МЕТОДИКА ОБРАБОТКИ ДАННЫХ ПОЛНОГЕНОМНОГО СЕКВЕНИРОВАНИЯ

**1. Метод секвенирования.**

Выделение ДНК осуществляется при помощи методов описанных в статье [Tyakht et al., 2013]. Подготовка paired-end библеотек и их баркодирование осуществляется из 2 мкг тотальной ДНК, выделенной из метагеномного образца, согласно рекомендациям производителя с использованием наборов NEBNext DNA Library Prep Master Mix Set for Illumina (New England Biolabs, USA) и NEBNext Multiplex Oligos for Illumina (96 Index Primers) (New England Biolabs, USA). Метагеномное секвенирование в формате shotgun проводиться на генетическом анализаторе HiSeq 2500 (Illumina, USA) согласно рекомендациям производителя с использованием следующих наборов реагентов: HiSeq Rapid PE Cluster Kit v2, HiSeq Rapid SBS Kit v2 (500 cycles), HiSeq Rapid PE FlowCell v2. Результатом является набор данных — парных ДНК-прочтений длиной 250 нуклеотидов.

**2. Анализ метагеномных данных включает следующие стадии:**

2.1. Оценка качества секвенирования проводиться с использованием программ FastQC [https://www.bioinformatics.babraham.ac.uk/projects/fastqc/] и Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic) [Bolger et al., 2014] (в дальнейшем возможно использование более новых биоинформатических решений с расширенным функционалом).

2.2 Таксономическое профилирование осуществляется при помощи MetaPhlan2 (https://bitbucket.org/biobakery/metaphlan2) [Truong et al., 2015] или Kraken (https://ccb.jhu.edu/software/kraken/) [Wood et al., 2014].

2.3. Описание функционального потенциала проводиться с использованием алгоритма HUMANn2 (https://bitbucket.org/biobakery/humann2/) [Abubucker et al., 2012]

Для решения некоторых других задач могут быть использованы биоинформатические алгоритмы и базы данных, отвечающих специфике исследования/анализа. Например, использование ShortBRED (https://bitbucket.org/biobakery/shortbred) [Kaminski et al., 2015] (или bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2) [Langmead et al., 2012) и базы данных генов резистентности CARD (https://card.mcmaster.ca/) [Jia et al., 2017] для расчета представленностей генов резистентности к антибиотикам в метагеномных образцах.

**3. Список использованной литературы**

1. Tyakht A. V. et al. Human gut microbiota community structures in urban and rural populations in Russia //Nature communications. – 2013. – Т. 4.

2. Bolger A. M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data //Bioinformatics. – 2014. – Т. 30. – №. 15. – С. 2114-2120.

3. Truong D. T. et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling //Nature methods. – 2015. – Т. 12. – №. 10. – С. 902-903.

4. Wood D. E., Salzberg S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments //Genome biology. – 2014. – Т. 15. – №. 3. – С. R46.

5. Abubucker S. et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome //PLoS computational biology. – 2012. – Т. 8. – №. 6. – С. E1002358.

6. Kaminski J. et al. High-specificity targeted functional profiling in microbial communities with ShortBRED //PLoS computational biology. – 2015. – Т. 11. – №. 12. – С. E1004557.

7. Langmead B., Salzberg S. L. Fast gapped-read alignment with Bowtie 2 //Nature methods. – 2012. – Т. 9. – №. 4. – С. 357-359.

8. Jia B. et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database //Nucleic acids research. – 2017. – Т. 45. – №. D1. – С. D566-D573.