DOI 10.18699/SBB-2022-118

Applicability of the Oxford Nanopore sequencing technology to the analysis of DNA modifications

Konanov D.N.[1]*, Babenko V.V.[2], Ilina E.N.[1, 2]

[1] *Institute of System Biology and Medicine of Rospotrebnadzor, Moscow, Russia*

[2] *Federal Research and Clinical Center of Physical and Chemical Medicine of FMBA, Moscow, Russia*

* *konanovdmitriy@gmail.com*

Key words: Oxford Nanopore sequencing, DNA modifications

Motivation and Aim: The Oxford Nanopore sequencing technology makes it possible not only to detect the four canonical bases, but also to additionally catch any of their modified forms that can significantly change the sequencing signal. However, the most popular tool used for DNA modification analysis called Tombo has a number of limitations, especially when processing low-represented DNA motifs. In this work, we attempt to improve existing methods for the analysis of DNA modifications and demonstrate their applicability on two highly methylated *H. pylori* genomes.

Methods and Algorithms: Using the Oxford Nanopore technology, both native and whole genome amplified (WGA) samples of two strains of *H. pylori* J99 and A45 were sequenced. The sequencing results were processed with Guppy and resquiggled with Tombo. To detect modified sites, the raw signal distributions of native and WGA samples were compared using the Kolmogorov-Smirnov test in combination with artificial data balancing to compensate the differences in the k-mers abundance. In addition, the Cohen's d-effect size was used to estimate the magnitude of the signal shift. A new greedy algorithm has been developed to extract a minimum set of modified sites based on observed signal shifts. The results obtained for J99 were compared with PacBio data.

Results: We have developed a new statistical approach for detecting modified positions and a new greedy algorithm for extracting potential modification motifs. Unlike Tombo, our method showed results close to the PacBio method on the *H. pylori* J99 strain. A new *H. pylori* strain A45 was analyzed by the same method and revealed the appearance of several methylation motifs other than J99.

Conclusion: Oxford Nanopore sequencing has a great potential for the analysis of epigenetic modifications and can successfully compete with PacBio technology. However, the currently existing software has certain drawbacks caused by the inconsistency of the statistical methods used. The authors hope that their efforts to improve the existing methods will be helpful to ONT users interested in DNA modifications analysis.

Acknowledgements: The study was supported by Government contract No. 122030900064-9.