

Methplotlib

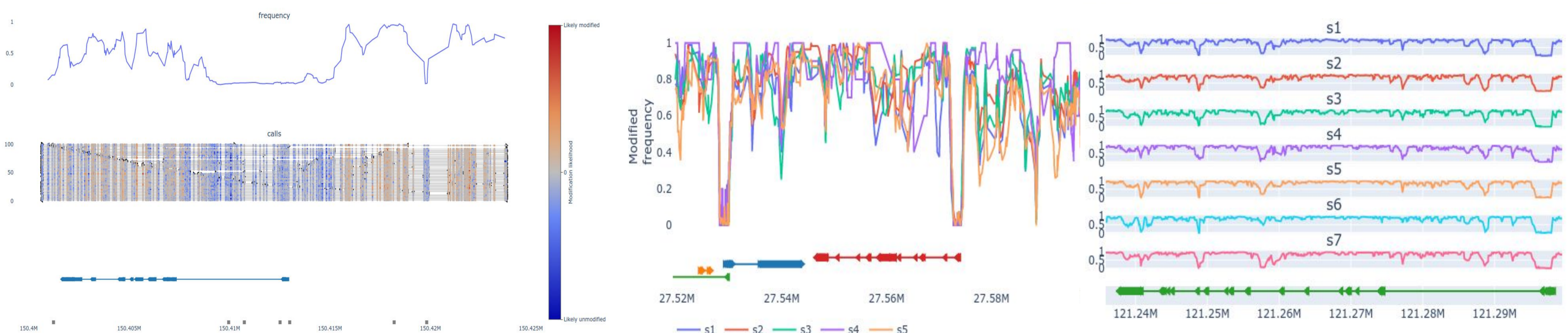
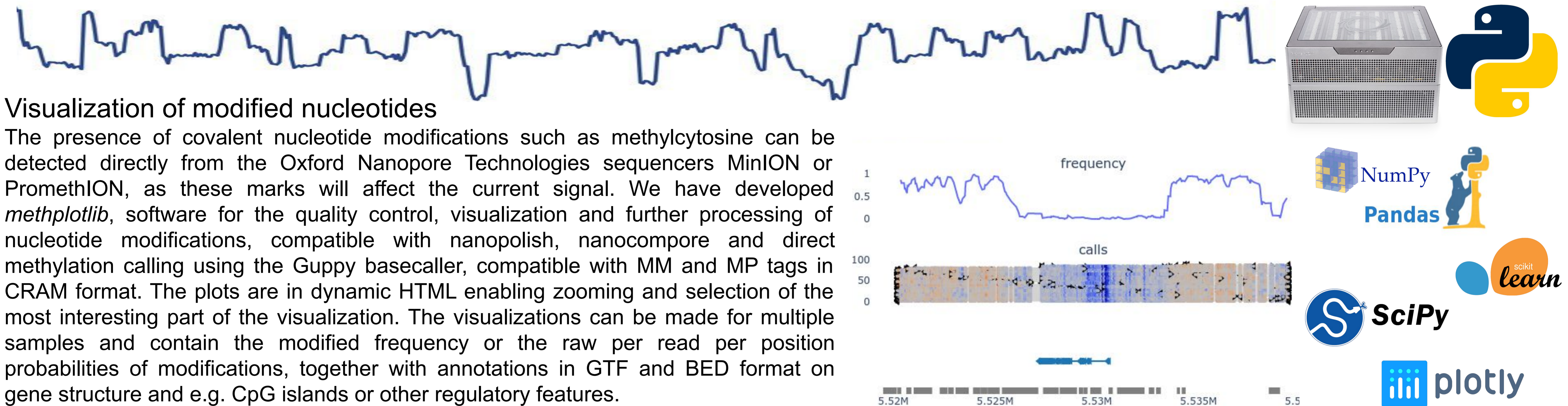
Analysis of modified nucleotides from nanopore sequencing

Wouter De Coster¹, Endre Bakken Stovner^{2,3}, Mojca Strazisar¹

¹ Center for Molecular Neurology, VIB, Antwerp, Belgium

² Department of Computer Science, Norwegian University of Science and Technology, Trondheim 7013, Norway

³ Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim 7013, Norway

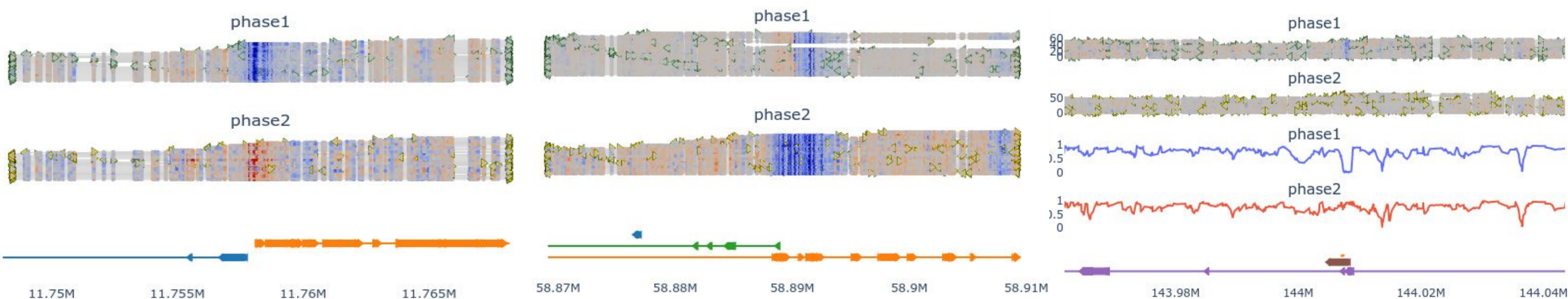


Quality assessment

In addition to the visualization of the signal also various QC plots are created, including barcharts of the number of assayed sites, box plots of the locus-wide methylation frequency, a pairwise correlation plot matrix of the methylation frequencies, a principal components analysis and overlapped histogram of methylated frequencies, all with the aim to provide a birds-eye overview of the data and to enable the detection of large, presumably technical differences. The plots to the right show differences in sequencing depth (s10, s12) and an outlier due to different tissue origin (s10).

Statistical evaluation

The methylation profile can be separated by parental haplotypes by phasing the reads based on SNV genotypes, using e.g. WhatsHap. Methplotlib comes with a couple of companion scripts to split nanopolish results by phase and search for allele-specific methylation deviations using a Fisher Exact test, identify multiple loci such as promoters affected by X chromosome silencing (below left) and non-coding RNAs with parental imprinting (below middle and right). In larger cohorts this will have powerful implications e.g. for the discovery of methylation quantitative trait loci (meQTL) and the evaluation the effect of regulatory variants on local modifications.



Wouter De Coster, Endre Bakken Stovner, Mojca Strazisar. Methplotlib: analysis of modified nucleotides from nanopore sequencing, Bioinformatics, 2020