

Кишечная микробиота при иммунотерапии меланомы: бактериофаги как прогностические биомаркеры и терапевтические агенты

Захаревич Н.*, Олехнович Е., Климина К.

Федеральный научно-клинический центр физико-химической медицины имени академика

Ю.М. Лопухина ФМБА, Москва, Россия

* zakharevich@yandex.ru

Ключевые слова: кишечная микробиота; меланома; виром; бактериофаги; биомаркеры; иммунотерапия; метагеном

Мотивация и цель: Роль бактериофагов по влиянию на структуру и функции кишечной микробиоты (КМ) недооценена. В связи с быстрой эволюцией устойчивости бактерий к антибиотикам, фаготерапия является перспективным и актуальным подходом. Многие заболевания, в том числе и злокачественные опухоли, возникают и/или сопровождаются конкретными изменениями в КМ. Так, например, в последнее время КМ признана новым значимым игроком в патогенезе и лечении злокачественной меланомы. В ряде исследований было показано, что состав КМ значительно отличается у пациентов с меланомой по сравнению с контрольной группой, а также что КМ может влиять на противоопухолевый иммунитет у пациентов с меланомой и на эффективность иммунотерапии. Несмотря на это, исследования взаимодействия различных компонентов КМ и злокачественной меланомы все еще находятся в зачаточном состоянии – в связи с чем их изучение является актуальной задачей. В частности, необходимо с большей точностью выявлять различные микробные сигнатуры (в том числе патогенных бактерий), предрасполагающие к развитию побочных эффектов при заболевании и/или снижению эффективности иммунотерапии. Успешно снизить патогенную нагрузку и улучшить течение заболевания возможно, например, при специфическом нацеливании бактериофагов на конкретные бактериальные виды или даже штаммы. Помимо этого, известно, что фаги, наравне с бактериями, способны усиливать ответ на иммунотерапию индуцируя Т-клетки, перекрестно-реагирующие с раковыми антигенами. Исследования КМ, а именно кишечного вирома пациентов с меланомой и результаты, полученные в ходе данного исследования, способны обогатить терапевтический потенциал модулирования КМ и стимуляции системных противоопухолевых иммунных реакций, а также предложить новые неинвазивные биомаркеры заболевания – бактериофаги.

Методы и алгоритмы: С помощью инструмента fastq-dump (NCBI SRA Toolkit) сырые метагеномные данные были загружены из базы данных NCBI. Оценка качества полученных данных была выполнена с помощью программы FastQC. Тримминг данных был проведён при помощи инструмента Trimmomatic, а фильтрация последовательностей человеческой ДНК с использованием инструмента bbmap (версия генома человека GRCh37). Для реконструкции фаговых геномов из метагеномных последовательностей была проведена сборка сырых метагеномных данных (прочтений) в контиги параллельно с помощью двух

инструментов *metaviralSPAdes* и *MEGANIT*. Для формирования качественного неизбыточного набора вирусных контигов (геномов) было выполнено предсказание последовательностей фагов и профагов в собранных метагеномных данных при помощи комбинации восьми алгоритмов: *viralVerify*, *VirSorter2*, *DeepVirFinder*, *Seeker*, *VIBRANT*, *geNomad*, *DBSCAN-SWA* и *Phigaro* – с последующим сравнением полученных результатов и отбором надежных предсказаний. Чтобы гарантировать качество и надежность идентифицированных вирусных контигов (геномов), был использован инструмент *CheckV*, позволяющий проверить качество и полноту предсказанных вирусных геномов. Последовательности с низким уровнем качества по результатам оценки инструмента *CheckV* были исключены из дальнейшего анализа. Затем с помощью инструмента *dRep* отобранные фаговые геномы были кластеризованы методом дерепликации. Полученный качественный неизбыточный каталог вирусных контигов был проанализирован с помощью библиотеки *Python Phage BOX* – являющейся набором инструментов для анализа фагов, включающим в себя: таксономическую классификацию вирусных контигов – инструмент *PhaGCN*, предсказание хозяев фагов – инструмент *CHERRY*, и определение типа взаимодействия бактериофага с бактериальной клеткой (вирулентный/умеренный) – инструмент *PhaTYP*. Для выявления патогенных бактерий в собранных метагеномных данных была применена следующая стратегия анализа: в качестве референса была использована база данных патогенных организмов *FDA-ARGOS*, картирование собранных метагеномов осуществляли при помощи алгоритма *HISAT2*; при помощи пакета программ *SAMtools*, результаты картирования были конвертированы из *SAM* формата в *BAM* формат. Расчет относительных представлений геномов патогенов производили при помощи программы *CoverM*. Обработка полученных результатов и расчет средней представленности бактерий проводили при помощи скриптов, написанных на языке *R* с использованием стандартных библиотек.

Результаты: В ходе проведенного исследования был сформирован каталог кишечных метагеномов от пациентов с меланомой с различным ответом на иммунотерапию. Этот каталог включает в себя 83 кишечных метагенома (образца) из двух опубликованных исследований [1, 2]. Анализируемые в проекте данные доступны в базе данных *NCBI* по следующим номерам доступа: *PRJNA397906* и *PRJNA399742*. По результатам идентификации последовательностей бактериофагов в собранном каталоге кишечных метагеномов при помощи восьми алгоритмов для 83 образцов был предсказан 204751 потенциальный вирусный геном. После проверки качества и полноты предсказанных геномов осталось 182423 последовательности; а после этапа дерепликации было получено 46380 вирусных контигов – из которых и был сформирован окончательный неизбыточный качественный набор фаговых геномов. После таксономической аннотации полученного набора неизбыточных вирусных контигов получилось, что все последовательности принадлежат к классу *Caudoviricetes* и распределены по 19 различным семействам: *Peduviridae* (10.38 %), *Straboviridae* (7.25 %), *Ackermannviridae* (3.71 %), *Salasmaviridae* (3.29 %), *Casjensviridae* (2.85 %), *Chaseviridae* (2.80 %), *Mesyazhinovviridae* (2.11 %), *Herelleviridae* (1.84 %), *Drexelviridae* (1.33 %), *Zierdtviridae* (0.94 %), *Kyanoviridae* (0.54 %), *Guelinviridae* (0.35 %), *Autographiviridae* (0.25 %), *Schitoviridae* (0.23 %), *Demereciviridae* (0.19 %), *Orlajensenviridae* (0.13 %), *Vilmaviridae* (0.11 %), *Rountreeviridae* (0.10 %),

Zobellviridae (0.01 %) и “no_family_available” (1.88 %). Также для всех предсказанных вирусных контигов был определён тип взаимодействия с бактериальной клеткой – в результате для данных из исследования Frankel et al. [1] получилось, что вирулентные фаги составляют 32.95 %, а умеренные фаги 67.05 %; для данных из исследования Matson et al. [2] были получены сходные значения – 34.02 % для вирулентных фагов и 65.98 % для умеренных. Полученные результаты указывают на преобладание умеренных фагов в КМ, что соответствует опубликованным на сегодняшний день исследованиям. Как видно из описанного выше, разницы в процентном соотношении между вирулентными и умеренными фагами не наблюдалось, ни между наборами данных, ни между образцами метабеномов ответчиков (R) и неответчиков (NR). Проанализировав ассоциации между фагами и бактериальной составляющей КМ, были сделаны следующие выводы – во всех исследуемых образцах, независимо от набора данных, доминирующими хозяевами были десять бактериальных родов: *Clostridium*, *Faecalibacterium*, *Mycoplasma*, *Bifidobacterium*, *Ruminococcus*, *Bacteroides*, *Blautia*, *Streptomyces*, *Eubacterium* и семейство *Lachnospiraceae*. Что касается семейств фагов, то увеличение численности наблюдалось для четырёх, а именно: для набора образцов из исследования Frankel et al. [1] это были семейства *Drexelviriidae* (для ответчиков (R)) и *Salasmaviridae* (для неответчиков (NR)); для набора образцов из исследования Matson et al. [2] это были семейства *Chaseviridae* и *Mesyanzhinovviridae* (оба для неответчиков (NR)). Для всех фаговых семейств с увеличением численности, также наблюдался и более широкий спектр бактериальных хозяев. По предварительным результатам поиска бактериофагов – кандидатов в прогностические биомаркеры ответа на иммунотерапию было выделено два фаговых семейства *Drexelviriidae* и *Chaseviridae* в качестве наиболее перспективных кандидатов в прогностические биомаркеры. Однако для окончательных выводов необходимо провести более глубокий анализ большего числа данных. Что касается бактериальной составляющей КМ пациентов с меланомой, в 'Топ 5' патогенных видов с наибольшей средней представленностью среди исследованных образцов вошли такие бактерии как *Enterocloster bolteae*, *Clostridium scindens*, *Bacteroides fragilis*, *Raoultella ornithinolytica*, *Enterocloster clostridioformis*.

Выводы: На сегодняшний день нет работ, в которых бы анализировался виром ответивших и не ответивших на иммунотерапию онкологических больных. Наше исследование по возможности применения бактериофагов КМ в качестве прогностических биомаркеров исхода иммунотерапии ещё не закончено и мы планируем проанализировать больший объём данных и более глубоко, в том числе использовать в дальнейшем для поиска воспроизводимых таксономических прогностических биомаркеров метод дифференциального ранжирования с применением уже зарекомендовавшего себя в анализе сложных метаданных инструмента Songbird – этот метод позволяет оценивать относительную дифференциальную численность, независимо от общей микробной нагрузки. Таким образом, текущее исследование имеет шанс существенно расширить представления о вовлеченности КМ, в частности вирусной её составляющей, в иммунотерапию, а также предложить новые подходы к усовершенствованию существующих методов борьбы с злокачественной меланомой.

Финансирование: Исследование выполнено за счет гранта Российского научного фонда (№ 23-75-10125, <https://rscf.ru/project/23-75-10125/>).

Gut microbiota in melanoma immunotherapy: bacteriophages as prognostic biomarkers and therapeutic agents

Zakharevich N.*, Olekhnovich E., Klimina K.

Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia

* zakharevich@yandex.ru

Key words: gut microbiota; melanoma; virome; bacteriophages; biomarkers; immunotherapy; metagenome

Motivation and Aim: The role of bacteriophages in influencing the structure and function of the gut microbiota (GM) is underestimated. Due to the rapid evolution of bacterial resistance to antibiotics, phage therapy is a promising and relevant approach. Many diseases, including malignant tumors, occur and/or are accompanied by specific changes in GM. For example, recently GM has been recognized as a new significant player in the pathogenesis and treatment of malignant melanoma. In a number of studies, it has been shown that the composition of GM differs significantly in patients with melanoma compared to the control group, and also that GM can affect antitumor immunity in patients with melanoma and the effectiveness of immunotherapy. Despite this, research of the interaction of various components of GM and malignant melanoma are still in their infancy – and therefore their study is an urgent task. In particular, it is necessary to identify with greater accuracy various microbial signatures (including pathogenic bacteria) predisposing to the development of side effects in the disease and/or a decrease in the effectiveness of immunotherapy. It is possible to successfully reduce the pathogenic load and improve the course of the disease, for example, by specifically targeting bacteriophages to specific bacterial species or even strains. In addition, it is known that phages, along with bacteria, are able to enhance the response to immunotherapy by inducing T cells that cross-react with cancer antigens. Studies of GM, namely intestinal virome in patients with melanoma and the results obtained during this study, can enrich the therapeutic potential of modulating GM and stimulating systemic antitumor immune responses, as well as offer new non-invasive biomarkers of the disease – bacteriophages.

Methods and Algorithms: Using the fastq-dump tool (NCBI SRA Toolkit), raw metagenomic data was downloaded from the NCBI database. The evaluation of the quality of the obtained data was performed using the FastQC program. Data trimming was performed using the Trimomatic tool, and filtering of human DNA sequences using the bbmap tool (GRCh37 version of the human genome). To reconstruct phage genomes from metagenomic sequences, raw metagenomic data were assembled into contigs in parallel using two tools metaviralSPAdes and MEGAHIT. To form a qualitative, non-redundant set of viral contigs (genomes), the prediction of phages and prophages sequences in the collected metagenomic data was performed using a combination of eight algorithms: viralVerify, VirSorter2, DeepVirFinder, Seeker, VIBRANT, geNomad, DBSCAN-SWA and Phigaro, followed by a comparison of the results obtained and the selection of reliable predictions. To ensure the quality and reliability of the identified viral contigs (genomes), the CheckV tool was used to check the quality and completeness of the predicted viral genomes. Sequences with a low level of quality according to the results of the CheckV tool evaluation were excluded from further

analysis. Then, using the dRep tool, the selected phage genomes were clustered using the dereplication method. The resulting high-quality, non-redundant catalog of viral contigs was analyzed using the Python Phage BOX library, which is a set of tools for phage analysis, including: The PhaGCN tool provides taxonomic classification of viral contigs, the CHERRY tool predicts phage hosts, and the PhaTYP tool determines the type of interaction between a bacteriophage and a bacterial cell (virulent/temperate). To identify pathogenic bacteria in the collected metagenomic data, the following analysis strategy was applied: the FDA-ARGOS database of pathogenic organisms was used as a reference, the mapping of the collected metagenomes was carried out using the HISAT2 algorithm, using the SAMtools software package, the mapping results were converted from SAM format to BAM format. The relative representation of pathogen genomes was calculated using the CoverM program. The processing of the obtained results and the calculation of the average bacterial representation were carried out using scripts written in the R language using standard libraries.

Results: During the study, a catalog of gut metagenomes from melanoma patients with different responses to immunotherapy was formed. This catalog includes 83 gut metagenomes (samples) from two published studies [1, 2]. The data analyzed in the project is available in the NCBI database using the following access numbers: PRJNA397906 and PRJNA399742. Based on the results of the identification of bacteriophage sequences in the collected catalog of gut metagenomes, 204,751 potential viral genomes were predicted using eight algorithms for 83 samples. After checking the quality and completeness of the predicted genomes, 182,423 sequences remained; and after the dereplication stage, 46,380 viral contigs were obtained – from which the final non-redundant qualitative set of phage genomes was formed. After taxonomic annotation of the obtained set of non-redundant viral contigs, it turned out that all sequences belong to the class *Caudoviricetes* and are distributed into 19 different families: *Peduviridae* (10.38 %), *Straboviridae* (7.25 %), *Ackermannviridae* (3.71 %), *Salasmaviridae* (3.29 %), *Casjensviridae* (2.85 %), *Chaseviridae* (2.80 %), *Mesyazhinovviridae* (2.11 %), *Herelleviridae* (1.84 %), *Drexelviriidae* (1.33 %), *Zierdtviridae* (0.94 %), *Kyanoviridae* (0.54 %), *Guelinviridae* (0.35 %), *Autographiviridae* (0.25 %), *Schitoviridae* (0.23 %), *Demereciviridae* (0.19 %), *Orlajensenviridae* (0.13 %), *Vilmaviridae* (0.11 %), *Rountreeviridae* (0.10 %), *Zobellviridae* (0.01 %) and “no_family_available” (1.88 %). Also, for all predicted viral contigs, the type of interaction with a bacterial cell was determined – as a result, for data from the study by Frankel et al. [1], it turned out that virulent phages account for 32.95 %, and temperate phages 67.05 %; for data from the study by Matson et al. [2], similar values were obtained – 34.02 % for virulent phages and 65.98 % for temperate ones. The results indicate a predominance of temperate phages in GM, which is consistent with studies published to date. As can be seen from the above, there was no difference in percentage between virulent and temperate phages, neither between datasets nor between responder (R) and non-responders (NR) metagenome samples. Having analyzed the associations between phages and the bacterial component of GM, the following conclusions were made – in all studied samples, regardless of the data set, the dominant hosts were ten bacterial genera: *Clostridium*, *Faecalibacterium*, *Mycoplasma*, *Bifidobacterium*, *Ruminococcus*, *Bacteroides*, *Blautia*, *Streptomyces*, *Eubacterium* and family *Lachnospiraceae*. Regarding phage families, an increase in abundance was observed for four, namely: for the set of samples from the study of Frankel et al. [1], these were the families *Drexelviriidae* (for responders (R)) and

Salasmaviridae (for non-responders (NR)); for the set of samples from the study by Matson et al. [2], these were the families *Chaseviridae* and *Mesyanzhinovviridae* (both for non-responders (NR)). For all phage families, as abundance increased, a wider range of bacterial hosts was also observed. Based on preliminary results of the search for bacteriophages – candidates for predictive biomarkers of response to immunotherapy, two phage families, *Drexelviriidae* and *Chaseviridae*, were identified as the most promising candidates for predictive biomarkers. However, for final conclusions it is necessary to conduct a more in-depth analysis of a larger amount of data. As for the bacterial component of GM from patients with melanoma, the “Top 5” pathogenic species with the highest average representation among the studied samples included bacteria such as *Enterocloster bolteae*, *Clostridium scindens*, *Bacteroides fragilis*, *Raoultella ornithinolytica*, *Enterocloster clostridioformis*.

Conclusion: To date, there are no studies that would analyze the virome of cancer patients who responded and did not respond to immunotherapy. Our study on the possibility of using GM bacteriophages as prognostic biomarkers of the outcome of immunotherapy has not yet been completed and we plan to analyze a larger volume of data and in more depth, in particular, in the future, to use the differential ranking method for searching for reproducible taxonomic prognostic biomarkers using the Songbird tool, which has already proven itself in the analysis of complex metadata – this method allows you to estimate the relative differential abundance, regardless of the total microbial load. Thus, the current study has the chance to significantly expand the understanding of the involvement of GM, in particular its viral component, in immunotherapy, as well as to propose new approaches to improving existing methods of combating malignant melanoma.

Funding: The research was funded by the Russian Science Foundation (No. 23-75-10125, <https://rscf.ru/project/23-75-10125/>).

Список литературы/References

1. Frankel A.E., Coughlin L.A., Kim J., Froehlich T.W., Xie Y., Frenkel E.P., Koh A.Y. Metagenomic shotgun sequencing and unbiased metabolomic profiling identify specific human gut microbiota and metabolites associated with immune checkpoint therapy efficacy in melanoma patients. *Neoplasia*. 2017;19(10):848-855. doi 10.1016/j.neo.2017.08.004
2. Matson V., Fessler J., Bao R., Chongsuwat T., Zha Y., Alegre M.-L., Luke J.J., Gajewski T.F. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science*. 2018;359(6371):104-108. doi 10.1126/science.aao3290