# Quantitative genome-wide methylation analysis of high-grade non-muscle invasive bladder cancer.

Mark O Kitchen*[1,2], Richard T Bryan[3], Richard D Emes[4], John R Glossop[1], Christopher Luscombe[2], KK Cheng[3], Maurice P Zeegers[3,5,6,7], Nicholas D James[8], Adam J Devall[3], Charles A Mein[9], Lyndon Gommersall[2], Anthony A Fryer[1], William E Farrell[1].

(1) Institute for Science and Technology in Medicine, Keele University, UK.

(2) Urology Department, University Hospitals of North Midlands NHS Trust, UK.

(3) Institute of Cancer and Genomic Sciences, University of Birmingham, UK.

(4) Advanced Data Analysis Centre, University of Nottingham, UK.

(5) Department of Complex Genetics, Maastricht University Medical Centre, The Netherlands.

(6) NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre, The Netherlands.

(7) CAPHRI School for Public Health and Primary Care, Maastricht University Medical Centre, The Netherlands.

(8) Cancer Research Unit, University of Warwick, UK.

(9) The Genome Centre, Barts and the London School of Medicine and Dentistry, London, UK.

* Corresponding author.

## Abstract

High-grade non-muscle invasive bladder cancer (HG-NMIBC) is a clinically unpredictable disease with greater risks of recurrence and progression relative to their low-intermediate-grade counterparts. The molecular events, including those affecting the epigenome, that characterise this disease entity in the context of tumour development, recurrence and progression, are incompletely understood. We therefore interrogated genome-wide DNA methylation using HumanMethylation450 BeadChip-arrays in 21 primary HG-NMIBC tumours relative to normal bladder controls. Using strict inclusion-exclusion criteria we identified 1,057 hypermethylated CpGs within gene promoter-associated CpG islands, representing 256 genes. Bisulphite Pyrosequencing validated the array data and examined 25 array-identified candidate genes in an independent cohort of 30 HG-NMIBC and 18 low-intermediate-grade NMIBC. These analyses revealed significantly higher methylation *frequencies* in high-grade tumours relative to low-intermediate-grade tumours for the *ATP5G2*, *IRX1* and *VAX2* genes ($p<0.05$), and similarly significant increases in *mean levels* of methylation in high-grade tumours for the *ATP5G2*, *VAX2*, *INSRR*, *PRDM14*, *VSX1*, *TFAP2b*, *PRRX1*, and *HIST1H4F* genes ($p<0.05$). Although inappropriate promoter methylation was not invariantly associated with reduced transcript expression, a significant association was apparent for the *ARHGEF4*, *PON3*, *STAT5a*, and *VAX2* gene transcripts ($p<0.05$). Herein, we present the first genome-wide DNA methylation analysis in a unique HG-NMIBC cohort, showing extensive and discrete methylation changes relative to normal bladder and low-intermediate-grade tumours. The genes we identified hold significant potential as targets for novel therapeutic intervention either alone, or in combination, with more conventional therapeutic options in the treatment of this clinically unpredictable disease.

**Key words:** *High-grade Non-Muscle Invasive Bladder Cancer, Epigenetics, Methylation, HumanMethylation450 BeadChip Array, Gene Expression*

49 **Introduction**

50 Bladder cancer is the ninth most common cancer worldwide.[1] The majority of bladder
51 cancers are transitional cell carcinomas (TCC), of which 70-80% are non-muscle invasive
52 (NMIBC) at presentation.[2] Poorly differentiated 'high-grade' (HG)-NMIBC is a clinically
53 important sub-type, accounting for approximately 10-15% of all NMIBCs at presentation.[3, 4],
54 These high-grade tumours are typically more aggressive than their low- and intermediate-
55 grade counterparts, manifest by higher rates of recurrence and progression to invasive and
56 metastatic disease despite intensive and prolonged intravesical treatment.[5, 6]

57

58 The majority of NMIBCs are thought to be consequent to, and represent initiation and
59 progression from, a complex interplay between sporadic, environmental, and heritable risk
60 factors, including those that impact upon genetic and epigenetic pathways. NMIBCs and
61 muscle invasive bladder cancers (MIBCs) have been shown to develop independently ('the
62 two pathway model') on the basis of gain of function fibroblast growth factor receptor 3
63 (*FGFR3*) mutations in NMIBC, and loss of function mutations in retinoblastoma 1 (*RB1*) and
64 tumour protein 53 (*p53*) in MIBC,[7-10] and have been shown to evolve from different cell
65 types. [11, 12] However, the molecular pathways responsible for the evolution, outgrowth and
66 progression of HG-NMIBC have not been subject to comprehensive study or investigation;
67 indeed, it is currently unclear whether HG-NMIBCs arise as a discrete disease entity,
68 whether they represent step-wise progression from low-intermediate-grade NMIBC tumours,
69 or whether they sit at a molecular crossroads between NMIBC and MIBC.[7, 13 11] This
70 uncertainty is illustrated by the findings that high-grade tumours harbour abnormalities in
71 common with low-intermediate-grade NMIBC, such as mutations of *FGFR3* and/or rat
72 sarcoma viral oncogene homolog (*RAS*) pathway genes[14, 15], but also display extensive
73 genetic instability and compromised regulation of vital cellular processes more in keeping
74 with MIBC.[14, 16]

75 Epigenetic modifications are frequently implicated in the development of human
76 malignancies, and in these cases, are typically apparent as inappropriate gene promoter
77 CpG island DNA methylation, histone tail modification(s), aberrant expression of micro- and
78 long non-coding-RNAs, and less frequently, loss of gene body/intergenic methylation.[17, 18]
79 These heritable modifications, or *epimutations*, impact upon gene expression either alone or
80 in combination, and promote tumour evolution and/or progression by suppressing the
81 expression of growth inhibiting and/or apoptosis promoting genes, and less frequently by
82 leading to relaxed control of expression of growth promoting genes.[17, 19, 20]

83 Epigenetic modifications and associated gene silencing have been shown in NMIBC, and
84 specific patterns of DNA methylation, histone modifications and microRNA expression have
85 been reported as associated with tumour growth characteristics, patient/clinical outcomes
86 and with field defect phenomena.[21, 22] However, the majority of these reports have described
87 epigenetic changes in heterogeneous populations of NMIBC, with an abundance of low- and
88 intermediate-grade tumours relative to high-grade tumours. With the exception of our recent
89 candidate-gene study[23] and a single report investigating the Myopodin A gene[24], HG-
90 NMIBCs have not been considered as a discrete entity for the investigation of epigenetic
91 modifications.

92 In this study, we interrogated DNA methylation on a genome-wide scale using methylation
93 BeadChip-array technology, in a unique cohort of HG-NMIBCs. Through comparisons with
94 methylation levels and gene-expression in low/intermediate-grade tumours, we extend the
95 current understanding of bladder cancer tumourigenesis and identify potential epigenetic
96 mechanisms implicated in the development of high-grade NMIBC, and those that might
97 represent novel therapeutic drug-targets.

98

99 **Results**

100 *Technical Validation of array by Pyrosequencing:*

101 Subsequent to array processing, normalisation and peak-based correction (see patients and
102 methods), a technical validation was performed by comparing array-derived β-values with
103 Pyrosequencing-derived methylation values. Across 120 data-points (5 CpGs, 24 samples)
104 encompassing a broad range of array β-values, a strong positive correlation was found
105 between the methylation values (Spearman's rank correlation $r=0.912$, $p<0.00001$;
106 **Supplemental Figure S1**).

107

108 *In-house filtering criteria:*

109 CpGs showing differential methylation in HG-NMIBC relative to normal bladder controls were
110 identified following a series of stringent filtering criteria, as described previously and shown
111 in **Figure 1**.[25, 26] On the basis of these criteria, a total of 1,057 CpGs, representing 256
112 genes, were identified as hypermethylated (≥0.4 β-value increase) in 15 or more of the 21
113 high-grade tumours, relative to their mean values in the normal bladder controls.

114

115 *Hierarchical clustering analyses:*

116 The filtered dataset was next subject to unsupervised hierarchical cluster analysis (**Figure**
117 **2)**: the high-grade tumours cluster independently from the normal bladder control samples.
118 In these cases, methylation is barely detectable within the normal bladder samples, whereas
119 15 or more of the high-grade tumours show inappropriate methylation across all 1,057 CpG
120 dinucleotides, spanning 256 gene-promoter-associated CpG islands (**Supplemental Table**
121 **S2**).

122

123 *Independent validation by Pyrosequencing:*

124 We next selected 25 genes for independent validation by Pyrosequencing on the basis of
125 their frequent methylation in the discovery cohort that comprised 21 high-grade tumours.
126 These analyses revealed similar frequencies and mean levels of methylation as those
127 apparent from the BeadChip array for 24 of the 25 genes. As further confirmation, we
128 extended the Pyrosequence analyses to an independent investigation cohort of 30 HG-
129 NMIBC tumours. Similar frequencies and mean levels of methylation between the discovery
130 and investigation cohorts reinforced our confidence in the array-derived data (**Supplemental**
131 **Table S3**). At this stage, and to assess for potential confounders, we assessed associations
132 between patient demographic data and methylation patterns across these 25 genes, using
133 separate multivariate models. No correlations were identified in these analyses, suggesting
134 demographic factors did not significantly impact upon the methylation patterns identified
135 (data not shown).

136

137 *Differential subtype-specific promoter methylation in NMIBC:*

138 We next determined methylation across the 25 genes described above in HG-NMIBC
139 relative to that apparent in low-intermediate-grade tumours and in comparison to normal
140 bladder controls (**Supplemental Table S4**). Similar to other groups[27 28], we displayed these
141 methylation data, across the high-grade and low-intermediate-grade tumours and normal
142 controls, by heatmap (**Figure 3**). This demonstrated heterogeneous patterns of methylation
143 across the 51 high- and 18 low-intermediate-grade tumours relative to the normal bladder
144 controls. Gene-specific differences in methylation were apparent between the high-grade
145 tumours and their low-intermediate-grade counterparts on visual inspection. Closer
146 examination of these data showed that the differences appeared to impact on either the
147 relative *frequency* and/or the *mean levels* of methylation between these tumour subtypes. As

148 examples of these differences, the ten most differentially methylated genes are shown in
149 **Table 1**.

150

151 *Methylation frequencies in high- and low-intermediate-grade tumours:*

152 For ten of the genes we took forward for further analyses (*ATP5G2*, *HIST1H4F*, *INSRR*,
153 *IRF8*, *IRX1*, *PRDM14*, *PRRX1*, *TFAP2b*, *VAX2* and *VSX1*), there was an higher frequency of
154 methylation in high-grade tumours versus low-intermediate grade tumours (**Table 1**).
155 Moreover, the increases were statistically significant for the *ATP5G2*, *VAX2* and *IRX1* genes
156 ($p<0.05$), and approached significance for the *INSRR*, *IRF8*, *PRDM14* and *VSX1* genes.

157

158 *Mean levels of methylation in high- and low-intermediate-grade tumours:*

159 The mean levels of methylation in the high-grade tumours were next assessed by
160 Pyrosequencing (right-sided panel of **Table 1**, and **Figure 4**); for eight of the ten genes,
161 mean levels of methylation were significantly greater in high-grade tumours relative to their
162 low-intermediate-grade counterparts. In addition, and as low-intermediate-grade tumours
163 were not subject to array analyses relative to normal bladder, further pairwise-testing was
164 performed. This analysis identified significant differences between mean levels of
165 methylation in the low-intermediate-grade tumours and normal bladder in four of the ten
166 genes assessed. The range, distribution and mean levels of methylation are shown in
167 **Figure 4**, and show for each of the genes, a stepwise trend toward increasing methylation
168 from normal bladder to low-intermediate and high-grade tumours.

169

170 *Methylation-Associated Changes in Gene Expression:*

171 Across the high-grade NMIBC tumours, sufficient sample was available for gene expression
172 analyses for 17 of the 25 genes. With the exception of the *ARHGEF4* gene, promoter-
173 associated CpG island methylation was negatively correlated with transcript expression for
174 all genes assessed (data not shown). Furthermore, the presence of promoter methylation
175 was significantly correlated with reduced transcript expression for the *PON3*, *STAT5a* and
176 *VAX2* genes (Spearman's correlation coefficients -0.60, -0.50 and -0.48 respectively, all
177 $p<0.05$). Conversely, promoter methylation was significantly positively correlated with gene
178 transcript expression for the *ARHGEF4* gene (Spearman's correlation coefficient 0.62,
179 $p<0.05$). **Figure 5** shows the expression levels for these four genes across the high-grade
180 tumours.

181

182 *Gene Ontology analysis of inappropriately methylated genes:*

183 Gene Ontology analyses of the 256 differentially methylated genes identified 'over-
184 representation' of multiple categories of biological processes, molecular functions and
185 pathways. In particular, highly significant over-representation was identified for specific
186 biological processes, including regulation of RNA polymerase II activity and DNA
187 transcription, and for pathways involving cell adhesion and PI3K-Akt signalling
188 (**Supplemental Table S5**).

189

190 **Discussion**

191 In common with most other tumour types, bladder cancers harbour epigenetic aberrations
192 which are frequently apparent as inappropriate DNA methylation.[8, 22, 29] However, reports are
193 limited and largely confined to heterogeneous patient cohorts of NMIBC or MIBC;[30] despite
194 their clinical importance, high-grade NMIBC tumours are rarely investigated as a discrete
195 entity in the context of disease and/or subtype-specific epigenetic modifications.[23] To
196 address this, we performed genome-wide analyses of DNA methylation using BeadChip
197 array technology in high-grade NMIBC, comprising a discrete cohort of tumours recruited at
198 initial presentation. This analysis, the first '450K array' interrogation in bladder cancer,
199 revealed multiple and novel frequently differentially methylated genes in these tumours
200 relative to normal bladder. Through Pyrosequence analysis of sodium bisulphite converted
201 DNA, we extended our analyses to include independent cohorts of high- and low-
202 intermediate-grade tumours. These investigations confirmed the array-derived data for the
203 high-grade tumours, and showed them as harbouring significantly increased frequencies
204 and/or mean levels of gene-specific methylation relative to low-intermediate-grade tumours.
205 Moreover, for some of the genes investigated, a significant inverse correlation between
206 promoter methylation and gene expression levels was apparent and suggests their potential
207 as targets for therapeutic intervention.[29 31 32]

208

209 Initially we performed a technical validation of the discovery cohort data by Pyrosequence
210 analysis of converted DNA.[25 33 34] In common with previous reports and across multiple
211 genes, these analyses confirmed and reinforced the array-derived data.[34 35 36] These
212 analyses also showed that for the majority of regions investigated, methylation extended to
213 include contiguous promoter-associated CpG sites. On the basis of previous reports from
214 our own and other groups,[37 38] we employed stringent criteria (β-value differences ≥0.4) to
215 identify differentially methylated genes across multiple CpG sites; such criteria are more

consistently associated with *bona fide* changes in methylation, and are more likely to show associations with gene expression.[37, 39 40, 41]

The analysis of the discovery cohort of high-grade NMIBC identified 1,057 CpGs, across 256 gene-promoter-associated CpG islands. Cluster analysis and heat map display of these regions revealed extensive and frequent differential methylation in the tumours relative to normal bladder controls. As our study represents the first 450K analysis of high-grade bladder cancer a direct 'like-for-like' comparisons of our findings with those of other groups was not possible; however, the number of differentially methylated sites we identified appeared to be lower than those previously reported in other tumour types.[42 43] Potential explanations for these findings are the tumour type *per se* and/or the stringency of our inclusion-exclusion criteria and definition of differential methylation.[44]

For the genes identified, we performed gene ontology and KEGG pathway analyses. In these cases we identified significant over-representation of genes in processes and pathways previously reported by other groups as subject to epigenetically-mediated dysregulation in tumour development. For examples, these included transcription and cell signalling and adhesion[45-47], suggesting possible similar roles in high-grade bladder tumours, and their validity as targets for further investigation.

We next extended our investigation of multiple novel genes to an independent cohort of high-grade tumours, and a cohort of low-intermediate-grade tumours for comparison. Similar frequencies and mean levels of methylation, as determined by Pyrosequence analysis, were apparent within the discovery and investigation cohorts of high-grade tumours, suggesting our approach for the identification of candidates by array analysis was robust. Interestingly, many of the genes identified as novel and differentially methylated were also inappropriately methylated in low-intermediate-grade tumours. However, and despite the absence of genes as being exclusively associated with either high- or low-intermediate-grade tumours, the frequency and mean levels of gene-promoter methylation in the high-grade tumours were

242 significantly higher than in the low-intermediate-grade tumours. Indeed, similar observations
243 with respect to differences in the frequencies of methylation between high- and low-grade
244 bladder tumours were first suggested by Ibragimova *et al*.[47] Similar subtype and/or grade-
245 associated differences have been reported in other tumour types including, pituitary, breast,
246 and colon cancer subtypes.[37, 48, 49] In our analysis of NMIBC it remains unclear whether the
247 increase in frequency and/or mean levels of methylation in the more aggressive tumours
248 represents a more rapid accumulation of epigenetic changes during tumour progression, or
249 reflects distinct epigenetic pathways of tumour development and outgrowth.[50, 51] Our findings
250 may therefore reflect either of the described scenarios in the more aggressive (high-grade)
251 tumours and suggests that these tumours are either consequent to progression from low-
252 intermediate-grade tumours, or are the progeny of aberrations in distinct epigenetic
253 pathways within these NMIBC subtypes. Moreover, the identification of different patterns of
254 methylation between tumours represents an important area for future investigation. In this
255 case, methylation may hold promise as an 'at diagnosis' biomarker of long-term tumour
256 outcome, similar to that described in colorectal, breast and lung cancers. [52-54]

257 Although many of the novel genes we identified have not been previously reported in
258 bladder cancer, their inappropriate methylation, accompanied with gene-silencing, has been
259 reported in the context of other tumour types and suggests potential roles as tumour
260 suppressor genes.[55, 56 57] To determine associations between methylation and gene
261 expression, we confined our studies to genes showing frequent and/or high mean levels of
262 methylation. For the majority of gene-transcripts we investigated, promoter methylation was
263 negatively correlated with reduced transcript expression, although not significantly so (data
264 not shown). However, as described by our own and other groups, this may reflect a
265 passenger-driver phenomenon where, in the 'passenger' context, gene expression is not
266 directly influenced by the observed epigenetic modification(s).[58 59] However, for four of
267 seventeen transcripts we examined, significant correlations between methylation and
268 transcript expression were apparent. In these cases, and for the *PON3*, *STAT5a* and *VAX2*

269 genes, promoter methylation was significantly associated with reduced gene expression,
270 whilst the converse was true for the *ARHGEF4* gene. Such associations are similar to those
271 described previously in multiple other cancers and in NMIBC. [20, 21 43] Indeed, for two of these
272 genes, *PON3* and *STAT5a*, previous studies in mice and cell-line models have described
273 potential tumour suppressor roles.[60 61] If this is the case, then these genes may represent
274 important targets for further studies of functional the significance of methylation and reduced
275 expression in a bladder tumour context, including *in-vitro* investigations of de-methylating
276 agents designed to restore gene expression.

277 In summary, we have presented the first comprehensive genome-wide DNA methylation
278 analysis of NMIBC in a unique cohort of high-grade tumours. The study has reported an
279 increase in the frequency and/or mean levels of methylation at gene promoter-associated
280 CpG islands in high-grade tumours relative to their low-intermediate-grade tumour
281 counterparts, that in some cases is associated with reduced gene expression. These
282 findings suggest that epigenetic modifications, alone or in combination with other
283 aberrations, are causal in the development and/or progression of this tumour type. Further
284 studies are required to assess the functional significance of epigenetic changes in HG-
285 NMIBC; however, we suggest that the genes identified hold significant potential as targets
286 for novel therapeutic interventions alone, or in combination, with conventional therapeutic
287 options in the treatment of this clinically unpredictable disease.

288

289

290

## Patients and methods

*Human tissue samples*

Primary tumour and normal bladder tissues used were provided by the Bladder Cancer Prognosis Programme (BCPP, National Research Ethics Service East Midlands - Derby 06/MRE04/65.)[62], the University of Birmingham Human Biomaterials Resource Centre (National Research Ethics Service (North West 5): 09/H1010/75), and the University Hospitals of North Midlands NHS Trust (National Research Ethics Service (South Central – Oxford C): 12/SC/0725). All samples were confirmed histologically as normal bladder urothelium (control, *n*=4), G3pT1 TCC (high-grade: discovery cohort *n*=21, investigation cohort *n*=30), and G1/2 pTa/1 TCC (low/intermediate-grade: *n*=18). As previously described[23], patients received repeat bladder tumour resection (TURBT), cystectomy and/or intra-vesical therapy as recommended by European Association of Urology guidelines.[63] All samples (details are provided in **Supplemental Table S1**) were stored at -80$^{o}$C prior to nucleic acid extraction, as described below.

*DNA extraction and bisulphite modification*

Genomic DNA was extracted from tumour and control tissues using a standard phenol-chloroform procedure [64], then bisulphite-converted using the EZ DNA Methylation Gold kit (Zymo Research) as we have previously described.[37] Bisulphite-conversion of DNA was confirmed in all cases by successful PCR using primers specific to bisulphite-converted DNA (primer sequences in **Supplemental Table S6**). To increase the relative amount and stability of bisulphite-converted DNA, whole-genome amplification (WGA) was performed as previously described.[37]

316 *Illumina 450K Methylation Bead-Array Analyses*

317 Bisulphite-converted DNA from 21 bladder tumours and three normal controls was
318 hybridised to Infinium-based HumanMethylation450 BeadChip arrays (Illumina, San Diego,
319 CA, USA) to quantify DNA methylation at approximately 480,000 CpG positions across the
320 genome, representing more than 21,000 RefSeq genes. In this case, normal bladder was
321 used as control for consistency with previous array analyses [35, 47, 65], and also to permit
322 comparisons with earlier reports of non-muscle invasive bladder cancer. Arrays were
323 processed according to the manufacturer's instructions (performed by Barts and the London
324 Genome Centre, UK), as described by us previously.[66]

325 Raw array data were processed using GenomeStudio software and the bioinformatical
326 platform 'NIMBL', as we [67, 68] and others[69] have described. For each probe, the methylation
327 status was reported as a methylation 'β-value', where 'β' is defined as the ratio of the
328 methylated signal intensity over the summed intensity of the methylated and unmethylated
329 signals + 100.[40] β-values range from 0 (unmethylated) to 1 (fully methylated). NIMBL was
330 used to perform 'peak-based' correction, to adjust for potential differences in array probe-
331 type sensitivity previously reported[33]; all comparative analyses of high-grade tumours to
332 normal bladder controls, were performed on peak-based corrected β-values, as described by
333 us previously.[68]

334 Each array passed quality control assessment based upon the performance of internal
335 controls and the distribution of β-values across all array CpGs. As previously described[68],
336 and represented by **step 1** of **Figure 1**, we excluded all CpGs for which any of the 24
337 samples displayed: (i) probe detection *p*-values >0.05 (unreliable probe data), or (ii) missing
338 β-values (preventing analyses of all samples). We also excluded all CpG loci on allosomes
339 (reducing confounding gender-based methylation differences). We used a series of stringent
340 filtering criteria, shown in **Figure 1** and described in the Results section, to identify

341 inappropriate methylation, defined as a β-value difference ≥0.4, in tumour samples relative
342 to the mean of the normal bladder controls.
343
344 Unsupervised hierarchical clustering using average linkage criteria was performed using
345 Genesis software (v1.7.6).[70] Gene Ontology (GO) analyses were performed using
346 http://geneontology.org/ and http://gather.genome.duke.edu/, and Kyoto Encyclopaedia of
347 Genes and Genomes (KEGG) analyses with http://www.genome.jp/kegg/ online platforms,
348 respectively. Bonferroni correction[71] was employed in all GO and KEGG pathway analyses.
349

350 *Technical validation of Methylation Bead-Chip Array Data*

351 Five CpG loci encompassing a broad range of β-values derived from 450k array analyses,
352 were assessed by Pyrosequencing (described below), using identical samples, to
353 independently validate the array data (β-values *vs.* methylation %). Correlation between the
354 methods was assessed across a total of 120 CpGs using Spearman's rank correlation, as
355 shown in **Supplemental Figure S1**. Primer sequences are provided in **Supplemental Table**
356 **S6**.
357

358 *Pyrosequencing™ of sodium bisulphite-converted DNA*

359 Validation of array data (discovery cohort) and further quantitative assessment of
360 methylation in the independent (investigation) tumour cohort were performed by
361 Pyrosequencing of sodium bisulfite-converted DNA, as previously described by us[66], using a
362 PyroMark Q24 Pyrosequencer, PyroMark Q24 Software 2.0 and PyroMark Gold Q24
363 Reagents. Dependent on the specific gene, and the density of CpGs within their promoter-
364 associated CpG island, between five and nine consecutive CpG sites were assessed.
365 Promoter methylation was defined in tumours if the mean level of methylation across the
366 assessed CpG island was greater either than four standard deviations (4SD), or 20% above,

the mean of the normal controls.[37] The number of tumours methylated for any given gene describes the *frequency* of methylation, whereas the mean percentage methylation *per se* of all of the CpGs surveyed within a gene describes the *mean level* of methylation.

*Quantitative RT-PCR*

Total RNA was extracted from control and tumour samples using a standard guanidinium thiocyanate-phenol-chloroform protocol [72]. Complementary DNA (cDNA) was synthesised as described previously[73]. Thermal cycling using SYBR Green was as previously described[74], with target genes normalised to glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as the endogenous control gene (**Supplemental Table S6**). Relative quantification of transcript expression was performed using the $2^{-\Delta\Delta}$ cycle threshold (CT) method[75], and as previously described[76]. Reduced transcript expression in a tumour was defined where expression was at least 3-fold lower than the mean level of expression observed in control samples; the converse was true for increased transcript expression.[37 38 77]

*Non-Array Informatics and statistics.*

STATA (version 8, Stata Corporation, College Station, TX) was used to analyse methylation and gene expression data in tumour and normal cohorts using Fisher's exact tests (frequency of methylation), Student's t-tests (mean level of methylation), and Spearman correlation coefficients (associations between methylation and gene expression). *p*-values <0.05 were considered statistically significant.

389 **Ethics Committee Approvals**

390 East Midlands - Derby: 06/MRE04/65.

391 The University of Birmingham Human Biomaterials Resource Centre (National Research
392 Ethics Service (North West 5): 09/H1010/75.

393 The University Hospitals of North Midlands NHS Trust (National Research Ethics Service
394 (South Central – Oxford C): 12/SC/0725.

395

396 **Reagents**

397 EZ DNA Methylation Gold kit, Zymo Research, D5005

398 HumanMethylation450 BeadChip arrays, Illumina, WG-314-1003

399 PyroMark Gold Q24 Reagents, Qiagen, 970802

400 SYBR III brilliant green, Agilent, 600882

401

402 **Acknowledgements.**

403 We would like to thank Dr Kim Haworth and Dr Kiren Yacqub-Usman for their support of the
404 laboratory work.

405 We would like to thank all the West Midlands Consultant Urologists and their units involved
406 with BCPP, as well as the BCPP research nurses and Margaret Grant, Deborah Bird,
407 Jennifer Barnwell, Duncan Nekeman and Eline van Roekel.

408

409

## References


1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer Journal international du cancer 2015; 136:E359-86.
2. Kirkali Z, Chan T, Manoharan M, Algaba F, Busch C, Cheng L, Kiemeney L, Kriegmair M, Montironi R, Murphy WM, et al. Bladder cancer: epidemiology, staging and grading, and diagnosis. Urology 2005; 66:4-34.
3. Sylvester RJ, van der Meijden AP, Oosterlinck W, Witjes JA, Bouffioux C, Denis L, Newling DW, Kurth K. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. European urology 2006; 49:466-5; discussion 75-7.
4. Boustead GB, Fowler S, Swamy R, Kocklebergh R, Hounsome L, Section of Oncology B. Stage, grade and pathological characteristics of bladder cancer in the UK: British Association of Urological Surgeons (BAUS) urological tumour registry. BJU international 2014; 113:924-30.
5. Cambier S, Sylvester RJ, Collette L, Gontero P, Brausi MA, van Andel G, Kirkels WJ, Silva FC, Oosterlinck W, Prescott S, et al. EORTC Nomograms and Risk Groups for Predicting Recurrence, Progression, and Disease-specific and Overall Survival in Non-Muscle-invasive Stage Ta-T1 Urothelial Bladder Cancer Patients Treated with 1-3 Years of Maintenance Bacillus Calmette-Guerin. European urology 2015.
6. Vedder MM, Marquez M, de Bekker-Grob EW, Calle ML, Dyrskjot L, Kogevinas M, Segersten U, Malmstrom PU, Algaba F, Beukers W, et al. Risk prediction scores for recurrence and progression of non-muscle invasive bladder cancer: an international validation in primary tumours. PloS one 2014; 9:e96849.
7. Knowles MA. Molecular pathogenesis of bladder cancer. International journal of clinical oncology 2008; 13:287-97.
8. Dudziec E, Goepel JR, Catto JW. Global epigenetic profiling in bladder cancer. Epigenomics 2011; 3:35-45.
9. Van Batavia J, Yamany T, Molotkov A, Dan H, Mansukhani M, Batourina E, Schneider K, Oyon D, Dunlop M, Wu XR, et al. Bladder cancers arise from distinct urothelial sub-populations. Nature cell biology 2014; 16:982-91.
10. Bakkar AA, Wallerand H, Radvanyi F, Lahaye JB, Pissard S, Lecerf L, Kouyoumdjian JC, Abbou CC, Pairon JC, Jaurand MC, et al. FGFR3 and TP53 gene mutations define two distinct pathways in urothelial cell carcinoma of the bladder. Cancer research 2003; 63:8108-12.
11. Van Batavia J, Yamany T, Molotkov A, Dan H, Mansukhani M, Batourina E, Schneider K, Oyon D, Dunlop M, Wu XR, et al. Bladder cancers arise from distinct urothelial sub-populations. Nature cell biology 2014; 16:982-91, 1-5.
12. Bryan RT, Ward DG. Words of wisdom. Bladder cancers arise from distinct urothelial sub-populations. European urology 2015; 67:590-1.
13. McConkey DJ, Lee S, Choi W, Tran M, Majewski T, Lee S, Siefker-Radtke A, Dinney C, Czerniak B. Molecular genetics of bladder cancer: Emerging mechanisms of tumor initiation and progression. Urologic oncology 2010; 28:429-40.
14. Castillo-Martin M, Domingo-Domenech J, Karni-Schmidt O, Matos T, Cordon-Cardo C. Molecular pathways of urothelial development and bladder tumorigenesis. Urologic oncology 2010; 28:401-8.
15. Knowles MA, Hurst CD. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. Nature reviews Cancer 2015; 15:25-41.
16. Majewski T, Lee S, Jeong J, Yoon DS, Kram A, Kim MS, Tuziak T, Bondaruk J, Lee S, Park WS, et al. Understanding the development of human bladder cancer by using a whole-organ genomic mapping strategy. Laboratory investigation; a journal of technical methods and pathology 2008; 88:694-721.

460 17. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis 2010; 31:27-36.
461 18. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and
462 translational implications. Nature reviews Cancer 2011; 11:726-34.
463 19. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell 2012;
464 150:12-27.
465 20. Kanwal R, Gupta S. Epigenetic modifications in cancer. Clinical genetics 2012; 81:303-11.
466 21. Wolff EM, Chihara Y, Pan F, Weisenberger DJ, Siegmund KD, Sugano K, Kawashima K, Laird
467 PW, Jones PA, Liang G. Unique DNA methylation patterns distinguish noninvasive and invasive
468 urothelial cancers and establish an epigenetic field defect in premalignant tissue. Cancer research
469 2010; 70:8169-78.
470 22. Catto JW, Azzouzi AR, Rehman I, Feeley KM, Cross SS, Amira N, Fromont G, Sibony M,
471 Cussenot O, Meuth M, et al. Promoter hypermethylation is associated with tumor location, stage,
472 and subsequent progression in transitional cell carcinoma. Journal of clinical oncology : official
473 journal of the American Society of Clinical Oncology 2005; 23:2903-10.
474 23. Kitchen MO, Bryan RT, Haworth KE, Emes RD, Luscombe C, Gommersall L, Cheng KK, Zeegers
475 MP, James ND, Devall AJ, et al. Methylation of HOXA9 and ISL1 Predicts Patient Outcome in High-
476 Grade Non-Invasive Bladder Cancer. PloS one 2015; 10:e0137003.
477 24. Alvarez-Mugica M, Cebrian V, Fernandez-Gomez JM, Fresno F, Escaf S, Sanchez-Carbayo M.
478 Myopodin methylation is associated with clinical outcome in patients with T1G3 bladder cancer. The
479 Journal of urology 2010; 184:1507-13.
480 25. Glossop JR, Nixon NB, Emes RD, Haworth KE, Packham JC, Dawes PT, Fryer AA, Mattey DL,
481 Farrell WE. Epigenome-wide profiling identifies significant differences in DNA methylation between
482 matched-pairs of T- and B-lymphocytes from healthy individuals. Epigenetics : official journal of the
483 DNA Methylation Society 2013; 8:1188-97.
484 26. de Araujo E, Marchi FA, Rodrigues TC, Vieira HC, Kuasne H, Achatz MI, Moredo LF, de Sa BC,
485 Duprat JP, Brentani HP, et al. Genome-wide DNA methylation profile of leukocytes from melanoma
486 patients with and without CDKN2A mutations. Experimental and molecular pathology 2014; 97:425-
487 32.
488 27. Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, Hilmarsdottir H,
489 Olafsdottir E, Tryggvadottir L, Jonasson JG, et al. A DNA methylation-based definition of biologically
490 distinct breast cancer subtypes. Molecular oncology 2015; 9:555-68.
491 28. Sanchez-Vega F, Gotea V, Petrykowska HM, Margolin G, Krivak TC, DeLoia JA, Bell DW,
492 Elnitski L. Recurrent patterns of DNA methylation in the ZNF154, CASP8, and VHL promoters across a
493 wide spectrum of human solid epithelial tumors and cancer cell lines. Epigenetics : official journal of
494 the DNA Methylation Society 2013; 8:1355-72.
495 29. Besaratinia A, Cockburn M, Tommasi S. Alterations of DNA methylome in human bladder
496 cancer. Epigenetics : official journal of the DNA Methylation Society 2013; 8:1013-22.
497 30. Bryan RT, Kirby R, Mostafid H. Does the nonurologic scientific community understand
498 urothelial bladder cancer? European urology 2014; 66:601-2.
499 31. Kelly TK, De Carvalho DD, Jones PA. Epigenetic modifications as therapeutic targets. Nature
500 biotechnology 2010; 28:1069-78.
501 32. Issa JP, Kantarjian HM. Targeting DNA methylation. Clinical cancer research : an official
502 journal of the American Association for Cancer Research 2009; 15:3938-46.
503 33. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the
504 Infinium Methylation 450K technology. Epigenomics 2011; 3:771-84.
505 34. Roessler J, Ammerpohl O, Gutwein J, Hasemeier B, Anwar SL, Kreipe H, Lehmann U.
506 Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina,
507 Inc. BMC research notes 2012; 5:210.
508 35. Reinert T, Modin C, Castano FM, Lamy P, Wojdacz TK, Hansen LL, Wiuf C, Borre M, Dyrskjot
509 L, Orntoft TF. Comprehensive genome methylation analysis in bladder cancer: identification and
510 validation of novel methylated genes and application of these as urinary tumor markers. Clinical

511 cancer research : an official journal of the American Association for Cancer Research 2011; 17:5582-
512 92.
513 36. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I, Rezola R, Alberdi MJ,
514 Dopazo J, Montaner D, et al. DNA methylation epigenotypes in breast cancer molecular subtypes.
515 Breast cancer research : BCR 2010; 12:R77.
516 37. Duong CV, Emes RD, Wessely F, Yacqub-Usman K, Clayton RN, Farrell WE. Quantitative,
517 genome-wide analysis of the DNA methylome in sporadic pituitary adenomas. Endocrine-related
518 cancer 2012; 19:805-16.
519 38. Kim YJ, Yoon HY, Kim JS, Kang HW, Min BD, Kim SK, Ha YS, Kim IY, Ryu KH, Lee SC, et al.
520 HOXA9, ISL1 and ALDH1A3 methylation patterns as prognostic markers for nonmuscle invasive
521 bladder cancer: array-based DNA methylation and expression profiling. International journal of
522 cancer Journal international du cancer 2013; 133:1135-42.
523 39. Schubeler D. Function and information content of DNA methylation. Nature 2015; 517:321-
524 6.
525 40. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP,
526 Gunderson KL, et al. High density DNA methylation array with single CpG site resolution. Genomics
527 2011; 98:288-95.
528 41. Barrera V, Peinado MA. Evaluation of single CpG sites as proxies of CpG island methylation
529 states at the genome scale. Nucleic acids research 2012; 40:11490-8.
530 42. Naumov VA, Generozov EV, Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov SV,
531 Alekseev MV, Shelygin YA, Govorun VM. Genome-scale analysis of DNA methylation in colorectal
532 cancer using Infinium HumanMethylation450 BeadChips. Epigenetics : official journal of the DNA
533 Methylation Society 2013; 8:921-34.
534 43. Karlsson A, Jonsson M, Lauss M, Brunnstrom H, Jonsson P, Borg A, Jonsson G, Ringner M,
535 Planck M, Staaf J. Genome-wide DNA methylation analysis of lung carcinoma reveals one
536 neuroendocrine and four adenocarcinoma epitypes associated with patient outcome. Clinical cancer
537 research : an official journal of the American Association for Cancer Research 2014; 20:6127-40.
538 44. Nagase H, Ghosh S. Epigenetics: differential DNA methylation in mammalian somatic tissues.
539 The FEBS journal 2008; 275:1617-23.
540 45. Cavallaro U, Christofori G. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.
541 Nature reviews Cancer 2004; 4:118-32.
542 46. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007; 128:683-92.
543 47. Ibragimova I, Dulaimi E, Slifker MJ, Chen DY, Uzzo RG, Cairns P. A global profile of gene
544 promoter methylation in treatment-naive urothelial cancer. Epigenetics : official journal of the DNA
545 Methylation Society 2014; 9:760-73.
546 48. Li Y, Li S, Chen J, Shao T, Jiang C, Wang Y, Chen H, Xu J, Li X. Comparative epigenetic analyses
547 reveal distinct patterns of oncogenic pathways activation in breast cancer subtypes. Human
548 molecular genetics 2014; 23:5378-93.
549 49. Gyparaki MT, Basdra EK, Papavassiliou AG. DNA methylation biomarkers as diagnostic and
550 prognostic tools in colorectal cancer. Journal of molecular medicine 2013; 91:1249-56.
551 50. Guo M, Ren J, House MG, Qi Y, Brock MV, Herman JG. Accumulation of promoter
552 methylation suggests epigenetic progression in squamous cell carcinoma of the esophagus. Clinical
553 cancer research : an official journal of the American Association for Cancer Research 2006; 12:4515-
554 22.
555 51. Kaneda A, Matsusaka K, Sakai E, Funata S. DNA methylation accumulation and its
556 predetermination of future cancer phenotypes. Journal of biochemistry 2014; 156:63-72.
557 52. Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, Midorikawa Y, Nishimura Y, Sakamoto
558 H, Seto Y, et al. Three DNA methylation epigenotypes in human colorectal cancer. Clinical cancer
559 research : an official journal of the American Association for Cancer Research 2010; 16:21-33.

560 53. Fackler MJ, Umbricht CB, Williams D, Argani P, Cruz LA, Merino VF, Teo WW, Zhang Z, Huang
561 P, Visvananthan K, et al. Genome-wide methylation analysis identifies genes specific to breast cancer
562 hormone receptor status and risk of recurrence. Cancer research 2011; 71:6195-207.
563 54. Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, Moran S, Heyn H,
564 Vizoso M, Gomez A, et al. A prognostic DNA methylation signature for stage I non-small-cell lung
565 cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology
566 2013; 31:4140-7.
567 55. Bennett KL, Karpenko M, Lin MT, Claus R, Arab K, Dyckhoff G, Plinkert P, Herpel E, Smiraglia
568 D, Plass C. Frequently methylated tumor suppressor genes in head and neck squamous cell
569 carcinoma. Cancer research 2008; 68:4494-9.
570 56. Snellenberg S, Cillessen SA, Van Criekinge W, Bosch L, Meijer CJ, Snijders PJ, Steenbergen RD.
571 Methylation-mediated repression of PRDM14 contributes to apoptosis evasion in HPV-positive
572 cancers. Carcinogenesis 2014; 35:2611-8.
573 57. Morris MR, Ricketts CJ, Gentle D, McRonald F, Carli N, Khalili H, Brown M, Kishida T, Yao M,
574 Banks RE, et al. Genome-wide methylation analysis identifies epigenetically inactivated candidate
575 tumour suppressor genes in renal cell carcinoma. Oncogene 2011; 30:1390-401.
576 58. Kalari S, Pfeifer GP. Identification of driver and passenger DNA methylation in cancer by
577 epigenomic analysis. Advances in genetics 2010; 70:277-308.
578 59. De Carvalho DD, Sharma S, You JS, Su SF, Taberlay PC, Kelly TK, Yang X, Liang G, Jones PA.
579 DNA methylation screening identifies driver epigenetic events of cancer cell survival. Cancer cell
580 2012; 21:655-67.
581 60. Schweikert EM, Devarajan A, Witte I, Wilgenbus P, Amort J, Forstermann U, Shabazian A,
582 Grijalva V, Shih DM, Farias-Eisner R, et al. PON3 is upregulated in cancer tissues and protects against
583 mitochondrial superoxide-mediated cell death. Cell death and differentiation 2012; 19:1549-60.
584 61. Sultan AS, Xie J, LeBaron MJ, Ealley EL, Nevalainen MT, Rui H. Stat5 promotes homotypic
585 adhesion and inhibits invasive characteristics of human breast cancer cells. Oncogene 2005; 24:746-
586 60.
587 62. Zeegers MP, Bryan RT, Langford C, Billingham L, Murray P, Deshmukh NS, Hussain S, James
588 N, Wallace DM, Cheng KK. The West Midlands Bladder Cancer Prognosis Programme: rationale and
589 design. BJU international 2010; 105:784-8.
590 63. Babjuk M, Burger M, Zigeuner R, Shariat SF, van Rhijn BW, Comperat E, Sylvester RJ,
591 Kaasinen E, Bohle A, Palou Redorta J, et al. EAU guidelines on non-muscle-invasive urothelial
592 carcinoma of the bladder: update 2013. European urology 2013; 64:639-53.
593 64. Blin N, Stafford DW. A general method for isolation of high molecular weight DNA from
594 eukaryotes. Nucleic acids research 1976; 3:2303-8.
595 65. Lauss M, Aine M, Sjodahl G, Veerla S, Patschan O, Gudjonsson S, Chebil G, Lovgren K, Ferno
596 M, Mansson W, et al. DNA methylation analyses of urothelial carcinoma reveal distinct epigenetic
597 subtypes and an association between gene copy number and methylation status. Epigenetics :
598 official journal of the DNA Methylation Society 2012; 7:858-67.
599 66. Fryer AA, Emes RD, Ismail KM, Haworth KE, Mein C, Carroll WD, Farrell WE. Quantitative,
600 high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma
601 homocysteine and birth weight in humans. Epigenetics : official journal of the DNA Methylation
602 Society 2011; 6:86-94.
603 67. Wessely F, Emes RD. Identification of DNA methylation biomarkers from Infinium arrays.
604 Frontiers in genetics 2012; 3:161.
605 68. Glossop JR, Emes RD, Nixon NB, Haworth KE, Packham JC, Dawes PT, Fryer AA, Mattey DL,
606 Farrell WE. Genome-wide DNA methylation profiling in rheumatoid arthritis identifies disease-
607 associated methylation changes that are distinct to individual T- and B-lymphocyte populations.
608 Epigenetics : official journal of the DNA Methylation Society 2014; 9:1228-37.

609 69. Sen A, Heredia N, Senut MC, Hess M, Land S, Qu W, Hollacher K, Dereski MO, Ruden DM.
610 Early life lead exposure causes gender-specific changes in the DNA methylation profile of DNA
611 extracted from dried blood spots. Epigenomics 2015; 7:379-93.
612 70. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data.
613 Bioinformatics 2002; 18:207-8.
614 71. Khatri P, Draghici S. Ontological analysis of gene expression data: current tools, limitations,
615 and open problems. Bioinformatics 2005; 21:3587-95.
616 72. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium
617 thiocyanate-phenol-chloroform extraction. Analytical biochemistry 1987; 162:156-9.
618 73. Dudley KJ, Revill K, Whitby P, Clayton RN, Farrell WE. Genome-wide analysis in a murine
619 Dnmt1 knockdown model identifies epigenetically silenced genes in primary human pituitary
620 tumors. Molecular cancer research : MCR 2008; 6:1567-74.
621 74. Al-Azzawi H, Yacqub-Usman K, Richardson A, Hofland LJ, Clayton RN, Farrell WE. Reversal of
622 endogenous dopamine receptor silencing in pituitary cells augments receptor-mediated apoptosis.
623 Endocrinology 2011; 152:364-73.
624 75. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time
625 quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25:402-8.
626 76. Kitchen MO, Yacqub-Usman K, Emes RD, Richardson A, Clayton RN, Farrell WE. Epidrug
627 mediated re-expression of miRNA targeting the HMGA transcripts in pituitary cells. Pituitary 2015.
628 77. Raponi M, Zhang Y, Yu J, Chen G, Lee G, Taylor JM, Macdonald J, Thomas D, Moskaluk C,
629 Wang Y, et al. Gene expression signatures for predicting prognosis of squamous cell and
630 adenocarcinomas of the lung. Cancer research 2006; 66:7466-72.
631 78. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. Journal of molecular
632 biology 1987; 196:261-82.

633

634

635

636 **Figure Legends**

637

638 **Figure 1. Array filtering steps.** Summary of the steps implemented for the identification of
639 CpGs hypermethylated in HG-NMIBC. The initial filtering steps (*) included exclusion of non-
640 significant probe data, probes with missing data and probes located on allsomes.
641 RefSeq (National Center for Biotechnology Information Reference Sequence Database).
642 CpG island based upon the UCSC genome browser definition from Gardiner-Garden and Frommer[78].

643

644 **Figure 2. Unsupervised hierarchical clustering analysis of the 1,057 gene promter-**
645 **associated hypermethylated CpGs in HG-NMIBC.** Heatmap and dendrogram of
646 differentially methylated gene promoter-associated CpG sites identified by array analysis.
647 The dendrogram above the heatmap separates normal bladder (green bar, *n*=3) and high-
648 grade-NMIBC bladder tumours (red bar, *n*=21). Each row represents an individual CpG
649 locus, and each column represents a normal control or tumour sample (listed beneath the
650 heatmap). The colour scale beneath the heatmap represents methylation status:
651 unmethylated is yellow (β-value=0.0), and fully methylated is blue (β-value=1.0).

652

653 **Figure 3. Heatmap for 25 hypermethylated gene promoter-associated CpG islands.**
654 Pyrosequencing validation of 25 gene promoter-associated CpG islands, identified as
655 frequently differentially methylated in high-grade tumours by 450k BeadChip-array analysis.
656 As indicated above the heatmap, the four normal bladder controls are presented to the left-
657 side of the heatmap, followed by 18 low-intermediate-grade tumours, and 51 high-grade
658 tumours (the combined discovery and investigation cohorts). Each row represents the
659 promoter-associated CpG island of the indicated gene, and each colour block the *mean level*
660 *of methylation* across the island. The colour scale beneath the heatmap represents
661 methylation status: unmethylated is green (0.0% methylation), and fully methylated is red
662 (100.0% methylated).

663 **Figure 4. Mean levels of methylation in high-grade tumours relative to low-**
664 **intermediate-grade tumours and normal bladder.** Top ten genes showing an increase in
665 mean level of methylation (solid red bar) in high-grade tumours (HG, *n*=51) relative to low-
666 intermediate-grade tumours (LG, *n*=18) and in comparison to normal bladder controls (C,
667 *n*=4). Each individual control or tumour sample is shown as an unfilled blue circle. Significant
668 differences in the mean levels of methylation between the low-intermediate- and high-grade
669 tumours, or between control and low-intermediate-grade tumours, are indicated by *, $p<0.05$,
670 or **, $p<0.005$ (Student's T-test).
671
672 **Figure 5. Association of methylation with gene transcript expression in HG-NMIBC.**
673 Tumour transcript expression in unmethylated (UM, unfilled circles) and methylated (M, filled
674 circles) high-grade tumours, relative to normal bladder control (C, unfilled triangles) for the
675 four genes showing significant Spearman's correlation coefficients between promoter
676 methylation and gene expression (*PON3*, *STAT5a*, *VAX2* and *ARHGEF4*; $p=0.0006$,
677 $p=0.005$, $p=0.013$ and $p=0.0007$, respectively). The double-headed arrow represents the
678 threshold for 3-fold reduced expression relative to the mean of the normal controls (solid
679 blue bar); expression at or below this threshold signifies reduced expression in tumour
680 samples.
681
682
683

684 **Tables**

685

| | METHYLATION FREQUENCY | | | MEAN LEVEL OF METHYLATION | | |
|---|---|---|---|---|---|---|
| **Gene Symbol** | **High-grade** | **Low-intermediate-grade** | | **High-grade** | **Low-intermediate-grade** | |
| | *Number (%)* | *Number (%)* | *P* value | *(%)* | *(%)* | *P* value |
| *ATP5G2* | 37/51 (72.5) | 6/18 (33.3) | **0.005** | 51.04 | 30.20 | **0.029** |
| *VAX2* | 13/51 (25.5) | 0/18 (0.0) | **0.015** | 32.31 | 19.56 | **0.004** |
| *IRX1* | 37/51 (72.5) | 8/18 (44.4) | **0.045** | 49.47 | 38.70 | 0.067 |
| *INSRR* | 29/51 (56.9) | 5/18 (27.8) | 0.054 | 24.06 | 24.06 | **0.028** |
| *IRF8* | 25/51 (49.0) | 4/18 (22.2) | 0.057 | 26.13 | 17.99 | 0.157 |
| *PRDM14* | 45/51 (88.2) | 12/18 (66.7) | 0.066 | 60.14 | 46.06 | **0.029** |
| *VSX1* | 44/51 (86.3) | 12/18 (66.7) | 0.086 | 56.37 | 38.26 | **0.0004** |
| *TFAP2b* | 22/51 (43.1) | 4/18 (22.2) | 0.160 | 32.25 | 17.68 | **0.047** |
| *PRRX1* | 27/51 (52.9) | 7/18 (38.9) | 0.413 | 47.03 | 34.36 | **0.041** |
| *HIST1H4F* | 42/51 (82.4) | 13/18 (72.2) | 0.496 | 59.46 | 41.91 | **0.017** |

686

687 **Table 1. Genes showing the greatest methylation increase in high-grade relative to**
688 **low-intermediate-grade NMIBC tumours.** Top ten genes showing an increase in frequency
689 of methylation (left side of table), and/or an increase in mean level of methylation (right side
690 of table) in high-grade tumours relative to low-intermediate-grade tumours. For the left side
691 of the table, the number and proportion of tumours methylated are displayed for the low-
692 intermediate- and high-grade cohorts, with *p*-value (Fishers exact, $p<0.05$ significant). For
693 the right side of the table, the mean level of methylation across the low-intermediate- and
694 high-grade tumour cohorts are displayed with p-value (Student's T-Test, $p<0.05$ significant).
695 Statistically significant p-values are displayed in bold.

696 **Supplemental Data**

697

698 **Figure S1. Technical validation of 450k BeadChip-array data.** Correlation between array-
699 derived β-values (*x*-axis) and methylation percentage as determined by Pyrosequencing (*y*-
700 axis) for 5 CpGs (cg07778029, cg14456683, cg01227537, cg05661282 and cg26465391)
701 across 24 samples is shown. Spearman-rank correlation coefficient r=0.912; $p<0.00001$.

702

703 **Table S1. Sample characteristics.**

704

705 **Table S2. List of 256 differentially methylated genes.**

706

707 **Table S3. Methylation in discovery and investigation high-grade tumour cohorts.**

708

709 **Table S4. Frequency and mean levels of methylation in 25 genes for high- and low-**
710 **intermediate-grade tumours.**

711

712 **Table S5. Gene Ontology and KEGG pathway annotation lists.**

713

714 **Table S6. Primer sequences.**

715

716

717