# DNA methylation analysis in human tumour samples using nanopore sequencing

Vahid Akbari[1,2], Kieran O'Neill[1], Richard Corbett[1], Vanessa L. Porter[1,2], Erin Pleasance[1], Pawan Pandoh[1], Richard Moore[1], Marco Marra[1,2], Martin Hirst[1,3], Steven J.M. Jones[1,2]

1- Canada's Michael Smith Genome Science Centre, British Columbia Cancer, Vancouver, British Columbia, Canada.

2- Department of medical genetics, University of British Columbia, Vancouver, British Columbia, Canada.

3- Department of Microbiology and Immunology, Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada

## INTRODUCTION

DNA methylation is critical for modulating gene expression, and its dysregulation is a common feature of cancers. For example, tumour suppressor genes may be inactivated by promoter hypermethylation, or global hypomethylation may lead to oncogene over-activation and transposon reactivation. There have also been reports of disruption of imprinting in tumours, but this is challenging to resolve using short-read sequencing. Nanopore sequencing's capability to assay methylation in parallel with long-read sequence data makes it an attractive platform for studying DNA methylation in the context of tumourogenesis.

As part of our work on examining the capabilities of Nanopore sequencing for personalized oncogenomics. we sequenced nine advanced-stage tumour samples on the PromethION and examined their methylation status in detail. These included glioma, squamous cell carcinoma, diffuse large B-cell lymphoma, Ewing sarcoma, and head and neck cancers. As comparators, we used nanopore data for normal heart, liver, and hippocampus, 16 normal blood samples from individuals with cancer, as well as 12 normal lymphocyte cell lines.

## Results

One case, a glioma. had a known *IDH1* Arg132His mutation. Gliomas with this mutation usually exhibit a CpG island methylator phenotype specifically at the PRC2 target genes' promoter due to inhibition of the TET demethylases. Comparison of this glioma with WGBS data from a normal brain tissue (GSM2877176) showed wide-spread hypermethylation, enriched at targets of the PRC2 complex and, to a lesser extent, PRC1 complex (Fig. 1). Using nanopore sequencing data from a blood sample (Blood15) of an individual with lung squamous cell carcinoma, we were able to detect the constitution epimutation as allele-specific methylation at the CpG island encompassing the promoter of MLH1 (Fig. 2). This epimutation silences the affected allele and predisposes the affected individual to Lynch syndrome–associated cancers.

We characterized genome-wide allele-specific methylation (ASM), and found it to be increased two- to four-fold, in tumour samples compared to normal samples. On average, tumour samples had 3900 allelic DMRs while normal blood, cell line, and tissue samples had 940 allelic DMRs. Allelic DMRs mapped to all genomic regions and mostly enriched at CpG islands (Fig. 3). Conversely, we observed wide-spread loss of ASM at imprinting control regions, particularly *IGF2R, IGF1R*, *RB1*, *BLCAP*, *ZNF597* (Fig. 4). We excluded allelic DMRs from tumour samples that overlapped to DMRs from normal blood, cell lines and tissue samples and also those that overlapped to known imprinted intervals and genes. We then mapped the remaining allelic DMRs to the genes' promoter. We detected 58 allelic DMRs that overlapped between 5 or more of the tumour samples and mapped to the promoter of 81 genes (Fig. 5a). Gene ontology analysis of these 81 genes demonstrated that they enrich for apoptosis and developmental terms (Fig. 5b).

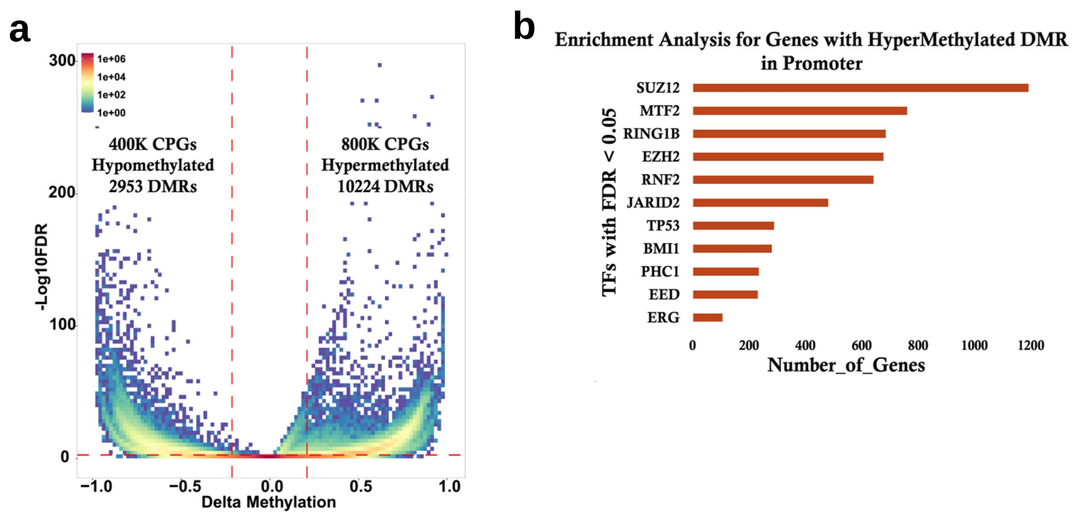


**Figure 1**. Differential methylation analysis (DMA) between nanopore-detected methylation data from a glioma sample with IDH1 mutation and a public WGBS data from brain (GSM2877176) . a) Volcano plot of DMA results. Glioma sample demonstrated 2953 hypomethylated regions compare to normal case while 10224 regions are hypermethylated compare to normal case. b) Enrichment transcription factor (TF) analysis of hypermethylated regions mapped to genes' promoter. Hypermethylated genes enrich mostly to PRC2 target genes and to lesser extend to PRC1 target genes. *EED*, *EZH2*, and *SUZ12* are PRC2 core subunits while *JARID2* and *MTF2* are non-core subunits. *BMI1*, *RING1B*, *PHC1, and* RNF2 are PRC1 complex subunits.

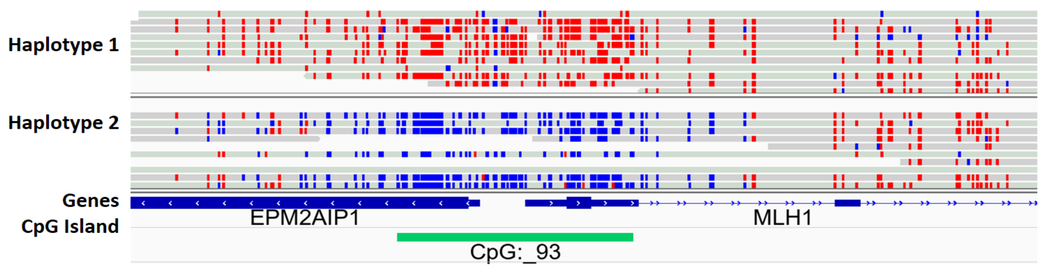


**Figure 2**. IGV screenshot of the constitutional allele-specific methylation of *MLH1* promoter in the blood sample from a case with lung cancer. Methylated CGs are red and unmethylated CGs are blue.

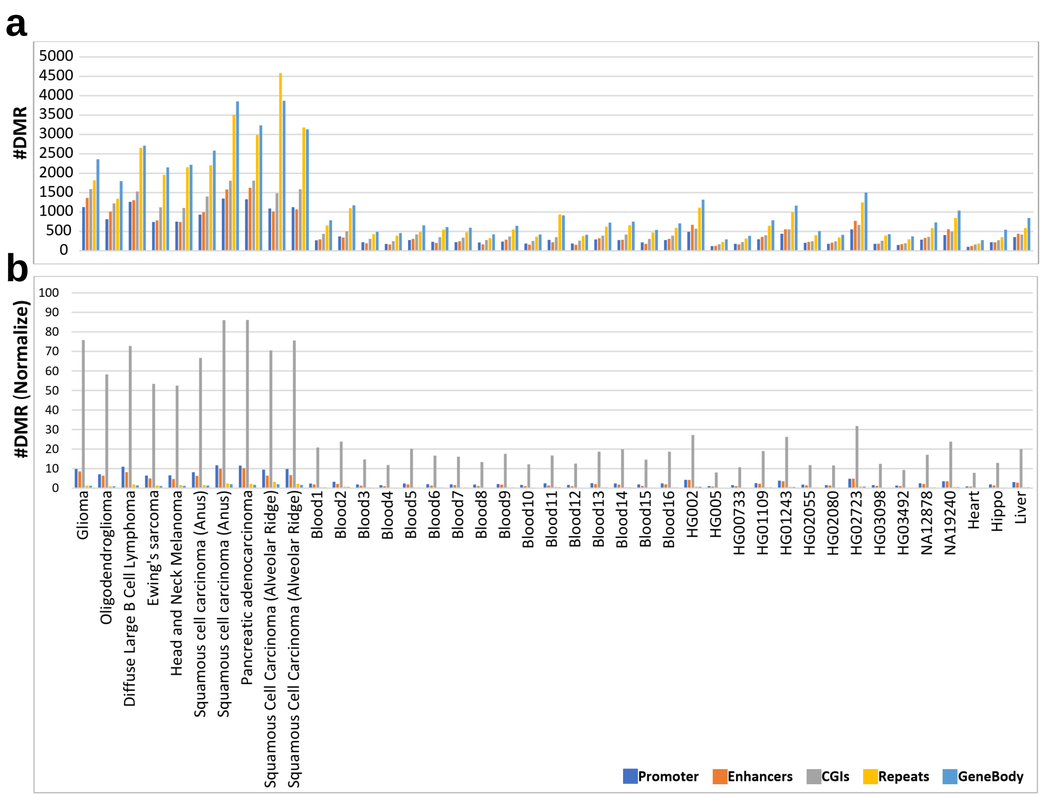


**Figure 3**. Allele-specific methylation in tumour and normal samples. a) number of allelic differentially methylated regions (DMRs) overlapping at various genomic regions in 10 tumour samples, 16 normal blood samples from individuals with cancer, 12 normal B-lymphocyte cell lines, and three normal tissues including liver, heart and hippocampus. b) Number of regions from part a normalized by the total length of each genomic region. Both squamous cell carcinoma (Alveolar ridge) samples are from the same case. Blood 4, blood10, and blood14 are the matched blood samples for diffuse large B cell lymphoma, pancreatic adenocarcinoma and squamous cell carcinoma (Alveolar ridge) cases, respectively. Allelic DMRs in these plots show redundancy, meaning some DMRs mapped to more than one genomic region.

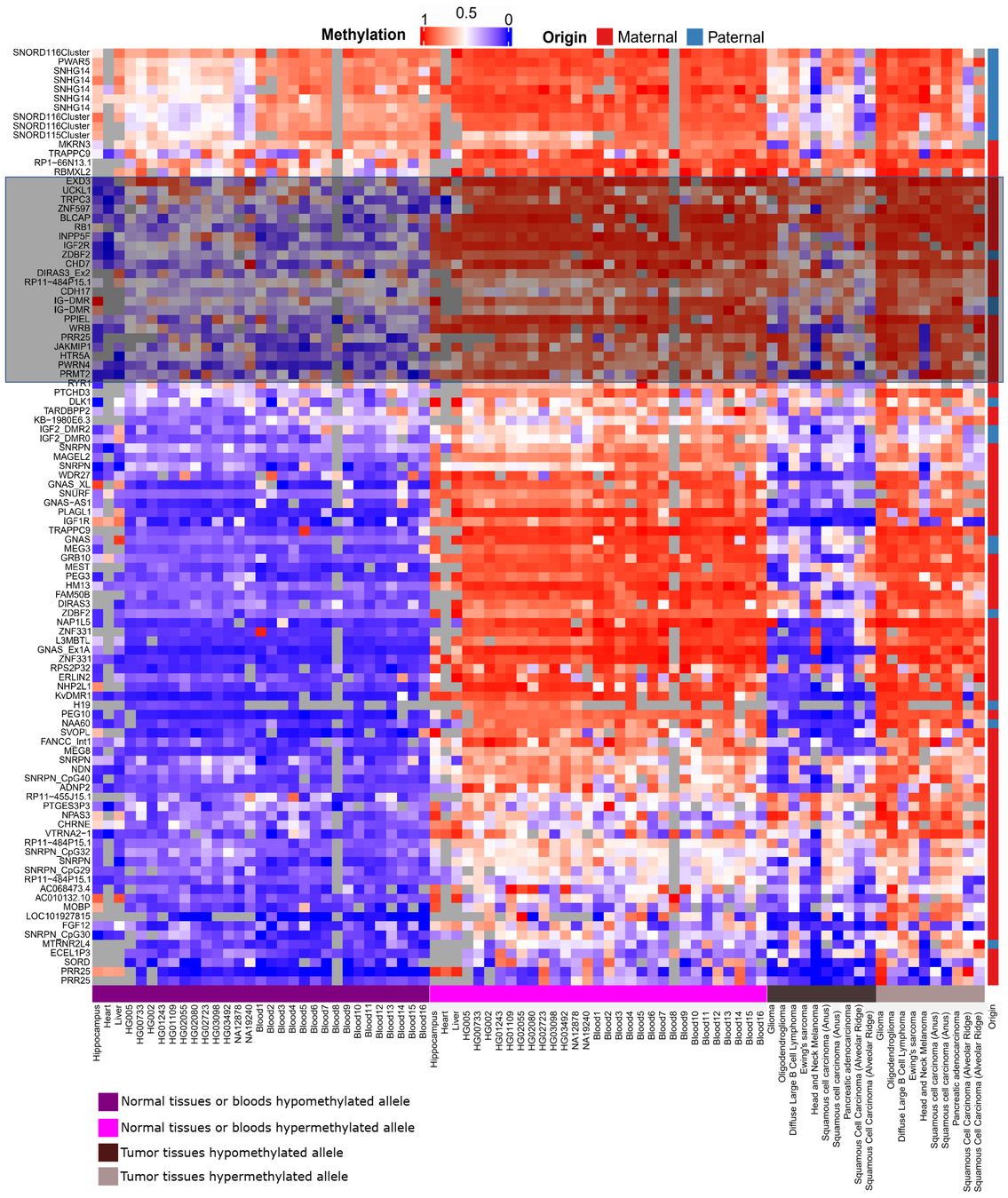


**Figure 4**. Allele-specific methylation at known imprinted intervals reported in 2 or more previous studies[1-5]. Heatmap representing average allelic methylation for 10 tumour samples, 16 normal blood samples from individuals with cancer, 12 normal B-lymphocyte cell lines, and three normal tissues including liver, heart and hippocampus. Allelic methylation at each imprinted region in the heatmap is the average of methylation frequencies from all CGs at each imprinted DMR from each allele. Both squamous cell carcinoma (Alveolar ridge) samples are from the same case. Blood 4, blood10, and blood14 are the matched blood samples for diffuse large B cell lymphoma, pancreatic adenocarcinoma and squamous cell carcinoma (Alveolar ridge) cases, respectively. Highlighted box represents most of the imprinted regions with loss of imprinting in tumours.

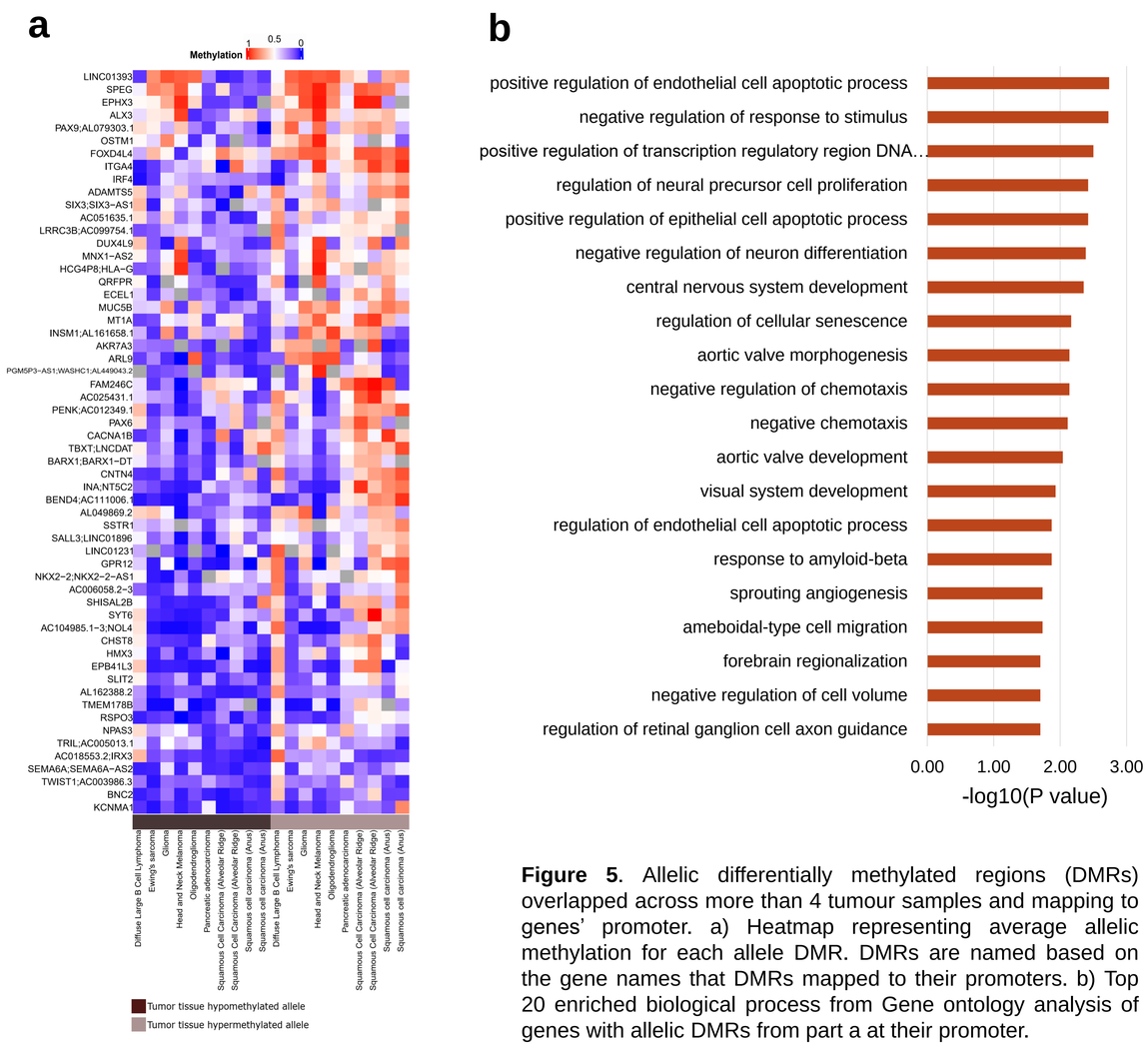


**Figure 5**. Allelic differentially methylated regions (DMRs) overlapped across more than 4 tumour samples and mapping to genes' promoter. a) Heatmap representing average allelic methylation for each allele DMR. DMRs are named based on the gene names that DMRs mapped to their promoters. b) Top 20 enriched biological process from Gene ontology analysis of genes with allelic DMRs from part a at their promoter.

## METHODS

We sequenced 10 tumour samples along with 16 normal blood samples from individuals with cancer. Moreover, we used public nanopore sequencing data from 12 B-lymphocyte cell lines and three normal tissues including heart, liver, and hippocampus. Nanopore fast5 files were basecalled using guppy v5.0.7 and dna_r9.4.1_450bps_hac_prom model. Reads were mapped to human reference genome hg38 using minimap2 v2.23 and methylations were called using nanopolish v13.3. We used Clair3 for variant calling and WhatsHap v0.18 for phasing of nanopore-called variants. Finally, we phased methylation data using our in-house tool NanoMethPhase (v1.0).

## References

**1.** Court, F. *et al.* Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res.* 24, 554–569 (2014). **2.** Joshi, R. S. *et al.* DNA Methylation Profiling of Uniparental Disomy Subjects Provides a Map of Parental Epigenetic Bias in the Human Genome. *Am. J. Hum. Genet.* 99, 555–566 (2016). **3.** Zink, F. *et al.* Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. *Nat. Genet.* 50, 1542–1552 (2018). **4.** Hernandez Mora JR. *et al*. Characterization of parent-of-origin methylation using the Illumina Infinium MethylationEPIC array platform. *Epigenomics.* 10:941–954 (2018). **5.** Akbari, V. *et al.* Genome-Wide Detection of Imprinted Differentially Methylated Regions Using Nanopore Sequencing. *bioRxiv* 2021.07.17.452734; doi: https://doi.org/10.1101/2021.07.17.452734.