Comparative methylome analysis of benign and malignant peripheral nerve sheath tumours

Andrew Feber, Gareth Wilson, Lu Zhang, et al.

Genome Res. published online February 1, 2011
Access the most recent version at doi:10.1101/gr.109678.110

Comparative methylome analysis of benign and malignant peripheral nerve sheath tumours

Andrew Feber[1], Gareth Wilson[1], Lu Zhang[2], Nadege Presneau [3], Bernadine Idowu[3,4], Thomas A. Down[5], Vardhman K. Rakyan[6], Elia Stupka[7], Vassia Schiza[1], Andrew E. Teschendorff[1], Gary P. Schroth[2], Adrienne Flanagan[3,4], Stephan Beck[1]

[1]Medical Genomics, UCL Cancer Institute, University College London, Huntley Street, London WC1E 6BT, UK.
[2]Illumina Inc. 9885 Towne Centre Drive, San Diego, CA 92121 USA
[3]Sarcoma Genetics, UCL Cancer Institute, University College London, Huntley Street, London WC1E 6BT, UK.
[4]Royal National Orthopaedic Hospital, Stanmore, Brockly Hill, Middlesex HA7 4LP, UK
[5]Wellcome Trust Cancer Research UK Gurdon Institute, and Department of Genetics, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK.
[6]Institute of Cell and Molecular Science, Barts and The London, 4 Newark Street, London, E1 2AT, UK.
[7]Comparative Genomics, UCL Cancer Institute, University College London, Huntley Street, London WC1E 6BT, UK.

Corresponding authors: Andrew Feber (a.feber@ucl.ac.uk), Adrienne Flanagan (a.flanagan@ucl.ac.uk) and Stephan Beck (s.beck@ucl.ac.uk)

Abstract

Aberrant DNA methylation (DNAm) was first linked to cancer over 25 years ago. Since then, many studies have associated hypermethylation of tumour suppressor genes and hypomethylation of oncogenes to the tumourigenic process. However, most of these studies have been limited to the analysis of promoters and CpG islands (CGIs). Recently, new technologies for whole-genome DNAm (methylome) analysis have been developed, enabling unbiased analysis of cancer methylomes. Using MeDIP-seq, we report a sequencing-based comparative methylome analysis of malignant peripheral nerve sheath tumours (MPNST), benign Neurofibromas and normal Schwann cells. Analysis of these methylomes revealed a complex landscape of DNAm alterations. In contrast to what has been reported for other tumour types, no significant global hypomethylation was observed in MPNST using methylome analysis by Medip-seq. However, a highly significant ($P<10^{-100}$) directional difference in DNAm was found in satellite repeats, suggesting these repeats to be the main target for hypomethylation in MPNST. Comparative analysis of the MPNST and Schwann cell methylomes identified 101,466 cancer-associated differentially methylated regions (cDMRs). Analysis showed these cDMRs to be significantly enriched for two satellite repeat types (SATR1 and ARLα) and suggests an association between aberrant DNAm of these sequences and transition from healthy cells to malignant disease. Significant enrichment of hypermethylated cDMRs in CGI shores ($P<10^{-60}$), non-CGI-associated promoters ($P<10^{-4}$) and hypomethylated cDMRs in SINE repeats ($P<10^{-100}$) was also identified. Integration of DNAm and gene expression data showed that the expression pattern of genes associated with CGI shore cDMRs was able to discriminate between disease phenotypes. This study establishes MeDIP-seq as an effective method to analyse cancer methylomes.

Introduction

The development of human cancer is driven by changes in the genetic and epigenetic landscape of the cell. It is now clear that epigenetic changes, such as the hypermethylation of tumour suppressor genes and the hypomethylation of oncogenes are hallmarks of cancer, and play a key role in the development and maintenance of the malignant phenotype (Baylin et al. 2006; Issa 2004; Laird 2005). Epigenetic alterations also appear to be some of the earliest abnormalities occurring in the tumourigenic process, and have the potential to predispose stem/progenitor cells to subsequent genetic alterations involved in tumour development and progression (Feinberg et al. 2006) and, therefore, may represent good markers for early disease diagnosis. The role of aberrant DNAm in the development and progression of human cancer has been studied in several tumours such as breast, prostate, lung and colorectal cancers (Barault et al. 2008; Ordway et al. 2007; Santourlidis et al. 1999; Vaissiere et al. 2009). However, until recently the majority of studies have focused on those regions where DNA methylation (DNAm) was assumed to have the greatest functional significance in the regulation of gene expression, such as promoters and CGIs. This has led to the identification of many genes where differential DNAm in CGIs results in aberrant gene expression and

association with disease phenotype. However, as only 7% of CpGs reside within CGIs, it is highly likely that many potentially informative sites have not yet been analysed (Rollins et al. 2006). This was recently confirmed by the identification of CGI shores as key DNAm gene regulatory sites (Doi et al. 2009; Irizarry et al. 2009), and suggests a truly comprehensive map of the DNAm landscape in human malignancies is yet to be defined.

Recently, next-generation sequencing technologies have emerged as powerful tools to allow, whole-genome profiling of epigenetic modifications, including DNAm (Down et al. 2008; Lister et al. 2009). MeDIP-seq, for instance, combines methylated DNA immunoprecipitation (MeDIP) with next-generation sequencing, provides an efficient approach for whole-genome DNAm (methylome) analysis and has been used to generate the methylome of human spermatozoa (Down et al. 2008). Similarly, whole-genome bisulphite sequencing (BS-seq) has recently been used to generate the first single-base resolution profiles of DNAm in human embryonic stem cells and differentiated cells (Laurent et al. 2010; Lister et al. 2009).

Neurofibromas are benign tumours that originate from Schwann cells in a peripheral nerve sheath. Neurofibromatosis type 1 (NF1) is predominantly an autosomal dominant neurocutaneous disorder affecting 1 in 3000 individuals world-wide and is caused by germline mutations of the *NF1* gene (Levy et al. 2004). However, it also occurs sporadically. Approximately 5-10% of patients with NF1 develop malignant peripheral nerve sheath tumours (MPNST) (Evans et al. 2002). Although the incidence of MPNST is relatively low, only 43% of patients with this diagnosis survive beyond 5 years (Evans et al. 2002; Zhou et al. 2003). We still do not understand the epigenetic or genetic alterations driving the development of MPNST, or the molecular markers which are associated with the development and progression of this disease. So far, few studies have assessed the DNAm state of MPNSTs and most of those have been restricted to the *NF1* gene itself (Harder et al. 2004) or selected CGIs and promoters (Gonzalez-Gomez et al. 2003). Nevertheless, these studies have identified differential DNAm of several genes thought to be involved in MPNST development and progression, including *CDKN2A* and *MSH2* (Kawaguchi et al. 2006; Titze et al. 2010). To provide a more comprehensive catalogue of DNAm differences in MPNST, we performed unbiased methylome analysis of benign and malignant NF samples as well as controls.

Results

Methylome analysis was conducted on three pools of genomic DNA extracted from 10 MPNST, 10 NF and 6 normal non-neoplastic Schwann cell (SC) samples. Schwann cell samples were derived from biopsies and expanded in culture. Using MeDIP-seq (Down et al. 2008), we generated on average 65 million paired-end reads per cohort pool of which 72% could be uniquely aligned to the human genome (build NCBI36) (Supplementary Table 1). Absolute DNAm values were inferred for 100 bp windows using the Batman algorithm (Down et al. 2008). To correct for potential bias in DNAm differences due to changes in copy number, sequence read counts were normalized for copy number variation (CNV) prior to Batman processing. CNV

profiles were generated using the Affymetrix Human SNP 6.0 microarray (Supplementary Figure 1) (Materials and Methods). Due to the study design which, in this case, involved sample pooling, it is difficult to estimate the contribution made by individual patients to specific CNVs. However, any CNV observed will have to be present in a significant proportion of the samples in order to be identified in the first place. The fact that we were able to identify known CNVs associated with MPNST (e.g. loss of *NF1* and *CDKN2A*) confirmed our ability to detect common CNVs. After normalisation of DNAm data for CNVs there was no significant difference in the number of DMRs found within regions of CNV gain or loss compared to control regions. Based on this methylome analysis, we covered on average 67% (range 65%-70%) of the approximately 26.7 million CpG sites in the haploid autosomal human genome reference sequence for each of the three methylomes (Supplementary Figure 2) (sex chromosomes were removed from the analysis because of the pooled study design). Furthermore, saturation analysis indicates we have sufficient reads to give reproducible genome-wide methylation profiles (Chavez et al. 2010) (Supplementary Figure 3). We validated our DNAm data for each cohort pool (R^2=0.78 for MPNST, R^2=0.80 for NF, R^2=0.77 for SC) using the Infinium Human Methylation 27K BeadArray (Bibikova et al. 2009) (Supplementary Figure 4). Although slightly lower, these results are consistent with previously reported correlations between BeadArray and bisulphite sequencing (R^2 = 0.82) and between MeDIP-seq and bisulphite sequencing (R^2 = 0.82) (Bibikova et al. 2009; Down et al. 2008; Gu et al. 2010; Smith et al. 2009). In addition, we further validated our MeDIP-seq data by confirming the DNAm status of 3 selected cDMRs (R^2 = 0.97) using targeted pyrosequencing (Supplementary Figure 5) and using bisulphite sequencing (BS-seq) which had a concordance of 77% with Batman methylation scores across 64943 CpG sites.

We first performed a comprehensive global analysis of the three methylomes to identify major directional changes in DNAm. This does not include balanced and un-directional changes in DNAm which will be described in the DMR section. Batman DNAm scores for individual CpG sites were categorised into three DNAm states: low DNAm (DNAm score <40%), intermediate (DNAm of 40-60%) and high DNAm (DNAm of >60%). As shown in Figure 1, there was only a 0.7% change in global DNAm in MPNST compared to NF and SC control. This finding is in contrast to our current understanding based on other cancer types, that cancer genomes undergo global hypomethylation with an average reduction in DNAm of between 10-20% between cancer and normal (Ehrlich 2002; Feinberg et al. 1983; Gama-Sosa et al. 1983). We then analysed CpG DNAm in a further 16 distinct genomic features, including CGI and other regulatory regions, exons, introns, and the repeat families annotated in the Ensembl database (Figure 1). This analysis revealed that the main target of differential DNAm in MPNST were satellite repeats (Figure 1), with a significant ($P < 10^{-100}$) increase in the percentage of CpG sites with low DNAm within satellite repeats between SC (33%) and NF (39%) and MPNST (49%). Compared to this, the level of global DNAm across the most commonly studied and usually unmethylated CGIs (Baylin et al. 2006) remained quite similar for all three phenotypes (MPNST (12%), NF (10%) and SC (11%)) (Figure 1B). Little change in the global DNAm of other regulatory features

such as CGI shores (Figure 1C) and promoters (Figure 1D) as well as genic features (Figure 1E-H) was observed.

One of the most commonly cited feature of cancer epigenomes concerns directional hypomethylation of repeats (Ehrlich 2002; Wilson et al. 2007). Historically, repeat sequences proved to be refractory to many analysis approaches (particularly array-based approaches) but this was not a problem for the MeDIP-seq approach used here, which provided excellent coverage of all major repeat types. In contrast to previous studies in other cancers (Dante et al. 1992; Jurgens et al. 1996; Santourlidis et al. 1999), we did not observe global hypomethylation of LINE repeats (Figure 1J). In fact, we saw a decrease (8%) in the proportion of low DNAm (<40%) at CpGs between MPNST and SC. Interestingly, analysis of the different subtypes of the LINE repeat family revealed different patterns of global DNAm profiles, with L2 and L1 repeats showing an increase in the proportion of highly methylated (>60% DNAm) CpGs, where as the L3 and L4 LINE repeats showed a decrease in the proportion of highly methylated (>60% DNAm) CpGs in MPNSTs (Figure 1Q-T). SINE repeat elements (Figure 1K) showed a small 6% decrease in the number of highly methylated CpGs between MPNST and SC (Table 1). However, the starkest difference in global DNAm was observed in satellite repeats (Figure 1L), with a large change in the proportion of methylated CpGs seen during transition from the normal to benign to malignant states. Although the pooling strategy employed in this study somewhat limited the sensitivity at which rare DNAm changes could be detected, we already showed above that we were able to detect common changes which hold the best potential for future diagnostic and prognostic markers.

Next, we carried out pair-wise comparisons between the three methylomes to identify DMRs associated with progression from non-neoplastic SC to benign NF (n2bDMRs), progression from benign NF to malignant MPNST (b2mDMRs), cDMRs between SC and MPNST and tissue-specific DMRs (tDMRs) between MPNST and previously determined tDMRs between germline and somatic tissues (Down et al. 2008; Irizarry et al. 2009). Since DNAm at neighbouring CpG sites are highly correlated up to 1000bp (Eckhardt et al. 2006), DMRs were defined as the average Batman DNAm score over 1kb windows with a minimum difference of ≥33% (equivalent to the 95th percentile of DNAm scores) in Batman DNAm score between samples (Down 2008). Each 1kb window was subsequently mapped back to its nearest corresponding genomic feature. Using this approach, we identified 101,466 unique cDMRs with an FDR of 3.7%. These cDMRs can further be subdivided into 48,391 hypermethylated and 53,075 hypomethylated cDMRs (Supplementary Figure 2, Table 1). Interestingly, the lowest number of cDMRs was identified in the NF/MPNST comparison (b2mDMRs), suggesting the benign methylome to be more closely related to the malignant than the normal methylome (Supplementary Figure 2). Changing this threshold from 95th to 90th and 99th percentile of DMR scores (difference in Batman DNAm score of ≥43% and ≥27%, respectively) showed that the numbers of DMRs associated with different genomic features changed proportionally with the change in threshold, suggesting that the noise is evenly distributed across different genomic features.

We subsequently mapped all DMRs to their nearest genomic feature (Table 2), and performed an enrichment analysis on the number of DMRs in unique genomic features. Analysis of these DMRs showed significant ($P > 10^{-3}$; range $10^{-3} – 10^{-60}$) enrichment of hypermethylated cDMRs (Figure 2) in all three sample types (defined by cDMRs n2bDMRs and b2mDMRs) in non-CGI associated promoters, CGI shores, LTRs and LINEs but not in CGIs. Conversely, hypomethylated DMRs were enriched in non-CGI associated promoters, CGI shores, SINEs and satellites but again, not in CGIs (Figure 2). Although the observed number of 412 CGI associated cDMRs (1.4% of total) was statistically not significant, this does not exclude them from having functional relevance. Of all the DMRs, only CGI shore-associated DMRs showed both highly significant gain and loss of DNAm, emphasizing the special role of DNAm at this genomic feature. The majority of the CGI-associated cDMRs were hypermethylated (367 or 89%). Interestingly only 11% (44 hypermethylated and 3 hypomethylated) of cDMRs were within 3kb of the nearest annotated gene. Analysis of gene promoter regions identified a similar number of hyper- and hypomethylated cDMRs (892 hypermethylated, 817 hypomethylated) (Table 2) (Supplementary Figure 6). When promoters were subdivided according to association or non-association with CGIs, we observed significant enrichment of both hypermethylated ($P = 9.9\text{x}10^{-4}$) and hypomethylated cDMRs in non-CGI associated promoters compared with no enrichment of CGI-associated promoters (Figure 2). No enrichment of promoter-associated DMRs in bivalent chromatin domains was observed. Similarly no enrichment of DMRs containing microRNAs was observed. These data are consistent with the recent finding that the majority of differential DNAm occurs outside CGIs (Irizarry et al. 2009). Consequently, many previous studies using CGI-biased profiling approaches have only assayed a fraction of the differential DNAm potential in the tumourigenic epigenome.

As expected, the majority of cDMRs mapped to repeat elements (Table 2). The propensity for hypermethylation of LINE repeat elements identified in the global analysis persisted (Figure 3), with the enrichment of hypermethylated cDMRs ($P < 10^{-100}$), n2bDMRs ($P = 6.47\text{x}10^{-32}$) and b2mDMRs ($P = 4.33\text{x}10^{-37}$) in both intronic and non-intronic LINE repeats (Figure 3). Intronic and non-intronic LINE cDMRs showed the largest level of enrichment and n2bDMRs the lowest, raising the possibility that hypermethylation of LINE repeats increases with disease progression. Contrary to LINE repeats, the other major class of repeats, SINEs, showed significant enrichment of hypomethylated DMRs (cDMRs ($P < 10^{-100}$), n2bDMRs ($P < 10^{-100}$) and b2mDMRs ($P = 2.48\text{x}10^{-8}$)) (Figure 3).

Satellite repeats showed significant enrichment of hypomethylated DMRs, particularly those repeats in non-intronic regions (cDMRs ($P < 10^{-100}$) and b2mDMRs ($P < 10^{-100}$)) (Figure 3 and Figure 4). Of the 1,447 satellite repeat cDMRs, 1,384 (107 hypermethylated, 1241 hypomethylated) mapped to regions outside of CNV. Only 13 of the 1241 hypomethylated satellite cDMRs mapped to regions of genomic loss. These data suggest that hypomethylation of satellite repeats observed in the MPNST epigenome is not due to changes

in the underlying genomic copy number of these loci. It is also possible that hypomethylation events, particularly in repeat regions, may be under called as a result of the removal of multi-mapping reads prior to methylome analysis. However, as part of our methylome analysis similar proportions of multi-mapping reads were removed from both the normal and malignant methylomes (20.43% and 23.23% respectively), suggesting that we did not under-call hypomethylation events. Further analysis of these DMRs according to satellite repeat subtypes showed significant enrichment of hypomethylated cDMRs in ARLα (P = 8.7x10^{-65}) and SATR1 (P = 3.7x10^{-9}) satellite repeats. Interestingly, we observed a progressive increase in the enrichment of hypomethylated ALRα DMRs during progression from non-neoplastic to malignant disease, whereas DMRs in SATR1 repeats appeared enriched to a similar level in n2bDMRs and cDMRs with no enrichment in b2mDMRs, suggesting that the hypomethylation of SATR1 repeats may be an early event in the tumourigenic process. These data suggest that specific subtypes of satellite repeats undergo preferential hypomethylation during tumourigenic development.

The DMR analysis revealed 3690 genes potentially involved in MPNST development and progression (Supplementary Table 2). For example, the hypermethylation of the promoter region of *MEST*, a target of both loss of imprinting and hypermethylation in multiple tumour types including breast cancer and glioblastoma (Pedersen et al. 2002)' (Martinez et al. 2009). Intriguingly, this study also identified the hypermethylation of the promoter region of the Wilms Tumour gene, *WT1*, a tumour suppressor gene associated with the development of several tumour types (Chen et al. 2003; Pelletier et al. 1991). Consistent with previous studies of DNAm in MPNST, we also found no difference in DNAm at the *NF1* promoter (Fishbein et al. 2005; Harder et al. 2004; Luijten et al. 2000), although we observed the heterozygous loss of the *NF1* locus in the MPNST samples (Supplementary Figure 1). It has recently been suggested that DNAm may impinge on the expression of microRNAs (Lujambio et al. 2007). Although we did not observe any enrichment of aberrantly methylated miRNAs, several interesting candidate miRNAs were shown to be differentially methylated. For example, *miR-124* mapped to a hypermethylated cDMR. This miRNA has recently been shown to be a predictive DNAm biomarker for high-grade disease in cervical cancer (Lujambio et al. 2007; Wilting et al. 2010). In addition, we also observed hypomethylation of *miR-30*, which has been shown to be up-regulated in a variety of both haematological and solid malignancies (Volinia et al. 2010).

Recently, it has been shown that tDMRs frequently coincide with cDMRs (Irizarry et al. 2009). Comparison with these previously determined DMRs showed that 15% of the tDMRs and 14% of cDMRs were also sites of differential DNAm in MPNST, including hypomethylated oncogenes *YES1, MATK, E2F3, EGFR* and *AFF1* and hypermethylated tumour suppressor genes *TSC2, RUNX1, FOXD3 and RASSF1A* among others.

To assess if the genes associated with cDMRs were enriched for certain pathways or networks, we performed gene ontology (GO) analysis (multiple

test corrected $P < 0.01$). This analysis showed hypermethylated promoter-associated cDMRs to be enriched for genes involved in the regulation of transcription factor activity, DNA binding and cell surface receptor signalling and hypomethylated cDMRs to be enriched for genes involved in differentiation and development, particularly brain and neural devolvement. Hypermethylated CGI shore cDMRs showed enrichment for genes involved in cell adhesion and regulation of transcription, along with genes involved in neurogenesis. The latter was also observed for CGI shore n2bDMRs, suggesting that genes in this ontology may be involved in the early development of MPNST.

Finally, we explored the relationship between differential DNAm and gene expression, by integrating our DNAm data with published gene expression data from NF and MPNST (Henderson et al. 2005; Miller et al. 2009). We next asked whether gene expression correlated sufficiently with aberrant DNAm to discriminate between benign and malignant disease phenotypes. We performed partition clustering of the MPNST and NF samples using the expression of genes associated with b2mDMRs in CGIs, CGI shores or promoters. The expression patterns of genes associated with CGI shore hypermethylated and hypomethylated b2mDMRs were able to discriminate between MPNST and NF phenotypes, significantly better ($P = 0.0001$) than expected by chance (Figure 5). We also observed significant discrimination between MPNST and NF samples based on the expression of genes associated with hypermethylated non-CGI associated promoter b2mDMRs ($P = 0.0003$). No association was seen between gene expression and CGI-associated promoter b2mDMRs (hypermethylation $P = 0.1057$, hypomethylation $P = 0.129$) or hypomethylated promoter-associated b2mDMRs ($P = 0.797$).

Integration of gene expression data also allowed the identification of several candidate genes involved in the development and progression of MPNST. For example, we observed hypermethylation and concomitant reduction in gene expression of the tumour suppressor genes *SOX10* ($P = 0.00054$) (Figure 6C) (Miller et al. 2006). We also observed reduced expression and hypermethylation of *CDKN2A* ($P = 0.00004$), implying that the suppression of the *CDKN2A* locus in these tumours may be driven by increased DNAm (Figure 6D) as well as genomic loss (Supplementary Figure 1) in MPNST(Nielsen et al. 1999).

Discussion

Many studies have shown that epigenetic changes play an important role in cancer development (Ehrlich 2002). However, our knowledge of how changes in DNAm impinge on the neoplastic genome is largely still limited to its effect on CGIs. In order to elucidate the impact of aberrant DNAm on cancer development, we have carried out a whole-genome comparative methylome analysis of malignant peripheral nerve sheath tumours, benign neurofibroma and normal non-neoplastic Schwann cells. This analysis reveals a complex pattern of epigenetic changes during disease progression, the majority of which occur outside of regions previously assumed to be the predominant regions involved in the epigenetic regulation of cancer.

Based on our methylome analysis we identified satellite repeats to be the predominant target of aberrant DNAm in MPNST. Although aberrant repeat DNAm has been observed before (Weisenberger et al. 2005; Wilson et al. 2007), it has previously been challenging to assess the DNAm state of individual repeats at the whole-genome level due to the technical and physical limitations of microarray- and PCR-based assays. Digressing from the current understanding of repeat DNAm in human malignancies, our data indicate that the statistically most significant change in repeat DNAm is seen in satellite repeats rather than LINE and SINE repeats. Furthermore, our analysis revealed two unique satellite subtypes to be the main targets of hypomethylation, ARLα and SATR1 which are predominantly located in non-intronic regions of the genome.

A similar increase in hypomethylation was also observed in benign neoplasms compared to non-neoplastic Schwann cells. Analysis of the DNAm of satellite repeats located in non-intronic compared to intronic regions, suggested that changes in global satellite DNAm occur predominately in non-intronic regions. These data suggest that aberrant DNAm of satellites may be an early and essential step in the tumourigenic process (Saito et al. 2001; Wong et al. 2001). The distinct DNAm patterns associated with different types of satellite repeats suggests that either regional or sequence-specific determinants influence the propensity of satellite repeats to undergo aberrant DNAm. It is therefore interesting to note that specific satellite repeat sequences appear to be targeted by specific members of the DNMT methyltransferase family (Robertson et al. 2004). However, as genomic instability does not appear to be a common feature in MPNSTs (Kobayashi et al. 2006; Serra et al. 1997), aberrant satellite DNAm may result in altered expression of nearby genes due to changes in chromatin accessibility by transcription factors (Cadieux et al. 2006). It will be interesting to investigate further whether the distinct pattern of satellite repeat DNAm observed in MPNSTs also occurs in other tumour types, whether it is under sequence-specific control, and how it affects the expression of nearby gene.

In addition to satellite repeats, the other major target for aberrant DNAm in MPNSTs are CGI shores, confirming the involvement of CGI shores in a second lineage of human cancers in addition to colorectal cancer (Irizarry et al. 2009). Previous studies have suggested that these regions surrounding CGIs play a more significant role in the regulation of gene expression than CGIs themselves (Doi et al. 2009; Irizarry et al. 2009). The balanced number of hypermethylated and hypomethylated cDMRs in CGI shores may further imply they are functionally implicated in MPNST development, allowing both the activation of oncogenes and suppression of tumour suppressor genes. In support of such a possible functional role for CGI shores, integration of gene expression data showed that the aberrant DNAm of CGI shores had a stronger association with gene expression than aberrant DNAm of CGIs. Although the independent expression data correlate well with our pooled methylation data, the conclusions drawn from the integration of the two data sets data must be viewed in the context that both data sets were derived from different individuals and other explanations may exist. Furthermore, the high

degree of overlap between DMRs from other cancer types (Irizarry et al. 2009) suggests that these regions may constitute core sites of aberrant DNAm involved in the development and progression of human malignancies, a concept consistent with the epigenetic progenitor model of cancer (Feinberg et al. 2006).

In summary, we have performed a comparative methylome analysis of normal Schwann cells along with benign and malignant nerve sheath tumours. Our data show the suitability of MeDIP-seq to investigate cancer methylomes and to identify novel sites of differential DNAm which constitute potential biomarkers for disease development and progression in human malignancies. However, further work is necessary to assess the biological function and clinical utility of such epigenetic biomarkers in the development and progression of MPNST.

Methods

Clinical samples

This study complies with Central Office for Research Ethics Committees standards (REC 07/Q0506/8). The material, snap-frozen tissue, was obtained from the histopathology department of The Royal National Orthopaedic Hospital. The clinical data was retrieved from the clinical notes. The cohort studied consisted of neurofibromas (NF) with NF1 from 10 patients, Malignant Peripheral Nerve Sheath tumours (MPNSTs) from 10 patients.

Primary human Schwann cells (SC) were isolated from 10-15 cm length of human sciatic/brachial nerve from amputation amputations specimens from patient without NF1 or nerve sheath tumours. The protocol was adapted from Muir et al., (1994). Briefly, after removal of the perineurium, the nerve was cut into ~ 2mm transverse sections with a scalpel and placed in cell culture plates with hSC medium, containing DMEM/F12 supplemented with 2uM of Forskolin (Calbiochem), recombinant GGF (20 ng/ml, R&D Systems), 3% FCS (Invitrogen), Kanamycin Monosulfate (80ug/ml) (Sigma) and Gentamicin (40ug/ml) (Gibco). The plates were incubated at 37°C and 5% CO_2 and fed every 2-3 days for 2-3 weeks. Then the nerve fragments were dissociated overnight at 37°C degrees with a mixture of 300U/ml collagenase Type XI (Sigma) and 1.25U/ml of Dispase (Roche) containing antibiotics kanamycin/gentamicin. The single cell suspension was then plated into 1xPoly-L-Lysin/Laminin-coated tissue culture plates (Sigma) and incubated at 37°C 5% CO2 for another few days before CD271 (low-affinity nerve growth factor receptor, LNGFR) –positive cells were selected by MACS according to the manufacturer's instructions (Miltenyi Biotec Ltd., Surrey, UK). After selection nucleic acid was extracted.

MeDIP-seq

5ug of pooled DNA from each sample cohort was sonicated to between 50-350bp. Sonicated DNA was then subjected to Illumina's paired-end library preparation and MeDIP enrichment as previously described by Down et al.(Down et al. 2008). Next-generation sequencing (50bp paired-end reads) was performed on pooled libraries (size selected to be between 150-200bp for

MPNST and between 100-150 for NF and SC) for each pooled sample type. Failure of the MPNST 100-150bp library resulted in the larger library (150-200bp) being the sequenced for MPNST. To assess if different library sizes effected enrichment, we assessed the relative representation of fragments containing the same number of CpG sites. This showed the enrichment of fragments with the same number CpGs to be similar across samples (Supplementary Figure 5). Reads were mapped to the human genome reference sequence (Build 36) using the alignment software 'maq' (http://maq.sf.net/). Reads with a maq score of <10 were removed form subsequent downstream analysis. Absolute DNAm states were inferred from raw reads using the Batman algorithm as previously described by Down et al.(Down et al. 2008). The MeDIP-seq data have been deposited in GEO under accession number GSE21714.

BS-Seq

BS-seq was performed as previously described (Lister et al. 2009). 1ug of pooled DNA from MPNST was sonicated to between 50-350bp. Sonicated DNA was then subjected to Illumina's paired-end library preparation, with the exception that 5-methyl-cytosine containing paired-end adapters were used. Adapter ligated fragments were bisulphite converted using the EZ DNA methylation kit (Zymo Research) and subsequently size selected by agarose gel electrophoresis. Size selected libraries between 250-350bp were then subjected to 18 rounds of amplification prior to sequencing on the Illumina GAIIx. 75bp long raw reads were mapped using the Bismarck software (http://www.bioinformatics.bbsrc.ac.uk/projects/bismark/). DNAm states were inferred from raw reads using the SeqMonk package (http://www.bioinformatics.bbsrc.ac.uk/projects/seqmonk/). Concordance with Batman methylation score was assessed for those CpG sites covered by at least 10 reads.

Pyrosequencing

1ug each of pooled DNA from MPNST, NF and SC samples were bisulphite treated using EZ DNA methylation kit (Zymo Research). Bisulphite treated DNA was used for generating PCR amplified templates for pyrosequencing using target specific primers (Supplementary Table 3). 10 ul of the biotinylated PCR product were the sequenced according to the manufacture's recommendation. Pyrosequencing was carried on the PSQ HS 96 System and PyroMark MD System using Pyro Gold Reagent kits (Biotage). Methylation was quantified using Pyro Q-CpG Software that calculates the ratio of converted C's (T's) to unconverted C's at each CpG and expresses this as a percentage DNAm. Average DNAm across the target region was determined for each sample and compared to the Batman and Infinium DNAm scores.

CNV profiling

Pooled DNA from each of the sample cohorts were interrogated for CNVs using the SNP 6.0 Arrays (Affymetrix). Arrays were processed in accordance with the manufacturer's instructions. Changes in CNVs for each sample were inferred using a Hidden Markov Model algorithm (Zhao et al. 2004). Regions of gain were called using cut off 2.3 copies, and regions of loss with a cut-off

of 1.7 copies, regions were only called altered if a minimum of 10 loci were above (or below) the threshold. Raw MeDIP-seq read counts were normalised to correct for any bias introduced by changes in genomic copy number by dividing the sequence read count number by the change in copy number. Sex chromosomes were removed from the analysis because of the pooled study design

Global DNAm analysis

Changes in global DNAm were assessed by binning CpG DNAm levels into three states: low DNAm (<40% DNAm), intermediate DNAm (40%-60% DNAm) and high DNAm (>60% DNAm). The significance of differences in global DNAm was assessed using the Fisher test in Bioconductor (Gentleman et al. 2004).

DMR identification

To detect regions of differential methylation between disease phenotypes and non-neoplastic Schwann cell controls, we screened genomic windows of 1Kb. For each window the Batman methylation scores (100pb) were averaged and compared between sample cohorts. Using the distribution of differences between sample methylation scores, we defined a conservative threshold for calling DMRs which was based on the 95th percentile of the difference in methylation score. This resulted in a difference in methylation score of ≥33% between samples being considered as DMRs.

To assess the false discovery rate (FDR) in the DMR calling, we randomised the Batman DNAm scores for each sample and repeated the DMR analysis (10,000 permutations). This analysis gave an FDR of 3.7% across the genome.

To investigate if DMRs were enriched among certain structural genomic features we used the epitools R-package to compute odds ratios for any given structural genomic feature (e.g. CGI) against all other feature types. Significance of odds ratio values was calculated using median-unbiased estimation (mid-p). Since feature types differed greatly in terms of the total number of 1Kb windows that mapped to each of them, thus resulting in large differences in the expected variance for each genomic feature type, we computed and displayed an enrichment score, which measured the number of 1.96 sigma deviations from the null value of 1. Specifically, the relative enrichment score for hypermethylated (or hypomethylated) DMRs to be of a given feature type *f* is given by:

$$ER_f = \frac{OR_f - 1}{D_f / 2}$$

where *OR* is the odds ratio for feature type *f,* and *D* is the estimated 95% confidence interval (-1.96*sigma,+1.96*sigma). Thus, the variance is scaled out allowing for a more suitable graphical representation of the significance of OR values.

Gene ontology analysis

Gene ontology (GO) analysis was performed to find enriched categories using the GOstats package in Bioconductor (Gentleman et al. 2004). *P*-values were multiple test corrected in order to reduce false positive rates. GO terms with adjusted *P*-values of <0.01 were considered significant.

DNAm profiling using Infinium Human 27K BeadArrays

1ug of pooled DNA from MPNST, NF and non-neoplastic Schwann cells were bisulphite converted using the EZ DNA methylation kit (Zymo Research). The samples were then hybridised to 27K BeadArrays according to the manufacture's recommendation. The BeadStudio software (Illumina Inc) was used to infer methylation scores from image intensities. Two technical replicates of each sample were run and data averaged for subsequent analysis.

Gene expression analysis

Gene expression analysis was performed using publically available data (MPNST n=10, NF n=27) (Henderson et al. 2005; Miller et al. 2009). Gene expression data were derived from raw intensities using RMA (Irizarry et al. 2003). As the two data sets were from two different Affymetrix gene expression platforms U95 (Henderson et al. 2005) and U133Plus2 (Miller et al. 2009), only the common probe sets (22218) between the two platforms were used in the feature analyses. DMRs and gene expression data were associated by Ensembl ID number. Significant differential gene expression was assessed using the LImma package in Bioconductor (Gentleman 2004) and *P*-values were corrected for multiple testing. Clustering of gene expression data by DMR was carried out using the Cluster package in BioConductor(Gentleman et al. 2004). Significance between expression and sample phenotype was determined by carrying out 10,000 permutations based on the same number of genes from each DMR category.

Abbreviations

MPNST – Malignant peripheral nerve sheath tumour
SC – Schwann cell
NF – Benign neurofibroma
CGI - CpG island
DNAm – DNA methylation
CNV - Copy number variation
MeDIP-seq- methylated DNA immunoprecipitation sequencing
BS-seq - Bisulphite sequencing
NF1 - Neurofibromatosis type 1
SINE - Short interspersed repeat element
LINE - Long interspersed repeat element
LTR - Long terminal repeat
DMR – Differentially methylated region
n2bDMRs - DMRs between non-neoplastic SC and benign NF
b2mDMRs - DMRs between benign NF and malignant MPNST
cDMRs - Cancer-specific DMRs between non-neoplastic SC and malignant MPNST
tDMRs – Tissue-specific DMRs

Acknowledgements
AF and SB were supported by the Wellcome Trust (084071) and SB was also supported by a Royal Society Wolfson Research Merit Award. NP was supported by Skeletal Action Cancer Trust, Stanmore, UK. GW was supported by an EU grant (Highthroughput Epigenetic Regulatory Organisation in Chromatin (HEROIC), LSHG-CT-2005-018883) under the 6th Framework Programme to SB. AEF was supported by a Heller Research Fellowship. The experimental work was supported in part by the UCLH/UCL Comprehensive Biomedical Research Centre, Illumina Inc and the Wellcome Trust. We thank UCL Genomics for excellent technical support. We are also grateful to the patients who donated their samples for research and to the staff of the London Sarcoma Service who were involved in the care of the patients.

References

Barault L, Charon-Barra C, Jooste V, de la Vega MF, Martin L, Roignot P, Rat P, Bouvier AM, Laurent-Puig P, Faivre J, et al. 2008. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. *Cancer Res.* **68**:8541-8546

Baylin SB and Ohm JE. 2006. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat.Rev.Cancer*. **6**:107-116

Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, and Gunderson KL. 2009. Genome-wide DNA methylation profiling using Infinium assay. *Epigenomics* **1**:177-200

Cadieux B, Ching TT, VandenBerg SR, and Costello JF. 2006. Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. *Cancer Res.* **66**:8469-8476

Chavez L, Jozefczuk J, Grimm C, Dietrich J, Timmermann B, Lehrach H, Herwig R, and Adjaye J. 2010. Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. *Genome Res.* **20**:1441-1450

Chen CM, Chen HL, Hsiau TH, Hsiau AH, Shi H, Brock GJ, Wei SH, Caldwell CW, Yan PS, and Huang TH. 2003. Methylation target array for rapid analysis of CpG island hypermethylation in multiple tissue genomes. *Am.J.Pathol.* **163**:37-45

Dante R, nte-Paire J, Rigal D, and Roizes G. 1992. Methylation patterns of long interspersed repeated DNA and alphoid repetitive DNA from human cell lines and tumors. *Anticancer Res.* **12**:559-563

Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S, et al. 2009. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat.Genet.* **41**:1350-1353

Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM, et al. 2008. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat.Biotechnol.* **26**:779-785

Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat.Genet.* **38**:1378-1385

Ehrlich M. 2002. DNA methylation in cancer: too much, but also too little. *Oncogene*. **21**:5400-5413

Evans DG, Baser ME, McGaughran J, Sharif S, Howard E, and Moran A. 2002. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J.Med.Genet.* **39**:311-314

Feinberg AP, Ohlsson R, and Henikoff S. 2006. The epigenetic progenitor origin of human cancer. *Nat.Rev.Genet.* **7**:21-33

Feinberg AP and Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature.* **301**:89-92

Fishbein L, Eady B, Sanek N, Muir D, and Wallace MR. 2005. Analysis of somatic NF1 promoter methylation in plexiform neurofibromas and Schwann cells. *Cancer Genet.Cytogenet.* **157**:181-186

Gama-Sosa MA, Slagel VA, Trewyn RW, Oxenhandler R, Kuo KC, Gehrke CW, and Ehrlich M. 1983. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res.* **11**:6883-6894

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**:R80

Gonzalez-Gomez P, Bello MJ, Arjona D, Alonso ME, Lomas J, De Campos JM, Kusak ME, Gutierrez M, Sarasa JL, and Rey JA. 2003. Aberrant CpG island methylation in neurofibromas and neurofibrosarcomas. *Oncol.Rep.* **10**:1519-1523

Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, Gnirke A, Lander ES, and Meissner A. 2010. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat.Methods* **7**:133-136

Harder A, Rosche M, Reuss DE, Holtkamp N, Uhlmann K, Friedrich R, Mautner VF, and von DA. 2004. Methylation analysis of the neurofibromatosis type 1 (NF1) promoter in peripheral nerve sheath tumours. *Eur.J.Cancer.* **40**:2820-2828

Henderson SR, Guiliano D, Presneau N, McLean S, Frow R, Vujovic S, Anderson J, Sebire N, Whelan J, Athanasou N, et al. 2005. A molecular map of mesenchymal tumors. *Genome Biol.* **6**:R76

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, and Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. **4**:249-264

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, et al. 2009. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat.Genet.* **41**:178-186

Issa JP. 2004. CpG island methylator phenotype in cancer. *Nat.Rev.Cancer* **4**:988-993

Jurgens B, Schmitz-Drager BJ, and Schulz WA. 1996. Hypomethylation of L1 LINE sequences prevailing in human urothelial carcinoma. *Cancer Res.* **56**:5698-5703

Kawaguchi K, Oda Y, Saito T, Yamamoto H, Takahira T, Kobayashi C, Tamiya S, Tateishi N, Iwamoto Y, and Tsuneyoshi M. 2006. DNA hypermethylation status of multiple genes in soft tissue sarcomas. *Mod.Pathol.* **19**:106-114

Kobayashi C, Oda Y, Takahira T, Izumi T, Kawaguchi K, Yamamoto H, Tamiya S, Yamada T, Oda S, Tanaka K, et al. 2006. Chromosomal aberrations and microsatellite instability of malignant peripheral nerve sheath tumors: a study of 10 tumors from nine patients. *Cancer Genet.Cytogenet.* **165**:98-105

Laird PW. 2005. Cancer epigenetics. *Hum.Mol.Genet.* **14 Spec No 1**:R65-R76

Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, et al. 2010. Dynamic changes in the human methylome during differentiation. *Genome Res.* **20**:320-331

Levy P, Vidaud D, Leroy K, Laurendeau I, Wechsler J, Bolasco G, Parfait B, Wolkenstein P, Vidaud M, and Bieche I. 2004. Molecular profiling of malignant peripheral nerve sheath tumors associated with neurofibromatosis type 1, based on large-scale real-time RT-PCR. *Mol.Cancer.* **3:20.**:20

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**:315-322

Luijten M, Redeker S, van Noesel MM, Troost D, Westerveld A, and Hulsebos TJ. 2000. Microsatellite instability and promoter methylation as possible causes of NF1 gene inactivation in neurofibromas. *Eur.J.Hum.Genet.* **8**:939-945

Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Git A, et al. 2007. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res.* **67**:1424-1429

Martinez R, Martin-Subero JI, Rohde V, Kirsch M, Alaminos M, Fernandez AF, Ropero S, Schackert G, and Esteller M. 2009. A microarray-based DNA methylation study of glioblastoma multiforme. *Epigenetics.* **4**:255-264

Miller SJ, Jessen WJ, Mehta T, Hardiman A, Sites E, Kaiser S, Jegga AG, Li H, Upadhyaya M, Giovannini M, et al. 2009. Integrative genomic analyses of neurofibromatosis tumours identify SOX9 as a biomarker and survival gene. *EMBO Mol.Med.* **1**:236-248

Miller SJ, Rangwala F, Williams J, Ackerman P, Kong S, Jegga AG, Kaiser S, Aronow BJ, Frahm S, Kluwe L, et al. 2006. Large-scale molecular comparison of human schwann cells to malignant peripheral nerve sheath tumor cell lines and tissues. *Cancer Res.* **66**:2584-2591

Nielsen GP, Stemmer-Rachamimov AO, Ino Y, Moller MB, Rosenberg AE, and Louis DN. 1999. Malignant transformation of neurofibromas in neurofibromatosis 1 is associated with CDKN2A/p16 inactivation. *Am.J.Pathol.* **155**:1879-1884

Ordway JM, Budiman MA, Korshunova Y, Maloney RK, Bedell JA, Citek RW, Bacher B, Peterson S, Rohlfing T, Hall J, et al. 2007. Identification of novel high-frequency DNA methylation changes in breast cancer. *PLoS.ONE.* **2**:e1314

Pedersen IS, Dervan P, McGoldrick A, Harrison M, Ponchel F, Speirs V, Isaacs JD, Gorey T, and McCann A. 2002. Promoter switch: a novel mechanism causing biallelic PEG1/MEST expression in invasive breast cancer. *Hum.Mol.Genet.* **11**:1449-1453

Pelletier J, Bruening W, Li FP, Haber DA, Glaser T, and Housman DE. 1991. WT1 mutations contribute to abnormal genital system development and hereditary Wilms' tumour. *Nature.* **353**:431-434

Robertson AK, Geiman TM, Sankpal UT, Hager GL, and Robertson KD. 2004. Effects of chromatin structure on the enzymatic and DNA binding functions of DNA methyltransferases DNMT1 and Dnmt3a in vitro. *Biochem.Biophys.Res.Commun.* **322**:110-118

Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, and Bestor TH. 2006. Large-scale structure of genomic methylation patterns. *Genome Res.* **16**:157-163

Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H, and Hirohashi S. 2001. Expression of mRNA for DNA methyltransferases and methyl-CpG-binding proteins and DNA methylation status on CpG islands and pericentromeric satellite regions during human hepatocarcinogenesis. *Hepatology.* **33**:561-568

Santourlidis S, Florl A, Ackermann R, Wirtz HC, and Schulz WA. 1999. High frequency of alterations in DNA methylation in adenocarcinoma of the prostate. *Prostate* **39**:166-174

Serra E, Puig S, Otero D, Gaona A, Kruyer H, Ars E, Estivill X, and Lazaro C. 1997. Confirmation of a double-hit model for the NF1 gene in benign neurofibromas. *Am.J.Hum.Genet.* **61**:512-519

Smith ZD, Gu H, Bock C, Gnirke A, and Meissner A. 2009. High-throughput bisulfite sequencing in mammalian genomes. *Methods* **48**:226-232

Titze S, Peters H, Wahrisch S, Harder T, Guse K, Buske A, Tinschert S, and Harder A. 2010. Differential MSH2 promoter methylation in blood cells of Neurofibromatosis type 1 (NF1) patients. *Eur.J.Hum.Genet.* **18**:81-87

Vaissiere T, Hung RJ, Zaridze D, Moukeria A, Cuenin C, Fasolo V, Ferro G, Paliwal A, Hainaut P, Brennan P, et al. 2009. Quantitative analysis of DNA methylation profiles in lung cancer identifies aberrant DNA methylation of

specific genes and its association with gender and cancer risk factors. *Cancer Res.* **69**:243-252

Volinia S, Galasso M, Costinean S, Tagliavini L, Gamberoni G, Drusco A, Marchesini J, Mascellani N, Sana ME, Abu JR, et al. 2010. Reprogramming of miRNA networks in cancer and leukemia. *Genome Res.* **20**:589-599

Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, Fiala E, Ehrlich M, and Laird PW. 2005. Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res.* **33**:6823-6836

Wilson AS, Power BE, and Molloy PL. 2007. DNA hypomethylation and human diseases. *Biochim.Biophys.Acta* **1775**:138-162

Wilting SM, van Boerdonk RA, Henken FE, Meijer CJ, Diosdado B, Meijer GA, le SC, Agami R, Snijders PJ, and Steenbergen RD. 2010. Methylation-mediated silencing and tumour suppressive function of hsa-miR-124 in cervical cancer. *Mol.Cancer* **9**:167

Wong N, Lam WC, Lai PB, Pang E, Lau WY, and Johnson PJ. 2001. Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma. *Am.J.Pathol.* **159**:465-471

Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen TH, Girard L, Minna J, Christiani D, Leo C, et al. 2004. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res.* **64**:3060-3071

Zhou H, Coffin CM, Perkins SL, Tripp SR, Liew M, and Viskochil DH. 2003. Malignant peripheral nerve sheath tumor: a comparison of grade, immunophenotype, and cell cycle/growth activation marker expression in sporadic and neurofibromatosis 1-related lesions. *Am.J.Surg.Pathol.* **27**:1337-1345

Figure and Table Legends

Figure 1
Global DNA methylation profiles. (A-T) Global patterns of methylation in unique genomic features from the Ensembl genome database (Ensembl 52). Shown are the colour-coded methylation states of CpGs for MPNST, NF and SC samples (x-axis). The mean Batman methylation states of CpGs within a 1Kb window were calculated and categorised into low methylation (yellow, 0-40%), intermediate methylation (green, 41-60%) and high methylation (blue, 61-100%) for each genomic feature. MeDIP-seq/Batman CpG coverage (y-axis) shows the proportion of CpGs covered with Batman methylation scores at low, intermediate or high methylation levels. Coordinates for genomic features were taken from the Ensembl genome database, with the exception of CGI shores which were defined as 2Kb either side of the CGI and promoters defined as 1kb upstream and 0.5kb downstream of the TSS.

Figure 2
Relative DMR enrichment. Relative enrichment of DMRs in CGI, CGI shore, promoter and repeat regions (SINE, Satellite, LTR and LINE), for (A) hypermethylated DMRs and (B) hypomethylated DMRs, for the three sample comparisons (n2bDMRs, b2mDMRs and cDMRs). Relative enrichment analysis was carried out on the number of DMRs in each genomic feature compared to the expected number. Features with significant relative enrichment (P<0.001) are highlight in red. X-axis shows the relative enrichment score and the y-axis the relevant genomic features.

Figure 3
Relative DMR enrichment of intronic and non-intronic repeats. Shown are the relative enrichment of unique repeats (SINE, Satellite, LINE and LTR) located within introns and outside introns for (A) hypermethylated DMRs and (B) hypomethylated DMRs. X-axis shows the level of relative enrichment and the y-axis the relevant repeat type. Features highlighted in red show significant enrichment (P<0.001).

Figure 4
DMR enrichment of satellite repeat types. Relative enrichment of 19 satellite repeat types for (A) hypermethylated DMRs and (B) hypomethylated DMRs. X-axis shows the level of enrichment and the y-axis the relevant genomic features. Features highlighted in red show significant enrichment (P<0.001)

Figure 5
Clustering of MPNST and NF samples using the expression of genes associated with b2mDMRs. (A) The expression level of 1056 genes which contain a hypermethylated b2mDMR in associated CGI shores were used to cluster MPNST and NF samples. (B) Clustering of MPNST and NF samples based on the expression of 667 genes associated with non-CGI associated promoter b2mDMRs. The two major branches of the dendrogram correspond to an MPNST cluster and NF cluster, respectively. (B) Clustering of MPNST and NF samples based on the expression of genes associated with hypermethylated non-CGI associated promoter b2mDMRs. Columns

represent individual samples (MPNST, red; NF, blue). (C) Clustering of MPNST and NF samples based on the expression of genes associated with hypermethylated CGI-associated promoter b2mDMRs.

Figure 6
Example of genes displaying differential gene expression and methylation. The upper panel shows gene expression in NF (blue) and MPNST (red) for (A) *SOX10*, (B) *CDKN2A*. Lower panel shows the genomic organisation of the gene (black) with genomic regulatory features (blue) and CGIs (red) and the Batman methylation profiles for the three disease states (MPNST, NF and SC) for (C) *SOX10* and (D) *CDKN2A,* respectively.

Table 1
Number of DMRs identified by comparison of MPNST, NF and SC methylation profiles. DMRs were defined as a change of >33% in Batman methylation score between samples.

Table 2
Number of DMRs in genomic features, comparing hypermethylated and hypomethylated cDMRs, b2mDMRs and n2bDMRs.

Figure 1

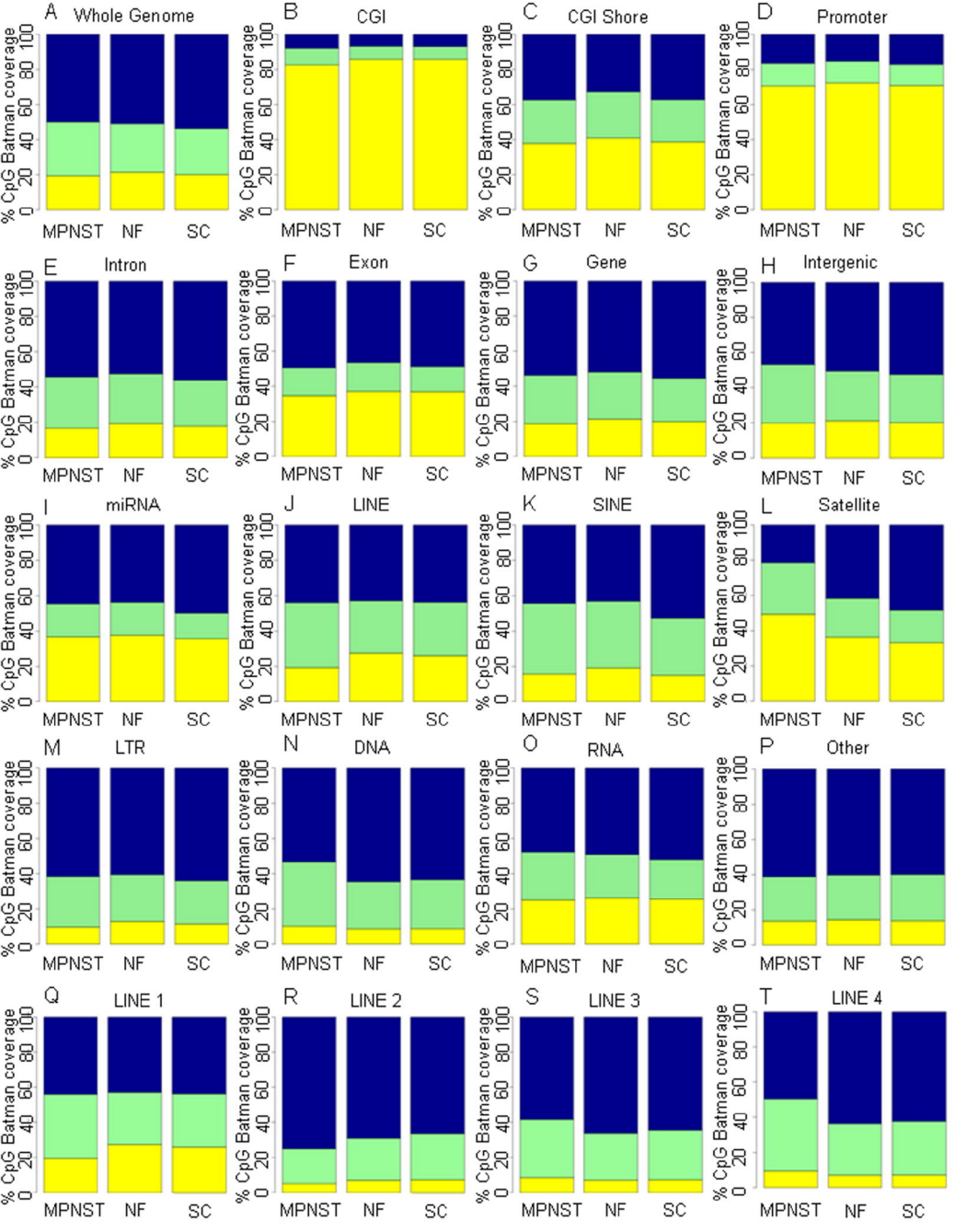

A Whole Genome
B CGI
C CGI Shore
D Promoter
E Intron
F Exon
G Gene
H Intergenic
I miRNA
J LINE
K SINE
L Satellite
M LTR
N DNA
O RNA
P Other
Q LINE 1
R LINE 2
S LINE 3
T LINE 4
% CpG Batman coverage
0
20
40
60
80
100
MPNST
NF
SC

Figure 2

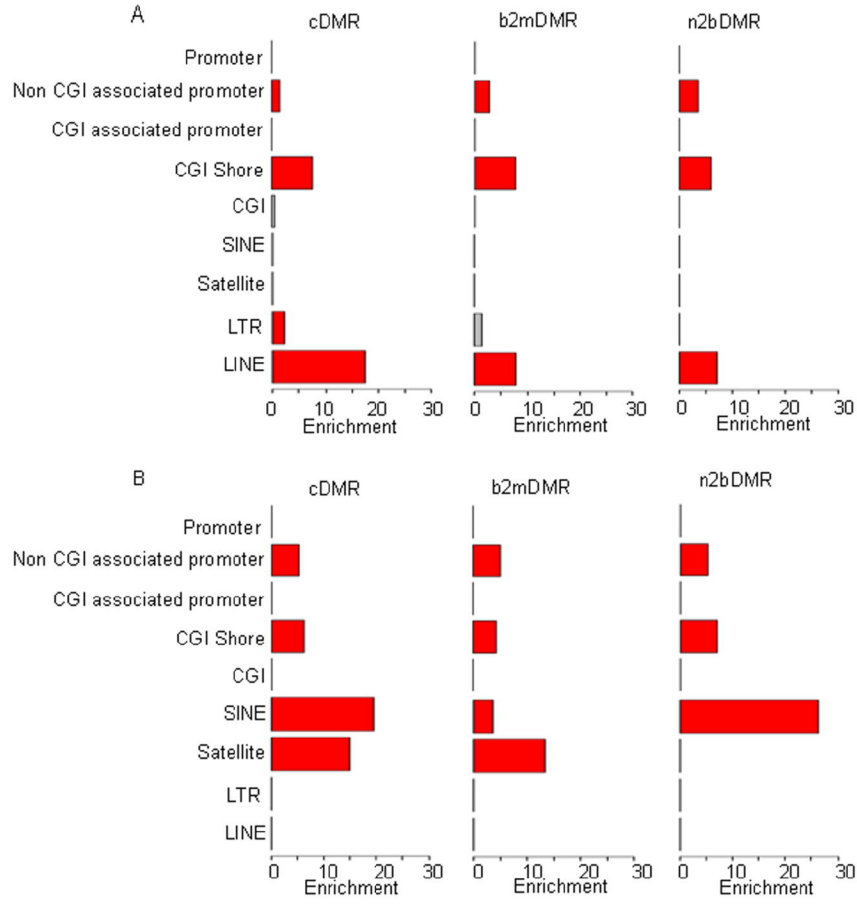

A
cDMR
b2mDMR
n2bDMR
Promoter
Non CGI associated promoter
CGI associated promoter
CGI Shore
CGI
SINE
Satellite
LTR
LINE
0 10 20 30
Enrichment
B
cDMR
b2mDMR
n2bDMR
Promoter
Non CGI associated promoter
CGI associated promoter
CGI Shore
CGI
SINE
Satellite
LTR
LINE
0 10 20 30
Enrichment

Figure 3

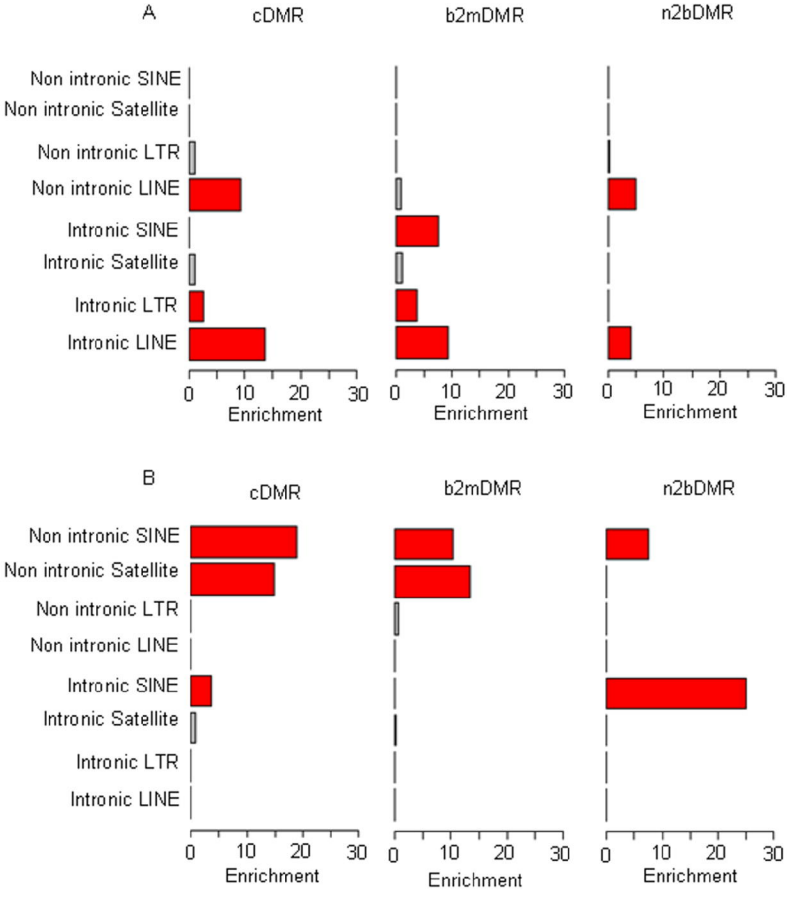

A
cDMR
b2mDMR
n2bDMR
Non intronic SINE
Non intronic Satellite
Non intronic LTR
Non intronic LINE
Intronic SINE
Intronic Satellite
Intronic LTR
Intronic LINE
0
10
20
30
Enrichment
B
cDMR
b2mDMR
n2bDMR
Non intronic SINE
Non intronic Satellite
Non intronic LTR
Non intronic LINE
Intronic SINE
Intronic Satellite
Intronic LTR
Intronic LINE
0
10
20
30
Enrichment

Figure 4

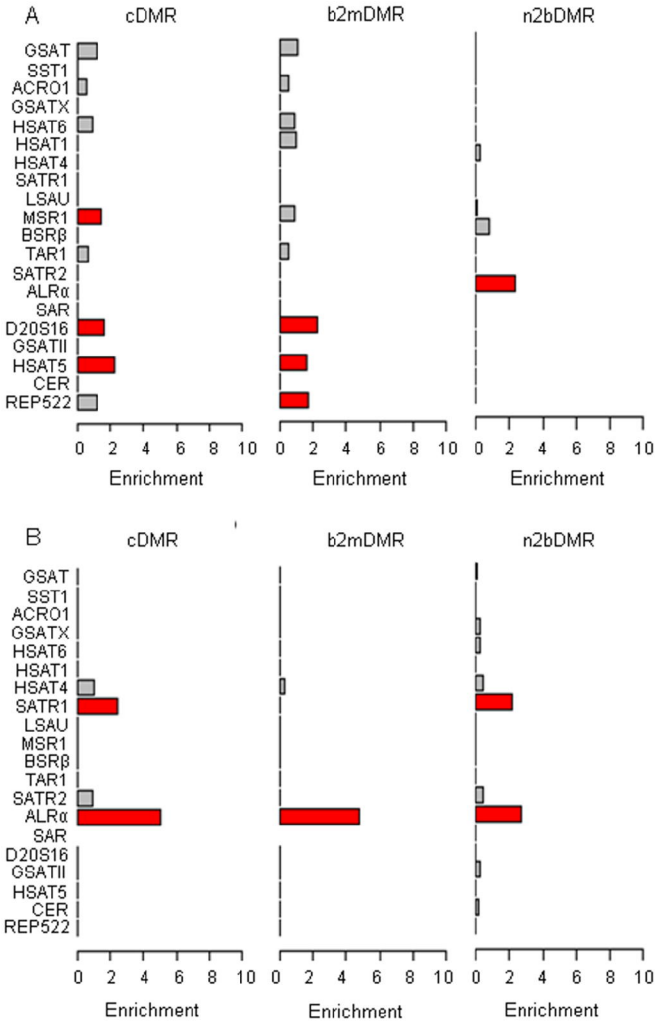

A
cDMR
b2mDMR
n2bDMR
GSAT
SST1
ACRO1
GSATX
HSAT6
HSAT1
HSAT4
SATR1
LSAU
MSR1
BSRβ
TAR1
SATR2
ALRα
SAR
D20S16
GSATII
HSAT5
CER
REP522
0 2 4 6 8 10
Enrichment
B
cDMR
b2mDMR
n2bDMR
GSAT
SST1
ACRO1
GSATX
HSAT6
HSAT1
HSAT4
SATR1
LSAU
MSR1
BSRβ
TAR1
SATR2
ALRα
SAR
D20S16
GSATII
HSAT5
CER
REP522
0 2 4 6 8 10
Enrichment

Figure 5

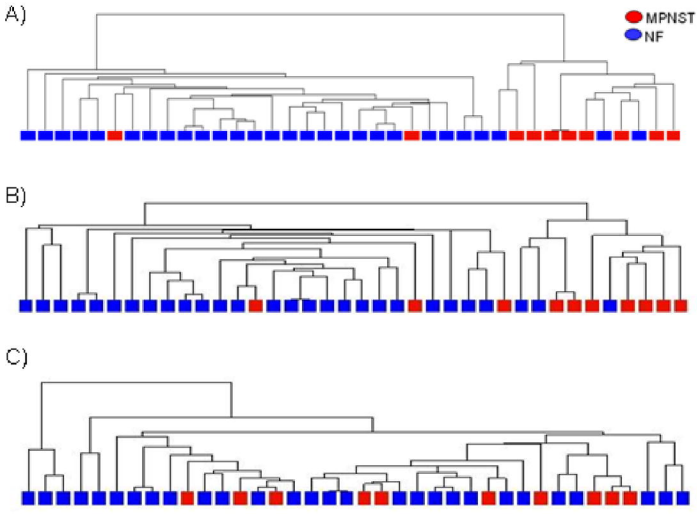
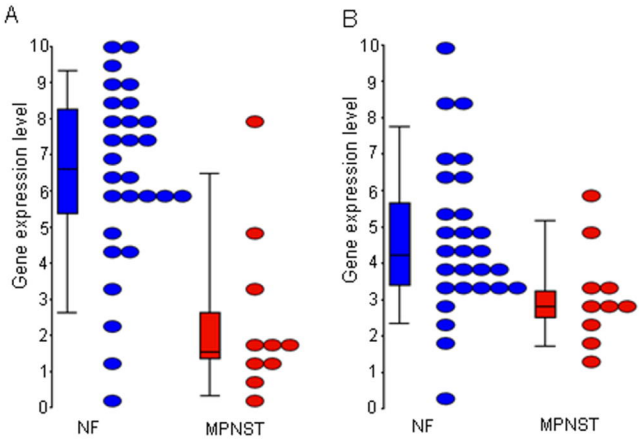
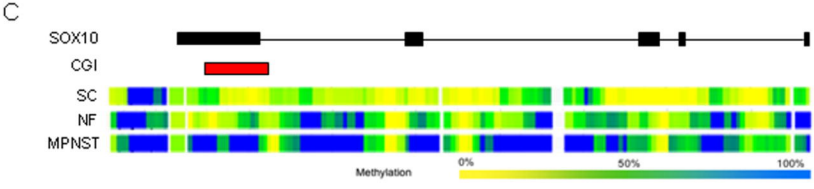
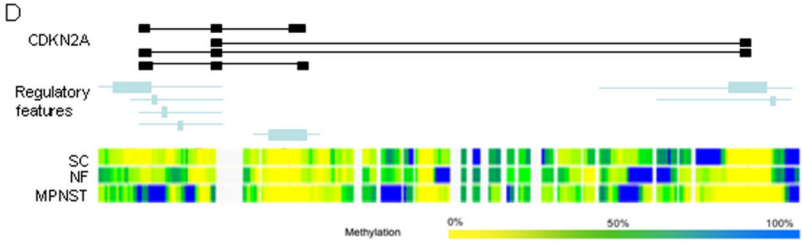


Figure 6


A
B
Gene expression level
0
1
2
3
4
5
6
7
8
9
10
NF
MPNST


C
SOX10
CGI
SC
NF
MPNST
Methylation
0%
50%
100%


D
CDKN2A
Regulatory features
SC
NF
MPNST
Methylation
0%
50%
100%