

Nanopore sequencing to analyse MGMT promoter methylation

Preliminary data analysis

Skarphedinn Halldorsson

Compiled on 14 December, 2022

0.1 Introduction

MGMT promoter methylation is an important diagnostic factor for glioblastoma treatment. Methylation on a CpG island in the MGMT promoter region predicts response to the alkylating agent temozolomide. MGMT promoter methylation is typically assessed by methylation specific PCR (MSP) of the first exon of MGMT or pyrosequencing of 4 CpG sites in the first exon. Both methods rely on bisulphate treatment of DNA prior to analysis. Nanopore sequencing allows methylation calling on native, genomic DNA. This circumvents the need for bisulphate treatment of DNA prior to sequencing and allows methylation analysis of far larger sequencing than either MSP or pyrosequencing.

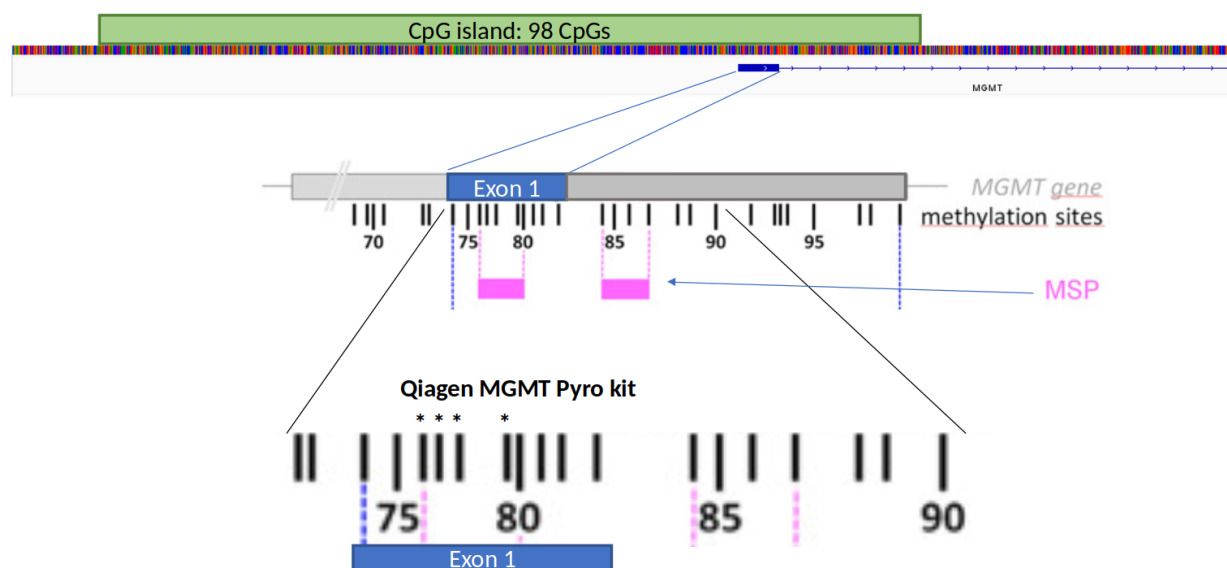


Figure 1: Organization of the MGMT promoter. MSP refers to the typical primer sites of methylation specific PCR to determine MGMT promoter methylation. Asterixes represent the 4 CpGs analysed by the Qiagen MGMT pyrosequencing kit.

Here, we present the results of nanopore sequencing of the whole CpG island in the promoter region of the MGMT gene. We gathered nanopore sequencing data of the MGMT promoter region from 148 patients, including 130 glioblastoma patients. Results were produced either by CRISPR/Cas9 targeted sequencing of the MGMT promoter region or as part of an adaptive sampling panel (RAPID-CNS). Samples were obtained from three separate cohorts: the DenStem project at Institute for Surgical Research, the glioblastoma DNA biobank at Radium Hospital and Rapid-CNS. Cas9 targeted sequencing samples were either run as single samples or as 5 barcoded samples run simultaneously. An overview of samples can be seen in Table 1.

Table 1: Overview of samples included in this study

Series	Astrocytoma	Astrocytoma HG	Glioblastoma	LGG_PA	Meningioma	Metastasis	Oligodendroglioma	Other
DenStem	3	0	13	0	0	0	0	0
Radium	1	4	29	0	12	7	2	10
Rapid-CNS	3	4	49	4	0	0	6	1
Total	7	8	91	4	12	7	8	11

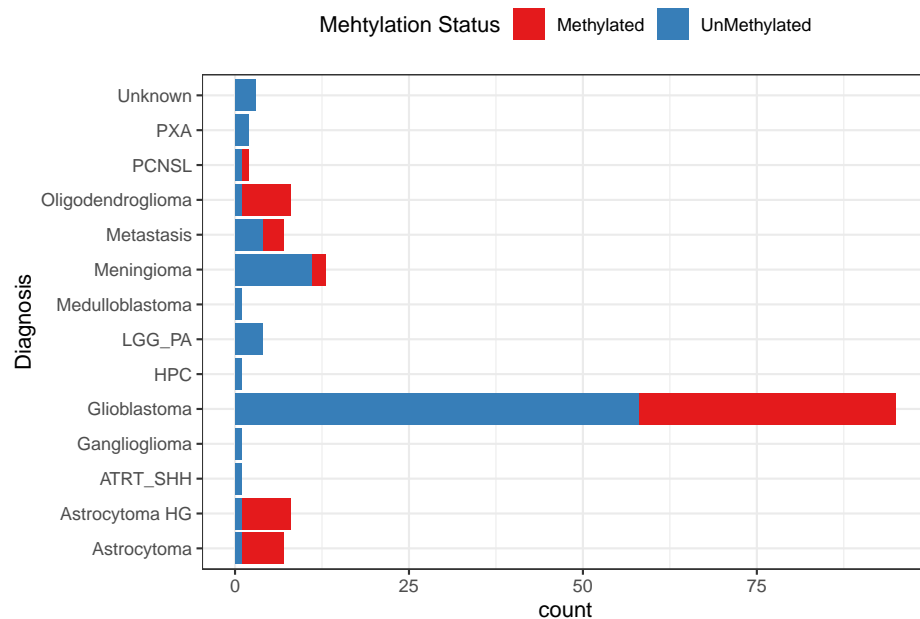


Figure 2: Distribution of methylated versus unmethylated samples as measured by pyrosequencing

0.2 Data acquisition

Figures 3 and 4 show the sequencing depth of samples in this study. Figure 3 shows that there is no implicit bias between methylated versus unmethylated samples when it comes to sequencing depth. Figure 3 shows that the range of depth is large. Single sample runs produce more sequences than barcoded runs, adaptive sampling similar sequencing depth as barcoded Cas9 targeted samples.

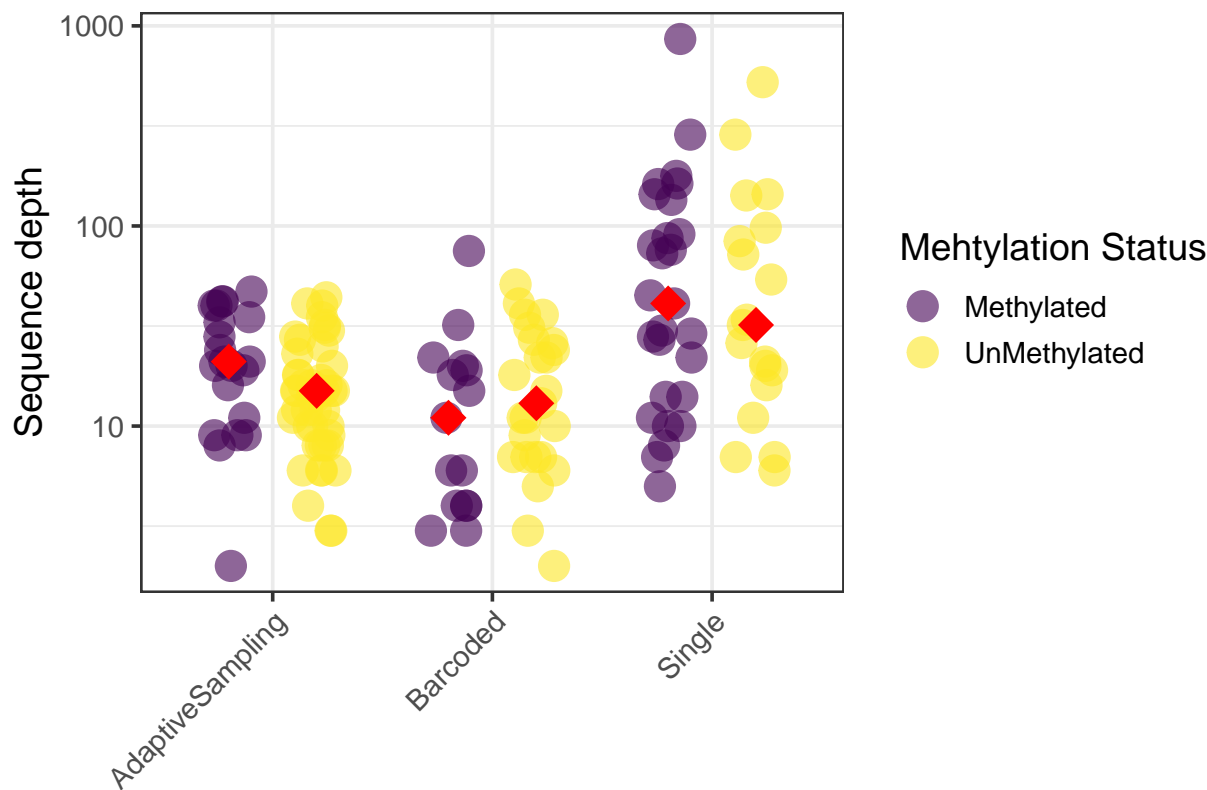


Figure 3: Sequencing depth of samples by method of acquisition, group median represented by red diamond. No bias in sequence depth observed between methylated and unmethylated samples.

0.3 Direct comparison of Nanopore sequencing and Pyrosequencing

At the Norwegian University hospital, MGMT promoter methylation is determined by the Qiagen MGMT pyrosequencing kit. Results are typically presented as an average methylation percentage across the four CpGs detected by the kit. A cut-off value of 9% determines if a given sample is classified as “methylated” or “unmethylated”. We currently have average methylation values for the DenStem samples and the Radium samples (Figure 5).

Plotting the average methylation percentage acquired via Nanopore sequencing against the values acquired by pyrosequencing shows reasonable correlation between the two methods, particularly for the Radium samples ($R^2=0.85$). It should be noted that the DenStem results are from the same tumor but not the same biopsy (one biopsy for nanopore, another biopsy for pyrosequencing) while the Radium samples are from the same biopsy (same biopsy analysed by nanopore sequencing and pyrosequencing). This is a likely explanation of why the correlation between nanopore sequencing and pyrosequencing is lower in this cohort. This also raises the question of heterogeneity of methylation status throughout a tumor entity, especially if methylation %

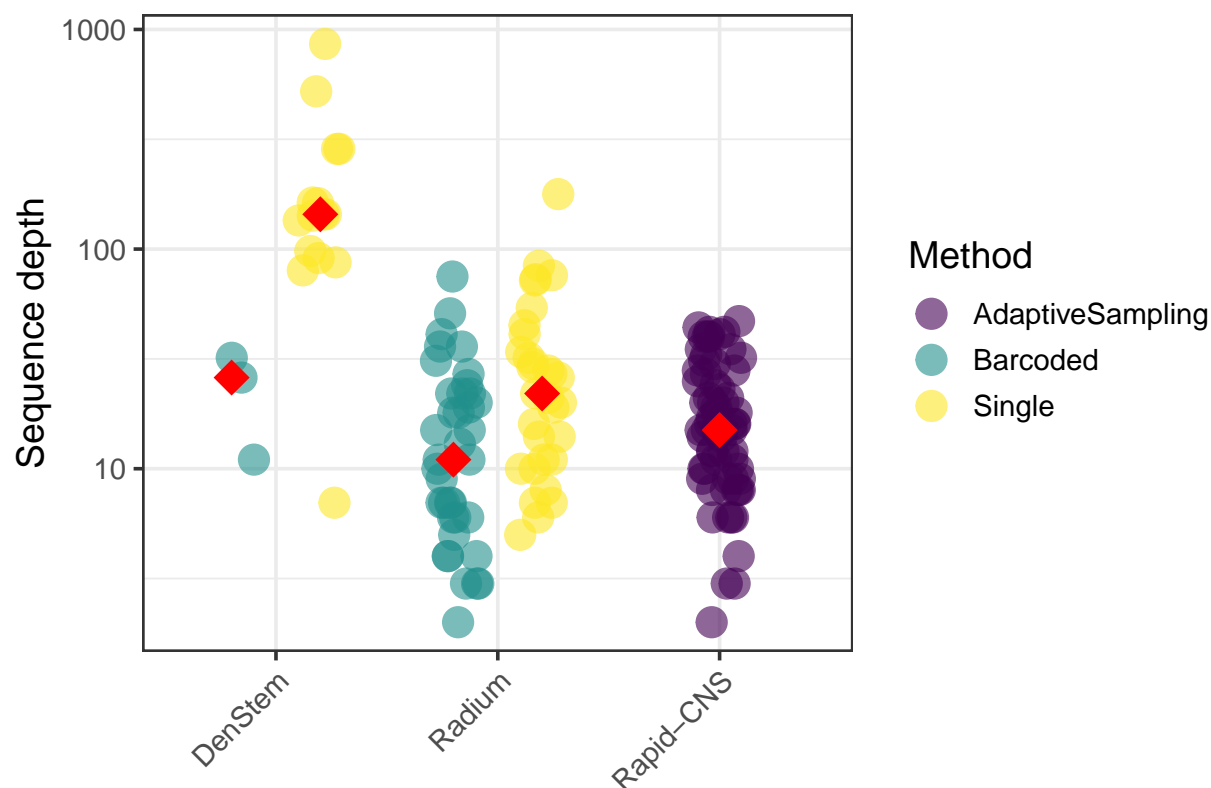


Figure 4: Sequencing depth of methods, group median represented by red diamond. Single sample runs generally have higher sequence depth than barcoded samples or adaptive sampling.

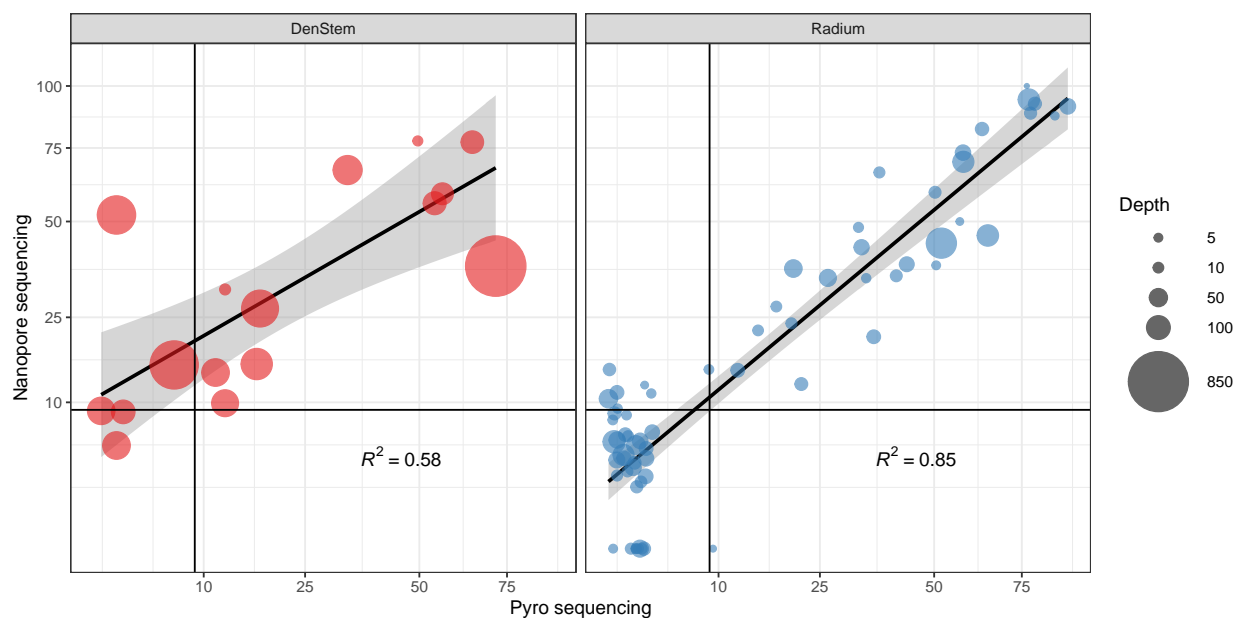


Figure 5: Comparison of nanopore sequencing and Pyrosequencing results of 4 CpGs in exon 1 of the MGMT promoter. Plotted values are average methylation of 4 CpGs. Black horizontal and vertical lines mark the 9 percent cut-off value between methylated and unmethylated samples, as determined by pyrosequencing.

is close to the cut-off value. Would a different biopsy have slightly higher or lower methylation values that might classify the tumor differently? I don't think the DenStem samples presented here can really answer that as they were analysed by different methods but it's worth a thought.

There appears to be a certain level of over-prediction of methylation % by nanopore sequencing compared to the pyrosequencing results, particularly in low methylation samples. This means that the % methylated cut-off value for nanopore sequencing is higher than the 9 % methylated threshold that is applied for pyrosequencing. These results only apply for the 4 CpGs sequenced by the Qiagen MGMT pyrosequencing kit. It should also be noted that % methylated values in very low coverage samples are unreliable.

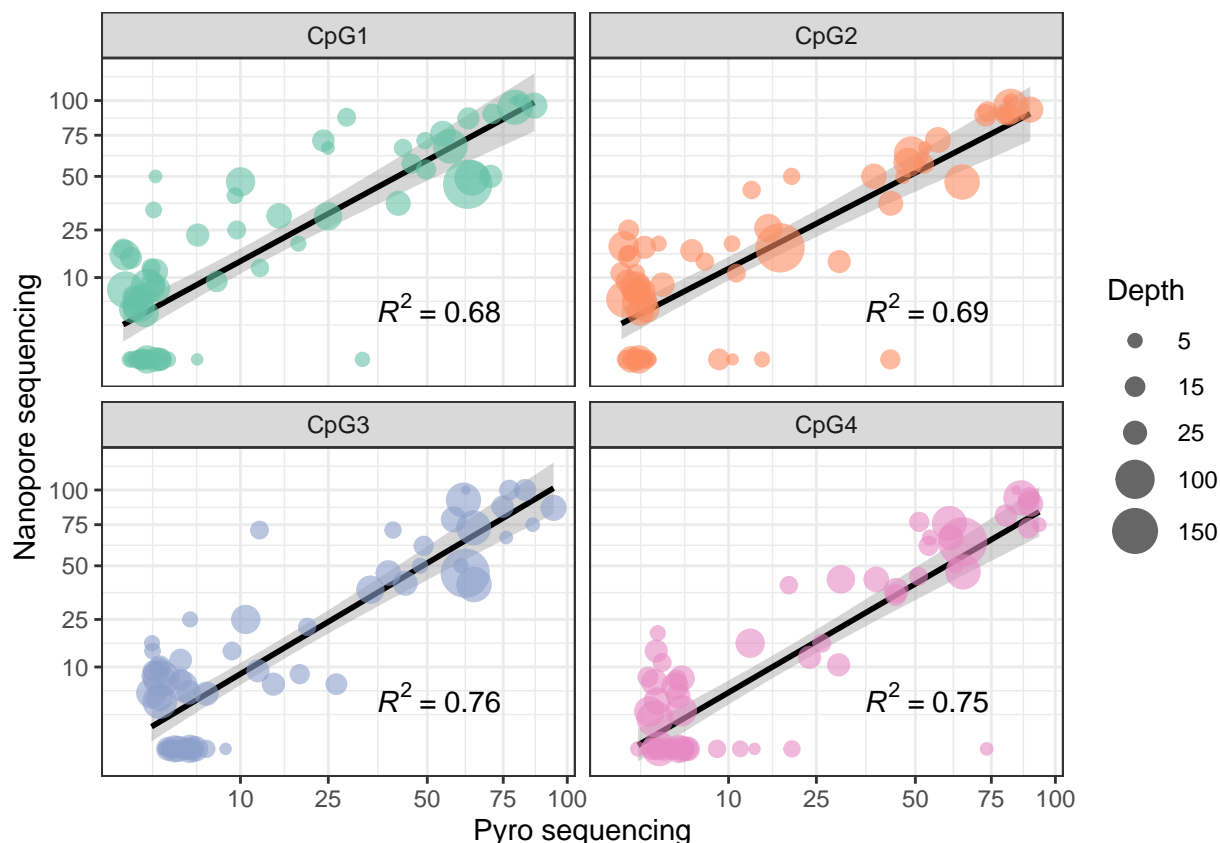


Figure 6: Comparison of individual CpGs within exon 1

We also have separate methylation percentage values for the four CpGs analysed by the Qiagen MGMT pyrosequencing kit for all the Radium samples. This affords a site by site comparison between nanopore sequencing and pyrosequencing (Figure 6). The correlation between nanopore and pyrosequencing is slightly less for the individual CpGs a difference that appears to be rounded out when the values are averaged (Figure 5, Radium).

Take together, nanopore sequencing of the same 4 CpGs analysed by the Qiagen MGMT pyrosequencing kit does a reasonable job of recreating the pyrosequencing results. However, this does not take advantage of the other 94+ CpGs that are included in the nanopore data.

0.4 Unsupervised clustering of nanopore results

Unsupervised clustering of all the samples included in the study shows very clear separation of methylated and unmethylated samples (Figure 7). The samples previously defined as unmethylated have very low

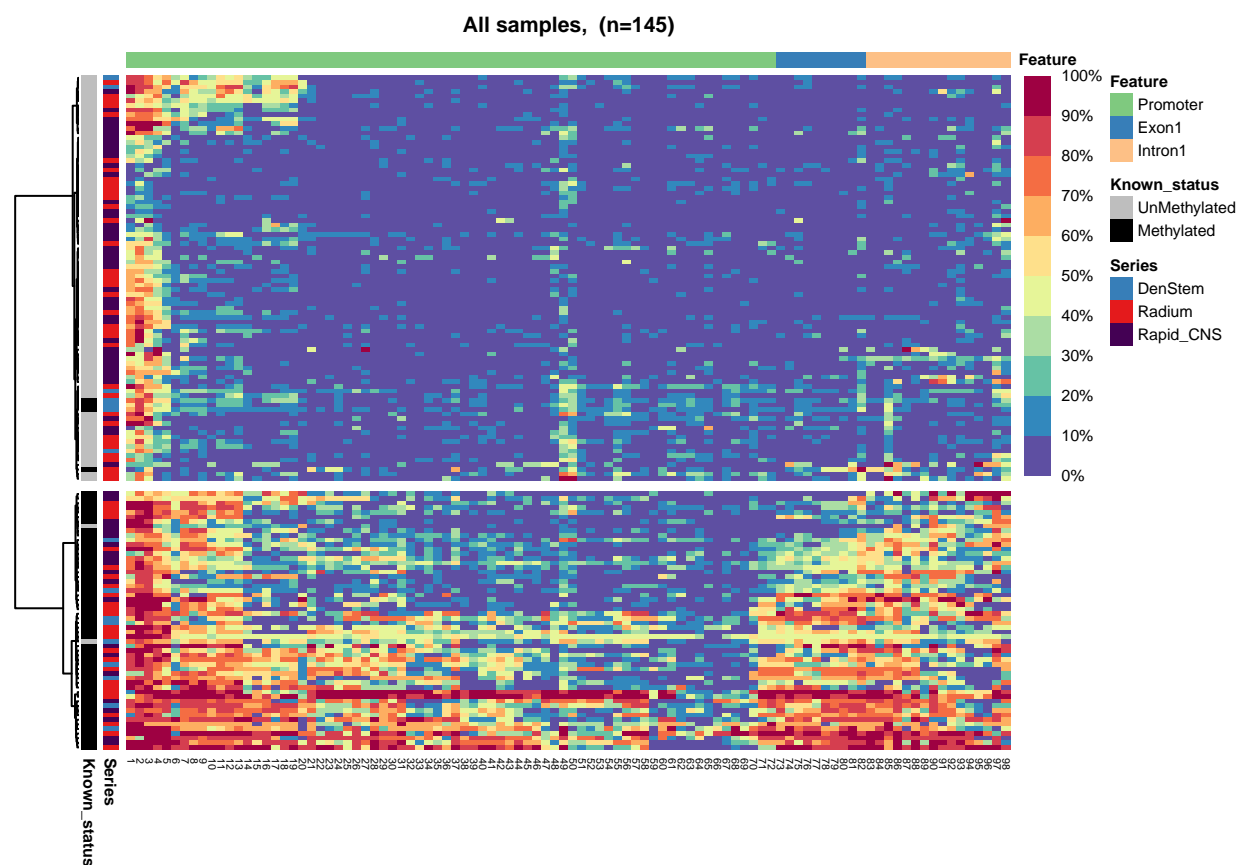


Figure 7: Clustered heatmap of all samples based on nanopore sequencing of CpG island of the MGMT promoter

methylation throughout the CpG island apart from the first 5 CpGs that are often methylated. Samples previously defined as methylated have a gradient of methylation which tends to show highest methylation values towards either end of the CpG island. Only six samples do not cluster according to their previously determined methylation status. Four of these samples belong to the DenStem cohort that is not directly comparable, as previously mentioned. A single sample from the Radium cohort was classified as methylated by pyrosequencing but clusters with the unmethylated samples. This sample is interesting as it has robust methylation in the first exon but very low methylation elsewhere in the CpG island.

We can also look at methylation patterns specifically in the glioblastoma samples (Table 1).

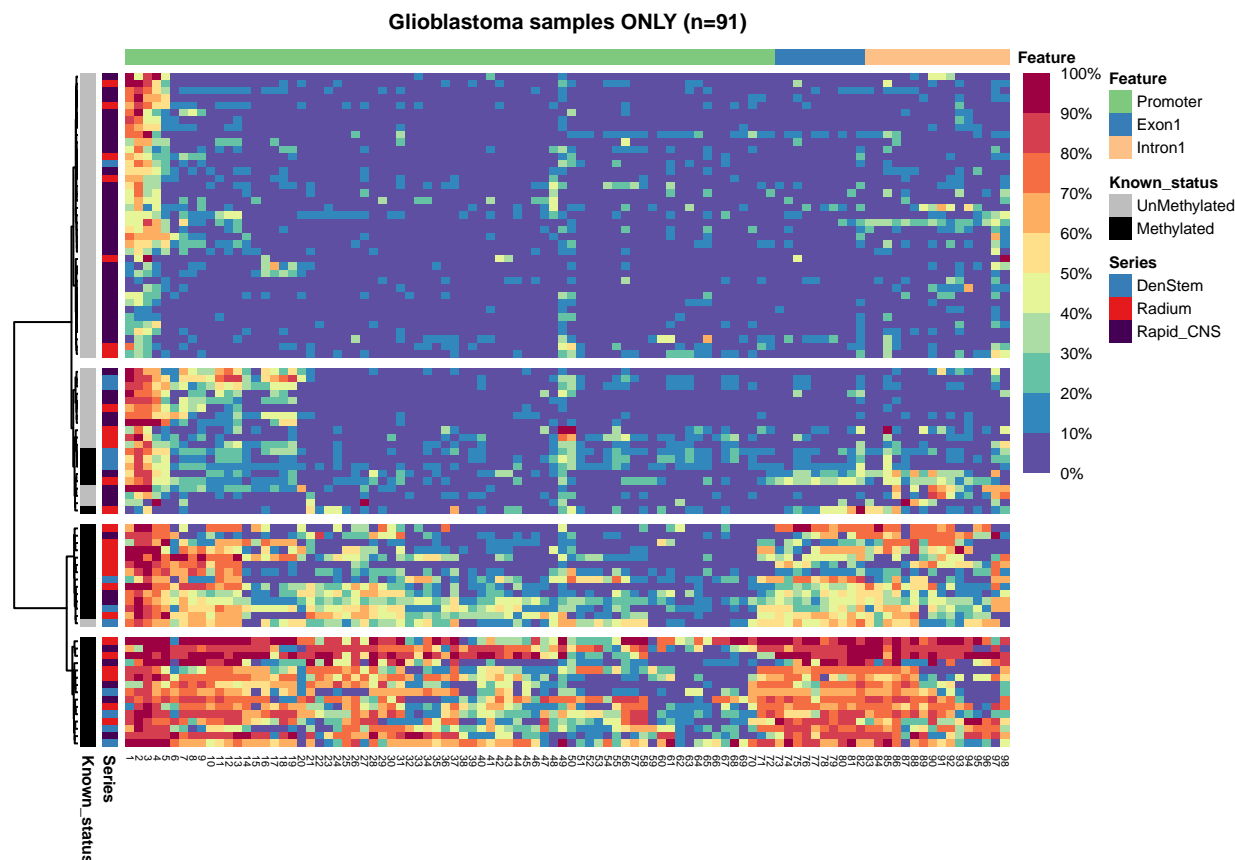


Figure 8: Heatmap showing unsupervised clustering of glioblastoma samples based on nanopore sequencing of the CpG island in the MGMT promoter

In Figure 8 we see that all of the Rapid-CNS GBM samples fall within their previously determined classes and only one of the Radium samples is “mis-classified”. We also see there are clusters within both the methylated and unmethylated samples. These clusters can also be represented by mean methylation values within each cluster (Figure 9).

0.5 Discussion

We have a fairly extensive dataset of 153 samples, including 91 GBMs. I’ve not seen data with this many samples that looks at methylation within the whole CpG island of MGMT. We have the option of looking for CpG methylation further upstream and downstream than the 98 CpG presented here, at least 10 CpGs in both directions without losing any samples due to lack of coverage. By focusing on high depth samples we can look much further. I have not included that here but will be looking into it and will report if I find anything interesting.

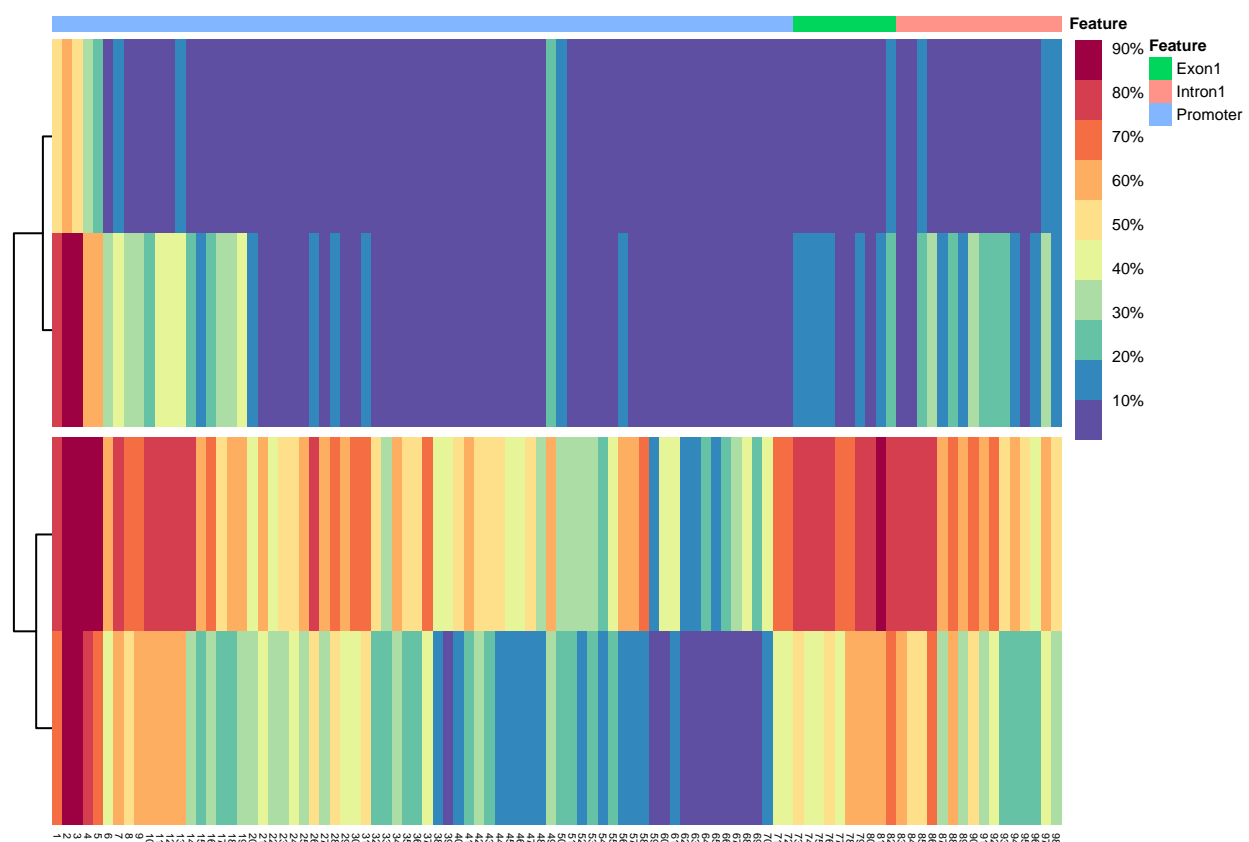


Figure 9: K-means clustering of glioblastoma samples

We can conclude that nanopore sequencing of the MGMT promoter region performs as well or better than standard methods such as pyrosequencing. This is true for both cas9 targetted sequencing of the MGMT promoter and inclusion of the MGMT promoter into an adaptive sequencing panel. Distinct subgroups within both methylated and unmethylated samples are captured via nanopore sequencing, it will be very interesting to see if there is a difference in patient outcome between these clusters.