

ГЕНЕТИЧЕСКИЕ ФАКТОРЫ РИСКА РАЗВИТИЯ КОРОНАВИРУСНОЙ ИНФЕКЦИИ COVID-19 У ЖИТЕЛЕЙ САНКТ-ПЕТЕРБУРГА

Чангалиди А.И. (Университет ИТМО, Институт биоинформатики), **Барбитов Ю.А.** (Институт биоинформатики, Санкт-Петербургский Государственный Университет, Отдел геномной медицины ФГБНУ “НИИ АГиР им. Д.О. Отта”), **Анисенкова А.Ю.** (Санкт-Петербургский Государственный Университет, Городская больница № 40 Курортного административного района Сестрорецк), **Щербак С.Г.** (Санкт-Петербургский Государственный Университет, Городская больница № 40 Курортного административного района Сестрорецк), **Глотов А.С.** (Санкт-Петербургский Государственный Университет, Отдел геномной медицины ФГБНУ “НИИ АГиР им. Д.О. Отта”), **Глотов О.С.** (Отдел геномной медицины ФГБНУ “НИИ АГиР им. Д.О. Отта”, Городская больница № 40 Курортного административного района Сестрорецк)

В ходе работы был проведен анализ ассоциаций больных коронавирусом между генетическими вариантами и клиническими признаками, связанными с болезнью. Были выявлены потенциальные генетические маркеры заболевания, дана предположительная интерпретация результатов.

Введение

Коронавирусы (CoV) — это группа РНК-вирусов, поражающих птиц и млекопитающих (включая людей). В конце 2019 года в китайском городе Ухань была зарегистрирована вспышка пневмонии, которая была вызвана новым штаммом коронавируса, SARS-CoV-2, и получила название COVID-19. Достаточно быстро эта вспышка переросла в пандемию, вследствие чего важно изучать факторы, влияющие на смертность и тяжесть течения заболевания, а также их связь с генетическими вариациями.

Основная часть

В данном исследовании были проанализированы данные экзомных секвенирований и клинические характеристики 537 пациентов, у которых был диагностирован COVID-19 (подтвержденные и возможные случаи).

Сначала нами был произведен поиск генетических вариантов в данных экзомного секвенирования. Для этого результаты выравнивания прочтений обрабатывали при помощи инструмента DeepVariant v.0.10.0, после чего проводили когортное генотипирование образцов при помощи GLnexus v.1.2.6. Затем производили фильтрацию полученных вариантов при помощи GATK v.4.1.2 по глубине покрытия вариативных позиций прочтениями (не менее 10 прочтений), а также по представленности в выборке (генотип определен не менее, чем у 95% пациентов). Варианты были проаннотированы при помощи Ensembl Variant Effect Predictor.

После этого производили предобработку клинической информации (смертность, тяжесть течения заболевания, а также биохимические показатели, такие как: уровень С-реактивного белка, лимфоцитов, ферритина, Д-димера, интерлейкина 6 (ИЛ-6), ПЦР-тестирование); недостающие биохимические данные были восстановлены при помощи метода kNN. Были исключены выбросы — значения, отстоящие больше, чем на 3 стандартных отклонения, и проведена нормализация данных при помощи метода INT. Была обнаружена статистическая корреляция тяжести и смертности больных COVID-19 с биохимическими маркерами (уровнем С-реактивного белка, ИЛ-6 и др.).

Последним шагом было обнаружение ассоциаций между генетическими вариантами и признаками, связанными с болезнью: были проведены полногеномный поиск ассоциаций (GWAS) и анализ ассоциаций редких вариантов (RVAS) с использованием библиотеки Nail для Python и скриптов собственной разработки.

При анализе ассоциаций отдельных частых вариантов были обнаружены сигналы ассоциации с уровнями Д-димера и ферритина, а также одиночные локусы, связанные с уровнем ИЛ-6. Для вариантов, связанных с уровнем ферритина, не удалось найти биологического обоснования, в случае ИЛ-6 ассоциированные локусы могут иметь связь с компонентами внеклеточного матрикса и метаболическим путем через рецептор ACE2. При анализе суммарной частоты встречаемости редких вариантов на уровне целых генов была обнаружена ассоциация миссенс-вариантов в гене *PAPSS2* с низким уровнем лимфоцитов. Также был обнаружен “слабый” сигнал ассоциации гена *ITGAM* с тяжестью течения и количеством С-реактивного белка.

Выводы

Результаты исследования могут помочь лучше понять пути воздействия вируса SARS-CoV-2 на организм, а также применять персонализированную медицину в серьезных случаях, чтобы избежать тяжелых последствий COVID-19. Полученные данные, однако, предварительны и требуют верификации на большей выборке и в других исследованиях, так как уже неоднократно в исследованиях геномных ассоциаций с течением заболевания COVID-19 были найдены варианты, роль которых не была подтверждена впоследствии.

Чангалиди А. И. (автор)

Барбитов Ю.А. (автор)

Анисенкова А.Ю. (автор)

Щербак С.Г. (автор)

Готов А.С. (научный руководитель)

Готов О.С. (научный руководитель)