagectomy for T1 esophageal cancer: outcomes in 100 patients and implications for endoscopic therapy. *The Annals of thoracic surgery*. 2009;87(4):1048-54; discussion 54-5. Epub 2009/03/28.
7. Smith KJ, O'Brien SM, Green AC, Webb PM, Whiteman DC, Study of Digestive H. Current and past smoking significantly increase risk for Barrett's esophagus. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2009;7(8):840-8. Epub 2009/05/05.
8. Coleman HG, Bhat S, Johnston BT, McManus D, Gavin AT, Murray LJ. Tobacco smoking increases the risk of high-grade dysplasia and cancer among patients with Barrett's esophagus. *Gastroenterology*. 2012;142(2):233-40. Epub 2011/11/09.
9. Cook MB, Kamangar F, Whiteman DC, Freedman ND, Gammon MD, Bernstein L, et al. Cigarette smoking and adenocarcinomas of the esophagus and esophagogastric junction: a pooled analysis from the international BEACON consortium. *J Natl Cancer Inst*. 2010;102(17):1344-53. Epub 2010/08/19.
10. Anderson LA, Watson RG, Murphy SJ, Johnston BT, Comber H, Mc Guigan J, et al. Risk factors for Barrett's oesophagus and oesophageal adenocarcinoma: results from the FINBAR study. *World journal of gastroenterology : WJG*. 2007;13(10):1585-94. Epub 2007/04/28.
11. Whiteman DC, Sadeghi S, Pandeya N, Smithers BM, Gotley DC, Bain CJ, et al. Combined effects of obesity, acid reflux and smoking on the risk of adenocarcinomas of the oesophagus. *Gut*. 2008;57(2):173-80. Epub 2007/10/13.
12. O'Doherty MG, Freedman ND, Hollenbeck AR, Schatzkin A, Abnet CC. A prospective cohort study of obesity and risk of oesophageal and gastric adenocarcinoma in the NIH-AARP Diet and Health Study. *Gut*. 2012;61(9):1261-8. Epub 2011/12/17.
13. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008;371(9612):569-78. Epub 2008/02/19.
14. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *The New England journal of medicine*. 2003;348(17):1625-38. Epub 2003/04/25.
15. Pentcheva R, Fichthorn KA, Scheffler M, Bernhard T, Pfandzelter R, Winter H. Non-arrhenius behavior of the island density in metal heteroepitaxy: Co on Cu(001). *Physical review letters*. 2003;90(7):076101. Epub 2003/03/14.

16. White A NG, Foster K , Browne F , Carey S (eds). Health Survey for England. London HMSO. 1991.
17. Lagergren J, Bergstrom R, Lindgren A, Nyren O. Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *The New England journal of medicine*. 1999;340(11):825-31. Epub 1999/03/18.
18. Lagergren J. Influence of obesity on the risk of esophageal disorders. *Nat Rev Gastroenterol Hepatol*. 2011;8(6):340-7. Epub 2011/06/07.
19. Hoyo C, Cook MB, Kamangar F, Freedman ND, Whiteman DC, Bernstein L, et al. Body mass index in relation to oesophageal and oesophagogastric junction adenocarcinomas: a pooled analysis from the International BEACON Consortium. *International journal of epidemiology*. 2012;41(6):1706-18. Epub 2012/11/14.
20. Corley DA, Kerlikowske K, Verma R, Buffler P. Protective association of aspirin/NSAIDs and esophageal cancer: a systematic review and meta-analysis. *Gastroenterology*. 2003;124(1):47-56. Epub 2003/01/04.
21. Alexandre L, Broughton T, Loke Y, Beales IL. Meta-analysis: risk of esophageal adenocarcinoma with medications which relax the lower esophageal sphincter. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / ISDE*. 2012;25(6):535-44. Epub 2011/12/02.
22. Whiteman DC, Parmar P, Fahey P, Moore SP, Stark M, Zhao ZZ, et al. Association of *Helicobacter pylori* infection with reduced risk for esophageal cancer is independent of environmental and genetic modifiers. *Gastroenterology*. 2010;139(1):73-83; quiz e11-2. Epub 2010/04/20.
23. Pohl H, Wrobel K, Bojarski C, Voderholzer W, Sonnenberg A, Rosch T, et al. Risk factors in the development of esophageal adenocarcinoma. *The American journal of gastroenterology*. 2013;108(2):200-7. Epub 2012/12/19.
24. Leeuwenburgh I, Scholten P, Alderliesten J, Tilanus HW, Looman CW, Steijgerberg EW, et al. Long-term esophageal cancer risk in patients with primary achalasia: a prospective study. *The American journal of gastroenterology*. 2010;105(10):2144-9. Epub 2010/07/01.
25. Siewert JR, Feith M, Stein HJ. Biologic and clinical variations of adenocarcinoma at the esophago-gastric junction: relevance of a topographic-anatomic subclassification. *Journal of surgical oncology*. 2005;90(3):139-46; discussion 46. Epub 2005/05/17.
26. Kajiyama Y, Tsurumaru M, Udagawa H, Tsutsumi K, Kinoshita Y, Ueno M, et al. Prognostic factors in adenocarcinoma of the gastric cardia: pathologic stage analysis and multivariate regression analysis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 1997;15(5):2015-21. Epub 1997/05/01.
27. Daly JM, Fry WA, Little AG, Winchester DP, McKee RF, Stewart AK, et al. Esophageal cancer: results of an American College of Surgeons Patient Care Evaluation Study. *Journal of the American College of Surgeons*. 2000;190(5):562-72; discussion 72-3. Epub 2000/05/09.
28. O'Farrell NJ, Malik V, Donohoe CL, Johnston C, Muldoon C, Reynolds JV, et al. Appraisal of Staging Endoscopic Ultrasonography in a Modern High-Volume Esophageal Program. *World journal of surgery*. 2013. Epub 2013/04/10.
29. Antoch G, Saoudi N, Kuehl H, Dahmen G, Mueller SP, Beyer T, et al. Accuracy of whole-body dual-modality fluorine-18-2-fluoro-2-deoxy-D-glucose positron emission tomography and computed tomography (FDG-PET/CT) for tumor staging in solid tumors: comparison with CT and PET. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2004;22(21):4357-68. Epub 2004/10/30.
30. de Graaf GW, Ayantunde AA, Parsons SL, Duffy JP, Welch NT. The role of staging laparoscopy in oesophagogastric cancers. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*. 2007;33(8):988-92. Epub 2007/03/09.

31. Mehta SP, Jose P, Mirza A, Pritchard SA, Hayden JD, Grabsch HI. Comparison of the prognostic value of the 6th and 7th editions of the Union for International Cancer Control TNM staging system in patients with lower esophageal cancer undergoing neoadjuvant chemotherapy followed by surgery. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / ISDE*. 2013;26(2):182-8. Epub 2012/05/18.
32. NCCN Clinical Practice Guidelines in Oncology. Esophageal and Esophagogastric Junction Cancers (Excluding the proximal 5 cm of the stomach). 2013.
33. Opstelten JL, de Wijkerslooth LR, Leenders M, Bac DJ, Brink MA, Loffeld BC, et al. Variation in palliative care of esophageal cancer in clinical practice: factors associated with treatment decisions. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / ISDE*. 2016. Epub 2016/02/27.
34. Pennathur A, Luketich JD, Landreneau RJ, Ward J, Christie NA, Gibson MK, et al. Long-term results of a phase II trial of neoadjuvant chemotherapy followed by esophagectomy for locally advanced esophageal neoplasm. *The Annals of thoracic surgery*. 2008;85(6):1930-6; discussion 6-7. Epub 2008/05/24.
35. Medical Research Council Oesophageal Cancer Working G. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet*. 2002;359(9319):1727-33. Epub 2002/06/07.
36. Uttley L, Campbell F, Rhodes M, Cantrell A, Stegenga H, Lloyd-Jones M. Minimally invasive oesophagectomy versus open surgery: is there an advantage? *Surgical endoscopy*. 2013;27(3):724-31. Epub 2012/10/12.
37. Tapias LF, Morse CR. Minimally invasive Ivor Lewis esophagectomy: description of a learning curve. *Journal of the American College of Surgeons*. 2014;218(6):1130-40. Epub 2014/04/05.
38. NOGCA. National Oesophago-Gastric Cancer Audit. 2013 [28/03/2016]; Available from: <http://www.hscic.gov.uk/catalogue/PUB11093/clin-audi-supp-prog-oeso-gast-2013-rep.pdf>.
39. Mamidanna R, Bottle A, Aylin P, Faiz O, Hanna GB. Short-term outcomes following open versus minimally invasive esophagectomy for cancer in England: a population-based national study. *Annals of surgery*. 2012;255(2):197-203. Epub 2011/12/17.
40. Schmidt HM, Mohiuddin K, Bodnar AM, El Lakis M, Kaplan S, Irani S, et al. Multidisciplinary treatment of T1a adenocarcinoma in Barrett's esophagus: contemporary comparison of endoscopic and surgical treatment in physiologically fit patients. *Surgical endoscopy*. 2015. Epub 2015/11/07.
41. Bergeron EJ, Lin J, Chang AC, Orringer MB, Reddy RM. Endoscopic ultrasound is inadequate to determine which T1/T2 esophageal tumors are candidates for endoluminal therapies. *The Journal of thoracic and cardiovascular surgery*. 2014;147(2):765-71: Discussion 71-3. Epub 2013/12/10.
42. Hunt BM, Louie BE, Dunst CM, Lipham JC, Farivar AS, Sharata A, et al. Esophagectomy for failed endoscopic therapy in patients with high-grade dysplasia or intramucosal carcinoma. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / ISDE*. 2014;27(4):362-7. Epub 2013/06/26.
43. Barrett NR. The oesophagus lined by columnar epithelium. *Gastroenterologia*. 1956;86(3):183-6.
44. Sun D, Wang X, Gai Z, Song X, Jia X, Tian H. Bile acids but not acidic acids induce Barrett's esophagus. *International journal of clinical and experimental pathology*. 2015;8(2):1384-92. Epub 2015/05/15.
45. Phillips WA, Lord RV, Nancarrow DJ, Watson DI, Whiteman DC. Barrett's esophagus. *Journal of gastroenterology and hepatology*. 2011;26(4):639-48. Epub 2010/12/21.
46. Hvid-Jensen F, Pedersen L, Drewes AM, Sorensen HT, Funch-Jensen P. Incidence of adenocarcinoma among patients with Barrett's esophagus. *The New England journal of medicine*. 2011;365(15):1375-83. Epub 2011/10/15.

47. Bhat S, Coleman HG, Yousef F, Johnston BT, McManus DT, Gavin AT, et al. Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study. *Journal of the National Cancer Institute*. 2011;103(13):1049-57. Epub 2011/06/18.
48. Choi WT, Emond MJ, Rabinovitch PS, Ahn J, Upton MP, Westerhoff M. "Indefinite for Dysplasia" in Barrett's Esophagus: Inflammation and DNA Content Abnormality are Significant Predictors of Early Detection of Neoplasia. *Clinical and translational gastroenterology*. 2015;6:e81. Epub 2015/03/13.
49. Gatenby P, Ramus J, Caygill C, Shepherd N, Winslet M, Watson A. Routinely diagnosed low-grade dysplasia in Barrett's oesophagus: a population-based study of natural history. *Histopathology*. 2009;54(7):814-9. Epub 2009/07/29.
50. Taylor JB, Rubenstein JH. Meta-analyses of the effect of symptoms of gastroesophageal reflux on the risk of Barrett's esophagus. *The American journal of gastroenterology*. 2010;105(8):1729, 30-7; quiz 38. Epub 2010/05/21.
51. Eloubeidi MA, Provenzale D. Clinical and demographic predictors of Barrett's esophagus among patients with gastroesophageal reflux disease: a multivariable analysis in veterans. *Journal of clinical gastroenterology*. 2001;33(4):306-9. Epub 2001/10/06.
52. Smith KJ, O'Brien SM, Smithers BM, Gotley DC, Webb PM, Green AC, et al. Interactions among smoking, obesity, and symptoms of acid reflux in Barrett's esophagus. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2005;14(11 Pt 1):2481-6. Epub 2005/11/15.
53. Edelstein ZR, Bronner MP, Rosen SN, Vaughan TL. Risk factors for Barrett's esophagus among patients with gastroesophageal reflux disease: a community clinic-based case-control study. *The American journal of gastroenterology*. 2009;104(4):834-42. Epub 2009/03/26.
54. Gerson LB, Edson R, Lavori PW, Triadafilopoulos G. Use of a simple symptom questionnaire to predict Barrett's esophagus in patients with symptoms of gastroesophageal reflux. *The American journal of gastroenterology*. 2001;96(7):2005-12. Epub 2001/07/27.
55. Cook MB, Wild CP, Forman D. A systematic review and meta-analysis of the sex ratio for Barrett's esophagus, erosive reflux disease, and nonerosive reflux disease. *Am J Epidemiol*. 2005;162(11):1050-61. Epub 2005/10/14.
56. Corley DA, Kubo A, Levin TR, Block G, Habel L, Zhao W, et al. Abdominal obesity and body mass index as risk factors for Barrett's esophagus. *Gastroenterology*. 2007;133(1):34-41; quiz 311. Epub 2007/07/17.
57. Edelstein ZR, Farrow DC, Bronner MP, Rosen SN, Vaughan TL. Central adiposity and risk of Barrett's esophagus. *Gastroenterology*. 2007;133(2):403-11. Epub 2007/08/08.
58. Brown CS, Ujiki MB. Risk factors affecting the Barrett's metaplasia-dysplasia-neoplasia sequence. *World journal of gastrointestinal endoscopy*. 2015;7(5):438-45. Epub 2015/05/21.
59. Bhat SK, McManus DT, Coleman HG, Johnston BT, Cardwell CR, McMenamin U, et al. Oesophageal adenocarcinoma and prior diagnosis of Barrett's oesophagus: a population-based study. *Gut*. 2014. Epub 2014/04/05.
60. Jankowski J, Barr H. Improving surveillance for Barrett's oesophagus: AspECT and BOSS trials provide an evidence base. *Bmj*. 2006;332(7556):1512. Epub 2006/06/24.
61. Kadri SR, Lao-Sirieix P, O'Donovan M, DeBiram I, Das M, Blazeby JM, et al. Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: cohort study. *BMJ*. 2010;341:c4372. Epub 2010/09/14.
62. Benaglia T, Sharples LD, Fitzgerald RC, Lyratzopoulos G. Health benefits and cost effectiveness of endoscopic and nonendoscopic cytosponge screening for Barrett's esophagus. *Gastroenterology*. 2013;144(1):62-73 e6. Epub 2012/10/09.
63. Kadri S, Lao-Sirieix P, Fitzgerald RC. Developing a nonendoscopic screening test for Barrett's esophagus. *Biomarkers in medicine*. 2011;5(3):397-404. Epub 2011/06/11.

64. Terheggen G, Horn EM, Vieth M, Gabbert H, Enderle M, Neugebauer A, et al. A randomised trial of endoscopic submucosal dissection versus endoscopic mucosal resection for early Barrett's neoplasia. *Gut*. 2016. Epub 2016/01/24.
65. Subramanian CR, Triadafilopoulos G. Endoscopic treatments for dysplastic Barrett's esophagus: resection, ablation, what else? *World journal of surgery*. 2015;39(3):597-605. Epub 2014/05/21.
66. Liu Y, Chen C, Xu Z, Scuoppo C, Rillahan CD, Gao J, et al. Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. *Nature*. 2016;531(7595):471-5. Epub 2016/03/17.
67. Vignot S, Lefebvre C, Frampton GM, Meurice G, Yelensky R, Palmer G, et al. Comparative analysis of primary tumour and matched metastases in colorectal cancer patients: evaluation of concordance between genomic and transcriptional profiles. *Eur J Cancer*. 2015;51(7):791-9. Epub 2015/03/24.
68. Busuttill RA, Zapparoli GV, Haupt S, Fennell C, Wong SQ, Pang JM, et al. Role of p53 in the progression of gastric cancer. *Oncotarget*. 2014;5(23):12016-26. Epub 2014/11/28.
69. Park JY, Mitrou PN, Keen J, Dahm CC, Gay LJ, Luben RN, et al. Lifestyle factors and p53 mutation patterns in colorectal cancer patients in the EPIC-Norfolk study. *Mutagenesis*. 2010;25(4):351-8. Epub 2010/03/17.
70. Giudice A, D'Arena G, Crispo A, Tecce MF, Nocerino F, Grimaldi M, et al. Role of Viral miRNAs and Epigenetic Modifications in Epstein-Barr Virus-Associated Gastric Carcinogenesis. *Oxidative medicine and cellular longevity*. 2016;2016:6021934. Epub 2016/03/16.
71. Walsh L, Gallagher WM, O'Connor DP, Ni Chonghaile T. Diagnostic and Therapeutic Implications of Histone Epigenetic Modulators in Breast Cancer. *Expert review of molecular diagnostics*. 2016:1-11. Epub 2016/02/20.
72. Sheaffer KL, Elliott EN, Kaestner KH. DNA hypomethylation contributes to genomic instability and intestinal cancer initiation. *Cancer Prev Res (Phila)*. 2016. Epub 2016/02/18.
73. Bird A. Perceptions of epigenetics. *Nature*. 2007;447(7143):396-8. Epub 2007/05/25.
74. Planello AC, Singhanian R, Kron KJ, Bailey SD, Roulois D, Lupien M, et al. Pre-neoplastic epigenetic disruption of transcriptional enhancers in chronic inflammation. *Oncotarget*. 2016. Epub 2016/02/26.
75. Rokavec M, Oner MG, Hermeking H. Inflammation-induced epigenetic switches in cancer. *Cellular and molecular life sciences : CMLS*. 2016;73(1):23-39. Epub 2015/09/24.
76. Rothbart SB, Dickson BM, Strahl BD. From histones to ribosomes: a chromatin regulator tangoes with translation. *Cancer discovery*. 2015;5(3):228-30. Epub 2015/03/10.
77. Amaro A, Chiara S, Pfeffer U. Molecular evolution of colorectal cancer: from multistep carcinogenesis to the big bang. *Cancer metastasis reviews*. 2016. Epub 2016/03/08.
78. Koh YW, Chun SM, Park YS, Song JS, Lee GK, Khang SK, et al. Association between the CpG island methylator phenotype and its prognostic significance in primary pulmonary adenocarcinoma. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2016. Epub 2016/02/13.
79. Otsubo T, Hagiwara T, Tamura-Nakano M, Sezaki T, Miyake O, Hinohara C, et al. Aberrant DNA hypermethylation reduces the expression of the desmosome-related molecule periplakin in esophageal squamous cell carcinoma. *Cancer medicine*. 2015;4(3):415-25. Epub 2015/01/15.
80. Kaz AM, Grady WM, Stachler MD, Bass AJ. Genetic and Epigenetic Alterations in Barrett's Esophagus and Esophageal Adenocarcinoma. *Gastroenterology clinics of North America*. 2015;44(2):473-89. Epub 2015/05/30.
81. Yu M, O'Leary RM, Kaz AM, Morris SM, Carter KT, Chak A, et al. Methylated B3GAT2 and ZNF793 Are Potential Detection Biomarkers for Barrett's Esophagus. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2015;24(12):1890-7. Epub 2015/11/08.

82. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
83. Cook MB, Corley DA, Murray LJ, Liao LM, Kamangar F, Ye W, et al. Gastroesophageal reflux in relation to adenocarcinomas of the esophagus: a pooled analysis from the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON). *PloS one*. 2014;9(7):e103508. Epub 2014/07/31.
84. Wang F, Lv ZS, Fu YK. Nonsteroidal anti-inflammatory drugs and esophageal inflammation - Barrett's esophagus - adenocarcinoma sequence: a meta-analysis. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / ISDE*. 2010.
85. Sadeghi S, Bain CJ, Pandeya N, Webb PM, Green AC, Whiteman DC, et al. Aspirin, nonsteroidal anti-inflammatory drugs, and the risks of cancers of the esophagus. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2008;17(5):1169-78.
86. Shanmugam MK, Sethi G. Role of epigenetics in inflammation-associated diseases. *Sub-cellular biochemistry*. 2013;61:627-57. Epub 2012/11/15.
87. Caldas C, Hahn SA, da Costa LT, Redston MS, Schutte M, Seymour AB, et al. Frequent somatic mutations and homozygous deletions of the p16 (MTS1) gene in pancreatic adenocarcinoma. *Nature genetics*. 1994;8(1):27-32. Epub 1994/09/01.
88. Paulson TG, Galipeau PC, Xu L, Kissel HD, Li X, Blount PL, et al. p16 mutation spectrum in the premalignant condition Barrett's esophagus. *PloS one*. 2008;3(11):e3809.
89. Hardie LJ, Darnton SJ, Wallis YL, Chauhan A, Hainaut P, Wild CP, et al. p16 expression in Barrett's esophagus and esophageal adenocarcinoma: association with genetic and epigenetic alterations. *Cancer Lett*. 2005;217(2):221-30.
90. Bian YS, Osterheld MC, Fontollet C, Bosman FT, Benhattar J. p16 inactivation by methylation of the CDKN2A promoter occurs early during neoplastic progression in Barrett's esophagus. *Gastroenterology*. 2002;122(4):1113-21.
91. Moinova H, Leidner RS, Ravi L, Lutterbaugh J, Barnholtz-Sloan JS, Chen Y, et al. Aberrant vimentin methylation is characteristic of upper gastrointestinal pathologies. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2012;21(4):594-600. Epub 2012/02/09.
92. Lind GE, Ahmed D, Lothe RA. Vimentin in upper gastrointestinal pathologies--letter. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2012;21(10):1889; author reply 90.
93. Huang Y, Peters CJ, Fitzgerald RC, Gjerset RA. Progressive silencing of p14ARF in oesophageal adenocarcinoma. *Journal of cellular and molecular medicine*. 2009;13(2):398-409. Epub 2008/04/16.
94. Werling RW, Yaziji H, Bacchi CE, Gown AM. CDX2, a highly sensitive and specific marker of adenocarcinomas of intestinal origin: an immunohistochemical survey of 476 primary and metastatic carcinomas. *The American journal of surgical pathology*. 2003;27(3):303-10. Epub 2003/02/27.
95. Choi BJ, Kim CJ, Cho YG, Song JH, Kim SY, Nam SW, et al. Altered expression of CDX2 in colorectal cancers. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica*. 2006;114(1):50-4.
96. Ee HC, Erler T, Bhathal PS, Young GP, James RJ. Cdx-2 homeodomain protein expression in human and rat colorectal adenoma and carcinoma. *The American journal of pathology*. 1995;147(3):586-92.
97. Mallo GV, Rechreche H, Frigerio JM, Rocha D, Zweibaum A, Lacasa M, et al. Molecular cloning, sequencing and expression of the mRNA encoding human Cdx1 and Cdx2 homeobox. Down-regulation of Cdx1 and Cdx2 mRNA expression during colorectal carcinogenesis. *International journal of cancer Journal international du cancer*. 1997;74(1):35-44.

98. Vaninetti N, Williams L, Geldenhuys L, Porter GA, Guernsey DL, Casson AG. Regulation of CDX2 expression in esophageal adenocarcinoma. *Molecular carcinogenesis*. 2009;48(10):965-74. Epub 2009/05/06.
99. Makita K, Kitazawa R, Semba S, Fujiishi K, Nakagawa M, Haraguchi R, et al. Cdx2 expression and its promoter methylation during metaplasia-dysplasia-carcinoma sequence in Barrett's esophagus. *World journal of gastroenterology : WJG*. 2013;19(4):536-41. Epub 2013/02/06.
100. Tischoff I, Hengge UR, Vieth M, Ell C, Stolte M, Weber A, et al. Methylation of SOCS-3 and SOCS-1 in the carcinogenesis of Barrett's adenocarcinoma. *Gut*. 2007;56(8):1047-53. Epub 2007/03/23.
101. Baron R, Kneissel M. WNT signaling in bone homeostasis and disease: from human mutations to treatments. *Nature medicine*. 2013;19(2):179-92. Epub 2013/02/08.
102. Zou H, Molina JR, Harrington JJ, Osborn NK, Klatt KK, Romero Y, et al. Aberrant methylation of secreted frizzled-related protein genes in esophageal adenocarcinoma and Barrett's esophagus. *International journal of cancer Journal international du cancer*. 2005;116(4):584-91. Epub 2005/04/13.
103. Clement G, Braunschweig R, Pasquier N, Bosman FT, Benhattar J. Alterations of the Wnt signaling pathway during the neoplastic progression of Barrett's esophagus. *Oncogene*. 2006;25(21):3084-92. Epub 2006/01/13.
104. Clement G, Guilleret I, He B, Yagui-Beltran A, Lin YC, You L, et al. Epigenetic alteration of the Wnt inhibitory factor-1 promoter occurs early in the carcinogenesis of Barrett's esophagus. *Cancer Sci*. 2008;99(1):46-53. Epub 2007/11/17.
105. Kawakami K, Brabender J, Lord RV, Groshen S, Greenwald BD, Krasna MJ, et al. Hypermethylated APC DNA in plasma and prognosis of patients with esophageal adenocarcinoma. *Journal of the National Cancer Institute*. 2000;92(22):1805-11. Epub 2000/11/18.
106. Zhai R, Zhao Y, Su L, Cassidy L, Liu G, Christiani DC. Genome-wide DNA methylation profiling of cell-free serum DNA in esophageal adenocarcinoma and Barrett esophagus. *Neoplasia*. 2012;14(1):29-33. Epub 2012/02/23.
107. Smith E, De Young NJ, Pavey SJ, Hayward NK, Nancarrow DJ, Whiteman DC, et al. Similarity of aberrant DNA methylation in Barrett's esophagus and esophageal adenocarcinoma. *Molecular cancer*. 2008;7:75. Epub 2008/10/04.
108. Baumann S, Keller G, Puhlinger F, Napieralski R, Feith M, Langer R, et al. The prognostic impact of O6-Methylguanine-DNA Methyltransferase (MGMT) promoter hypermethylation in esophageal adenocarcinoma. *International journal of cancer Journal international du cancer*. 2006;119(2):264-8. Epub 2006/02/16.
109. Kaz AM, Wong CJ, Luo Y, Virgin JB, Washington MK, Willis JE, et al. DNA methylation profiling in Barrett's esophagus and esophageal adenocarcinoma reveals unique methylation signatures and molecular subclasses. *Epigenetics : official journal of the DNA Methylation Society*. 2011;6(12):1403-12. Epub 2011/12/06.
110. Wang JS, Guo M, Montgomery EA, Thompson RE, Cosby H, Hicks L, et al. DNA promoter hypermethylation of p16 and APC predicts neoplastic progression in Barrett's esophagus. *The American journal of gastroenterology*. 2009;104(9):2153-60. Epub 2009/07/09.
111. Kuester D, El-Rifai W, Peng D, Ruemmele P, Kroeckel I, Peters B, et al. Silencing of MGMT expression by promoter hypermethylation in the metaplasia-dysplasia-carcinoma sequence of Barrett's esophagus. *Cancer letters*. 2009;275(1):117-26. Epub 2008/11/26.
112. Kaz AM, Luo Y, Dzieciatkowski S, Chak A, Willis JE, Upton MP, et al. Aberrantly methylated PKP1 in the progression of Barrett's esophagus to esophageal adenocarcinoma. *Genes, chromosomes & cancer*. 2012;51(4):384-93. Epub 2011/12/16.
113. Clement G, Braunschweig R, Pasquier N, Bosman FT, Benhattar J. Methylation of APC, TIMP3, and TERT: a new predictive marker to distinguish Barrett's oesophagus patients at risk for malignant transformation. *The Journal of pathology*. 2006;208(1):100-7. Epub 2005/11/10.

114. Schulmann K, Sterian A, Berki A, Yin J, Sato F, Xu Y, et al. Inactivation of p16, RUNX3, and HPP1 occurs early in Barrett's-associated neoplastic progression and predicts progression risk. *Oncogene*. 2005;24(25):4138-48. Epub 2005/04/13.
115. Jin Z, Cheng Y, Gu W, Zheng Y, Sato F, Mori Y, et al. A multicenter, double-blinded validation study of methylation biomarkers for progression prediction in Barrett's esophagus. *Cancer research*. 2009;69(10):4112-5. Epub 2009/05/14.
116. Jin Z, Hamilton JP, Yang J, Mori Y, Olaru A, Sato F, et al. Hypermethylation of the AKAP12 promoter is a biomarker of Barrett's-associated esophageal neoplastic progression. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2008;17(1):111-7. Epub 2008/01/18.
117. Agarwal R, Jin Z, Yang J, Mori Y, Song JH, Kumar S, et al. Epigenomic program of Barrett's-associated neoplastic progression reveals possible involvement of insulin signaling pathways. *Endocrine-related cancer*. 2012;19(1):L5-9. Epub 2011/12/24.
118. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protocols*. 2007;2(9):2265-75.
119. Illumina. 450 array main page. Available from: http://www.illumina.com/products/methylation_450_beadchip_kits.html.
120. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98(4):288-95. Epub 2011/08/16.
121. Illumina FFPE restore manual. 2015; Available from: http://support.illumina.com/content/dam/illumina-support/documents/myillumina/5c3d90a3-793c-4a8b-932b-0434590f98ef/infinium_ffpe_sample_restore_booklet_15014614_c.pdf.
122. Illumina methylation 450 array manual. 2015; Available from: http://support.illumina.com/documents/myillumina/2173547c-5fc1-4aa8-8769-8ce4e49c8942/infinium_hd_ffpe_assay_guide_15021525_b.pdf
123. ChAMP package. 2015; Available from: <http://www.bioconductor.org/packages/release/bioc/html/ChAMP.html>.
124. Habets GG, van der Kammen RA, Stam JC, Michiels F, Collard JG. Sequence of the human invasion-inducing TIAM1 gene, its conservation in evolution and its expression in tumor cell lines of different tissue origin. *Oncogene*. 1995;10(7):1371-6. Epub 1995/04/06.
125. Duron O, Boureux A, Echaubard P, Berthomieu A, Berticat C, Fort P, et al. Variability and expression of ankyrin domain genes in Wolbachia variants infecting the mosquito *Culex pipiens*. *Journal of bacteriology*. 2007;189(12):4442-8. Epub 2007/04/24.
126. Bustelo XR, Sauzeau V, Berenjano IM. GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2007;29(4):356-70. Epub 2007/03/22.
127. Wu QY, Wang Y, Tong JC, Zhang M, Zhang K. Expression and clinical significance of Tiam1 gene in esophageal carcinoma. *International journal of clinical and experimental medicine*. 2015;8(11):21229-34. Epub 2016/02/18.
128. GM M. ENDCaP-C Test Accuracy Study. Available from: <http://public.ukcrn.org.uk/search/StudyDetail.aspx?StudyID=17739>.
129. Uchil PD, Hinz A, Siegel S, Coenen-Stass A, Pertel T, Luban J, et al. TRIM protein-mediated regulation of inflammatory and innate immune signaling and its association with antiretroviral activity. *Journal of virology*. 2013;87(1):257-72. Epub 2012/10/19.
130. Kawai T, Akira S. Regulation of innate immune signalling pathways by the tripartite motif (TRIM) family proteins. *EMBO molecular medicine*. 2011;3(9):513-27. Epub 2011/08/10.
131. Nilsson E, Lendahl U. Transient expression of a human beta-actin promoter/lacZ gene introduced into mouse embryos correlates with a low degree of methylation. *Molecular reproduction and development*. 1993;34(2):149-57. Epub 1993/02/01.

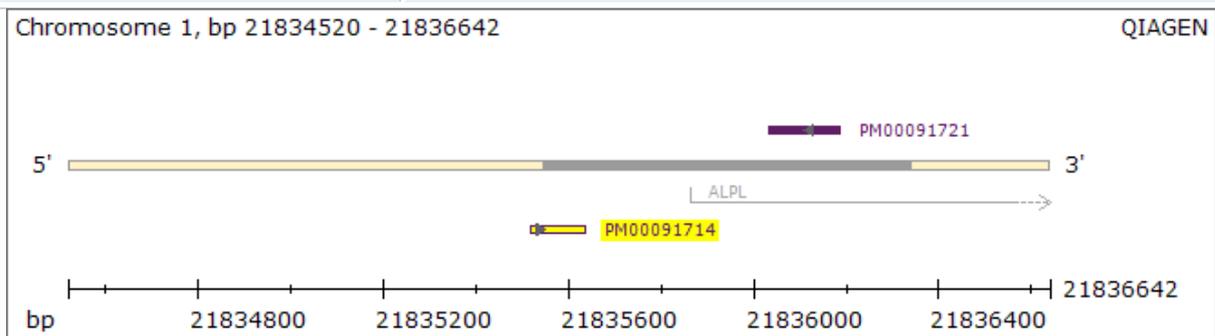
132. Peters CJ, Rees JR, Hardwick RH, Hardwick JS, Vowler SL, Ong CA, et al. A 4-gene signature predicts survival of patients with resected adenocarcinoma of the esophagus, junction, and gastric cardia. *Gastroenterology*. 2010;139(6):1995-2004 e15. Epub 2010/07/14.
133. Ong CA, Shapiro J, Nason KS, Davison JM, Liu X, Ross-Innes C, et al. Three-gene immunohistochemical panel adds to clinical staging algorithms to predict prognosis for patients with esophageal adenocarcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2013;31(12):1576-82. Epub 2013/03/20.
134. Ong CA, Shannon NB, Ross-Innes CS, O'Donovan M, Rueda OM, Hu DE, et al. Amplification of TRIM44: pairing a prognostic target with potential therapeutic strategy. *Journal of the National Cancer Institute*. 2014;106(5). Epub 2014/04/30.
135. Spechler SJ, Goyal RK. The columnar-lined esophagus, intestinal metaplasia, and Norman Barrett. *Gastroenterology*. 1996;110(2):614-21. Epub 1996/02/01.
136. McDougall NI, Johnston BT, Collins JS, McFarland RJ, Love AH. Three- to 4.5-year prospective study of prognostic indicators in gastro-oesophageal reflux disease. *Scandinavian journal of gastroenterology*. 1998;33(10):1016-22. Epub 1998/11/26.
137. Fitzgerald RC, di Pietro M, Ragnauth K, Ang Y, Kang JY, Watson P, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut*. 2014;63(1):7-42. Epub 2013/10/30.
138. Gatenby P, Soon Y. Barrett's oesophagus: Evidence from the current meta-analyses. *World journal of gastrointestinal pathophysiology*. 2014;5(3):178-87. Epub 2014/08/19.
139. Dulak AM, Schumacher SE, van Lieshout J, Imamura Y, Fox C, Shim B, et al. Gastrointestinal adenocarcinomas of the esophagus, stomach, and colon exhibit distinct patterns of genome instability and oncogenesis. *Cancer research*. 2012;72(17):4383-93. Epub 2012/07/04.
140. Paterson AL, O'Donovan M, Provenzano E, Murray LJ, Coleman HG, Johnson BT, et al. Characterization of the timing and prevalence of receptor tyrosine kinase expression changes in oesophageal carcinogenesis. *The Journal of pathology*. 2013;230(1):118-28. Epub 2012/06/27.
141. Katoh M, Nakagama H. FGF receptors: cancer biology and therapeutics. *Medicinal research reviews*. 2014;34(2):280-300. Epub 2013/05/23.

9. Appendix:

9.1 Gene/primer information

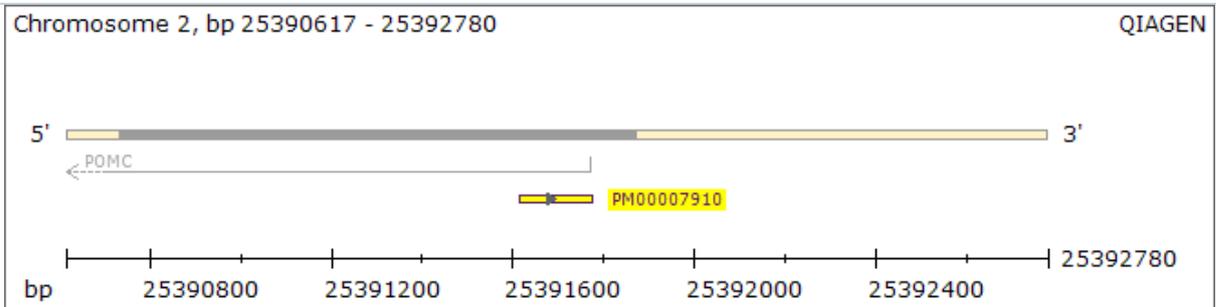
ALPL

| | |
|--|--|
| Amplicon length | 122 bp |
| Sequenced strand | sense |
| Biotin modification on ² | reverse PCR primer |
| PyroMark Assay setup file ³ | Download |
| Chromosomal location ⁴ | Chromosome 1, BP 218355XX-218355XX |
| Sequence to Analyse | GGAAGATGCCGACGGTAAAGACAAAACAGGAGACGCGCGC |



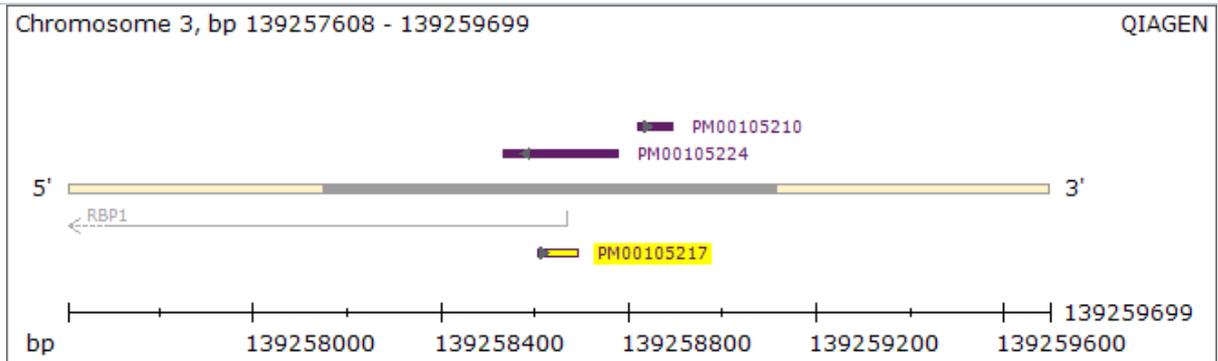
POMC

| | |
|--|------------------------------------|
| Amplicon length | 163 bp |
| Sequenced strand | sense |
| Biotin modification on ² | reverse PCR primer |
| PyroMark Assay setup file ³ | Download |
| Chromosomal location ⁴ | Chromosome 2, BP 253916XX-253917XX |
| Sequence to Analyze | GCGTTGGGTCTCCGCTTAGAACGGGCGGGA |



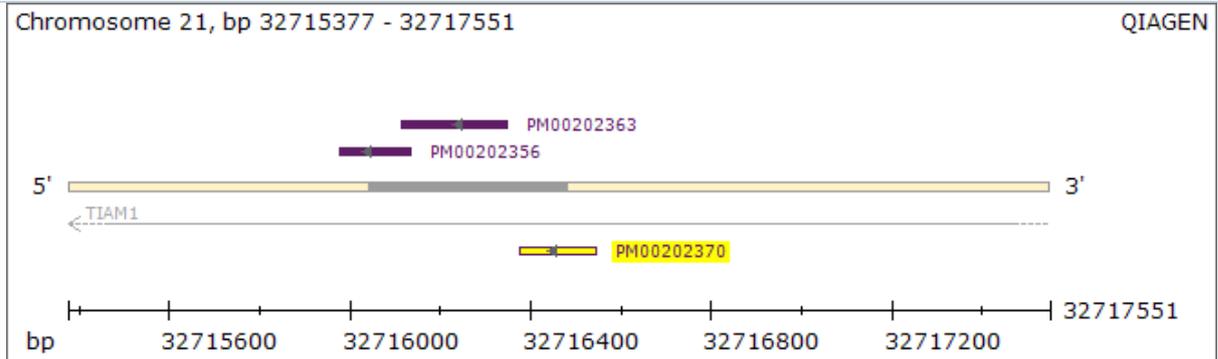
RBP1

| | |
|--|--------------------------------------|
| Amplicon length | 91 bp |
| Sequenced strand | sense |
| Biotin modification on ² | reverse PCR primer |
| PyroMark Assay setup file ³ | Download |
| Chromosomal location ⁴ | Chromosome 3, BP 1392586XX-1392586XX |
| Sequence to Analyze | CCGCGCCACCTTCGTTTCAGACGTTTCAGTTCGT |



TIAM1

| | |
|--|---|
| Amplicon length | 174 bp |
| Sequenced strand | antisense |
| Biotin modification on ² | reverse PCR primer |
| PyroMark Assay setup file ³ | Download |
| Chromosomal location ⁴ | Chromosome 21, BP 327164XX-327164XX |
| Sequence to Analyze | AC G TG C G C CTTGTGAGCAGSGGTG |



VCAN

| | |
|--|--|
| Amplicon length | 157 bp |
| Sequenced strand | sense |
| Biotin modification on ² | reverse PCR primer |
| PyroMark Assay setup file ³ | Download |
| Chromosomal location ⁴ | Chromosome 5, BP 827685XX-827685XX |
| Sequence to Analyze | TCACGGTTCTCTCCCGGC GCCCCTGGGTTCTTTTGTCTGCACGC GGT |

