Quantitative, genome-wide analysis of the DNA methylome in sporadic pituitary adenomas

Cuong V. Duong¹, Richard D. Emes², Frank Wessely², Kiren Yac qb-Usman¹, Richard N. Clayton¹ and William E. Farrell¹,³

Short title: Analysis of the pituitary adenoma methylome

Key words: pituitary, adenoma, genome-wide, epigenetic

¹Institute of Science and Technology in Medicine, Keele University School of Medicine, Stoke-on-Trent, Staffordshire, ST4 7QB. UK.

²School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, College Road, Sutton Bonington, Leicestershire, LE12 5RD. UK

³Address correspondence to Professor William E. Farrell at the Institute of Science and Technology in Medicine, Keele University School of Medicine, Stoke-on-Trent ST4 7QB. Telephone +44 (0)1782 555225, e-mail: w.e.farrell@keele.ac.uk

Conflict of Interest:
All authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported

Funding: This work was partially supported by the award of a research scholarship by the Vietnamese Government to Cuong V Doung.
ABSTRACT

DNA methylation is one of several epigenetic modifications that together with genetic aberrations are hallmarks of tumourigenesis including those emanating from the pituitary gland. In this study we examined DNA methylation across 27578 CpG sites spanning more than 14000 genes in the major pituitary adenoma subtypes. Genome-wide changes were first determined in a discovery cohort comprising, non-functioning (NF), growth hormone (GH), prolactin secreting (PRL) and corticotroph (CT) adenoma relative to post-mortem pituitaries. Using stringent cut-off criteria we validated increased methylation by pyrosequencing in 12 of 16 (75%) genes. Overall these criteria identified 40 genes in NF, 21 in GH, 6 in PRL and 2 in CT that were differentially methylated relative to controls. In a larger independent cohort of adenomas, for genes where hypermethylation had been validated, different frequencies of hypermethylation were apparent; where the KIAA1822 and TFAP2E genes were hypermethylated in 12 of 13 NF adenomas whereas the COLIA2 gene showed increase in 2 of 13 adenomas. For genes showing differential methylation across and between adenoma subtypes pyrosequencing confirmed these findings. For 3 of 12 genes investigated, an inverse relationship between methylation and transcript expression was observed where increased methylation of, EML2, RHOD and HOXB1 is associated with significantly reduced transcript expression. This study provides the first, genome-wide survey of adenoma, subtype-specific epigenomic changes and will prove useful for identification of biomarkers that perhaps predict or characterise growth patterns. The functional characterisation of identified genes will also provide insight of tumour aetiology and identification of new therapeutic targets.
INTRODUCTION

The genesis and outgrowth of sporadic pituitary adenomas is, in common with most other tumour types, characterised by inappropriate expression of hormone and growth factor receptors, mediators in their associated signal transduction pathways, transcription factors and cell cycle regulators. (Melmed. 2003, Dudley et al. 2009) The progenitor cells that give rise to each of the differentiated cell types within this gland are also the precursor population of their cognate subtype-specific adenomas. (Melmed. 2003, Melmed. 2011)

Although important exceptions exist, genetic aberrations leading to inappropriate expression, activation or silencing of regulatory genes are infrequent in pituitary adenomas. (Vandeva et al. 2010) However, in these tumours, multiple studies have described epigenetic change within gene associated CpG islands and/or those that lead to modification of histone tails. (Yacqub-Usman et al. 2012b, Ezzat. 2008) In these cases inappropriate methylation of CpG islands is frequently associated with gene silencing whereas particular combinatorial modifications to histone tails are associated with either silenced or expressed genes. (Dudley et al. 2009, Yacqub-Usman et al. 2012b)

In common with specific genetic aberrations, that frequently show subtype specificity, epigenetic modification termed epimutations also show a degree of subtype specificity. (Zhao et al. 2005, Zhang et al. 2002, Simpson et al. 2000) As example, subtype-specific genetic changes are apparent for the gsp oncogene and are also apparent for inappropriate expression of HMGA2 in somatotroph and lactotroph derived adenomas, respectively. (Fedele et al. 2006, Hayward et al. 2001, Farrell and Clayton. 2000) For specific epimutations, methylation mediated silencing of the MEG3 gene is a frequent finding in pituitary adenomas of gonadotroph origin whereas methylation mediated silencing of the p16 gene (CDKN2A) is infrequent in somatotrophinomas but is a common finding in most other pituitary adenoma subtypes. (Zhao et al. 2005, Simpson et al. 1999) Equally, for epigenetic modifications that are manifest as histone tail modification a degree of subtype specificity is also apparent. In this case, and in a recent report, we showed specific histone modification to the BMP-4 gene, that are indicative of either gene silencing or modifications that are
permissive for expression and furthermore that these modification show adenoma subtype specificity. (Yacqub-Usman et al. 2012a)

For epimutations that characterise pituitary adenomas their identification, in most cases, has been through candidate gene approaches. (Dudley et al. 2009) Less frequently, for identification of novel genes, investigators have used differential display techniques. These approaches have employed either, tumour derived cDNA or DNA that has been subject to prior methylation sensitive digestion respectively. (Zhao et al. 2005, Bahar et al. 2004) Although these techniques have identified novel genes a limitation is that identification is frequently confined to a single transcript or gene, that is, on a gene-by-gene basis. Relatively few studies have adopted genome-wide investigation. In cases where these approaches have been used they have relied upon siRNA knock-down or pharmacological unmasking strategies. (Dudley et al. 2008, Al-Azzawi et al. 2011). However, a current limitation of these techniques is that they are reliant upon actively dividing cells for the effective reversal and or erasure of epimutations. Moreover, primary pituitary adenomas show limited proliferative potential in vitro and investigations are, therefore, reliant upon pituitary adenoma cell lines of mouse or rat origin. While these approaches have uncovered novel genes they are constrained by the issue of extrapolation across species boundaries. (Dudley et al. 2008)

Recent technological advances in high throughput array-based techniques now provide opportunity for the comprehensive genome-wide investigation of DNA methylation in normal and disease states. (Wilop et al. 2011, Irizarry et al. 2008, Fryer et al. 2011) For pituitary adenomas these high throughput techniques have the potential to not only increase our understanding of the molecular biology of this tumour type but also to identify biomarker for diagnosis, prognosis and for therapeutic intervention. Therefore, and as a first step toward these outcomes we analysed the methylation profile of 27,578 CpG sites spanning more than 14,000 genes in each of the major pituitary adenoma subtypes relative to post-mortem normal pituitary.
MATERIALS AND METHODS

Human tissue sample:

The primary sporadic human pituitary tumours used in this study comprised a discovery and an independent cohort. The discovery cohort comprised three each of the major adenoma subtypes. Tumours in the discovery and independent cohort were classified and graded as we described previously. (Yacqub-Usman et al. 2012a) The investigation cohort comprised, 7 growth hormone (GH) secreting tumours all of which were grade 2 macroadenomas; 6 corticotrophinomas (CT) four of which were grade 1 microadenomas and 2 were grade 2 macroadenomas; 6 prolactinomas (PRL) all of which were grade 1 microadenomas and 13 non-functioning adenomas all of which were grade 3 macroadenomas (NF). Adenoma subtype classifications was on the basis of staining for mature hormone (GH, ACTH, FSH, LH and prolactin but not for alpha subunit) as previously described. (Yacqub-Usman et al. 2012a). Since tumours were not stained for alpha subunit it was not possible to detect subunit producing gonadotrophinomas or null cell adenomas, however, none stained for mature LH or FSH and were therefore classified as non-functioning adenomas. As controls we used four post-mortem normal pituitaries, acquired within 12hr of deaths with no evidence of any endocrine disease. Primary human tissues were stored at -80°C prior to their use. Tumour tissues were obtained with informed consent and all of studies performed with Regional Ethics Committee (South Birmingham Committee: REC reference number: 10/H1207/406) and institutional approval.

Only those adenomas where tumour cells comprised ≥80% of the specimen, as determined at surgery and confirmed by neuropathological assessment, were used in the study. Prior to the described extraction protocols tumours and normal pituitaries were freeze fractured using a biopulveriser (Biospec, Bartesville, USA) to achieve a homogenous mixture of cells.

DNA extraction and Bisulphite modification

High molecular weight DNA was extracted from adenoma tissue and post-mortem pituitaries (pmPits) using a standard phenol-chloroform extraction procedure. DNA was quantified using a NanoDrop 1000® (Thermo Scientific, Wilmington, USA) and diluted in molecular biology grade water (Sigma). Bisulphite
modification of 1µg of DNA was performed using either the EpiTect® Plus DNA bisulphite kit (Qiagen, UK) or Zymo EZ DNA Methylation Gold kit (Zymo Research) as previously described. (Al-Azzawi et al. 2011, Fryer et al. 2011) In some cases, where we used whole genome amplification (WGA) of bisulphite converted DNA this was performed as described by Mill et al. (Mill et al. 2006)

**Illumina BeadArray analysis**

For the discovery cohort DNA was extracted from three of each of the major pituitary adenoma subtypes (NF, GH, PRL, CT) and from three pmPits. 500 ng of total DNA from each of the samples was sodium bisulphite converted as described above. Post conversion DNA was eluted in 12µL of elution buffer and 4µL of converted DNA was used as template on the Infinium Methylation 27K Arrays (Illumina, San Diego, CA) and was processed as per the manufacturers instruction as previously described. (Fryer et al. 2011) This array examines the methylation status of 27,578 CpG sites across 14,496 genes. Data was collected using the Illumina Bead Array (Illumina, San Diego, CA) reader and analysed with GenomeStudio V2009.1 methylation module 1.1.1 (Illumina, San Diego, CA). This assigns a ‘beta value’ which is a qualitative measure of methylation for each CpG site and ranges from 0 (no methylation) to 1.0 (100% methylation of both alleles).

Of the 27,578 sites on the array, and as we described previously we removed all probes targeted on X and Y chromosomes (n=1092) and this resulted in 26,486 probes. (Fryer et al. 2011) To reduce the number of non-variable sites from subsequent analyses, probes where beta values in all samples were ≥0.8 or ≤0.2 and as described previously by us and other groups (Byun et al. 2009, Fryer et al. 2011) were excluded (n=13,945) from the analysis and this reduced the probe count to 12,541. From this dataset we also eliminated all sites where one or more of the samples demonstrated (i) detection P values of >0.05 (internal quality control) or (ii) null (missing) beta values. We took the decision to remove those probes which failed in any one of the samples rather than excluding probe values for individual tumour samples as we deemed this a more robust way to identify probes which varied between tumours in a limited number of samples. This resulted in a final data set of 10,667 probes for further analysis.

**Technical validation of Array generated data**

For the technical validation of the BeadArray, 4 single CpGs (cg08047457, cg20289949, cg06197492, and cg24739326) were selected that had been interrogated on the array and that demonstrated a
range of beta values across the tumour samples. The tumour samples, post sodium bisulphite conversion, were subject to pyrosequence analysis (as described in a subsequent section) Primer sequences are provided in supplemental data (Table S1).

**DNA pyrosequencing analysis of Sodium bisulphite converted DNA**

The pyrosequencing experiments were first performed on the discovery adenoma cohort for the validation of the BeadArray generated data. Subsequent experiments were performed on the larger investigation cohort of adenoma. The dataset initially used in this study comprised genes that showed an increase in $\beta$ value of $\geq 0.4$ in two or more adenomas within a subtype (as described in the results section) relative to the mean $\beta$ values of the pmPits. From the BeadArray dataset we first selected 22 genes at random (Table 1). For each of these genes their promoter region associated CpG islands were identified and downloaded from the UCSC Genome browser (http://genome.ucsc.edu/) and imported into PyroMark Assay Design 2.0 software for primer design of sodium bisulphite converted DNA. Dependent upon the specific gene and the density of CpGs within their promoter region the amplicons encompassed between 5 and 11 CpGs (Table S1, supplementary data).

In these cases, 2 $\mu$L of sodium bisulphite converted and WGA DNA was used as template in the first round of a nested PCR reaction. The product of the first round was diluted 100 fold in dH$_2$O and 2 $\mu$L used as template in a second round nested PCR reaction using the same cycling condition. After initial denaturation at 98°C for 10 minutes, we employed touch-down cycling for 14 cycles in which each cycle was touched down 0.5°C. The subsequent 35 amplification cycles includes denaturing at 98°C for 30 seconds, annealing at 55°C for 30 seconds, and elongation at 72°C for 30 seconds. Additional elongation was set at 72°C for 10 minutes.

The capture of the single stranded and biotin-labelled DNA was carried out using a Pyromark Q96 ‘Vacuum Prep’ workstation as recommended by the manufacturer and Pyrosequencing was performed using a PSQ96A Pyrosequencer using Pyro Q CpG software (version 1.0.9; Qiagen). For each set of independent bisulphite modification reactions, a control DNA sample was run to ensure efficient sodium bisulphite-mediated conversion. The Pyrosequencing data from the control DNA sample was used to standardise levels across experimental runs.
Quantitative RT-PCR:

Total RNA was extracted from primary human tumour tissue and pmPits, as previously described. (Bahar et al. 2004) cDNA was synthesised using 200U of M-MLV reverse transcriptase (Promega, Southampton, UK) according to the manufactures instructions and as previously described. (Bahar et al. 2004) The primers sequences used for real time quantification are shown Table S1 (supplementary data) using conditions described previously. (Al-Azzawi et al. 2011) The target genes were normalised to an endogenous control and relative quantification was carried out using relative-standard curve and \(2^{-\Delta\Delta CT}\) methods, where, \(-\Delta\Delta CT = CT_{\text{Gene of interest of tumour}} - GAPDH_{\text{of tumour}} - CT_{\text{Gene of interest of normal pituitary}} - GAPDH_{\text{of normal pituitary}}\). Loss or significantly reduced transcript expression in individual adenomas was assigned using the following criteria: We first determined the mean and SD of expression in four normal pituitaries. We next set limits of four times that SD, where expression, in individual adenomas at or less than the lower limit was deemed loss or significantly reduced expression.

RESULTS

Methylation Profiling of Sporadic Pituitary Adenoma. To evaluate whether sporadic human pituitary tumours display subtype-specific methylation profiles we used Illumina Infinium BeadArray technology to determine the methylation status of 27,578 CpG sites that span more than 14,000 genes. The discovery cohort comprised adenomas that represented each of the major pituitary adenoma subtypes, Non-functioning (NF), Somatotrophinoma (GH), Prolactinoma (PR) and Corticotrophinoma (CT). Quantitative differential methylation was assessed relative to post-mortem normal pituitaries.

Prior to determining differential methylation we first excluded probes in chromosome X and Y chromosomes and low quality probes (detection probability > 0.05). In addition, we removed individual probes where the \(\beta\) values were either \(\leq 0.2\) or \(\geq 0.8\) across the 15 samples as we and others have previously described. (Byun et al. 2009, Fryer et al. 2011) Through application of these criteria a total of 10667 probes
spanning 7956 genes were included in the analysis. Within this probe set, 5287 CpG were within and 5280 were outside of CpG islands. Analysis of the CpG methylation profiles across individual CpG sites showed these to be different between adenoma subtypes, however, individual adenomas, within a subtype, showed similar methylation profiles to each other (data not shown). A Kolmogorov–Smirnov test was used to compare the distribution of beta values between samples using the NIMBL software (Wessely and Emes, 2012). A single pmPit sample was identified as having a significantly different ($p < 0.05$) distribution to the other samples and was excluded from further analysis. The remaining pmPits showed different methylation profiles to those apparent in each of the adenoma subtypes (data not shown).

To characterise the differential methylation profiles of the adenoma subtypes we first performed a class comparison between each of the adenoma subtypes and relative to that apparent in the pmPits. For the comparison we initially filtered for a minimum increase of $\Delta\beta$ of 0.4 relative to pmPits at single gene-specific CpG sites in at least two of three subtype specific adenomas. This analysis identified 326 genes in the NF adenomas, 49 in PRL, 97 in GH and 4 in CT that were hypermethylated relative to pmPits supplemental data (Table S2. In the majority of cases and irrespective of subtype the CpGs were within CpG islands.

To validate the bead array generated data we performed pyrosequence analysis of 48 CpG sites, in four gene-associated CpG islands from the 12 samples (discovery cohort) that had been interrogated on the array. The analysis (Fig. 1) for the quantitative measurement of CpG methylation, showed the two techniques to be significantly correlated (Spearman r: $0.85 \ p = < 0.0001$). Importantly, the correlation involves samples across the range of beta values and hence is unlikely to be an artefact due to dominance of extreme values. (Roessler et al. 2012) In addition, inspection of the 20 imprinted genes interrogated on the array showed a mean $\beta$ value of $0.51 \pm 0.1$ across a total of 40 CpG sites. These $\beta$ values are consistent with monoallelic methylation apparent in most normal tissues including pituitaries and providing further confidence in the methodological approach.

CpG Island Methylation Profiles in Primary Adenoma Subtypes: To determine the relationship between methylation at array-interrogated CpG sites and methylation across the gene-associated CpG islands
we used pyrosequence analysis that encompassed, in each case, between 5 and 11 CpGs. We initially selected 22 genes on the basis that they each showed a $\Delta \beta$ of $\geq 0.4$ (in at least two of the three NF adenomas) relative to pmPits, and where the array-interrogated CpG are within promoter-associated CpG islands (Table 1). The analysis was initially performed in the discovery cohort of adenomas and relative to four pmPits.

Mean methylation across these CpGs, as determined by pyrosequencing, was able to confirm the bead-array methylation status for only 12 of the 22 genes (55%). However, we noted that where concordance between the techniques is observed (12 of 22 genes) it was for genes that were represented by more than one island-associated CpG on the array, and where the $\Delta \beta$ for the second CpG also showed increase. For the 22 genes that we had selected for pyrosequence analysis, on the basis of a single CpG site, 16 genes also showed increase in a second CpG site ($\beta \geq 0.25$), in at least two of the three NF adenomas that had also been interrogated on the array (Table 1). Of these 16 array-identified genes pyrosequence analysis confirmed increase in mean methylation (across 5 to 11 CpG sites) for 12 of 16 (75%) genes in the discovery cohort of NF adenomas.

On the basis of the finding described above we modified our filter criteria to include genes where one of two CpGs showed a $\Delta \beta$ of $\geq 0.4$ and a second CpG a $\Delta \beta$ of $\geq 0.25$ and where these criteria were fulfilled in at least two of the three adenomas (within a subtype) interrogated on the array. The revised, more stringent, filtering criteria identified 40 genes in NF adenomas, 21 in GH, 6 in PRL and 2 in CT that showed increase in methylation relative pmPit (Table 2). Table 2 also shows genes that are methylated in common across adenoma subtypes and those that segregate with one or more pituitary adenoma subtypes. In addition, and for two of the adenoma subtypes, NF and GH, where these criterion were met in multiple genes we generated a heatmap displaying $\beta$ values across individual adenomas within these subtypes. (Figure 2)

**Gene Specific Methylation Frequencies in Non-Functioning Adenomas:** Through constraints imposed by methodologies (sodium bisulphite conversion, pyrosequencing and transcript expression analyses) the availability of some adenoma subtypes (number and amount of tumour tissue) with the exception of NF adenomas, was limited. We, therefore, focused this part of the study primarily, but not exclusively, on the NF adenoma cohort. Through analysis of the 12 genes identified by bead-array and confirmed by pyrosequencing in the discovery cohort we determine the frequency of aberrant methylation
across an independent cohort of 13 NF adenomas relative to 4 pmPits. Figure 3 shows the analysis where each gene in the pmPits, with the exception of HOXB1, shows low but variable levels of methylation. We set a stringent cut-off criterion for increased methylation in individual adenomas of ≥ 4 SD higher than the standard deviation of the mean (for each of the genes) that was apparent in the four pmPits. The proportion of adenomas showing increase, above each of the gene-specific thresholds, is shown in Figure 3 and is summarised in Table 3. Within the NF cohort the number of adenomas that show gene specific increase in methylation is variable, where the CpG island methylation of the KIAA1822 and TFAP2E hypermethylated is found in 12 of 13 NF adenomas whereas the COL1A2 CpG island shows increased methylated in only 2 of 13 adenomas.

**Subtype Specific Gene Methylation Frequencies:** Of the 12 genes subject to validation in the NF adenomas 6 of these genes also fulfilled the Δβ criterion in two or more of the other adenoma subtypes (Table 2). To gain insight into the subtype specificity of hypermethylation we performed pyrosequence analysis for three of these genes where, in addition to the NF adenomas, we included each of the other subtypes (Fig. 4). For EML2, and confirming the bead-array finding, methylation is confined to NF adenomas. Cross subtype specific methylation is apparent for the RHOD and KIAA1822 genes where the array (Table 2) and pyrosequencing shows increased methylation in a proportion of the NF and GH secreting adenomas and in the NF, GH and PRL secreting adenomas respectively (Figure 4).

**Methylation Associated Gene Silencing:** For the 12 genes with pyrosequencing confirmed methylation status, we used RT-qPCR to determine association between CpG island methylation and gene silencing. In these cases loss or significantly reduced transcript expression was assigned on the basis of the stringent criterion described for gene specific methylation. However, in these cases it was ≤ 4 SD of the standard deviation on the mean of the pmPits. In the 13 NF adenomas, an inverse relationship between increased methylation and reciprocal significant decrease in transcript expression is apparent for 3 of the 12 genes examined. Figure 5 shows that relative to pmPit all of the methylated tumour (filled circles) show significantly reduced transcript expression. Conversely, in adenomas that were not hypermethylated (unfilled circles) expression levels are similar to or in some cases higher than that apparent in the pmPits. A single exception to these findings is apparent for the HOXB1 gene, where a single adenoma that was not methylated
also shows reduced transcript expression. In addition, since multiple studies using candidate gene approaches have shown frequent methylation associated loss of CDKN2A (Woloschak et al. 1997, Simpson et al. 1999, Vandeva et al. 2010) we included RT-qPCR analysis of this transcript. These studies show significantly reduced expression of this transcript in 12 of 13 adenomas relative to pmPits (data not shown) further confirming and reinforcing the robustness of our approach.

Gene Ontology Analysis of Hypermethylated genes: Gene ontology analysis of the hypermethylated genes was performed to determine the biological processes where these genes are known to contribute together with their molecular function and finally their associated KEGG pathways. These analyses are shown in Supplemental Table S3-5 respectively. The aberrantly methylated, promoter-associated CpG islands were distributed across various categories of biological processes, molecular function and pathways. However, multiple genes associated with particular biological processes showed frequent methylation in the analysis. In this context, we noted that 11 genes in intracellular signalling cascades (Supplemental Table S4) were hypermethylated in pituitary adenomas. Interestingly, among these genes, the majority (SOCS2, RAC2, ERBB2, RASSF1, SOCS1, COL1A2) have also been shown to exhibit tumour suppressor properties in multiple other tumour types.(Sasi et al. 2010, Mizukawa et al. 2011, Walter et al. 2012, Qian et al. 2005, Mottok et al. 2007, Bonazzi et al. 2011)

DISCUSSION

The study of the epigenome in normal and disease states, and in particular as it relates to tumorigenesis is an area of growing interest and research focus. In this report, the role of changes in DNA methylation at CpG dinucleotides has been analysed in each of the major pituitary adenoma subtypes. Through use of an Infinium BeadArray approach we performed a genome-wide analysis, thus providing a quantitative measure of DNA methylation at CpG sites across the genome. The analysis revealed aberrant methylation in each of the adenoma subtypes. Furthermore, it showed that hypermethylation, at CpG sites within gene-associated CpG islands is more frequent compared to CpG sites outside of CpG islands and similar findings in other tumour types have been reported elsewhere. (Kim et al. 2011) The epigenetic changes were more frequent in
non-functioning adenomas than other pituitary adenoma subtypes; however, we recognise that the absolute 
numbers of sites identified as hypermethylated is dependent on the cut-off criteria employed.

As in our previous studies and those of other groups the technical validation of the BeadArray approach 
shows it to be a robust approach for the quantitative determination of methylation (Fryer et al. 2011, Lima et 
al. 2011). Furthermore, and although not reported by others we found, in pmPits, that the imprinted genes 
show β values that reflect silencing of one of two alleles and provides further confidence in the derived data.

For the majority of genes interrogated by the 27K BeadArray technology most are represented by one or two 
CpG sites on the array (Wilop et al. 2011, Irizarry et al. 2008). In these cases, although the number of CpG 
sites interrogated, as representations of a CpG islands, are few in number, the a priori of the technology is 
that CpG sites in close proximity are often co-methylated (Irizarry et al. 2008, Ogoshi et al. 2011). However, 
our validation studies on the basis of an elevated β value at a single CpG site failed to show a robust 
correlation. In these cases and by pyrosequence analysis we had determined mean methylation across 5-11 
CpG within the promoter-associated CpG islands. Using modified and more stringent criteria, which 
considered two promoter-associated CpGs on the BeadArray we found significant concordance between the 
techniques. In this case, 12 of 16 (75%) genes were validated by pyrosequencing. Other studies, where they 
report validation of BeadArray data, frequently use techniques such as methylation sensitive PCR (Wilop et 
al. 2011, Yoon et al. 2010). A limitation of this technique is that they determine methylation at a limited 
number of CpG sites. Equally, this approach is reliant upon an amplification step and may, not necessarily 
accurately reflect the methylation status of the CpG islands that is representative of the majority of cells 

The gene ontogeny analysis, on the basis of the more stringent cut-off criteria, shows the genes to be 
distributed across various categories of biological processes, molecular function and pathways. Although no 
biological processes are specifically enriched we found that multiple genes involved in intracellular 
signalling pathways are hypermethylated relative to pmPit. Interestingly, many of these genes are silenced in 
association with CpG methylation in other tumour types (Sasi et al. 2010, Mizukawa et al. 2011, Walter et al. 
2012, Qian et al. 2005, Mottok et al. 2007, Bonazzi et al. 2011) and interestingly the RASSF1 gene has been
identified in earlier studies using a candidate gene approach of sporadic pituitary adenomas. (Qian et al. 2005)

Although the BeadArray allowed us to examine each of the major pituitary adenoma subtypes in a single experiment, that is on the same array, we were aware of the constraints and limitations of any conclusions reached through examining a limited number of each of the adenoma subtypes. Equally the amount and absolute number of each adenoma subtype was also a limitation for all of the analyses we wished to undertake. However, through focusing our studies, post BeadArray analysis, principally but not exclusively on non-functioning (NF) adenomas we gained further insight into aspects of epigenetic change. In this case, in an independent cohort that comprised 13 NF adenomas, and for 12 genes (identified in the discovery cohort) we determined methylation frequencies relative to pmPits. Again using stringent criteria (four times the SD apparent in pmPits) we determined the frequencies of differential methylation across the cohorts for each of these genes. The genes show variable frequencies of methylation and have not been previously described in pituitary adenomas. Indeed, for many of the genes identified in our discovery cohort, as example, SOCS1, SEPT9, PDLIM4, TFAP2E, MT1G, HAAO, TFAP2E, CRIP1 and COL1A2 methylation mediated gene silencing, together with known or putative tumour suppressor characteristics have been reported in other tumour types. (Sasi et al. 2010, Mizukawa et al. 2011, Walter et al. 2012, Qian et al. 2005, Mottok et al. 2007, Bonazzi et al. 2011)

Although the absolute number and amount of tumour tissue, for the other adenoma subtypes was limiting we considered it important to validate the subtype-specificity of aberrant methylation. To that end, we examined hypermethylation of CpG islands that were methylated in common and those where this change is apparent in two or more subtypes as determined by BeadArray analysis. The analysis shows, by an independent technique, namely pyrosequencing, that the identified changes and differences in hypermethylation were indeed subtype specific.

In this first report of unbiased genome-wide changes in methylation we focused our studies on increase in methylation of CpG islands associated with gene promoter. In these cases, these regions are considered the most relevant for expression of the corresponding gene. (Irizarry et al. 2008, Moran et al. 2012) In the 13 NF
adenomas and across the 12 genes studied only three showed an association between hypermethylation and reduced transcript expression. Similar findings have been reported by others and specifically where associations between increased BeadArray-identified CpG methylation and loss or reduced transcript expression has been investigated. (Kim et al. 2011, Zeller et al. 2012, Irizarry et al. 2008, Wilop et al. 2011, Noushmehr et al. 2010) Multiple factors might account for this lack of association, including density of methylation across the CpG islands, histone modification or the milieu of transcription factor within a cell or cell type. (Wilop et al. 2011, Irizarry et al. 2008) These findings have led some authors to conclude that mRNA expression is more likely to be down regulated when the promoter region is hypermethylated even through it is not statistically significant. (Kim et al. 2011) Therefore, and by extension this likely reflects the view, put forward by other investigators, that many epigenetic and genetic alterations in cancer are “passenger” as opposed to “driver” events in the evolution and progression of a tumour (Kim et al. 2011, Zeller et al. 2012).

This study provides a first, unbiased survey of the pituitary tumour epigenome across different adenoma subtype. Further studies will be needed to corroborate identification of these CpG islands, characterised on the basis of their hypermethylation and relative to that seen in normal pituitaries. These types of studies will be useful for identification of biomarker that perhaps predict or characterise tumours likely to show aggressive or recurrent growth characteristics. In addition the functional characterisation of down regulated genes will provide new understanding of tumour aetiology and biology and perhaps identify novel genes or pathways that provide us with new therapeutic targets or options.

Table and Figure Legends:

**Table 1:** Identity of 22 random selected genes from BeadArray dataset selected on the basis that they fulfilled the following criteria: (A), $\beta$ value of $\geq 0.4$ in a single CpG in at least two of three specimens of an adenoma subtype and relative to the values seen in pmPits. (B), Identity of 16 genes selected on the basis of two CpG within the gene (in at least two of three specimens of an adenoma subtype) where the $\beta$ values are $\geq 0.4$ in one of two CpGs and $\geq 0.25$ in the second and relative to the values seen in pmPits. Genes, where pyrosequencing validated their methylation status are shown in bold.
Table 2: Genes identified on the BeadArray as hypermethylated on the basis of one of two CpG (in two of
three adenomas within a subtype) showing a $\beta$ value of $\geq 0.4$ and a second CpG showed a $\beta \geq 0.25$ relative to
the mean of the pmPits. The table shows the genes fulfilling these criteria in each of adenoma subtypes and
where the CpG is within a CpG island. Genes shown in bold and underlined were subject to the additional
analyses, described in Figure 3 and Table 3.

Table 3: Frequencies of gene specific methylation in an independent cohort of NF adenomas. Genes were
first identified in the discovery cohort and using the criteria described in the text and in Fig. 2 and 3.

FIGURES

Figure 1: Correlation between BeadArray derived $\beta$ values (expressed as a percentage) and pyrosequencing
across a range on $\beta$ values from 12 adenomas for four CpGs (cg08047457, cg20289949, cg06197492, and
cg24739326).

Figure 2: Heatmap of genes in individual NF and GH adenomas that fulfilled the criteria of elevated $\beta$ in
two gene-associated CpG relative to the pmPits (NP). For each gene that satisfied the criteria the sum of the
two $\beta$ values is used and displayed as a graded colour scale that is shown in the figure.

Figure 3: Pyrosequence analysis of promoter-associated CpG island methylation of genes first identified as
hypermethylated by BeadArray analysis. The figure shows methylation in an independent cohort of 13 NF
adenomas relative to the mean methylation for each gene examined in the 4 pmPits. For each gene, in the
pmPits, the mean percentage methylation and four times the SD of the mean are shown as horizontal bars,
either without or with arrows respectively. Only adenomas equal to or greater than this upper threshold were
deemed hypermethylated and are shown as filled circles. Adenomas where methylation is not increased
relative to pmPits are shown as unfilled circles. The pmPits are shown as diamonds to the left of the
adenomas (circles) in each case.

Figure 4: Adenoma subtype specificity of hypermethylation for three genes where pyrosequencing
confirmed their methylation in NF adenomas and shown in Fig. 3. For subtype specificity, in addition to the
13 NF adenomas the methylation status of each gene was also determined in 7 GH, 6 PRL and 6 CT (CUSH)
adenoma. Symbols and cut-off criteria are as described in Fig. 3. The pmPits are shown as diamonds on the right side of each of the gene investigated.

**Figure 5:** Quantitative RT-PCR analysis for three of twelve gene transcripts that show an inverse relationship between hypermethylation and decrease in transcript expression. For each gene, their methylation status in NF adenomas was first validated in the discovery cohort (Table 2) and in the independent cohort shown in Fig. 3. Expression is shown relative to four pmPits and the same cut-off criteria of four SD, but in this case lower than the mean level shown in pmPits. Filled and unfilled circles represent adenomas that are hypermethylated and those that were not for individual NF adenoma respectively. The pmPits are shown as diamonds on the left side of each the gene investigated.

**Supplemental Tables**

**S1:** (a), Pyrosequencing primers designed to amplify sodium bisulphite converted DNA. The primers shown are for CpGs used in the technical validation and determined methylation at the single CpG interrogated on the array: cg08047457, cg20289949, cg06197492, and cg24739326. Primer sequences are also shown for the 22 gene-associated CpG islands investigated in this study. (b), Primers used for RT-qPCR analyses.

**S2:** (a), Genes identified on the BeadArray as hypermethylation on the basis of one CpG (in two of three adenomas within a subtype) showing a $\beta$ value of $\geq 0.4$ relative to the mean in the pmPits. The table shows the genes fulfilling these criteria in each of adenoma subtypes, and where the CpG is within or (b), outside of a CpG island.

**S3:** Genes identified on the BeadArray as hypermethylation on the basis of one of two CpG (in two of three adenomas within a subtype) showing a $\beta$ value of $\geq 0.4$ and where the second CpG showed a $\beta \geq 0.25$ relative to the mean in the pmPits. The table shows the genes fulfilling these criteria in each of adenoma subtypes and where the CpG is outside of CpG island.

**S4, 5 and 6:** Gene ontogeny analysis of genes described in Table 2 and that were identified on the basis of their CpGs within the promoter associated CpG islands. Each table shows the gene associated with,
Biological process (S4), Molecular function (S5) and KEGG pathway (S6) analysis associated with each of the genes respectively.

References


Bonazzi VF, Nancarrow DJ, Stark MS, Moser RJ, Boyle GM, Aoude LG, Schmidt C & Hayward NK 2011 Cross-platform array screening identifies COL1A2, THBS1, TNFRSF10D and UCHL1 as genes frequently silenced by methylation in melanoma. *PloS one* **6** e26121.


Sasi W, Jiang WG, Sharma A & Mokbel K 2010 Higher expression levels of SOCS 1,3,4,7 are associated with earlier tumour stage and better clinical outcome in human breast cancer. *BMC cancer* 10 178.


Wessely F, & Emes RD 2012 Identification of DNA methylation biomarkers from Infinium arrays. *Front genet* 3 161


Figure 1: Correlation between BeadArray derived β values (expressed as a percentage) and pyrosequencing across a range on β values from 12 adenomas for four CpGs (cg08047457, cg20289949, cg06197492, and cg24739326).

Spearman $r$: 0.85, p value < 0.0001
Figure 2: Heatmap of genes in individual NF and GH adenomas that fulfilled the criteria of elevated β in two gene-associated CpG relative to the pmPits (NP). For each gene that satisfied the criteria the sum of the two β values is used and displayed as a graded colour scale that is shown in the figure.
Figure 3: Pyrosequence analysis of promoter-associated CpG island methylation of genes first identified as hypermethylated by BeadArray analysis. The figure shows methylation in an independent cohort of 13 NF adenomas relative to the mean methylation for each gene examined in the 4 pmPits. For each gene, in the pmPits, the mean percentage methylation and four times the SD of the mean are shown as horizontal bars, either without or with arrows respectively. Only adenomas equal to or great than this upper threshold were deemed hypermethylated and are shown as filled circles. Adenomas where methylation is not increased relative to pmPits are shown as unfilled circles. The pmPits are shown as diamonds to the left of the adenomas (circles) in each case.

141x85mm (300 x 300 DPI)
Figure 4: Adenoma subtype specificity of hypermethylation for three genes where pyrosequencing confirmed their methylation in NF adenomas and shown in Fig. 3. For subtype specificity, in addition to the 13 NF adenomas the methylation status of each gene was also determined in 7 GH, 6 PRL and 6 CT (CUSH) adenoma. Symbols and cut-off criteria are as described in Fig. 3. The pmPits are shown as diamonds on the right side of each of the gene investigated.
Figure 5: Quantitative RT-PCR analysis for three of twelve gene transcripts that show an inverse relationship between hypermethylation and decrease in transcript expression. For each gene, their methylation status in NF adenomas was first validated in the discovery cohort (Table 2) and in the independent cohort shown in Fig. 3. Expression is shown relative to four pmPits and the same cut-off criteria of four SD, but in this case lower than the mean level shown in pmPits. Filled and unfilled circles represent adenomas that are hypermethylated and those that were not for individual NF adenoma respectively. The pmPits are shown as diamonds on the left side of each the gene investigated.
Table 1:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st CpG ≥ 0.4</td>
<td>EML2, RHOD, HOXB1, FLJ46380, MT1G, HAAO, KIAA1822, KCNE3, TFAP2E, BCL9L, PGLYRP1, COL1A2, ST6GALNAC6, UBD1, KIAA0676, PDLIM4, CXCL14, FEM1C, MAL, RAMP1, D4ST1, SAMD11</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; CpG ≥ 0.25</td>
<td></td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; CpG ≤ 0.25</td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>GH</td>
</tr>
<tr>
<td>---------</td>
<td>----</td>
</tr>
<tr>
<td>40057</td>
<td></td>
</tr>
<tr>
<td>ACTA1</td>
<td></td>
</tr>
<tr>
<td>ADRA1A</td>
<td></td>
</tr>
<tr>
<td>ALX4</td>
<td></td>
</tr>
<tr>
<td>BC19L</td>
<td></td>
</tr>
<tr>
<td>BIK</td>
<td></td>
</tr>
<tr>
<td>BMP8A</td>
<td></td>
</tr>
<tr>
<td>C1QTNF5</td>
<td></td>
</tr>
<tr>
<td>C6orf150</td>
<td></td>
</tr>
<tr>
<td>CHST8</td>
<td></td>
</tr>
<tr>
<td>CNFN</td>
<td></td>
</tr>
<tr>
<td>COL1A2</td>
<td></td>
</tr>
<tr>
<td>CRIP1</td>
<td></td>
</tr>
<tr>
<td>DLX5</td>
<td></td>
</tr>
<tr>
<td>EFEMP1</td>
<td></td>
</tr>
<tr>
<td>EFS</td>
<td></td>
</tr>
<tr>
<td>ELN</td>
<td></td>
</tr>
<tr>
<td>EML2</td>
<td></td>
</tr>
<tr>
<td>ENTPD2</td>
<td></td>
</tr>
<tr>
<td>ERBB2</td>
<td></td>
</tr>
<tr>
<td>FLJ32569</td>
<td></td>
</tr>
<tr>
<td>FLJ32569</td>
<td></td>
</tr>
<tr>
<td>FLJ46380</td>
<td></td>
</tr>
<tr>
<td>FLJ90166</td>
<td></td>
</tr>
<tr>
<td>HAAD</td>
<td></td>
</tr>
<tr>
<td>HAT1</td>
<td></td>
</tr>
<tr>
<td>HDAC11</td>
<td></td>
</tr>
<tr>
<td>HOXB1</td>
<td></td>
</tr>
<tr>
<td>KCNE3</td>
<td></td>
</tr>
<tr>
<td>KCNQ1</td>
<td></td>
</tr>
<tr>
<td>KIAA0676</td>
<td></td>
</tr>
<tr>
<td>KIAA1822</td>
<td></td>
</tr>
<tr>
<td>LDHC</td>
<td></td>
</tr>
<tr>
<td>MT1G</td>
<td></td>
</tr>
<tr>
<td>NPPB</td>
<td></td>
</tr>
<tr>
<td>PGLYRP1</td>
<td></td>
</tr>
<tr>
<td>PON3</td>
<td></td>
</tr>
<tr>
<td>RAB34</td>
<td></td>
</tr>
<tr>
<td>RAC2</td>
<td></td>
</tr>
<tr>
<td>RASSF1</td>
<td></td>
</tr>
<tr>
<td>RHOD</td>
<td></td>
</tr>
<tr>
<td>SECTM1</td>
<td></td>
</tr>
<tr>
<td>SIPA1</td>
<td></td>
</tr>
<tr>
<td>SLC5A1</td>
<td></td>
</tr>
<tr>
<td>SLITL2</td>
<td></td>
</tr>
<tr>
<td>SOCS1</td>
<td></td>
</tr>
<tr>
<td>SOCS2</td>
<td></td>
</tr>
<tr>
<td>ST6GALNAC6</td>
<td></td>
</tr>
<tr>
<td>TCF15</td>
<td></td>
</tr>
<tr>
<td>TFAP2E</td>
<td></td>
</tr>
<tr>
<td>TNFRSF10D</td>
<td></td>
</tr>
<tr>
<td>UBD1</td>
<td></td>
</tr>
<tr>
<td>VILL</td>
<td></td>
</tr>
<tr>
<td>WFIKN2</td>
<td></td>
</tr>
</tbody>
</table>

**Total**  21  40  6  2
<table>
<thead>
<tr>
<th>Gene</th>
<th>Number of hypermethylated adenoma</th>
<th>Number of adenomas showing methylation similar to pmPits</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCL9L</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>COL1A2</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>EML2</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>FLJ46380</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>HAAO</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>HOXB1</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>KCNE3</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>KIAA1822</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>MT1G</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>PGLYRP1</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>RHOD</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>TFAP2E</td>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>