

## **Новые подходы к изучению микробиома кожи и слизистых человека**

*Бакунчик Илья Дмитриевич, Пастерняк Владимир Владимирович*

*Белорусский государственный медицинский университет, Минск*

*Научный(-е) руководитель(-и) – кандидат медицинских наук, доцент Слизень Вероника Вячеславовна, Белорусский государственный медицинский университет, Минск*

### **Введение**

В настоящее время уже расшифровано более 3,3 миллионов генов микроорганизмов нормальной микрофлоры человека, что свидетельствует о том, что кишечный микробиом примерно в 150 раз превышает геном человека. Актуальность представляют исследования, направленные на оценку взаимосвязи между составом микрофлоры человека в норме и патологии.

### **Цель исследования**

Провести анализ существующих подходов к изучению микробиома человека и разработать алгоритм его расшифровки с использованием высокоскоростного секвенирования метагеномных библиотек ДНК.

### **Материалы и методы**

Проведен анализ алгоритма изучения микробиома человека. В качестве источников для составления каталога геномов использовались база проекта Human Microbiome Project (<http://www.hmpdacc.org>) и NCBI (<http://www.ncbi.nlm.nih.gov>). Проведен анализ программного обеспечения для оценки метагеномных баз, а также для идентификации состава полимикробных ассоциаций: MUMMER 3.0, bowtie, rdp, TopHat, GLIMMER 3.0, GeneMarkS+, MetaGeneMark.

### **Результаты**

Короткие сиквенсы ДНК, или риды, отфильтровывают от мелких (до 35 п.о.) и эукариотических сиквенсов. Риды выравниваются с помощью высокопроизводительного алгоритма выравнивания – bowtie - по отношению к референс наборам генов и геномов в базах «Human Microbiome Project» и «NCBI». Геномы, которые не похожи ни на один другой геном более чем на 80%, включаются в каталог. Оставшиеся геномы кластрируют по порогу сходства 80% на 80% длины. Из каждого кластера схожих геномов выбирают по одной репрезентативной последовательности, которые добавляют в каталог. По результатам выравнивания, для каждой последовательности подсчитываются два типа данных: глубина покрытия референс- последовательности, или суммарную длину ридов легших на данную референс- последовательность, и ширину покрытия – процент длины референс-генома, покрытого хотя бы единожды. По первому типу данных подсчитывается уровень представленности каждого микроорганизма или функциональной группы генов, а по второму – задается порог идентификации микроорганизма, то есть если ридом покрыт, по крайней мере, 1% длины референс последовательности, она учитывается. Видовая принадлежность последовательностей из составленного каталога

определяется программой RDP по сходству с базой 16S рРНК генов референс-геномов микроорганизмов.

### **Выводы**

Составленный алгоритм *in silico* анализа результатов высокоскоростного секвенирования метагеномных библиотек ДНК, позволит охарактеризовать микробиом