

## Вклад методов молекулярной филогенетики в контроль разнообразия микробиома легких у пациентов с муковисцидозом при проведении комплексной терапии

Воронина О.Л.<sup>1\*</sup>, Рыжова Н.Н.<sup>1</sup>, Кунда М.С.<sup>1</sup>, Ермолова Е.И.<sup>1</sup>, Кагазеев Р.У.<sup>2</sup>,  
Амелина Е.Л.<sup>2</sup>, Гинцбург А.Л.<sup>1</sup>

<sup>1</sup> ФГБУ НИЦЭМ им. Н.Ф. Гамалеи Минздрава России, Москва, Россия

<sup>2</sup> ФГБУ НИИ пульмонологии ФМБА России, Москва, Россия

\* olv550@gmail.com

**Ключевые слова:** микробиом; *Pseudomonadota*; муковисцидоз; филогенетическое разнообразие; биомаркеры микробиома

**Мотивация и цель:** Применение методов секвенирования ДНК в изучении разнообразия микробиоты различных биотопов человеческого организма, создание баз данных и программ для аннотирования полученных прочтений позволили охарактеризовать разнообразные микробиоценозы условно здоровых людей [1]. При заболевании происходят изменения в составе микробного сообщества. Для пациентов с муковисцидозом (МВ) наиболее важным отслеживаемым биотопом является респираторный тракт, главным образом нижние дыхательные пути. С 2016 г. в терапии российских пациентов с МВ начали использовать модуляторы CFTR (cystic fibrosis transmembrane conductance regulator), способствующие фолдингу и поддержанию в активном состоянии белка хлорного канала, что нормализует функции эпителиоцитов и приводит к улучшению общего состояния пациентов [2]. Для хронически инфицированных пациентов путь восстановления микробиоты респираторного тракта долг и требует подбора комплексной терапии. В анализе эффективности протокола лечения помогает сравнение микробиомов мокроты на этапах терапии. Ключевыми в сравнительном исследовании образцов являются методы молекулярной филогенетики и основанные на них филогенетические протоколы [3]. Мы применили такой протокол для многофакторного исследования образцов взрослых пациентов с МВ, хронически инфицированных клинически значимыми *Pseudomonadota*, для определения эффективности комплексной терапии, включающей антибактериальные и таргетные препараты.

**Методы и алгоритмы:** Выборка включала 103 образца мокроты от 52 взрослых пациентов с МВ, из которых 56 % составили женщины и 44 % – мужчины. Образцы, собранные в течение 11 месяцев, группировали по следующим параметрам: возраст пациентов (18–23, 24–30, 31–56); мутации в гене *cftr* (F508del/F508del, F508del/неF508de, неF508del/неF508del); фенотип (мягкий, тяжелый); степень тяжести болезни легких: легкое течение (mild lung disease, MiLD, ОФВ1  $\geq$  70 %), средней тяжести (moderate lung disease, MoLD, 40 %  $\leq$  ОФВ1 < 70 %), тяжелое течение (severe lung disease, SLD, ОФВ1 < 40 %); клиническое состояние (вне обострения, baseline, В; обострение, exacerbation, Е; эпизодическая антибиотикотерапия, treatment, АТ), инфицирующие микроорганизмы (РСН – хронически инфицированные клинически значимыми *Pseudomonadota*; Р – с появляющимися

клинически значимыми *Pseudomonadota*; NP – без клинически значимых *Pseudomonadota* в микробиоме легких); таргетная терапия (наличие – ТТ, отсутствие – НТТ).

Тотальную ДНК из образцов выделяли с помощью MagAttract HMW DNA Kit (QIAGEN, США). ДНК анализировали с помощью экспресс-диагностики, включающей амплификацию и секвенирование по Сэнгеру следующих мишеней: 16S rDNA, *trpE* – *P. aeruginosa*, *gltB*, *gyrB* – *Burkholderia*, *nrdA* – *Achromobacter*, *adkA* – *E. coli*, *rpoB* – *K. pneumonia*, а также с помощью массового параллельного секвенирования ампликонов 16S rDNA (V1-V5). Результаты депонировали в GenBank (биопроект PRJNA717158). Для анализа прочтений использовали Microbial Genomics Module пакета CLC Genomic Workbench v.21.0.1. (QIAGEN, США), для определения операционных таксономических единиц (Operational Taxonomic Unit, OTU) – базу данных Greengenes v13\_8 с уровнем сходства 97 %. Для оценки альфа-разнообразия применяли индексы Shannon entropy и коэффициент филогенетического разнообразия (PD). Достоверность отличия альфа-разнообразия по группам образцов рассчитывали с помощью H-критерия Крускала–Уоллиса и U-критерия Манна–Уитни, а также подтверждали анализом PERMANOVA (p-Value < 0.05). Бета-разнообразие оценивали с помощью индекса Euclidean distance, а также метрик UniFrac. Для подробного анализа сходства между образцами использовали Principal Coordinate Analysis (PCoA) и статистические функции MS Excel. Для обнаружения таксонов, являющихся маркерами различий между анализируемыми группами, использовали алгоритм LEfSe (Linear discriminant analysis Effect Size) [4]. В расчет брали только достоверные значения LDA Score (log 10) >4–5.

**Результаты:** Анализ всей выборки образцов показал отсутствие зависимости филогенетического разнообразия (PD) от фенотипа мутации и наличия/отсутствия таргетной терапии. Вместе с тем PD достоверно снижалось в образцах РСН пациентов по сравнению с Р и NP группами, что было характерно как для наиболее массовой подгруппы инфицированных *P. aeruginosa*, так и для подгрупп с *Burkholderia* spp., *Achromobacter* spp. и *E. coli*. Наиболее низкие показатели PD наблюдали в образцах SLD пациентов, достоверно отличавшиеся от образцов MoLD и MiLD групп. Сравнение PD образцов в разных состояниях пациентов выявило наиболее высокие показатели вне обострения и достоверно более низкие в состояниях Е и АТ. Оценка бета-разнообразия с помощью индекса Euclidean distance продемонстрировала группировку образцов по основным возбудителям инфекции с миграцией в направлении групп Р и NP образцов РСН *P. aeruginosa* при доле *Staphylococcus* 9–48 %. Использование метрики «взвешенный не нормализованный UniFrac», учитывающей значимость редких таксонов, позволило выявить больше мигрирующих РСН образцов, в том числе тех, в которых нарастала доля *Streptococcus* (45 %) или *Capnocytophaga* (41 %) на этапах терапии. Алгоритм LEfSe применяли для выявления маркерных микроорганизмов для групп многофакторного анализа. Так, образцы SLD пациентов отличали *Burkholderiales*, а группы MoLD и MiLD – представители *Actinomycetota* и *Bacillota* соответственно. Маркерами образцов, сгруппированных по мутациям пациентов, были *Bacillota* при мутации F508del/F508del, *Pseudomonas* при F508del/неF508del и *Enterobacteriaceae* при неF508del/неF508del. Маркерами образцов в состоянии вне обострения были *Streptococcus* и *Bacteroidota* (прежде всего, *Prevotella*), в обострении – *Flavobacteriia* и *Capnocytophaga*, после

эпизодической антибиотикотерапии – *Bacilli* (*Staphylococcus*) и *Achromobacter*. В группах образцов по времени приема таргетного препарата *Burkholderia* была маркером двух лет терапии, а *Achromobacter* – 5-летнего лечения модуляторами CFTR. Таким образом, сокращение доли *Burkholderiales* происходит очень медленно даже при длительной таргетной терапии на фоне поддерживающей и эпизодической антибиотикотерапии. Таргетная терапия не исключает обострений, после которых доля *Burkholderiales* возрастает. Снижение доли другого представителя *Pseudomonadota*, *P. aeruginosa*, приводит к другой проблеме – росту доли *Staphylococcus*, конкурента *P. aeruginosa* [5]. Анализ образцов показал, что PD сохраняется на высоком уровне при 4–29 % *Staphylococcus* в микробиоме и достоверно снижается в образцах как с низкой долей *Staphylococcus* (<4 %), так и с высокой (>30 %). При восстановлении микробиома наблюдается дисбаланс для многих комменсалов ротовой полости, поступающих в нижние дыхательные пути посредством микроаспирации, например, *Streptococcus*, *Prevotella*, *Rothia*. Предполагается роль этих бактерий и их метаболитов в воспалительных процессах в легких [6].

**Выводы:** Исследования показали, что возвращение сбалансированного разнообразия микробиома легких у хронически инфицированных взрослых пациентов с МВ возможно при длительной комплексной терапии и внимании к представителям всех таксонов, составляющих биотоп нижних дыхательных путей, выявить которые позволяют методы молекулярной филогенетики.

**Финансирование:** Исследование поддержано Государственным заданием НИЦЭМ им. Н.Ф. Гамалеи (№ 056-00066-23-00).

## **The contribution of molecular phylogenetic methods to the control of the diversity of the lung microbiome in patients with cystic fibrosis during complex therapy**

Voronina O.<sup>1\*</sup>, Ryzhova N.<sup>1</sup>, Kunda M.<sup>1</sup>, Ermolova E.<sup>1</sup>, Kagazetchev R.<sup>2</sup>, Amelina E.<sup>2</sup>, Gintsburg A.<sup>1</sup>

<sup>1</sup> N.F. Gamaleya National Research Center for Epidemiology and Microbiology, Ministry of Health of Russia, Moscow, Russia

<sup>2</sup> Pulmonology Research Institute under FMBA of Russia, Moscow, Russia

\* olv550@gmail.com

**Key words:** microbiome; *Pseudomonadota*; cystic fibrosis; phylogenetic diversity; microbiome biomarkers

**Motivation and Aim:** The use of DNA sequencing methods in studying the diversity of the microbiota of various biotopes of the human body, the creation of databases and programs for annotating the obtained reads made it possible to characterize the various microbiocenoses of relatively healthy people [1]. During disease, changes occur in the composition of the microbial community. For patients with cystic fibrosis (CF), the most important biotope monitored is the respiratory tract, mainly the lower respiratory tract. Since 2016, CFTR (cystic fibrosis transmembrane conductance regulator) modulators have been used in the treatment of Russian patients with CF, which promote the folding and maintenance of the chloride channel protein in an active state, which normalizes the functions of epithelial cells and leads to an improvement in the general condition of

patients [2]. For chronically infected patients, the path to restoring the microbiota of the respiratory tract is long and requires the selection of complex therapy. Comparison of sputum microbiomes at therapy stages helps in analyzing the effectiveness of a treatment protocol. The key to the comparative study of samples are the methods of molecular phylogenetics and phylogenetic protocols based on them [3]. We applied such a protocol to a multifactorial study of samples from adult CF patients chronically infected with clinically significant *Pseudomonadota* to determine the effectiveness of combination therapy, including antibacterial and targeted drugs.

**Methods and Algorithms:** 103 sputum samples from 52 adult patients with CF, of whom 56 % were women and 44 % were men, were included in the analysis. Samples collected over 11 months were grouped according to the following parameters: patient age (18–23, 24–30, 31–56); mutations in the *cfr* gene (F508del/F508del, F508del/non-F508del, non-F508del/non-F508del); phenotype (mild, severe); severity of lung disease: mild (mild lung disease, MiLD, FEV1  $\geq$  70 %), moderate (moderate lung disease, MoLD, 40 %  $\leq$  FEV1 < 70 %), severe (severe lung disease, SLD, FEV1 < 40 %); clinical condition (out of exacerbation, baseline, B; exacerbation, E; episodic antibiotic therapy, treatment, AT), infecting microorganisms (PCH – chronically infected with clinically significant *Pseudomonadota*; P – with emerging clinically significant *Pseudomonadota*; NP – without clinically significant *Pseudomonadota* in lung microbiome); targeted therapy (presence – TT, absence – NTT). Total DNA from the samples was isolated using MagAttract HMW DNA Kit (QIAGEN, USA). DNA was analyzed using express diagnostics, including amplification and Sanger sequencing of the following targets: 16S rDNA, *trpE* – *P. aeruginosa*, *gltB*, *gyrB* – *Burkholderia*, *nrdA* – *Achromobacter*, *adkA* – *E. coli*, *rpoB* – *K. pneumonia*, and using massively parallel sequencing of 16S rDNA amplicons (V1-V5).

The results were deposited in GenBank (bioproject PRJNA717158). Microbial Genomics Module of the CLC Genomic Workbench v.21.0.1 package (QIAGEN, USA) was used for analysis of reads. Greengenes v13\_8 database with a similarity level of 97 % was used for determination Operational Taxonomic Units (OTU). Shannon entropy index and the coefficient of phylogenetic diversity (PD) were used to assess alpha diversity. The significance of the difference in alpha diversity between groups of samples was calculated using the Kruskal–Wallis H test and the Mann–Whitney U test, and was also confirmed by PERMANOVA analysis (p-Value < 0.05). Beta diversity was assessed using the Euclidean distance index, as well as UniFrac metrics. For a detailed analysis of the similarity between samples, Principal Coordinate Analysis (PCoA) and MS Excel statistical functions were used. To detect taxa that are markers of differences between the analyzed groups, the LEfSe (Linear discriminant analysis Effect Size) algorithm was used [4]. Only reliable values of LDA Score (log 10) >4–5 were taken into account.

**Results:** Analysis of the entire sample collection showed no dependence of phylogenetic diversity (PD) on the phenotype of the mutation and the presence/absence of targeted therapy. At the same time, PD significantly decreased in the samples of PCH patients compared to the P and NP groups, which was typical both for the most widespread subgroup of those infected with *P. aeruginosa*, and for the subgroups with *Burkholderia* spp., *Achromobacter* spp. and *E. coli*. The lowest PD rates were observed in the samples of SLD patients, which were significantly different from the samples of the MoLD and MiLD groups. Comparison of PD samples in different patient states revealed the highest rates outside of exacerbation and significantly lower ones in states E and AT.

Assessment of beta diversity using the Euclidean distance index demonstrated grouping of samples according to the main pathogens with migration towards the P and NP groups of PCH *P. aeruginosa* samples with a proportion of *Staphylococcus* of 9–48 %. The use of the “weighted non-normalized UniFrac” metric, which takes into account the significance of rare taxa, made it possible to identify more migrating PCH samples, including those in which the proportion of *Streptococcus* (45 %) or *Capnocytophaga* (41 %) increased during the treatment stages. The LEfSe algorithm was used to identify marker microorganisms for multivariate analysis groups. Thus, the SLD patient samples were distinguished by *Burkholderiales*, and the MoLD and MiLD groups were distinguished by representatives of *Actinomycetota* and *Bacillota*, respectively. Markers for samples grouped by patient mutations were *Bacillota* for F508del/F508del mutation, *Pseudomonas* for F508del/non-F508del, and *Enterobacteriaceae* for non-F508del/non-F508del. The markers of the samples in the non-exacerbation state were *Streptococcus* and *Bacteroidota* (primarily *Prevotella*), in the exacerbation – *Flavobacteriia* and *Capnocytophaga*, after episodic antibiotic therapy – *Bacilli* (*Staphylococcus*) and *Achromobacter*. In groups of samples based on the time of taking the targeted drug, *Burkholderia* was a marker of two years of therapy, and *Achromobacter* was a marker of 5 years of treatment with CFTR modulators. Thus, the reduction in the proportion of *Burkholderiales* occurs very slowly even with long-term targeted therapy against the background of maintenance and episodic antibiotic therapy. Targeted therapy does not exclude exacerbations, after which the proportion of *Burkholderiales* increases. A decrease in the share of another representative of *Pseudomonadota*, *P. aeruginosa*, leads to another problem – an increase in the share of *Staphylococcus*, a competitor of *P. aeruginosa* [5]. Analysis of the samples showed that PD remains at a high level with 4–29 % *Staphylococcus* in the microbiome and is significantly reduced in both samples with a low proportion of *Staphylococcus* (<4 %) and high (>30 %). When the microbiome is restored, an imbalance is observed for many oral commensals entering the lower respiratory tract through microaspiration, for example, *Streptococcus*, *Prevotella*, *Rothia*. The role of these bacteria and their metabolites in inflammatory processes in the lungs is assumed [6].

**Conclusion:** Studies have shown that the return of balanced diversity of the lung microbiome in chronically infected adult patients with CF is possible with long-term complex therapy and attention to representatives of all taxa that make up the biotope of the lower respiratory tract, which can be identified by molecular phylogenetics methods. **Funding:** The study is supported by the State assignment of the N.F. Gamaleya National Research Center for Epidemiology and Microbiology (No. 056-00066-23-00).

### Список литературы/References

1. Dickson R.P., Erb-Downward J.R., Freeman C.M., McCloskey L., Beck J.M., Huffnagle G.B., Curtis J.L. Spatial variation in the healthy human lung microbiome and the adapted island model of lung biogeography. *Ann Am Thorac Soc*. 2015;12(6):821-830. doi 10.1513/AnnalsATS.201501-029OC
2. Jia S., Taylor-Cousar J.L. Cystic fibrosis modulator therapies. *Annu Rev Med*. 2023;74:413-426. doi 10.1146/annurev-med-042921-021447
3. Jermiin L.S., Catullo R.A., Holland B.R. A new phylogenetic protocol: dealing with model misspecification and confirmation bias in molecular phylogenetics. *NAR Genom Bioinf*. 2020;2(2):lqaa041. doi 10.1093/nargab/lqaa041
4. Segata N., Izard J., Waldron L., Gevers D., Miropolsky L., Garrett W.S., Huttenhower C. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60. doi 10.1186/gb-2011-12-6-r60
5. Воронина О.Л., Рыжова Н.Н., Кунда М.С., Аксенова Е.И., Зигангирова Н.А. и др. *Pseudomonas aeruginosa*. Ассистенты и конкуренты в микробиоме инфицированных легких больных

муковисцидозом. *Медицинский вестник Северного Кавказа*. 2020;15(2):186-191. doi 10.14300/mnnc.2020.15045

[Voronina O.L., Ryzhova N.N., Kunda M.S., Aksenova E.I., Zigangirova N.A. et al. *Pseudomonas aeruginosa*. Assistants and competitors in the microbiome of infected CF patients' lungs. *Meditsinskiy vestnik Severnogo Kavkaza*. 2020;15(2):186-191. doi 10.14300/mnnc.2020.15045 (in Russian)]

6. Natalini J.G., Singh S., Segal L.N. The dynamic lung microbiome in health and disease. *Nat Rev Microbiol*. 2023;21:222-235. doi 10.1038/s41579-022-00821-x