

МУТАЦИИ В ГЕНОМАХ SARS-COV-2 БИОЛОГИЧЕСКИХ ОБРАЗЦОВ, ПОЛУЧЕННЫХ В КОНЦЕ МАРТА-НАЧАЛЕ АПРЕЛЯ, ОТ ПАЦИЕНТОВ ГОРОДА МОСКВЫ

В.В. Каптелова*†, А.С. Сперанская†, А.Е. Самойлов, А.В. Валдохина, В.П. Буланенко, Е.В. Корнеенко, О.Ю. Шипулина, В.Г. Акимкин.

ФБУН «ЦНИИ эпидемиологии» Роспотребнадзора, Москва, Россия.

† Равный вклад

Ключевые слова: SARS-CoV-2, мутации, D614G, P323L, NGS

Неоднократно высказывались предположения, что мутация D614G в S-белке SARS-CoV-2 может влиять на способность передачи вируса. В недавней работе [1] было отмечено, что на скорость передачи заболевания и на скорость репликации вируса оказывает влияние не одиночная мутация D614G, а ее комбинация с мутацией P323L в белке РНК-зависимой РНК-полимеразы. Мы определили последовательности полных геномов для 28 вирусов SARS-CoV-2, полученных из образцов, представляющих собой клинический материал мазок/ отделяемое ротоглотки от пациентов разных возрастов. Проанализированные изоляты принадлежат к линиям B.1 (GH) и B.1.1 (GR). Во всех геномах были обнаружены сочетания мутаций P323L в области РНК-зависимой РНК-полимеразы и D614G в гене S-белка. Обнаруженные отличия могут объяснены выборкой: для полногеномного секвенирования отбирались только образцы с высокой вирусной нагрузкой, которая, могла быть обусловлена именно высокой скоростью размножения вируса в организме хозяина, что, в свою очередь, может иметь связь с наличием в геноме вируса мутаций P323L/ D614G.

MUTATIONS IN THE GENOMES OF SARS-COV-2 FROM CLINICAL SAMPLES OBTAINED IN LATE MARCH-EARLY APRIL FROM PATIENTS IN MOSCOW

**V.V. Kaptelova*†, A.S. Speranskaya†, A.E. Samoilov, A.V. Valdokhina, V.P. Bulanenko,
E.V. Korneenko, O.Y. Shipulina, V.G. Akimkin.**

Central Research Institute of Epidemiology of the Federal Service on Customers' Rights Protection and Human Well-being Surveillance, Moscow, Russia.

† Contributed equally

Key words: SARS-CoV-2, mutations, D614G, P323L, NGS

Many papers suggested that D614G mutation in the viral spike (S) protein SARS-CoV-2 can influence the ability of virus transmission. In recent work [1], it was shown D614G influences the rate of disease transmission only in combination with the P323L mutation in the viral polymerase. We have sequenced 28 full genomes of SARS-CoV-2, obtained from clinical material from patients of different ages. The analyzed isolates belong to clades B1 (GH) and B1.1. (GR). Combinations of mutations P323L and D614G were found in all genomes. These differences can be explained by sampling: the samples for the sequencing of the whole genome were selected with high viral load, it can be related to the rate of viral replication in intra-host. That, in turn, can be dependent on the presence of P323L/D614G mutations in the virus genome.

Введение. Новое респираторное заболевание COVID-19, вызванное вирусом SARS-CoV-2, охватило весь мир и получило статус пандемии. Заражение нулевого пациента произошло в Ухане и далее инфекция распространилась по всему миру. На сегодняшний день ученые из разных регионов мира пытаются рассмотреть и проанализировать геномные изменения SARS-CoV-2 с помощью молекулярно-генетических методов и выяснить, какие именно генетические особенности вируса могут быть связаны с различной тяжестью течения заболевания. Трансмембранный гликопротеин S-белок SARS-CoV-2 является белком вирусной оболочки, он способствует прикреплению вируса к рецепторам на поверхности клетки-хозяина и отвечает за слияние с ней [1]. Неоднократно высказывались предположения, что мутация D614G в S-белке может улучшать адгезию и, как следствие, увеличивать вероятность передачи вируса (от хозяина к хозяину) [2]. По данным всемирной базы GISAID, мутация D614G в S-белке встречается в 76,53% последовательностей (по данным на 20 августа 2020); первый полногеномный сиквенс, в котором она была зарегистрирована - hCoV-19/England/NORW-E82AE/2020 [3]. В недавней работе [1] было высказано предположение, что на скорость передачи заболевания и на репликацию вируса оказывает влияние не одиночная мутация D614G, а ее комбинация с мутацией P323L в белке РНК-зависимой РНК-полимеразы [1]. Те же авторы предполагают коэволюцию участков генома SARS-CoV-2, что приводит к независимому неоднократному формированию обеих аминокислотных замен P323L и D614G в разных эволюционных линиях.

Цель работы. Определить полные последовательности геномов SARS-CoV-2 из образцов, полученных из КДЛ, от пациентов города Москвы, с подтвержденным наличием вируса в конце марта - начале апреля, описать обнаруженные мутации и определить, как часто встречается комбинация мутаций P323L/D614G.

Материалы и методы. Образцы мазков от пациентов с симптомами ОРВИ или подозрениями на SARS-CoV-2 были обработаны в лаборатории молекулярной диагностики ФБУН ЦНИИ Эпидемиологии Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека (лаборатория имеет разрешение для работы с ПБА) с использованием набора реагентов AmpliSens® CoV-Bat-FL для обнаружения SARS-CoV-Family, включающий новый SARS-CoV-2 в соответствии с инструкцией производителя. Для амплификации геномов мы использовали остатки образцов РНК, хранившихся в клинической лаборатории. Были отобраны образцы, в которых была выявлена РНК SARS-CoV-2 со значением Ct ≤ 20 . Нами была разработана праймерная панель для амплификации полного генома, состоящая из 17 пар праймеров [4]. Реакцию обратной транскрипции проводили набором Reverta-L kit (AmpliSens, Russia) согласно инструкции производителя. Полученную кДНК использовали в качестве матрицы для амплификации фрагментов генома. Фрагменты генома SARS-CoV-2 размером около 2000 пар нуклеотидов амплифицировали с использованием разработанных праймеров [4].

Предварительно смешанные и очищенные ПЦР-продукты (17-300 нг для образца) были фрагментированы в microTUBE-50 AFA Fiber Screw-Cap (PN 520166) с использованием Covaris M220 (Covaris, Woburn, MA). Парное лигирование было произведено набором NEBNext® Ultra™ End Repair/dA-Tailing Module (NEB E7442L), NEBNext® Ultra™ Ligation Module (NEB E7445L) and Y-shaped adapters compatible with Nextera XT Index Kit. Амплификацию библиотеки проводили Q5 High-Fidelity ДНК-полимеразой в 25 мкл согласно инструкции производителя (New England BioLabs, NEB) с 10 циклами амплификации. Так же использовали другой метод приготовления библиотек, называемый Nextera™ Library Prep (Illumina, USA) согласно с инструкцией производителя. Выбор размера конечных библиотек производился с использованием Agencourt AMPure XP (Beckman Coulter, Danvers, MA, USA). Длину фрагментов полученных библиотек оценивали на Agilent Bioanalyzer 2100 (Agilent Technologies, USA). Сиквенс был выполнен на Illumina HiSeq 1500 с использованием набора реагентов HiSeq PE Rapid Cluster Kit v2 и HiSeq Rapid SBS Kit v2 (500 циклов) или Illumina MiSeq с набором MiSeq Reagent Kit V2 (500 циклов). Обработка данных производилась методом, описанным ранее [4].

Результаты. Мы определили последовательности полных геномов для 28 вирусов SARS-CoV-2 из образцов, представляющих собой клинический материал (мазок/ отделяемое ротоглотки) от пациентов разных полов, в возрасте от 20 до 82 лет. Данные доступны в базе данных GISAD (Таблица № 1). Был проведен филогенетический анализ последовательностей (из-за ограничения по формату, данные не представлены), мы

обнаружили, что последовательности вирусов не выделяются в отдельную кладу, а демонстрируют родство с вирусами европейского и американского происхождения. Проанализированные изоляты принадлежат к линиям В.1 (GH) (к ней относятся 6 (21,4%) проанализированных последовательностей) и В.1.1 (GR) (22 (78,6%) проанализированных последовательностей). Во всех геномах были обнаружены сочетания мутаций P323L в области РНК-зависимой РНК-полимеразы и D614G в гене S-белка (см. также таблицу №1). Также часто встречались мутации R203K и G204R, в гене, кодирующем нуклеокапсидный белок, они были найдены в 20 (71%) и 19 (67%) проанализированных образцов, соответственно. Обе мутации по данным GISAID относятся к достаточно распространенным. Их встречаемость среди проанализированных последовательностей SARS-CoV-2 во всем мире составляет 32,30% и 32,21%, соответственно. Также были найдены редкие варианты SNP мутаций, такие как: T16A (помимо последовательности, описываемой в настоящей работе, такая мутация наблюдалась во всего двух последовательностях SARS-CoV-2 из базы GISAID), V359I (наблюдалась также в двух последовательностях из базы GISAID) и E39D (также найдены всего в двух последовательностях, по данным GISAD). В одном из образцов (GISAID ID EPI_ISL_462149) был обнаружен сдвиг рамки считывания в позиции 27330 (ORF6).

Таблица

№	Название	GISAD ID	Линия	Мутации	Предварительный диагноз/ по МКБ	Возраст пациента	Пол
1	hCoV-19/ Russia/ CRIE187963/2020	EPI_ISL_486829	В.1.1 (GR)	P323L; D614G; R203K, G204R	Контакт с больным / Z20.8	27	М
2	hCoV-19/ Russia/ CRIE176675/2020	EPI_ISL_486828	В.1.1 (GR)	P323L; D614G; P199S, R203K, G204R	ОРВИ / J06.9	20	М
3	hCoV-19/ Russia/ CRIE175960/2020	EPI_ISL_486827	В.1.1 (GR)	D218E; P323L; D614G; R203K, G204R	ОРВИ / J06.9	43	М
4	hCoV-19/ Russia/ CRIE170180/2020	EPI_ISL_486826	В.1.1 (GR)	P323L; D614G; R203K, G204R	ОРВИ / J06.9	36	Ж
5	hCoV-19/ Russia/ CRIE168693/2020	EPI_ISL_486825	В.1.1 (GR)	P323L; D614G; R203K, G204R	Пневмония / J18.9	48	Ж
6	hCoV-19/	EPI_ISL_4	В.1.	D218E, P1326L	ОРВИ /	25	Ж

	Russia/ CRIE164980/2 020	86824	1 (GR)	; P323L; D614G; R203K, G204R	J06.9		
7	hCoV-19/ Russia/ CRIE164399/2 020	EPI_ISL_4 86823	B.1. 1 (GR)	P323L; T16A; D614G; R203K	ОРВИ / J06.9	40	М
8	hCoV-19/ Russia/ CRIE164298/2 020	EPI_ISL_4 86822	B.1. 1 (GR)	T176A; P323L; D614G; R134H	ОРВИ / J06.9	29	Ж
9	hCoV-19/ Russia/ CRIE162398/2 020	EPI_ISL_4 86821	B.1. 1 (GR)	D218E; P323L; D614G; R203K, G204R	Острый бронхит / J20.9	54	М
10	hCoV-19/ Russia/ CRIE161143/2 020	EPI_ISL_4 86820	B.1. 1 (GR)	P323L; D614G; R203K, G204R	ОРВИ / J06.9	63	Ж
11	hCoV-19/ Russia/ CRIE160790/2 020	EPI_ISL_4 86819	B.1. 1 (GR)	P323L; D614G; R203K, G204R	Контакт с больным / Z20.8	32	М
12	hCoV-19/ Russia/ CRIE160277/2 020	EPI_ISL_4 86818	B.1. 1 (GR)	P323L; D614G; R203K, G204R, Q418H	ОРВИ / J06.9	39	Ж
13	hCoV-19/ Russia/ CRIE159932/2 020	EPI_ISL_4 86817	B.1. 1 (GR)	P323L; D614G; R203K, G204R	ОРВИ / J06.9	79	М
14	hCoV-19/ Russia/ CRIE159875/2 020	EPI_ISL_4 86816	B.1. 1 (GR)	P323L; D614G; A152S, R203K, G204R	Контакт с больным / Z20.8	54	N/A
15	hCoV-19/ Russia/ CRIE159560/2 020	EPI_ISL_4 86815	B.1. 1 (GR)	P323L; D614G; R203K, G204R	Контакт с больным / Z20.8	30	Ж
16	hCoV-19/ Russia/ CRIE155506/2 020	EPI_ISL_4 79624	B.1. 1 (GR)	P323L, V359I; D614G; R203K, G204R	ОРВИ / J06.9	57	М
17	hCoV-19/ Russia/ CRIE180117/2 020	EPI_ISL_4 79623	B.1 (GH)	Q1884H; P323L; D614G; Q57H	Наблюдени е при подозрении на другие болезни / Z03.8	64	N/A
18	hCoV-19/ Russia/ CRIE164776/2	EPI_ISL_4 79622	B.1 (GH)	Q1884H; P323L; D614G; Q57H	Наблюдени е при подозрении	82	М

	020				на другие болезни / Z03.8		
19	hCoV-19/ Russia/ CRIE161454/2 020	EPI_ISL_4 79621	B.1. 1 (GR)	P323L; D614G; R203K, G204R	ОРВИ / J06.9	28	М
20	hCoV-19/ Russia/ CRIE182061/2 020	EPI_ISL_4 79620	B.1. 1 (GR)	P395L; P323L; D614G, H1101 Y; R203K, G204R	ОРВИ / J06.9	21	М
21	hCoV-19/ Russia/ CRIE130014/2 020	EPI_ISL_4 70539	B.1 (GH)	T85I; P323L; D614G; Q57H; E39D	N/A	N/A	Ж
22	hCoV-19/ Russia/ CRIE139642/2 020	EPI_ISL_4 67775	B.1. 1 (GR)	P323L; D614G; R203K, G204R	ОРВИ / J06.9	N/A	М
23	hCoV-19/ Russia/ CRIE163844/2 020	EPI_ISL_4 67774	B.1 (GH)	Q1884H; P323L; D614G; Q57H	N/A	N/A	Ж
24	hCoV-19/ Russia/ CRIE162784/2 020	EPI_ISL_4 62150	B.1. 1 (GR)	D218E; P323L; D614G; R203K, G204R	N/A	N/A	N/A
25	hCoV-19/ Russia/ CRIE160583/2 020	EPI_ISL_4 62149	B.1 (G)	A318V; P323L; D614G	N/A	N/A	Ж
26	hCoV-19/ Russia/ CRIE140762/2 020	EPI_ISL_4 60605	B.1 (GH)	T85I; P323L; D614G; Q57H	N/A	N/A	Ж
27	hCoV-19/ Russia/ CRIE145922/2 020	EPI_ISL_4 60604	B.1. 1 (GR)	P323L; D614G; R203K, G204R	N/A	N/A	М
28	hCoV-19/ Russia/ CRIE156278/2 020	EPI_ISL_4 60096	B.1 (GH)	T85I; P323L; G13E;S33R, I1 12V; D614G ; Q57H	N/A	N/A	Ж

Выводы. Мы обнаружили сочетание мутации P323L и D614G во всех проанализированных геномах, образцы были получены от пациентов самых разных возрастных групп, для анализа отбирали материал только с высокой вирусной нагрузкой (Ct ≤20). Сочетание мутаций P323L и D614G было обнаружено во всех проанализированных последовательностях, в отличие от данных GISAD, в которой

показано, что мутация D614G встречается в 76,53% последовательностей, мутация P323L - в 76,24%. Обнаруженные отличия могут быть объяснены выборкой: для полногеномного секвенирования отбирались только образцы с высокой вирусной нагрузкой, которая, могла быть обусловлена именно высокой скоростью размножения вируса в организме хозяина, что, в свою очередь, может иметь связь с наличием в геноме вируса мутаций P323L/D614G.

Литература

1. Sten Ilmjärv, Fabien Abdul, Silvia Acosta-Gutiérrez, Carolina Estarellas, Ioannis Galdadas *et al.* Epidemiologically most successful SARS-CoV-2 variant: concurrent mutations in RNA-dependent RNA polymerase and spike protein. medRxiv 2020.08.23.20180281;doi: <https://doi.org/10.1101/2020.08.23.20180281>.
2. Zhang, Lizhou *et al.* “The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity.” bioRxiv : the preprint server for biology 2020.06.12.148726. 12 Jun. 2020, doi:10.1101/2020.06.12.148726.
3. Isabel S, Graña-Miraglia L, Gutierrez JM, *et al.* Evolutionary and structural analyses of SARS-CoV-2 D614G spike protein mutation now documented worldwide. Sci Rep. 2020;10(1):14031. Published 2020 Aug 20. doi:10.1038/s41598-020-70827-z.
4. AS Speranskaya, VV Kapteleva *et al.* SCV-2000bp: a primer panel for SARS-CoV-2 full-genome sequencing. bioRxiv 2020.08.04.234880; doi: <https://doi.org/10.1101/2020.08.04.234880>.