

УДК 575.116.4:601.4]:579

DOI: <https://doi.org/10.51523/2708-6011.2021-18-3-20>

## Опыт использования современных геномных технологий для изучения микроорганизмов и их сообществ

Е. В. Воропаев, И. О. Стома, Д. В. Тапальский

Гомельский государственный медицинский университет, г. Гомель, Беларусь

### РЕЗЮМЕ

**Цель исследования.** Целью данной работы явился обзор основных результатов геномных исследований микроорганизмов и их сообществ, выполненных на базе научно-исследовательской лаборатории Гомельского государственного медицинского университета.

**Материалы и методы.** Геномный, транскриптомный и метагеномный анализ микроорганизмов желудка и респираторного тракта.

**Результаты.** Проведен анализ возможностей современных платформ высокопроизводительного секвенирования ДНК, описаны собственные результаты использования геномного, транскриптомного и метагеномного анализа для изучения микробиоты у пациентов с различной патологией желудка и респираторного тракта.

**Заключение.** В ходе анализа полученных данных выявлены особенности структуры микробных сообществ желудка пациентов, инфицированных *H. pylori* с различной патологией желудка: доленое участие *H. pylori* в составе метагенома образцов, с различным формами рака желудка не превысило 25 %, гастрита — 6 %, язвы желудка — 1 %. При этом, минимальное количество *H. pylori* во всех случаях могло достигать 0,1 %. Выявлена значительная степень вариабельности *CagA* и *CagY* локусов *H. pylori*. В бактериальном микробиоме пациента с диагнозом «коронавирусная инфекция» установлено доминирование бактерий рода *Streptococcus* (36 %), в вирусном — SARS-CoV-2 составляет 59 % от общего количества вирусов в данном материале. При анализе 13 штаммов *Klebsiella pneumoniae* с множественной и экстремальной устойчивостью к антибиотикам установлена принадлежность изученных штаммов к пяти MLST-типам, три из которых относятся к группам высокого эпидемического риска.

**Ключевые слова:** высокопроизводительное секвенирование ДНК, метагеномный анализ, микрогеном желудка, *Helicobacter pylori*, *Klebsiella pneumoniae*, респираторный микробиом, коронавирусная инфекция, SARS-CoV-2.

**Вклад авторов.** Воропаев Е.В., Стома И.О., Тапальский Д.В.: проведение поисково-аналитической работы и подготовка статьи, редактирование, утверждение рукописи для публикации.

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Источники финансирования.** Работа выполнена в рамках НИР «Изучение молекулярно-генетических механизмов реализации инфекционного канцерогенного потенциала на модели заболеваний органов пищеварения, ассоциированных с *Helicobacter pylori*» № ГР 20190388 от 29.03.2019 г. ГПНИ «Трансляционная медицина» подпрограмма 4.2 «Фундаментальные аспекты медицинской науки».

**Для цитирования.** Воропаев ЕВ, Стома ИО, Тапальский ДВ. Опыт использования современных молекулярно-генетических технологий для изучения структуры ДНК микроорганизмов, вирусов и их сообществ. *Проблемы здоровья и экологии.* 2021;18(3):159–167. DOI: <https://doi.org/10.51523/2708-6011.2021-18-3-20>

## Experience of using modern genomic technologies to study microorganisms and their communities

Evgenii V. Voropaev, Igor O. Stoma, Dmitry V. Tapalski

Gomel State Medical University, Gomel, Belarus

### ABSTRACT

**Objective.** The aim of this work was to review the main results of genomic studies of microorganisms and their communities performed on the base of the Research Laboratory of Gomel State Medical University.

**Materials and methods.** Genomic, transcriptomic and metagenomic analysis of the microorganisms of the stomach and respiratory tract.

**Results.** The capabilities of modern next-generation sequencing platforms have been analyzed, and the authors' own results of the use of genomic, transcriptomic and metagenomic analysis of the microbiota in patients with various gastric and respiratory pathologies have been described.

**Conclusion.** The analysis of the obtained data has revealed peculiarities of the structure of the microbial communities of the stomach of the patients infected with *H. pylori* with different gastric pathology: the proportion participation of *H. pylori* in the metagenome of the samples with different forms of gastric cancer did not exceed 25 %, of gastritis — 6 %, of peptic ulcer — 1 %. At the same time, the minimal amount of *H. pylori* in all the cases could reach 0.1 %. A significant degree of CagA and CagY loci variability of *H. pylori* was detected. *Streptococcus* genus bacteria dominated (36 %) in the bacterial microbiome of a patient diagnosed with the coronavirus disease, and in the viral microbiome, SARS-CoV-2 constituted 59 % of the total number of viruses in the material. The analysis of 13 strains of *Klebsiella pneumoniae* with multiple and extreme resistance to antibiotics has found that the studied strains belong to five MLST-types, three of which are classified as high epidemic risk groups.

**Keywords:** next-generation DNA sequencing, metagenomic analysis, gastric microgenome, *Helicobacter pylori*, *Klebsiella pneumoniae*, respiratory microbiome, coronavirus disease, SARS-CoV-2.

**Author contributions.** Voropaev E.V., Stoma I.O., Tapalski D.V.: search and analytical work and preparation of the article, editing and approval of the manuscript for publication.

**Conflict of interests.** The authors declare no conflicts of interest.

**Funding.** The work was performed within the framework of the research project “Study of the molecular and genetic mechanisms of the implementation of infectious carcinogenic potential on the model of digestive diseases associated with *Helicobacter pylori*” State Registration No. 20190388 dated 29.03.2019 State Research Program “Translational Medicine”, Subprogram 4.2 “Fundamental aspects of medical science”.

**For citation:** Voropaev EV, Stoma IO, Tapalsky DV. Experience of using modern genomic technologies to study microorganisms and their communities. *Health and Ecology Issues*. 2021;18(3):159–167. (In Russ.). DOI: <https://doi.org/10.51523/2708-6011.2021-18-3-20>

## Введение

Высокопроизводительное секвенирование или секвенирование нового поколения (NGS, от англ. Next generation sequence) представляет собой группу технологий установления первичной структуры нуклеиновых кислот, основанных на множественном параллельном анализе молекул ДНК или РНК [18]. Благодаря значительной производительности данного подхода и огромному массиву получаемых данных, NGS-технологии нашли широкое применение для решения широкого спектра вопросов, связанных с изучением различных биологических объектов, в том числе и человека. Так, в современной медицинской науке среди рассматриваемых проблем можно отметить полный анализ геномов этиологических агентов различных инфекционных заболеваний, установление значимых особенностей патогенеза широкого спектра патологии инфекционного и неинфекционного генеза, включая и онкологические заболевания, разработку высокоэффективных таргетных препаратов и методов лечения болезней, оценку наследственной предрасположенности формирования патофизиологических состояний [18]. Одним из наиболее результативных аспектов активного внедрения технологий высокопроизводительного секвенирования являются исследования

в области клинической микробиологии: от расшифровки структурно-функциональных особенностей геномов отдельных патогенных микроорганизмов до оценки микробных сообществ различных биотопов человека [1, 4, 11, 17]. Так, биоинформатический анализ данных, получаемых в результате проведения высокопроизводительного секвенирования как чистых культур, так и сообществ микроорганизмов, позволяет установить таксономическую принадлежность входящих в них видов, идентифицировать генетические детерминанты, ассоциированные с протеканием патогенетических процессов, а также оценить их функциональную активность, в том числе и антибиотикорезистентность, а в случае патосистем — идентифицировать различные виды молекулярно-физиологических aberrаций у организма-хозяина. Кроме того, применение NGS-технологий способствует развитию фундаментальных исследований в области изучения патогенеза инфекционных и неинфекционных заболеваний, включая разработку молекулярно-генетических основ персонализированной медицины [18].

Разработанные к настоящему времени технологии высокопроизводительного секвенирования отличаются между собой, как по используемым молекулярно-генетическим принципам их исполнения, так и по ряду ха-

рактических, отражающих их производительность, точность, информативность, стоимость за один анализ или в пересчете на единицу информации, что и обуславливает предпочтительность выбора того или иного подхода в зависимости от характера проводимых исследований и решаемых в их ходе задач [1].

### Цель исследования

Обзор основных направлений исследований и результатов, полученных с использованием различных подходов для геномного анализа микроорганизмов, а также их сообществ на базе научно-исследовательской лаборатории Гомельского государственного медицинского университета.

### Материалы и методы

Для проведения исследований использовали геномный, транскриптомный и метаболомный анализ по методикам, описанным нами ранее [27, 28]. Аннотация последовательностей проводилась с помощью программного обеспечения PGAP NCBI [26].

### Результаты исследований

#### Установление структурно-функциональной организации геномов бактерий и вирусов

В качестве объекта исследования был выбран изолят *Helicobacter pylori* HP42K, полученный из атрофического очага слизистой оболочки антрального отдела желудка пациентки (48 лет) с диагнозом «эритематозная гиперпластическая антральная гастропатия». Проведенное секвенирование генома изолята *H. pylori* HP42K показало, что его размер составляет 1645783 нуклеотидных оснований (н.о.) и представляет собой среднее значение данного параметра, характерного для различных штаммов вида — 1,49–1,91 млн н.о. [18].

По результатам аннотации генома изолята *H. pylori* HP42K было идентифицировано 1590 генов, 1545 из которых имели открытую рамку считывания и кодировали различные типы полипептидов. Среди оставшихся генов 36 детерминировали тРНК, шесть — рРНК, один — транспортно-матричную РНК. Функция еще двух генов не была установлена.

Расположение белок-кодирующих генов в геноме изолята *H. pylori* HP42K носило дисперсный характер и в целом соответствовало структурно функциональной организации, установленной для данного вида. В то же время локализация ряда локусов отличалась от таковой, установленной для других штам-

мов *H. pylori*, что демонстрирует наличие процессов транслокации в геноме бактерии. Полученные результаты подтверждают выводы, сделанные рядом авторов, о возможности использования признака внутригеномного расположения локусов в качестве дополнительного диагностического критерия для генетической паспортизации и классификации изолятов бактерий.

Характер пространственного распределения генов тРНК носил кластерно-дисперсный характер: 25 локусов были сгруппированы в семь различающихся по размеру кластеров, а оставшиеся 11 были расположены по отдельности. Применительно к генам рРНК была выявлена выраженная внутривидовая консервативность — все гены (5S, 16S и 23S) были дублированы, единичные копии 23S РНК и 5S РНК группировались в два кластера, а единичные копии 16S располагались в геноме по отдельности. При этом расположение между кластерами 23S-5S относительно единичных копий 16S не являлось симметричным.

Функциональная принадлежность белок-кодирующих генов была разнообразной: 59 были ассоциированы с процессами биосинтеза аминокислот, 83 — с биосинтезом кофакторов, простетических групп и молекул-переносчиков, 303 — участвовали в формировании клеточной оболочки, 123 — относились к клеточным процессам, 34 — детерминировали процессы промежуточного метаболизма, 135 — были задействованы в энергетических реакциях процессы, 89 — контролировали обмен жирных кислот и фосфолипидов, 34 — детерминировали регуляторные белки, 124 — ассоциированы с процессами репликации ДНК, 14 и 295 — вовлечены в механизмы транскрипции и трансляции соответственно. Оставшиеся 353 локуса детерминировали различные структурно-функциональные полипептиды, а также участвовали в формировании факторов патогенности и вирулентности. Среди последних у изолята *H. pylori* HP42K клиническую значимость представляли гены цитотоксин-ассоциированного антигена (CagA) и вакуолизирующего цитотоксина (VacA). Ген CagA был расположен в островке патогенности CagPAI, а VacA находился отдельно в геноме. Кроме перечисленных генов, в геноме изолята *H. pylori* HP42K было идентифицировано более 20 кодирующих регионов, относящихся к процессам анаболизма липополисахаридов, формирующих наружную мембрану клеточной стенки и вовлеченных



в механизмы адгезии патогена к цитоплазматической мембране клеток слизистой желудка. Также среди идентифицированных факторов вирулентности были аннотированы гены системы утилизации железа (*fecA*, *exbB*, *exbD*, *frnB* и др.), регуляции, синтеза и сборки элементов жгутиков (*flaA*, *flaB*, *flg*, *flh* и др.) и др.

Кроме полноценных кодирующих последовательностей, в геноме изолята *H. pylori* HP42K было выявлено 108 псевдогенов, происхождение большей части которых было связано с делециями/инсерциями, приводящих к смещению рамки считывания, последующему перераспределению структуры триплетов и образованию стоп-кодонов. Для 24 локусов были идентифицированы однонуклеотидные полиморфизмы (SNP), определяющие формирование стоп-кодонов, а 14 и 12 локусов представляли собой результаты фрагментации или комплексных аббераций, соответственно.

Внехромосомные элементы были представлены двумя плазмидами [21] размером 10 013 н.о и 2 658 н.о. В плазмиде большего размера было идентифицировано 13 генов, имеющих открытые рамки считывания, в меньшей — два, оба из которых кодировали полипептиды. Функционально плазмидные полипептиды представляли собой транспозазы, токсины, антитоксины, инициаторы репликации и др. При этом следует отметить, что некоторые элементы большей плазмиды были идентифицированы и в геномной ДНК, что указывает на возможность ее нахождения как в автономном, так и интегрированном состоянии.

Еще одним объектом исследований было изучение структурно-функциональной организации генома изолята 724 Gomel Belarus/2021 коронавируса SARS-CoV-2 [8, 10]. Проведенное секвенирование генома изолята вируса показало, что его размер составляет 29879 н.о. Не смотря на наличие полиморфизма в первичной структуре нуклеиновой кислоты (по отношению к изолятам, представленным в международных базах данных), определяемого как делетированными участками (9 и 15 п.о.) в ORF1ab и ORF7b соответственно, так и расположенными дисперсно по всему геному заменами, структурно-функциональная организация генома изолята 724 Gomel Belarus/2021 коронавируса SARS-CoV-2 была сходной с описанным референсным изолятом Wuhan-Hu-1. Геном SARS-CoV-2 представлял собой одноцепочечную (+)-цепь РНК, содержащую

12 функциональных открытых рамок считывания. Последовательное расположение генов (с установленной функцией) в рамке считывания являлось следующим: репликаза, протеазы (1a – 1b) и основные белки S, E, M и N. Наибольший размер имела первая открытая рамка считывания ORF1ab (21281 н.о.), кодирующая полипептид, состоящий из множества структурных и функциональных белков. Еще одной особенностью открытой рамки считывания ORF1ab, являясь тот факт, что она функционально (применительно к процессам трансляции) разделена на два фрагмента, непрерывная реализация генетической информации с которых возможна лишь в случае активизации механизма сдвига молекулы РНК на один нуклеотид в ходе трансляции на рибосоме. В ином случае биосинтез целостного полипептида не представляется возможным, что связано с наличием стоп-кодона во втором функциональном фрагменте ORF1ab. В данном случае будет образовываться только укороченный полипептид размером 4406 аминокислотных остатков (а.о.). Процессинг полноценного полипептида ab, осуществляемый кодируемыми вирусами протеиназами, позволяет получить 16 структурных и функциональных белков.

#### **Выявление и изучение клинически значимых генов**

Для описания особенностей клинически значимых генов патогенных микроорганизмов в качестве объектов исследования были выбраны ранее описанный изолят *H. pylori* HP42K [18] и еще один секвенированный нами изолят *H. pylori* HP45K [20]. Предметом исследований явились структурные особенности Cag-ассоциированных генов, расположенных в островке патогенности CagPAI. Размер CagPAI для обоих изолятов составил порядка 40 тыс н.о. и включал в своем составе тридцать четыре гена, в том числе *cagA* ген и ряд ассоциированных с ним ортологичных генов. Согласно литературным данным, CagPAI кодирует систему секреции Cag типа IV (CagT4SS), которая доставляет онкопротеин CagA и другие эффекторные молекулы в эпителиальные клетки желудка человека.

В перечне аннотированных генов Cag-PAI были идентифицированы семь локусов, детерминирующих систему секреции IV типа, определяющую биосинтез и транспортировку цитотоксин-ассоциированного антигена: Cag1, Cag3, CagX, CagY, CagM, CagT. При этом, гены Cag3, CagT и CagM, CagX, CagY детерминируют полипептиды, формирующие

внешнюю и внутреннюю структуру транспортного канала системы секреции IV типа, ген *Cag1* кодирует мембранный белок, определяющий цитотоксическую активность *H. pylori* по отношению к клеткам человека, а *CagA* представляет собой ген патогенности, отвечающий за выработку цитотоксин-ассоциированного антигена.

Проведенный анализ локуса *Cag1* показал, что генетические различия (3,74 %) между изолятами *H. pylori* HP42K и *H. pylori* HP45K обусловлены исключительно нуклеотидными замещениями, без изменения размера гена — он составил в обоих случаях 348 н.о. Основной характер замещений был связан с транзициями А↔G и С↔T, на долю которых приходилось 84,6 % от всех изменений. Изучение транскрибируемых последовательностей (115 а.о.) выявило только четыре аминокислотные замены, что свидетельствовало о синонимичном характере большинства (69,2 %) нуклеотидных вариаций между изолятами по локусу *Cag1*. При этом, в трех из четырех случаев, несинонимичные замены были связаны с миссенс-мутациями, приводящими к замене аминокислотного остатка на остаток с другими физико-химическими свойствами. В то же время, наблюдающиеся изменения в полипептидной структуре не приводили к значительным изменениям третичной структуры белковой молекулы, что было установлено на основании построения пространственных моделей. Стоп-кодены в гене *Cag1* были сходны у изолятов и представлены триплетом TGA.

Уровень различий между нуклеотидными последовательностями изолятов 42K и 45K по гену *CagA* был выше, чем для *Cag1*, и составил 7,26 %. Выявляемый полиморфизм был связан как с нуклеотидными заменами, так и делециями/инсерциями. Размер *CagA*-локуса у изолята 45K составил 3552 н.о. и был больше на 9 н.о. по сравнению с изолятом 42K. Следует отметить, что инсерции/делеции были сгруппированы в начале локуса, представляя собой участки размером в один, пять и три нуклеотидов. Соотношение транзиций к трансверсиям было сходно с *Cag1* и равнялось 1:2,96. Последовательность стоп-кодонов у изолятов была сходной и представлена триплетом TAA. Величина отличий транскрибируемых последовательностей составила 12,09 %, указывая что большая часть (>55 %) замен являлась несинонимичной. Несмотря на наличие инсерций/делеций нуклеотидов с размером, некрatным

трех, их суммарная протяженность равная 9 н.о. и кластерный характер расположения обусловил отсутствие сдвига рамки считывания в целом, и соответственно определил сходство аминокислотных последовательностей в значительной части полипептида. Вследствие сохранения консервативности структуры белка цитотоксин-ассоциированного антигена, у обоих изолятов в его структурной организации были идентифицированы функциональные домены SMC\_prok\_B (суперсемейство TIGR02168) и *CagA* (суперсемейство rfam03507), определяющие токсикологические свойства полипептида.

Среди генов, детерминирующих структуру транспортного канала системы секреции, наименьшие различия между изолятами были выявлены для *CagT* — 1,54 %, а наибольшие — для *CagY* — 9,36 % (без учета делеций/инсерций) и 27,0 % (с учетом делеций/инсерций). Уровень несоответствия между изолятами для остальных локусов (*Cag3*, *CagX*, *CagM*) не превысил 3,5 %. Сравнительный анализ аминокислотных последовательностей также показал высокий уровень консерватизма для детерминирующих их локусов *CagT*, *CagX*, *CagM*, *Cag3* — степень дифференциации составила 0,06, 1,53, 1,60 и 2,49 % соответственно. Также следует отметить, что выявленный полиморфизм данных генов был связан только с заменами азотистых оснований, не приводящими к изменению длины первичной последовательности нуклеотидной и кодируемой полипептидной цепи. В то же время различия в гене *CagY* между изолятами были также связаны и с делециями/инсерциями (5 фрагментов, длиной 552, 384, 114, 2 и 112 н.о.). Размер *CagY*-локуса у изолята 42K составил 5766 н.о. и превышал свой ортолог у изолята 45K (4602 н.о.). Данные различия в гене *CagY* также определили и различия в структуре белковой молекулы — 18,27 % (без учета делеций/инсерций) и 25,80 % (с учетом делеций/инсерций), и ее функциональных особенностей — число копий мотивов функционального домена *CagY* 1 Repeat у изолята 42K составило 3, у изолята 45K — 1, DC-EC Repeat — 3 и 2 соответственно. Также у изолята 45K отсутствовали функциональные домены MSCRAMM (адгезивные свойства) и SbcCD (экзонуклеазная активность). В то же время у обоих изолятов в *CagY* были идентифицированы функциональные домены PTZ00121 и VIRB10.

Проведенный сравнительный анализ выявил высокую степень консервативности

локусов CagT, Cag3, CagX, CagM, что указывает на их ведущую роль в формировании патогенных свойств *H. pylori* и нахождение под действием стабилизирующего отбора. В то же время локусы CagA и CagY характеризуются значительной степенью вариабельности, что, по всей видимости, определяет вирулентность штаммов *H. pylori*, и делает их информативными молекулярными маркерами для оценки патогенетического потенциала бактерий.

Кроме изолятов *H. pylori*, объектами исследования по данному направлению выступали геномы 13 клинических изолятов *Klebsiella pneumoniae* с множественной и экстремальной устойчивостью к антибиотикам, выделенным от госпитализированных пациентов. Предметом изучения являлись особенности первичной структуры генов, ассоциированных с патогенетическими характеристиками штаммов, а также локусов, детерминирующих резистентность к некоторым антибиотикам (и в частности, к колистину).

Исходя из результатов типирования генов «домашнего хозяйства» *gapA*, *infB*, *mdh*, *pgi*, *phoE*, *rpoB*, *tonB*, была установлена принадлежность изученных штаммов к пяти MLST-типам: ST23 (два штамма), ST512 (один штамм), ST11 (два штамма), ST147 (два штамма), ST395 (шесть штаммов), три последних из которых относились к группам высокого эпидемического риска.

Анализ генов вирулентности (*rmpA*, *kfu*, *allS*, *iucA*, *tagA* и др.) показал, что ряд из них может иметь как «хромосомную» локализацию, так располагаться и в плазмидах (например, в pLVK).

По отношению к антибиотикам, среди изученных штаммов *K. pneumoniae* идентифицированы гены, определяющие устойчивость к препаратам различных групп, включая β-лактамы, аминогликозиды, фторхинолоны, фосфомицин, хлорамфеникол, полимиксины. При этом, устойчивость для ряда антибиотиков могла иметь полигенный характер. Так, например, среди детерминант карбапенемаз, идентифицирован спектр генов — *blaCTX-M*, *blaKPC-2*, *blaKPC-3*, *blaLEN7*, *blaOXA-9*, *blaOXA-48*, *blaTEM*, *blaSHV-182* и др. При этом, в каждом штамме могли выявляться одновременно от 4 до 6 локусов. Кроме того, ряд наследственных детерминант, также могли иметь как «хромосомную», так и плазмидную локализацию.

Спектр диагностированных плазмид и изученных штаммов был широким и пред-

ставлен следующими типами: Col(IRGK), Col(pHAD28), Col440I, IncFIB(pQil), IncFII(K), IncFII(pKP91), IncN, IncR, ColRNAI, ColpVC, IncFIB(K), IncFIB(pNDM-Mar), IncHI1B(pNDM-MAR), IncX3, IncFII(K), IncL и др. При этом, у всех штаммов идентифицировано наличие более двух плазмид одновременно.

Анализ генов устойчивости к колистину показал, что наиболее частым механизмом формирования резистентности являлась инсерционная инактивация или делеция генов белков-мишеней, вызванная действием мигрирующих генетических элементов (например, у изогенных штаммов применительно к гену *mgrB*). Среди циркулирующих штаммов *K. pneumoniae* также выявлено наличие генетического полиморфизма, вызывающего смещение рамки считывания, или формирование аминокислотных замен (D150Y, T157P и G207S в гене *mgrB*), определяющих пониженную восприимчивость к полимиксинам.

#### Анализ метагеномов

Проведенный анализ микробиома желудка пациентки со смешанной формой хронического гастрита и полипом антрального отдела желудка показал, что долевое участие *H. pylori* в метагеноме желудка было менее 0,1%. В то же время, среди идентифицированных бактерий, доминирующим являлся род *Halomonas* (19,4 %). Долевое участие ниже 5 % было отмечено для *Prevotella* (3,6 %), *Veillonella* (3,3 %), *Streptococcus* (1,3 %), *Bacillus* (0,7 %), *Neisseria* (1,1 %), *Leptotrichia* (0,8 %), *Fusobacterium* (0,6 %) и др.

На основании полученных данных был проведен дополнительный анализ образцов биоптатов слизистой желудка с различными заболеваниями желудка с использованием фрагментного анализа ДНК. Было установлено, что долевое участие *H. pylori* в составе метагенома образцов, с различным формами рака желудка не превысило 25 %, гастрита — 6 %, язвы желудка — 1 %.

Также интересными оказались результаты сравнительной оценки долевого участия *H. pylori* в микробиомах биоптатов, взятых у одних и тех же пациентов из различных (с наличием и отсутствием признаков патологии) участков желудка. Установлено, что в случае рака желудка наблюдалось значительное (до 30 раз) снижение содержания *H. pylori* в микрофлоре патологических образцов. В то же время, в случае язвенной патологии существенного колебания удельного веса *H. pylori* в микробиоме патологических и нормальных образцов не наблюдалось.



Диапазон вариаций частоты встречаемости при диагнозах «гастрит» и «эрозия желудка» был значительным — 3–40 раз, однако не имел выраженной направленности по отношению к нормальным и патологическим образцам.

В ходе анализа назальных образцов у пациента с диагнозом «коронавирусная инфекция», кроме представленного ранее генома SARS-CoV-2, также был изучен микробиом на основании анализа не баркодинговых маркеров, а экспрессирующихся белок-кодирующих последовательностей (транскриптов). На основании результатов анализа метатранскриптома было установлено, что в доминирующем количестве был представлен генетический материал бактерий рода *Streptococcus* — 36 %. В меньшем количестве были диагностированы представители рода *Trichinella* (9 %), *Acinetobacter* (7 %), *Veillonella* (5 %), *Prevotella* (5 %) и др. На долю SARS-CoV-2 приходилось менее 2 % от всей полученной генетической информации. В разрезе вирусной инфекции 59 % составлял SARS-

CoV-2, 9 % — *Rhinolophus affinis coronavirus*, 6 % — *Bat SARS coronavirus HKU3*, 3 % — *Mitovirus* sp. и др.

## Заключение

Проведенные молекулярно-генетические исследования показали, что значительная часть наследственной изменчивости микроорганизмов связана не с внутригенным полиморфизмом, а с особенностями структурной организации геномов, что может оказывать существенное влияние на вирулентные свойства патогенов, и должно учитываться в качестве дополнительного критерия при проведении диагностики инфекционных заболеваний. Описание свойств микробиомов и функциональной оценки метаболомов сообществ микроорганизмов должно осуществляться не только на основе типирования таксон-специфических маркеров, а с использованием дополнительных критериев — генов, детерминирующих факторы патогенности и вирулентности, являющихся молекулярным признаком патосистем.

## Список литературы

1. Бородинов АГ, Манойлов ВВ, Заруцкий ИВ, Петров АИ, Курочкин ВЕ. Поколения методов секвенирования ДНК (обзор). *Научное приборостроение*. 2020;30(4):3-20.
2. Воропаев ЕВ, Осипкина ОВ, Баранов ОЮ, Рузанов ДЮ, Лызикова ЮА, Платошкин ЭН, Зятьков АА, Шафарост АС, Зайцева ВИ. Метод определения видового состава микробных сообществ с использованием Т-ПДРФ анализа рДНК-маркеров. Инструкция по применению. Рег. №: 120-1118. Дата утверждения: 30.11.2018.
3. Стома ИО. Микробиом в медицине: руководство для врачей. М.: ГЭОТАР-Медиа, 2020. 320 с.
4. Стома ИО, Ющук НД. Микробиом человека на стыке инфектологии и других разделов медицины: современное состояние проблемы и переоценка взглядов на патогенез заболеваний. *Инфекционные болезни: новости, мнения, обучение*. 2019;8(3):78-84. DOI: <https://doi.org/10.24411/2305-3496-2019-13012>
5. Воропаев ЕВ, Баранов ОЮ, Валентович ЛН, Осипкина ОВ, Зятьков АА, Бонда НА, Воропаева АВ, Платошкин ЭН, Мицура ВМ, Шафарост АС. Структурно-функциональная организация генома клинически значимого штамма *Helicobacter pylori* HP42k, изолированного от пациентки учреждения здравоохранения Республики Беларусь. *Молекулярная медицина*. 2020;18(2):39-43.
6. Allali I, Arnold JW, Roach J, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol*. 2017;17(1):194. DOI: <https://doi.org/10.1186/s12866-017-1101-8>
7. Ardui S, Ameur A, Vermeesch JR, Hestand MS. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res*. 2018;46(5):2159-2168. DOI: <https://doi.org/10.1093/nar/gky066>
8. Stoma IO, Baranov OY, Voropaev EV, Osipkina OV, Ziatskov AA, Shaforost AS, Padutov VE, Pantelev SV, Kiryanov PS, Mozharovskaya LV. Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/BLR/Sars-Cov-2 isolate 724 Gomel Belarus/2021, complete genome. [Electronic resource]. NCBI Nucleotide. GenBank NCBI: MW674675.1 [date of access 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/nucleotide/MW674675.1>
9. Gomel State Medical University. Stomach metagenome[Electronic resource].NCBI.SRAPRJNA700813 1. [date of access 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA700813>
10. GISAID: hCoV-19/Belarus/Gomel/2021 Accession ID: EPI\_ISL\_1222766. [Electronic resource]. [date of access 2021 Aug 12]. Available from: <https://www.gisaid.org>
11. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front Microbiol*. 2017 Nov;15(8):2224. DOI: <https://doi.org/10.3389/fmicb.2017.02224>
12. Goodwin S, McPherson J, McCombie W. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*. 2016 May 17;17(6):333-351. DOI: <https://doi.org/10.1038/nrg.2016.49>
13. Hengyun Lu, Francesca Giordano, Zemin Ning, Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*, 2016;14(5):265-279. DOI: <https://doi.org/10.1016/j.gpb.2016.05.004>
14. TRAPID. Online tool for the fast, reliable and user-friendly analysis of de novo transcriptomes. [Electronic resource]. [date of access 2021 Aug 12]. Available from: [http://bioinformatics.psb.ugent.be/trapid\\_02/](http://bioinformatics.psb.ugent.be/trapid_02/)

15. Воропаев ЕВ, Осипкина ОВ. Алгоритм функциональной аннотации транскрибируемых последовательностей в геноме клинически значимых для Беларуси штаммов *Helicobacter pylori* [Электронный ресурс]. Достижения медицинской науки Беларуси. [дата обращения 2021 Aug 12]. Режим доступа: [http://med.by/dmn/book.php?book=19-4\\_2](http://med.by/dmn/book.php?book=19-4_2)
16. One Codex. A fast, easy-to-use platform for microbiome sequencing and analysis. [Electronic resource]. [date of access 2021 August 12]. Available from: <https://www.onecodex.com>
17. Li J, Du P, Ye AY, Zhang Y, Song C, Zeng H, Chen C. GPA: A Microbial Genetic Polymorphisms Assignments Tool in Metagenomic Analysis by Bayesian Estimation. *Genomics Proteomics Bioinformatics*. 2019 Feb;17(1):106-117. DOI: <https://doi.org/10.1016/j.gpb.2018.12.005>
18. Ребриков ДВ, Коростин ДО, Шубина ЕС, Ильинский ИВ. NGS: высокопроизводительное секвенирование: монография. 3-е изд. Москва: Лаборатория знаний, 2020. 235 с.
19. Voropaev EV, Osipkina OV, Baranov OU, Ziatskov AA, Voropaev AV, Platoshkin EN, Bonda NA, Valentovich LN. *Helicobacter pylori* complete genome in Belarus. [Electronic resource]. NCBI Nucleotide. GenBank NCBI: NZ\_CP034314.1 [date of access 2021 August 12]. Available from: <https://www.ncbi.nlm.nih.gov/nuccore/1535583272>
20. Voropaev EV. New pathogenic *Helicobacter pylori* strain in Belarus. [Electronic resource]. NCBI Nucleotide. GenBank NCBI: NZ\_SZUB00000000.1. [date of access 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/nuccore/1635704349>
22. Shintani M, Sanchez ZK, Kimbara K. Genomics of

- microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. (англ.). *Frontiers In Microbiology*. 2015;6:242-242. DOI: <https://doi.org/10.3389/fmicb.2015.00242>
23. Stoma I, Littmann ER, Peled JU, Giralt S, van den Brink MRM, Pamer EG, Taur Y. Compositional flux within the intestinal microbiota and risk for bloodstream infection with gram-negative bacteria. *Clin Infect Dis*. 2020 Jan 24:ciaa068. DOI: <https://doi.org/10.1093/cid/ciaa068>
24. Stoma I, Moduleva E; Gubanov T, Iskrou I, Karpov I, Krivenko S, Shturich I, Kalachyk O, Uss A., Rummo O. Gut microbiome dynamics over a course of kidney transplantation: preliminary results of a pilot study. *Transplantation*: 2020; 104(Is S3):S191 DOI: <https://doi.org/10.1097/01.tp.0000699336.48313.1e>
25. Sun J, Liao XP, D'Souza AW, et al. Environmental remodeling of human gut microbiota and antibiotic resistance in livestock farms. *Nat Commun*. 2020 Mar 18;11(1):1427. DOI: <https://doi.org/10.1038/s41467-020-15222-y>
26. Toyokawa M, Taniguchi M, Uesaka K, Nishimura K. Complete Genome Sequence of Multidrug-Resistant Strain *Nocardia wallacei* FMUON74, Isolated from a Sputum Culture. *Microbiol Resour Announc*. 2020 Nov 19;9(47):e01022-20. DOI: <https://doi.org/10.1128/MRA.01022-20>
27. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res*. 2016; 44(14): 6614-6624. DOI: <https://doi.org/10.1093/nar/gkw569>

## References

1. Borodinov AG, Manoilo BB, Zarutsky IV, Petrov AI, Kurochkin VE. Generations of DNA sequencing methods (review). *Scientific instrumentation*. 2020;30(4):3-20. (in Rus.).
2. Voropaev EV, Osipkina OV, Baranov OY, Ruzanov DU, Lyzikova YuA, Platoshkin EN, Ziatskov AA, Shafarost AS, Zaitseva VI. A method for determining the species composition of microbial communities using T-PDRF analysis of rDNA markers. Instruction for use. In: 120-1118. Date of approval: 30.11.2018. (in Rus.).
3. Stoma IO. Microbiome in medicine: a guide for doctors. Moscow: GEOTAR-Media, 2020. 320 с. (in Rus.).
4. Stoma IO, Yushchuk ND. Human microbiome at the interface of infectiology and other sections of medicine: the current state of the problem and reassessment of views on disease pathogenesis. *Infectious diseases: news, opinions, and teaching*. 2019;8(3):78-84. (in Rus.). DOI: <https://doi.org/10.24411/2305-3496-2019-13012>
5. Voropaev EV, Baranov OY, Valentovich LN, Osipkina OV, Ziatskov AA, Bond NA, Voropaeva AV, Platoshkin EN, Mitsura VM, Shafarost AS. Structural and functional organization of the genome of a clinically relevant *Helicobacter pylori* HP42k strain isolated from a patient of a Belarusian healthcare institution. *Molecular Medicine*. 2020;18(2):39-43. (in Rus.).
6. Allali I, Arnold JW, Roach J, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol*. 2017;17(1):194. DOI: <https://doi.org/10.1186/s12866-017-1101-8>
7. Ardui S, Ameer A, Vermeesch JR, Hestand MS. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res*. 2018;46(5):2159-2168. DOI: <https://doi.org/10.1093/nar/gky066>
8. Stoma IO, Baranov OY, Voropaev EV, Osipkina OV, Ziatskov AA, Shafarost AS, Padutov VE, Panteleev SV, Kiryanov PS, Mozharovskaya LV. Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/BLR/Sars-Cov-2 isolate 724 Gomel Belarus/2021, complete genome. [Electronic resource]. NCBI Nucleotide. GenBank NCBI: MW674675.1 [cited 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/nuccore/MW674675.1>
9. Gomel State Medical University. Stomach metagenome [Electronic resource]. NCBI. SRA PRJNA7008131. [date of access 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA7008131>; GISAID: hCoV-19/Belarus/Gomel/2021 Accession ID: EPI\_ISL\_1222766. [Electronic resource]. [date of access 2021 Aug 12]. Available from: <https://www.gisaid.org>
11. GISAID: hCoV-19/Belarus/Gomel/2021 Accession ID: EPI\_ISL\_1222766. [Electronic resource]. [date of access 2021 Aug 12]. Available from: <https://www.gisaid.org>
12. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front Microbiol*. 2017 Nov;15(8):2224. DOI: <https://doi.org/10.3389/fmicb.2017.02224>
13. Goodwin S, McPherson J, McCombie W. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*. 2016 May 17;17(6):333-351. DOI: <https://doi.org/10.1038/nrg.2016.49>
14. Hengyun Lu, Francesca Giordano, Zemin Ning, Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*. 2016;14(5):265-279. DOI: <https://doi.org/10.1016/j.gpb.2016.05.004>



15. TRAPID. Online tool for the fast, reliable and user-friendly analysis of de novo transcriptomes. [Electronic resource]. [date of access 2021 Aug 12]. Available from: [http://bioinformatics.psb.ugent.be/trapid\\_02/](http://bioinformatics.psb.ugent.be/trapid_02/)
16. Voropaev EV, Osipkina OV. Algorithm of functional annotation of translated sequences in the genome of *Helicobacter pylori* strains clinically significant for Belarus [Electronic resource]. [date of access 2021 August 12]. Available from: [http://med.by/dmn/book.php?book=19-4\\_2](http://med.by/dmn/book.php?book=19-4_2) (in Rus.).
17. One Codex. A fast, easy-to-use platform for microbiome sequencing and analysis. [Electronic resource]. [date of access 2021 August 12]. Available from: <https://www.onecodex.com>
18. Li J, Du P, Ye AY, Zhang Y, Song C, Zeng H, Chen C. GPA: A Microbial Genetic Polymorphisms Assignments Tool in Metagenomic Analysis by Bayesian Estimation. *Genomics Proteomics Bioinformatics*. 2019 Feb;17(1):106-117. DOI: <https://doi.org/10.1016/j.gpb.2018.12.005>
19. Rebrikov DV, Korostin DO, Shubina ES, Ilyinsky BB. *NGS: high-throughput sequencing: a monograph*. Moscow: Laboratory of Knowledge, 2020. 235 c. (in Rus.).
20. Voropaev EV, Osipkina OV, Baranov OU, Ziatskov AA, Voropaev AV, Platoshkin EN, Bonda NA, Valentovich LN. *Helicobacter pylori* complete genome in Belarus. [Electronic resource]. NCBI Nucleotide. GenBank NCBI: NZ\_CP034314.1 [date of access 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/nucleotide/1535583272>
21. Voropaev EV. New pathogenic *Helicobacter pylori* strain in Belarus. [Electronic resource]. NCBI Nucleotide. GenBank NCBI: NZ\_SZUB000000000.1. [date of access 2021 Aug 12]. Available from: <https://www.ncbi.nlm.nih.gov/nucleotide/1635704349>
22. Shintani M, Sanchez ZK, Kimbara K. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. (англ.). *Frontiers In Microbiology*. 2015;6:242-242. DOI: <https://doi.org/10.3389/fmicb.2015.00242>
23. Stoma I, Littmann ER, Peled JU, Giralt S, van den Brink MRM, Pamer EG, Taur Y. Compositional flux within the intestinal microbiota and risk for bloodstream infection with gram-negative bacteria. *Clin Infect Dis*. 2020 Jan 24:ciaa068. DOI: <https://doi.org/10.1093/cid/ciaa068>
24. Stoma I, Moduleva E; Gubanova T, Iskrou I, Karpov I, Krivenko S, Shturich I, Kalachyk O, Uss A., Rummo O. Gut microbiome dynamics over a course of kidney transplantation: preliminary results of a pilot study. *Transplantation*: 2020; 104(Is S3):S191 DOI: <https://doi.org/10.1097/01.tp.0000699336.48313.1e>
25. Sun J, Liao XP, D'Souza AW, et al. Environmental remodeling of human gut microbiota and antibiotic resistance in livestock farms. *Nat Commun*. 2020 Mar 18;11(1):1427. DOI: <https://doi.org/10.1038/s41467-020-15222-y>
26. Toyokawa M, Taniguchi M, Uesaka K, Nishimura K. Complete Genome Sequence of Multidrug-Resistant Strain *Nocardia wallacei* FMUON74, Isolated from a Sputum Culture. *Microbiol Resour Announc*. 2020 Nov 19;9(47):e01022-20. DOI: <https://doi.org/10.1128/MRA.01022-20>
27. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res*. 2016;44(14):6614-6624. DOI: <https://doi.org/10.1093/nar/gkw569>

## Информация об авторах / Information about the authors

**Воропаев Евгений Викторович**, к.м.н., доцент, проректор по научной работе УО «Гомельский государственный медицинский университет»;

ORCID: <https://orcid.org/0000-0002-9435-6109>

e-mail: [voropaev.evgenii@gmail.com](mailto:voropaev.evgenii@gmail.com)

**Стома Игорь Олегович**, д.м.н., доцент, ректор УО «Гомельский государственный медицинский университет»;

ORCID: <https://orcid.org/0000-0003-0483-7329>

e-mail: [rektor@gsmu.by](mailto:rektor@gsmu.by)

**Тапальский Дмитрий Викторович**, д.м.н., доцент, заведующий кафедрой микробиологии, вирусологии и иммунологии УО «Гомельский государственный медицинский университет»;

ORCID: <https://orcid.org/0000-0002-9484-7848>

e-mail: [tapalskiy@yandex.by](mailto:tapalskiy@yandex.by)

**Evgenii V. Voropaev**, PhD (Med), Associate Professor, Vice-Rector in charge of scientific work of Gomel State Medical University;

ORCID: <https://orcid.org/0000-0002-9435-6109>

e-mail: [voropaev.evgenii@gmail.com](mailto:voropaev.evgenii@gmail.com)

**Igor O. Stoma**, DMedSc, Associate Professor, Rector of Gomel State Medical University;

ORCID: <https://orcid.org/0000-0003-0483-7329>

e-mail: [rektor@gsmu.by](mailto:rektor@gsmu.by)

**Dzmitry V. Tapalski**, DMedSc, Associate Professor, Head of the Department of Microbiology, Virology and Immunology, Gomel State Medical University;

ORCID: <https://orcid.org/0000-0002-9484-7848>

e-mail: [tapalskiy@yandex.by](mailto:tapalskiy@yandex.by)

## Автор, ответственный за переписку / Corresponding author

**Воропаев Евгений Викторович**

e-mail: [voropaev.evgenii@gmail.com](mailto:voropaev.evgenii@gmail.com)

**Evgenii V. Voropaev**

e-mail: [voropaev.evgenii@gmail.com](mailto:voropaev.evgenii@gmail.com)

Received / Поступила в редакцию 14.07.2021

Revised / Поступила после рецензирования 23.08.2021

Accepted / Принята к публикации 20.09.2021