NaNoMe-seq: Nanopore implementation of Nome-seq protocol

Jussis idea of using GpC methyltransferase M.CviPI to mark open chromatin regions in intact nuclei and sequence with Oxford Nanopore PromethION, obtaining information on Structural Variants, CpG methylation and open chromatin in single sequencing experiment. This would provide long range information from single molecules, similar to NOMe-seq method(Lay et al. 2018; Kelly et al. 2012)

Lab protocols are being developed by Kimmo, Sahu and Saija. Alternative for GpC methyltransferase would be Adenine methyltransferase M. EcoGII which could improve resolution and help decoding information at GpCpG sites.

For training data to recognize the nanopore signal, Sahu used GP5d cell lines and following 6 conditions barcoded and sequenced on one flowcell on PromethION

1. PCR amplified naked DNA,
2. PCR amplified naked DNA with enzymatic CpG methylation,
3. PCR amplified naked DNA with GpC methylation,
4. PCR amplified naked DNA with CpG and GpC methylation
5. PCR amplified naked DNA with A methylation
6. PCR amplified naked DNA with CpG with A methylation

The analysis pipeline needs to be developed for this. The two alternatives are nanopolish and Tombo.

For Tombo, there are fairly good instructions for learning the new models https://nanoporetech.github.io/tombo/model_training.html and recent versions include the necessary per-read statistics and also some sort of all contexts 5mC and 6mA models.

Nanopolish is my favorite because I kind of understand what it does and it's very clean code base. Learning the model might be tricky though:

- https://github.com/jts/nanopolish/issues/400
    - I've implemented GmC methylation contexts.
- https://github.com/jts/nanopolish/issues/338
- https://github.com/jts/methylation-analysis
-

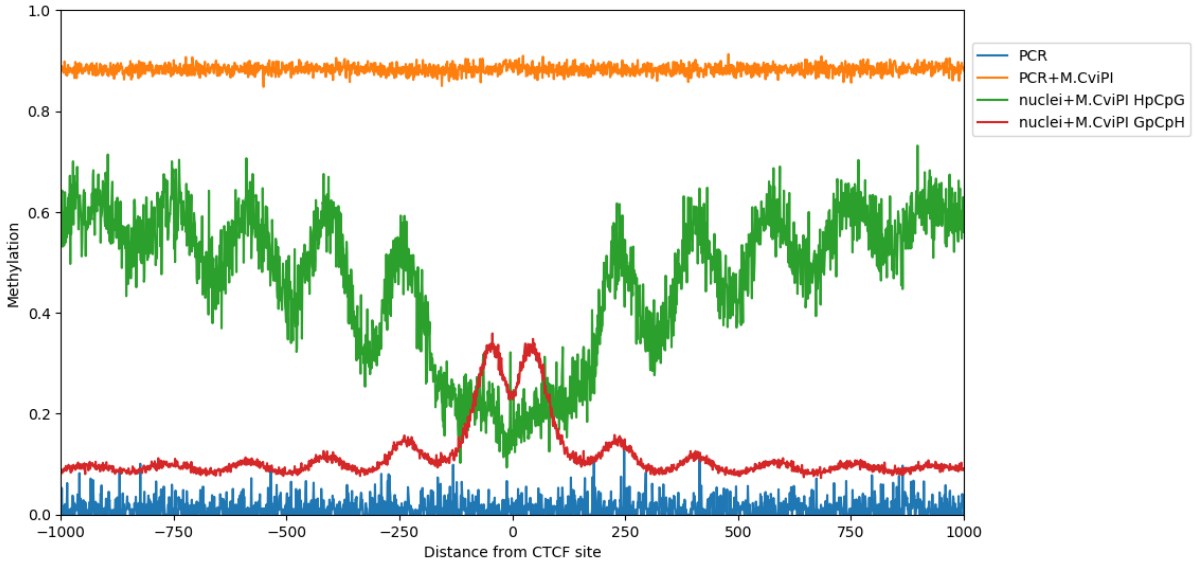
Results

Barcoded training run with name “6pool_meth” on 15. August 2018.
Total basecalled yield for pooled data was 73.4GBp and read lenght N50 around 4-5kbp

Sample	Barcode	Basecalled Yield (Gbp)	Aligned yield
GP5d_PCR	BC01	0.42513	0.279387
GP5d_mCpG	BC02	28.694	23.9027
GP5d_mCpG_mA	BC06	11.1584	8.96248
GP5d_mCpG_GmC	BC05	2.90544	1.88254
GP5d_mA	BC04	1.73859	1.35333
GP5d_GmC	BC03	15.9587	11.1017
	none	10.47	4.19

Apparently ligating the barcodes in BC01, or the subsequent PCR failed badly. The barcode demultiplexing discovered loads of BC01 concatemers. Important is that CpG and GmC datasets are high abundance.

Experiments seem to have worked fine. Below are methylation proportions around CTCF sites for naked PCRed DNA with and without enzymatic methylation and live GP5d nuclear extraction treated with GC methyltransferase and measured for CpG and GpC(artifical) methylation. These are mean fraction methylated on sites with coverage less than 5x (robust) standard deviations above median.


PCR
PCR+M.CviPI
nuclei+M.CviPI HpCpG
nuclei+M.CviPI GpCpH
Methylation
Distance from CTCF site
1.0
0.8
0.6
0.4
0.2
0.0
−1000
−750
−500
−250
0
250
500
750
1000

Others work

Phased read analysis with nanopore, including spaghetti plots

Gigante, S., Gouil, Q., Lucattini, A., Keniry, A., Beck, T., Tinning, M., Gordon, L., Woodruff, C., Speed, T.P., Blewitt, M. and Ritchie, M., 2018. Using long-read sequencing to detect imprinted DNA methylation. *bioRxiv*, p.445924. https://www.biorxiv.org/content/early/2018/10/17/445924

Extremely similar ideas to ours from Winston Timp's lab

https://www.timplab.org/wp-content/uploads/2018/05/180501_CSHL_nanoNOMe_v6post.pdf

References

Sahu, Biswajyoti, Päivi Pihlajamaa, Kaiyang Zhang, Kimmo Palin, Saija Ahonen, Alejandra Cervera, Ari Ristimäki, Lauri A. Aaltonen, Sampsa Hautaniemi, and Jussi Taipale. "Cellular transformation by combined lineage conversion and oncogene expression." *Oncogene* (2021)

Lee, Isac, Roham Razaghi, Timothy Gilpatrick, Michael Molnar, Ariel Gershman, Norah Sadowski, Fritz J. Sedlazeck, Kasper D. Hansen, Jared T. Simpson, and Winston Timp. "Simultaneous profiling of chromatin accessibility and methylation on human cell lines with nanopore sequencing." *Nature Methods* 17, no. 12 (2020): 1191-1199.

Kelly, T. K., Liu, Y., Lay, F. D., Liang, G., Berman, B. P., & Jones, P. A. (2012). Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome research*.

Clark, Stephen J., et al. "scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells." *Nature communications* 9.1 (2018): 781.