

Новый метод распознавания иммуоэпитопов, маркеры долговременного иммунитета, иммуносупрессивные домены и вакцины против COVID-19

Е. П. Харченко*

ФГБУН «Институт эволюционной физиологии и биохимии им. И. М. Сеченова» РАН, Санкт-Петербург

Резюме

Актуальность. Поиск компьютерных методов с высокой эффективностью распознавания иммуоэпитопов и прогнозирование долговременности индуцируемого ими иммунитета определяется прежде всего необходимостью быстрого создания вакцин против вновь возникших инфекций, особенно в периоды пандемий. **Цель.** Разработка нового иммуоинформационного метода распознавания иммуоэпитопов, выявление в первичной структуре вирусных белков возможных маркеров их потенциала индуцировать долговременный иммунитет и в оценке ими вакцин против COVID-19. **Материалы и методы.** Для компьютерного анализа использовались доступные в Интернете базы данных иммуоэпитопов длиной в 15 и 9 аминокислот, рестриктированных соответственно по MHC I и MHC II, и пептидов, не связывающихся с MHC, а также белков человека и вирусов. Алгоритм дискриминации иммуоэпитопов основывался на позиционном различии в их первичных структурах специфичных коротких пептидов. **Результаты.** «Инвентаризация» в обучающих выборках ди- и трипептидов или пентапептидов иммуоэпитопов и неиммуоэпитопов позволяет безошибочно распознать в контрольных выборках до 93–97% иммуоэпитопов, рестриктированных по MHC I и MHC II. В белках разных субъединичных вакцин, вызывающих длительный иммунитет, доминируют аминокислоты (особенно пролина), составляющие основу внутренне дезорганизованных областей, и пролин-содержащие дипептиды, что позволяет рассматривать их как биомаркеры потенциала вирусного белка формировать долговременную иммунную память. **Вывод.** Метод распознавания иммуоэпитопов и биомаркер индуцирования долговременной иммунной памяти могут быть использованы как биоинформативные инструменты вычислительной вакцинологии. Обеспечение долговременного иммунитета вакцинами на основе белка S коронавируса SARS-CoV-2 маловероятно.

Ключевые слова: иммуоэпитопы, биомаркеры, иммунная память, иммуносупрессивные домены, коронавирусы, прогнозирование Конфликт интересов не заявлен.

Для цитирования: Харченко Е. П. Новый метод распознавания иммуоэпитопов, маркеры долговременного иммунитета, иммуносупрессивные домены и вакцины против COVID-19. Эпидемиология и Вакцинопрофилактика. 2022;21(1): 4–20. <https://doi.org/10.31631/2073-3046-2022-21-1-4-20>.

Novel Method of Immunoepitope Recognition, Long-Term Immunity Markers, Immunosuppressive Domens and Vaccines against COVID-19

EP Kharchenko

Sechenov Institute of Evolutionary Physiology and Biochemistry, Russian Federation, St. Petersburg

Abstract

Relevance of searching for computer methods with high efficiency of immunoepitopes recognition and predicting the longevity of the immunity they induce is determined primarily by the need to quickly create vaccines against newly emerging infections, especially during pandemic periods. **Aim.** To develop a new immunoinformation method for recognizing immunoepitopes, to identify in the viral proteins possible potential markers to induce long-term immunity and to evaluate by them the vaccines against Covid-19.

Materials and methods. For computer analysis, an Internet-accessible databases of immunoepitopes 15 and 9 amino acids long, restricted respectively by MHC I and MHC II, and peptides not binding to MHC, as well as human and virus proteins, were used. The algorithm for discriminating immunoepitopes was based on positional distinction of specific short peptides in their primary structures. **Results.** The «inventory» in the training samples of di- and tripeptides or pentapeptides of immunoepitopes and non-immunoepitopes makes it possible to accurately recognize in the control samples up to 93–97% of immunoepitopes restricted by MHC I and MHC II. Comparison of the amino acid composition of proteins of subunit vaccines causing long-term immunity revealed dominance of amino acids (especially proline), which form the basis of internally disorganized regions, and proline-containing

* Для переписки: Харченко Евгений Петрович, д. б. н., ведущий научный сотрудник Института эволюционной физиологии и биохимии им. И. М. Сеченова, 194223, Санкт-Петербург, пр. Тореза, 44. 8(904)338-22-80, neuro.children@mail.ru. ©Харченко Е. П.

* For correspondence: Kharchenko Eugene P., Dr. Sci. (Biol.), leader researcher Sechenov Institute of Evolutionary Physiology and Biochemistry, 44 Torza pr., St. Petersburg, Russian Federation, 194223, 8(904)338-22-80, neuro.children@mail.ru. ©Kharchenko EP.

dipeptides, that allowed them to be considered as biomarkers of the potential of a viral protein to form a long-term immune memory. In the S-protein of coronavirus SARS-CoV-2 two candidates for immunosuppressive domains are present and the dominance of proline and dipeptides containing it is absent. **Conclusion.** The immunopeptide recognition method and the biomarker for inducing long-term immune memory can be used as immunoinformative tools of computational vaccinology. Providing long-term immunity by vaccines based on the coronavirus SARS-CoV-2 protein S is unlikely.

Keywords: immunopeptides, biomarkers, immune memory, immunosuppressive domains, coronaviruses, prediction
No conflict of interest to declare.

For citation: Kharchenko EP. Novel Method of Immunopeptide Recognition, Long-Term Immunity Markers, Immunosuppressive Domains and Vaccines against COVID-19. *Epidemiology and Vaccinal Prevention*. 2022;21(1): 4–20 (In Russ.). [https://doi: 10.31631/2073-3046-2022-21-1-4-20](https://doi.org/10.31631/2073-3046-2022-21-1-4-20).

Введение

Распознавание «своего» и «не своего» на уровне адаптивной иммунной системы (ИС) реализуется Т-клетками через сканирование ими презентруемых главными комплексами гистосовместимости (МНС) фрагментов белков и липидов. Короткие фрагменты белков (чаще всего инфекционных агентов), которые комплексируются с МНС и распознаются рецепторами Т-клеток (запуская реакцию ИС), именуются иммуноэпитопом (ИЭ). Сам процесс распознавания ИЭ предполагает связывание его первоначально с двумя разными молекулами, составляющими МНС, а затем с двумя разными молекулами рецептора Т-клеток.

В первом приближении априорно можно предположить, что комплексообразование ИЭ с МНС требует трех его контактов с МНС (по обоим концам и центральной части ИЭ), еще два контакта необходимы на взаимодействие ИЭ с разными цепями рецептора Т-клеток. Если принять длину ИЭ, связывающегося с МНС класса I (МНС I), в 9 аминокислот, а для ИЭ, связывающегося с МНС класса II (МНС II), в 15 аминокислот, то становится очевидным, что каждый из 5 контактов в иммунном узнавании опосредуется участием 1–2 аминокислот в случае ИЭ, рестриктированного МНС I, и не более 2–3 аминокислот в случае ИЭ, рестриктированного МНС II, вовлекая в контакты практически всю последовательность ИЭ. Этим можно было бы объяснить очень слабую аффинность взаимодействий компонентов в комплексах МНС-ИЭ-Т-клеточный рецептор и необходимость тонких методов обнаружения их немногочисленных циркулирующих экзепляров [1].

Использование столь коротких фрагментов в ИЭ, как ди- и трипептиды для иммунного распознавания в комплексе МНС-ИЭ-Т-клеточный рецептор, является достаточным благодаря динамичности самого процесса. В качестве других примеров использования природой минимальных «образов» для распознавания можно привести обонятельную систему или генетический код. Как известно, последний основан на 64 триплетях и является вырожденным, и в 32 триплетях определяющими являются лишь первые два нуклеотида. Природа не роскошествует.

Три основных экспериментальных метода, направленных на распознавание кандидатов

Т-клеточных ИЭ (без прогнозирования длительности индуцируемого ими иммунитета), включают определение прямого связывания синтезированного пептида с МНС, элюцию связанных с МНС лигандов с последующей их идентификацией и, наконец, анализ способности Т-клетки реагировать на соответствующий ИЭ-кандидат. С накоплением базы данных по первичным структурам идентифицированных ИЭ и не связывающихся с МНС пептидов (далее в тексте они будут именоваться как неИЭ) возникла возможность компьютерного прогнозирования Т-клеточных ИЭ, что стало неотъемлемым элементом формирующейся ныне компьютерной (вычислительной) вакцинологии.

Динамичность структуры субъединиц МНС I и МНС II, как и конформационные переходы связывающихся с ними ИЭ позволяют реализовать взаимодействие МНС-ИЭ множеством, однако ограниченным числом, способов. Искусственные нейронные сети дают возможность создать целый ряд компьютерных программ по распознаванию Т-клеточных ИЭ. Особенно программы эффективны по распознаванию ИЭ, рестриктированным МНС I, однако в отношении ИЭ, рестриктированных МНС II, таких результатов пока достичь не удалось [2–4]. Этим, по-видимому, отчасти объясняется наличие в базе данных значительно меньшего (более чем на порядок) количества ИЭ, рестриктированных МНС II, по сравнению с ИЭ, рестриктированными МНС I.

Актуальность поисков методов с высокой эффективностью распознавания кандидатов в ИЭ, рестриктированных по МНС I и МНС II, определяется прежде всего необходимостью их идентификации в белках при аутоиммунных, аллергических, онкологических заболеваниях и при трансплантации органов. Важны идентификация ИЭ и прогнозирование долговременности индуцируемого ими иммунитета для быстрого создания вакцин, например, пептидных или субъединичных, против вновь возникших инфекций (особенно в период пандемии) и возбудителей сезонных эпидемий или при дизайне индивидуальных вакцин для лечения рака. С другой стороны, при появлении новых мутантов возбудителей инфекций расшифровка состава ИЭ белков позволила бы прогнозировать чувствительность мутантов к существующим вакцинам.

В не стихающей уже более двух лет коронавирусной пандемии мир столкнулся с быстрым рождением в разных регионах новых мутантов SARS-CoV-2, снижающих усилия по созданию коллективного иммунитета. Поскольку мутации в новых штаммах SARS-CoV-2 многочисленны, то происходит изменение иммуногенности вируса, что ведет к снижению эффективности используемых вакцин и, соответственно, к проявлениям гетерологичного иммунитета.

Цели исследования – разработка нового метода с высокой эффективностью распознавания кандидатов в ИЭ, выявление в первичной структуре вирусных белков маркеров потенциала индукции долговременного иммунитета для использования их в разработке новых вакцин и прогнозирования их эффективности; оценка потенциала существующих вакцин против новых штаммов SARS-CoV-2.

Материалы и методы

Для компьютерного анализа были использованы первичные структуры ИЭ белков человека из доступной в Интернете базы данных (www.iedb.org). Размеры ИЭ, распознаваемых МНС II, составляют 13–24 аминокислот. Поскольку полость молекулы МНС II вмещает лишь ИЭ длиной в 13–15 аминокислот и концы ИЭ большей длины провисают вне торцов полости, не участвуя в иммунном узнавании, то анализировались лишь ИЭ длиной в 15 аминокислот. По ИЭ, рестриктированным МНС I, в анализ были включены наиболее часто встречающиеся среди них пептиды длиной в 9 аминокислот. Анализировались линейные ИЭ и неИЭ безотносительно к способу их идентификации и привязки к определенному аллелю МНС I и МНС II. Обучающие и тестовые выборки формировались (по 1000 пептидов) случайным набором пептидов. Их необходимые объемы определялись путем последовательного подбора на основе анализа результатов дискриминации. В общей сложности анализ охватил около 20 000 ИЭ, рестриктированных по МНС I и МНС II.

Следует отметить, что объем данных по ИЭ и неИЭ отличается для разных организмов, в частности вирусов. Кроме того, для многих ИЭ не идентифицирована рестрикция по определенному гаплотипу МНС II и их аллелям. Поэтому при разработке метода распознавания ИЭ сделан акцент на универсальности, не уточняя его рестрикцию по конкретному гаплотипу МНС II либо его аллелю.

Поскольку контакты между компонентами в комплексе МНС-ИЭ-Т-клеточный рецептор, как отмечено выше, ограничены короткими последовательностями ИЭ (1–3 аминокислоты), анализ включал определение частоты встречающихся как в ИЭ, так и в неИЭ аминокислот, ди-, три- и пентапептидов по каждой из позиций (а применительно к дипептидам – и по нескольким позициям одновременно) с записью их в отдельные многомерные

массивы компьютерной программы. Такая инвентаризация репертуара коротких пептидов по каждой позиции аминокислотной последовательности ИЭ и неИЭ обучающих выборок позволила использовать принцип запрета (невозможности [5]) для различения ИЭ от неИЭ. Применительно к ИЭ и неИЭ его можно было бы выразить в следующей формулировке: ИЭ и неИЭ не могут иметь идентичные первичные структуры и должны хотя бы минимально отличаться друг от друга, что и наблюдается в реальности. Поэтому для различения их (при сильном преобладании данных по ИЭ) необходимо по ИЭ сформировать репрезентативную обучающую выборку, содержащую наиболее полный состав встречающихся в них, например, ди- и трипептидов. Инвентаризация позволяет разделять ИЭ от неИЭ по следующему правилу: если при пошаговом просмотре всей последовательности пептида в нем присутствовали разрешенные по каждой позиции ди- и трипептиды, зафиксированные в массивах данных по обучающей выборке ИЭ, но которые отсутствовали в массивах данных по обучающей выборке неИЭ, то пептид относился к ИЭ, а при обратной ситуации – к неИЭ. Поясним это на конкретном примере. Так, если при последовательной проверке каждой позиции аминокислотной последовательности его пептида на приходящий на нее дипептид XY (трипептид XYZ или пентапептид XYZQR) выявляется отсутствие в этой позиции у неИЭ, то пептид причисляется к ИЭ, поскольку, по данным инвентаризации дипептидного (трипептидного, пентапептидного) состава обучающих выборок ИЭ и неИЭ, дипептид XY (трипептид XYZ или пентапептид XYZQR) присутствует в данной позиции только у ИЭ и запрещен для данной позиции у неИЭ (X,Y,Z,Q,R –любая из 20 аминокислот). Используемый метод распознавания кандидатов ИЭ от неИЭ принципиально отличается от других иммуноинформативных инструментов определения Т-клеточных ИЭ, основанных на вероятностных подходах, и позволяет с высокой эффективностью безошибочно выявить ИЭ. Возможно применять две стратегии дискриминации ИЭ от неИЭ: прямую и обратную. В случае прямой стратегии ИЭ отсеиваются от множества неИЭ, а при обратной стратегии исключаются неИЭ. Нами использовались обе стратегии на основе базы данных ИЭ белков человека (www.iedb.org).

Для иллюстрации изменения в белках состава ИЭ и неИЭ у мутантных вирусов были применены последовательности S-белка трех штаммов пандемического коронавируса (уханьского, «дельта, δ» и «омикрон, о»). Дополнительно были проанализированы ИЭ у белков вирусов гепатита В, С и Е, папилломы человека, краснухи, гриппа А, кори и Эбола, а также у гистона H4, инсулина и тиреопероксидазы человека. Источником первичных структур пандемического коронавируса SARS-CoV-2, сезонных коронавирусов человека HCoV-NL63, HCoV-229E, HCoV-NKU1 и HCoV-OC43 и других

вирусов, а также белков человека служили общедоступные в Интернете базы данных (<http://www.ncbi.nlm.nih.gov> <http://www.platform.gisaid.org>, www.nextprot.org и <http://viralzone.expasy.org>).

В статье используется международный код аминокислот: А – аланин, С – цистеин, D – аспарагиновая кислота, Е – глутаминовая кислота, F – фенилаланин, G – глицин, H – гистидин, I – изолейцин, K – лизин, L – лейцин, M – метионин, N – аспарагин, P – пролин, Q – глутамин, R – аргинин, S – серин, T – треонин, V – валин, W – триптофан, Y – тирозин.

Результаты и обсуждение

Формально разработка метода распознавания кандидатов ИЭ, связывающихся с МНС I и МНС II, относится к задачам, сопряженным с распознаванием образов на основе различных подходов. Если кратко охарактеризовать особенности нашего метода дискриминации связывающихся с МНС II пептидов от неИЭ, то они заключаются в обнаружении тонких различий в их первичных структурах, в частности в выявлении по каждой позиции в их аминокислотных последовательностях разрешенных и запрещенных коротких пептидов (ди-, три- или пентапептидов) [5].

Возможность разработки метода распознавания ИЭ, связывающихся с МНС I и МНС II, подкрепляется прежде всего тем, что, несмотря на большую численность ИЭ для МНС I, для МНС II их значительно меньше. В этой связи следует заметить, что реализованное в процессе эволюции многообразие белковых последовательностей существенно меньше потенциально возможного, что, по-видимому, объясняется происхождением ныне существующих белков из сравнительно небольшого числа предковых генов. Принципиально важным моментом в механизмах возникновения разнообразия белков является то, что в числе основных способов увеличения размеров и числа белков оказались генные дубликации и мозаичные комбинации, причем большинство генов белков являются разорванными и составленными из разного числа экзонов и интронов. Сопоставление первичных структур белков показало, что они обнаруживают блочное родство, т.е. их последовательности родственны не по всей длине, а лишь по отдельным протяженным блокам, причем разветвленная сеть блочного родства охватывает белки, глубоко различающиеся по своим функциям, образуя континуум пептидного родства белков. Соответствие размеров экзонов в разорванных генах функциональным доменам белков дает основание полагать, что полипептиды представляют собой наборы фрагментов с различными функциями, которые эволюция по непонятным пока правилам собирает как одно структурно-функциональное целое [6].

Адаптивная ИС возникла на уровне позвоночных, намного позднее формирования основных функциональных систем и вовлеченных в них

белков. Предсуществование глобального пептидного континуума родства белков, охватывающего организмы разных ступеней эволюционного развития, обрело адаптивную ИС формироваться и функционировать на предуготовленном ей иммуноэпитопном континууме родства белков [6].

ИЭ не являются каким-то обособленным подмножеством пептидных фрагментов белков, и их рассеянность и повторяемость в разных белках лишь отражает особенности возникновения и эволюции белков. Достаточно взглянуть на таблицу 1, в которой приводится частота встречаемости аминокислот в каждой позиции ИЭ, связывающихся с гаплотипами DP, DQ и DR МНС II, чтобы убедиться в пестрой мозаике их возможных сочетаний. Комплексообразование ИЭ с МНС обоих классов является результатом возникновения между ними стерической конгруэнтности в диапазоне допустимых для них конформационных изменений, что разрешает взаимодействие МНС с разными по своей аминокислотной последовательности ИЭ, т.е. при ограниченности числа аллелей МНС у каждого индивидуума и возможности образования огромного числа пептидов узнавание МНС обоих классов пептидов априорно обречено быть вырожденным, т.е. один и тот же аллель МНС будет распознавать множество ИЭ. Но последнее не бесконечно (что следует прежде всего из существования иммуноэпитопного континуума родства белков [6]), и задача исследования состояла в выявлении этих ограничений на уровне аминокислотных последовательностей ИЭ и неИЭ по каждой их позиции.

Другое ограничение, налагаемое на численность ИЭ, связывающихся с МНС II, исходит из структурной специфики самих гаплотипов DP, DQ и DR, допускающих посадки на них лишь ограниченного множества пептидов со специфическим аминокислотным и соответственно ди-, три- и пентапептидным составом, что вытекает, например, из данных в таблицах 2, в которых представлены обобщенные частоты встречаемости дипептидов в ИЭ (табл. 2а) и в одной (для примера приводятся данные по первой позиции) из 15 позиций ИЭ (табл. 2б). Заметим, что хотя таблица 2а отображает разрешенные и запрещенные дипептиды в структуре ИЭ, однако не позволяет судить о полном репертуаре ди- и трипептидов в каждой позиции ИЭ. Эта информация заключена в многомерных (точнее – в 15-мерных) массивах данных по ди- и трипептидам сравниваемых ИЭ и неИЭ. Расчеты показали, что по каждой позиции в ИЭ не используется в среднем 307, в то время как в неИЭ – 144 дипептидов из 400. Еще большие различия между ИЭ и неИЭ проявляются по каждой позиции на уровне трипептидов.

Уже по общей сводке дипептидов в ИЭ (таблица 2а) просматривается запрет на 7 дипептидов, связанных с редко встречающимися в белках аминокислотами: цистеином, триптофаном и метионином (для сравнения: у неИЭ запрещены лишь два

Таблица 1. Распределение частот аминокислот в 15 позициях ИЭ, рестриктированных тремя гаплотипами MHC II
Table 1. Distribution of amino acid frequencies in 15 immunopeptides positions restricted by three MHC II haplotypes

DP MHC II																				
	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
1.	75	45	17	97	73	10	2	88	65	68	85	14	6	26	94	15	54	24	73	69
2.	62	44	22	86	74	129	0	51	38	46	77	18	4	24	105	11	47	47	68	47
3.	84	87	23	52	88	58	2	59	45	72	78	24	13	36	65	15	37	63	54	45
4.	101	134	10	30	63	26	2	106	52	65	50	49	21	91	27	17	30	32	46	48
5.	65	100	13	40	67	8	2	98	60	81	98	58	24	86	28	18	18	34	49	53
6.	50	44	30	26	48	17	3	70	55	81	129	63	25	63	55	16	28	35	92	70
7.	55	23	20	62	107	17	2	48	55	86	125	45	13	43	54	12	17	25	107	84
8.	62	18	20	62	107	47	3	78	52	77	111	24	12	45	37	13	32	39	71	90
9.	42	21	18	35	70	61	2	82	47	66	161	39	17	125	35	9	25	28	63	54
10.	41	20	16	25	93	37	2	145	38	69	183	34	7	93	17	10	37	48	50	35
11.	66	56	15	44	96	54	5	94	34	88	142	22	7	26	30	18	31	65	65	42
12.	36	39	26	63	74	41	10	101	61	88	111	21	7	20	49	12	70	37	91	43
13.	55	47	25	58	89	57	4	72	46	69	96	18	2	16	76	12	64	49	93	52
14.	83	40	35	65	77	112	1	66	53	73	91	26	3	18	33	12	39	57	68	48
15.	109	87	25	57	79	45	2	80	18	28	65	16	6	21	95	19	57	70	65	56
DQ MHC II																				
	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
1.	21	20	13	233	182	11	8	41	35	41	190	5	3	36	48	6	21	16	38	32
2.	35	29	29	90	90	177	2	84	46	65	66	15	11	19	84	9	30	36	44	39
3.	50	54	20	58	93	62	5	68	42	73	95	30	19	32	76	11	28	61	72	51
4.	36	37	27	68	88	39	7	83	59	70	91	31	26	49	61	18	28	48	78	56
5.	35	30	25	35	74	34	9	69	52	88	124	60	19	46	71	19	36	37	77	60
6.	32	37	26	38	59	33	9	59	53	59	141	39	13	34	99	13	21	44	104	87
7.	20	35	13	52	68	27	13	63	38	72	177	21	9	23	107	19	19	32	110	82
8.	35	33	10	66	81	34	12	77	48	80	161	15	7	31	90	10	21	39	85	65
9.	27	28	15	77	84	68	11	57	58	83	164	23	7	16	51	10	35	37	84	65
10.	27	42	24	69	109	56	10	83	59	84	132	20	3	26	51	11	38	41	65	50
11.	27	40	24	65	121	63	8	85	34	63	143	30	8	20	48	14	29	52	75	51
12.	30	53	14	61	104	74	7	85	41	54	130	22	13	21	87	11	33	42	74	44
13.	33	65	19	66	86	62	3	82	45	56	100	19	10	28	92	16	42	42	89	45
14.	56	56	23	64	90	107	4	73	45	79	81	31	12	38	43	24	25	54	61	34
15.	83	95	24	60	89	64	4	80	14	17	77	16	11	25	96	9	52	62	70	52
DR MHC II																				
	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
1.	68	25	8	115	70	7	0	83	76	83	83	18	6	37	93	10	54	27	78	58
2.	57	44	22	91	59	154	0	35	31	62	77	13	1	21	105	5	56	41	75	50
3.	87	72	19	53	71	76	0	64	46	75	43	37	6	33	92	9	47	57	65	47
4.	52	29	16	37	74	23	0	108	110	75	43	116	18	84	34	21	25	46	42	47
5.	39	42	15	19	52	14	1	95	134	105	50	120	14	141	24	8	19	38	37	33
6.	71	54	21	29	58	22	2	97	100	90	67	66	13	48	36	21	33	54	60	57
7.	40	41	20	54	59	32	6	128	88	84	94	46	8	48	31	39	40	49	46	46
8.	58	31	19	91	65	48	0	110	41	54	98	37	7	30	32	27	67	48	85	51
9.	92	46	20	50	43	61	1	45	31	69	109	13	3	15	84	12	64	29	116	96
10.	76	35	28	57	45	60	1	70	33	61	130	19	2	12	62	12	101	36	93	66
11.	61	68	25	47	75	60	1	95	35	70	110	21	4	15	57	19	47	57	75	57
12.	71	76	28	36	41	42	3	96	49	84	111	25	5	12	85	15	38	51	80	51
13.	78	85	22	55	41	36	1	75	43	91	105	22	8	26	68	11	50	46	99	37
14.	86	55	32	72	72	113	2	69	50	71	78	37	7	36	38	13	26	42	65	35
15.	90	86	33	69	77	34	1	70	10	28	68	24	6	19	126	12	43	67	80	5

Таблица 2а. Сводка дипептидов в ИЭ, рестриктированных МНС II
Table 2a. Summary of dipeptides in immunopeptides restricted by MHC II

	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
K	56	41	12	43	60	34	3	57	52	76	77	22	9	35	49	10	34	23	62	45
R	39	34	13	33	49	28	2	56	51	39	74	32	5	51	37	14	28	32	55	30
H	16	13	9	10	18	26	0	25	3	22	21	9	3	10	30	2	10	17	24	13
D	38	31	12	40	38	35	6	62	43	50	79	32	10	31	34	17	27	32	58	55
E	77	67	19	52	83	45	2	77	44	53	84	37	10	50	62	18	31	43	52	41
P	42	52	10	43	52	36	4	49	31	48	103	13	4	31	60	15	32	29	75	35
C	3	8	1	4	4	2	3	3	0	2	3	1	1	1	3	0	5	2	7	1
L	59	62	28	56	72	95	3	104	49	55	103	36	4	26	63	18	52	67	118	58
I	50	23	16	29	47	41	1	65	48	40	82	32	2	29	45	13	42	34	56	40
V	49	65	29	69	73	70	4	70	61	73	90	42	13	39	58	16	48	31	62	71
A	81	84	26	50	109	108	1	120	61	125	340	38	12	43	108	18	45	75	147	67
Y	38	22	13	22	39	20	2	52	25	27	33	20	11	19	17	7	17	20	38	34
W	15	7	2	12	17	3	1	7	8	11	8	1	0	5	14	0	14	6	4	14
F	36	33	19	32	52	12	1	56	19	56	38	24	8	5	18	10	18	37	38	25
G	63	51	16	43	55	49	1	53	41	71	112	23	3	17	68	5	33	35	59	30
M	6	12	5	5	14	13	0	13	7	17	25	4	2	9	17	2	17	11	12	9
N	32	29	24	29	36	39	5	48	37	43	43	21	2	31	29	8	21	22	43	26
Q	34	45	12	33	47	27	0	44	33	42	56	20	9	39	42	9	20	27	34	32
S	56	48	23	57	67	54	7	106	49	54	80	59	26	42	85	11	47	59	92	59
T	27	31	9	29	47	67	6	60	38	65	83	28	6	27	51	7	19	44	50	30

дипептида – MC и WM), Визуальному просмотру доступны лишь таблицы распределения дипептидов. При этом выявляется, что в отдельных позициях ИЭ число запретов на дипептиды сильно возрастает и проявляется их специфичность. Например, при наличии в 14-й позиции цистеина в 15-й позиции может быть только лейцин, причем сам дипептид CL выявлен лишь у двух из 1000 ИЭ и цистеин не сочетается с остальными 19 аминокислотами. Диапазон запретов дипептидов в позициях весьма широкий. Общность запретов на дипептиды во всех 15 позициях ИЭ связана с редко встречающимися аминокислотами в белках. Но помимо дипептидов с цистеином, триптофаном и метионином запреты распространяются и на дипептиды с гистидином, тирозином и фенилаланином, встречающиеся в белках чаще, чем первые, затрагивая и те аминокислоты, которые не относятся к разряду редких в белках, и в каждой позиции запреты на них обнаруживают специфичность, что в совокупности обеспечило эффективность разрабатываемого подхода дискриминации ИЭ от неИЭ.

ИЭ являются короткими пептидами, и каждая аминокислота в них (и тем более ди-

три- и пентапептиды) определяет потенциал пептида быть ИЭ. Замена в последовательности ИЭ одной аминокислоты может привести к снижению аффинности связывания и даже к утрате способности связываться с МНС, но ИЭ и полученный из него пептидный аналог по своим интегральным физико-химическим свойствам практически не различимы, что обрекает на низкую эффективность распознавания их подходами, основанными на использовании только физико-химических параметров аминокислот [2,3] без учета их последовательности в пептиде. Ранее такой подход оказался успешным в случае небольших выборок ИЭ [7,8], но он мало эффективен при анализе выборки, исчисляемой десятками и сотнями тысяч ИЭ.

Объективно трудности распознавания кандидатов в ИЭ по особенностям аминокислотных последовательностей пептидов связаны с сильной количественной асимметричностью данных по самим ИЭ и неИЭ. Если ИЭ, рестриктированные по МНС II, исчислялись несколькими десятками тысяч, а ИЭ, рестриктированные по МНС I, несколькими сотнями тысяч, то на неИЭ по каждому классу МНС в базе данных (www.iedb.org)

Таблица 2б. Пример сводки дипептидов в первой позиции ИЭ, рестриктированных МНС II
Table 2b. Example of a summary of dipeptides in the first position of immunopeptides restricted by MHC II

	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
K	8	5	0	4	8	8	0	2	6	4	5	1	0	3	5	2	2	3	3	5
R	4	1	1	2	7	3	0	4	4	1	5	0	0	0	8	0	3	2	0	2
H	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0
D	7	0	1	15	5	5	0	5	1	5	21	0	2	3	3	0	5	1	3	8
E	1	7	0	4	2	4	0	2	0	4	1	0	0	2	6	2	1	4	2	1
P	1	1	0	1	1	2	1	0	0	1	1	0	0	0	0	0	0	0	1	1
C	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
L	7	4	2	3	3	15	0	91	0	1	9	0	0	2	8	0	6	6	4	2
I	4	2	0	2	4	13	0	0	0	4	3	1	0	1	2	0	4	4	6	2
V	3	4	2	9	9	7	0	4	5	7	6	2	0	3	9	1	9	2	4	5
A	8	5	7	12	12	44	0	15	6	13	28	2	2	3	15	0	1	8	14	6
Y	0	0	0	1	0	2	0	0	0	0	0	1	0	0	0	0	0	0	1	0
W	1	0	1	0	0	1	0	1	0	1	0	0	0	0	4	0	1	0	0	0
F	1	2	0	0	2	0	0	0	1	2	1	2	0	0	2	0	0	1	3	0
G	3	1	1	3	1	10	0	1	4	3	5	1	0	0	8	0	2	2	2	2
M	0	1	2	1	2	2	0	0	0	0	3	0	0	0	0	0	2	0	0	0
N	4	4	1	2	2	10	0	4	1	7	4	2	0	1	4	1	2	2	4	5
Q	1	2	1	4	0	1	0	0	1	0	4	1	0	3	5	0	0	0	1	1
S	2	3	4	6	5	8	0	3	2	3	1	1	1	1	8	0	6	3	6	2
T	0	0	0	8	1	15	0	1	2	4	3	0	1	0	10	0	0	3	7	2

приходится немногим более 1000 пептидов. Такая асимметричность в данных чревата ситуацией, описываемой в области искусственных нейронных сетей, с одной стороны, как «переобученность», а с другой стороны, «недообученность», т.е. избыточностью одной группы данных и недостаточностью сравниваемой с ней другой группы. Другая трудность в разработке методов распознавания ИЭ связана с особенностями первичных структур пептидов: пептиды образуют огромное множество (особенно в случае пептидов длиной в 15 аминокислот) разных последовательностей из одного и того же состава аминокислот, что невозможно учесть при использовании методов распознавания, основанных на использовании физико-химических характеристик аминокислот и малом числе пептидов в обучающих выборках.

Поскольку основу развиваемого нами подхода составляла дискриминация ИЭ от неИЭ по их первичным структурам, разнообразие которых определяется как составом входящих в них аминокислот, так и их последовательностью, то важным моментом в разработке метода было выявление размера обучающей выборки. Как и ожидалось, первичный анализ результатов дискриминации показал, что наибольшее число правильных распознаваний по ИЭ, рестриктированным по каждому классу МНС, было связано с различиями по трипептидам, так как по численности они значительно преобладают над дипептидами. Поэтому в расчетах при определении необходимого числа ИЭ в обучающей

выборке мы ориентировались на необходимое множество трипептидов.

Если исходить из принципов комбинаторики и принять возможное число трипептидов как размещения из 20 аминокислот по 3, то с учетом повторов обучающая выборка должна составлять не менее 20^3 трипептидов, а их суммарная длина – 24 000 аминокислот. Деля последнюю величину на 15, получаем 1600 ИЭ, связывающихся с МНС II – минимальное число ИЭ для обучающей выборки для идеального случая. Но в реальности ИЭ длиной в 15 аминокислот не являются уникальными по пептидному составу, и в своем множестве в их первичных структурах многократно повторяются короткие пептиды. Поэтому численность обучающей выборки ИЭ априорно должна быть значительно выше величины для идеальной ситуации. Пошаговым увеличением ее объема на 1000 пептидов выяснено, что 4000–5000 ИЭ в обучающей выборке позволяют (при последовательной фильтрации всех 15 позиций пептида на наличие в них разрешенных ди- и трипептидов) достичь в среднем 96% безошибочного распознавания ИЭ, рестриктированных по МНС II, и полного исключения неИЭ, что является, насколько нам известно, наиболее высоким уровнем правильного прогнозирования для МНС II. Среди особенностей дискриминации ИЭ отметим, что, по сравнению с ИЭ, рестриктированными по DP и DR гаплотипам МНС II, ИЭ, рестриктированные по DQ гаплотипу распознавались с меньшей эффективностью. Естественно,

асимметричность в размерах обучающих выборок ИЭ и неИЭ (последняя из-за ограниченного числа их в базе данных составляла 1000 пептидов), с одной стороны, и сосуществование в реальных ИЭ и неИЭ идентичных пептидных фрагментов в одной и той же позиции, с другой стороны, предопределили отказ от требования наличия в ИЭ уникальности по разрешенным пептидным фрагментам во всех позициях.

Таким образом, используя обучающую выборку ИЭ большого размера и оптимизируя уровень разрешенных трипептидных фрагментов, удалось при ограниченной обучающей выборке неИЭ достичь высокого уровня безошибочной (относительно базы данных ИЭ (www.iedb.org)) дискриминации кандидатов ИЭ, рестриктированных по МНС II, от неИЭ. Приведенные показатели уровня распознавания ИЭ не являются предельными. С увеличением репрезентативности обучающей выборки ИЭ, что предполагает включение в нее еще несколько тысяч ИЭ, распознавание приблизится к максимально возможному минимуму потери информации. Кроме того, следует подчеркнуть, что метод потенциально позволяет распознавать новые кандидаты ИЭ, не входящие в существующие базы данных.

В случае ИЭ, рестриктированных по МНС I, потребовалась еще большая обучающая выборка из-за меньшей их размерности (9 аминокислот) по сравнению с ИЭ, рестриктированными по МНС II (15 аминокислот), и из-за многократной повторяемости разных коротких пептидных фрагментов в ИЭ. При 9000 и 1000 пептидов соответственно в обучающей выборке ИЭ и неИЭ была достигнута правильная дискриминация 92% ИЭ тестируемой выборки при полном исключении неИЭ. Отметим, что уровень правильного распознавания ИЭ в отдельных тестовых выборках из 1000 пептидов превышал 97%.

Распознавание ИЭ представляется ценным в нескольких аспектах. Один из них связан, как отмечалось выше, с дизайном пептидных или субъединичных вакцин и позволяет посредством анализа состава ИЭ и неИЭ в первичной структуре пептида или белка прогнозировать их иммуногенность и тем самым ускорить создание вакцин. Первый этап в таком прогнозировании – определение способности, например, пептида или фрагмента белка комплексоваться с МНС. На основе разработанного нами подхода были построены последовательности распределения потенциальных ИЭ в S-белке трех штаммов коронавируса SARS-CoV-2 (уханьского, «δ» и «ο»), отличающихся по возникшим в них мутациям. Неожиданным оказалось то, что последовательность S-белка всех трех штаммов плотно усеяна ИЭ белков человека, и эту иммуоэпитопную мимикрию вирусных белков можно было бы охарактеризовать, следуя терминологии De Groot et al., как иммунный камуфляж [9], ослепляющий ИС инфицируемого

и обеспечивающий вирусу высокую контагиозность с индуцированием им множественной патологии при проникновении в организм. Выявленное в S-белке SARS-CoV-2 множество ИЭ, рестриктированных по МНС II, по численности превосходит те, что выявлены другими авторами другими методами [10,11], и свидетельствует о том, что разработанный нами метод открывает новые кандидаты ИЭ, обнаружение которых недоступно другими методами. Как продукты деградации S-белка, отдельные ИЭ, будучи гомологичными разным белкам человека, способны индуцировать к ним антитела, вызывая многочисленные аутоиммунные поражения, чем так характерен патогенез инфекции SARS-CoV-2 и что послужило поводом сам вирус назвать как аутоиммунный [12].

Уместно заметить, что одного иммуоэпитопного камуфляжа S-белка SARS-CoV-2 недостаточно для проявления высокой контагиозности вируса. (Он предшествует и свойственен, по-видимому, белкам всех инфекционных патогенов, обеспечивая на начальных стадиях инфекции ускользание от ИС). Необходимо также обретение структурой S-белка конгруэнтности клеточным белкам поражаемого им хозяина. Оно «оттачивается» множеством мутаций, разные комбинации которых порождают новые пандемические штаммы SARS-CoV-2. Для них характерно увеличение в S1-субъединице S-белка количества аминокислот, образующих в белках внутренне дезорганизованные области, и повышенное количество аргинина и лизина по отношению к глутаминовой и аспарагиновой кислотам [13].

Следующий этап в дизайне вакцин – прогнозирование узнавания кандидата ИЭ Т-клетками ИС. Не все пептиды, комплексирующие с МНС, будут распознаны Т-клетками с запуском иммунной реакции. Из организма элиминируются большинство аутореактивных Т-клеток, и основной репертуар Т-лимфоцитов в циркуляции представлен теми, которые распознают «не своего». К сожалению, база данных по ИЭ, рестриктированным МНС II и длиной в 15 аминокислот, как и по неИЭ, не многочисленна (и первых, и вторых около 1000), и применить в полной мере подход, использованный для распознавания пептидов, связывающихся с МНС I и МНС II, не представляется возможным. Из-за недостаточной репрезентативности обеих выборок полностью распознаются неактивирующие Т-клетки пептиды, а безошибочно выявляется до 51% ИЭ, распознаваемых Т-клетками. Но безошибочность выявления последних искупает неполноту их распознавания.

В таблице 3 представлено распределение в S-белке SARS-CoV-2 кандидатов Т-клеточных ИЭ. При внимательном просмотре таблицы 3 видно, как от штамма к штамму перекраивается череда следования ИЭ и неИЭ (за каждым знаком «+» или «-» стоит последовательность из 14 аминокислот), что позволяет ретроспективно спрогнозировать изменение как патогенности, так и чувствительности

штаммов «δ» и «ο» к вакцинам, разработанным против уханьского штамма. Примечательна высокая плотность и протяженность блоков ИЭ в белках. Столь же высокой плотностью характеризуется S-белок и по В-клеточным ИЭ (распознавание их не описывается в данной публикации). Это свидетельствует о том, что образование антител против S-белка может быть направлено к разным его сайтам в обеих субъединицах, однако их доступность антителам в цельном вирионе будет разной. Протективной активностью будут обладать те антитела, что направлены к обнаженным фрагментам молекулы. В совокупности информация по распределению Т-клеточных ИЭ в S-белке SARS-CoV-2 в таблице 3 свидетельствует о том, что S-белок разных штаммов SARS-CoV-2 будет обладать специфичной иммуногенностью, что и подтверждается сообщениями о резком снижении чувствительности штаммов «δ» и «ο» к вакцинам на основе S-белка уханьского штамма SARS-CoV-2.

Выявление мутаций в S-белке необходимо, но не достаточно для прогнозирования их влияния на его иммуногенность. Развиваемый нами метод позволяет выяснить, как изменилась природа фрагмента белка, в котором произошла мутация. Она может проявить себя тремя исходами: не изменить прежнего состояния фрагмента, который она затрагивает, относительно принадлежности к ИЭ или неИЭ; привести к утрате принадлежности фрагмента к ИЭ либо к обретению им принадлежности к ИЭ, если ранее он соответственно являлся ИЭ или неИЭ.

При множественности Т- и В-клеточных ИЭ вакцины (уханьского) штамма SARS-CoV-2 первой волны пандемии могут сохранять в значительной степени их протективную способность против новых мутантов, что подтверждено в отношении «δ» штамма, имеющего, по сравнению с штаммом «ο», значительно меньше сдвигов в составе ИЭ. Однако вирус штамма «ο», SARS-CoV-2, оказывается, заражает привитых вакциной из S-белка уханьского штамма SARS-CoV-2. Более 30 мутаций в S-белке «ο» штамма привели к сокрытию части последовательностей, узнаваемых антителами, индуцируемыми вакцинами к уханьскому штамму, и к появлению новых ИЭ, не узнаваемых ими, что, вероятно, обернулось в отношении штамма «ο» реализацией (в рамках гетерологичного иммунитета) феномена антитело-зависимого усиления вирусной инфекции. Последний может обуславливаться парадоксальным эффектом моноштаммспецифичной вакцины, когда после ее введения вакцинированный подвергается инфицированию новым мутантным вирусом, который по механизму иммунного импринга («первородного греха») вызывает образование в организме антител, реактивируя клетки иммунной памяти, сформировавшиеся в результате предшествующей вакцинации. Влияние активированных Т- и В-клеток памяти может быть двояким. Если

индуцированные их активацией антитела способны нейтрализовать мутантный вирус, то это ослабляет инфекционный процесс. При отсутствии у антител нейтрализующей активности наблюдается иммунная коллизия: антитело-зависимое усиление инфекции, цитокиновый шторм, усиление воспаления и вызванная им гибель клеток. В итоге – повышенная контагиозность инфекции, вызванной мутантным вирусом. В разных странах распространение штамма «ο» протекает по-разному, но особенно сильной его волна оказалась, например, в Израиле и Франции, охватив значительную долю вакцинированных и свидетельствуя, по существу, о том, что предшествовавшая иммунизация вакцинами против уханьского штамма не может гарантировать высокую защиту от штамма «ο».

Среди используемых против COVID-19 вакцин пептидные и мРНК-вакцины являются самыми простыми по молекулярному составу, а самые сложные – векторные вакцины, в частности, вакцина Спутник V, в которой носителями гена S-белка служат аденовирусы HAd5 и HAd26, способные проявлять себя и как адъюванты. Оба вектора содержат гомологичные последовательности к S-белку коронавируса SARS-CoV-2, что наделяет их потенциалом индуцировать антитела с перекрестной активностью против SARS-CoV-2 [14]. Априорно три разных источника ИЭ в векторной вакцине «Спутник V» должны обеспечивать большую стабильность ее протективного эффекта против разных штаммов SARS-CoV-2. Поскольку вакцина Спутник V является сложной по своим молекулярным компонентам, ее протективный эффект реализуется с привлечением множества механизмов врожденной ИС, активированных ее сенсорами, и сформированным ею натренированным иммунитетом, и он более длительный по сравнению с мРНК-вакцинами. Как показало сравнение штаммов «δ» и «ο», S-белок последнего содержит больше и притом разных последовательностей, гомологичных белкам аденовирусов HAd5 и HAd26. Даже при рекордном числе мутаций у штамма «ο» подавляющая часть его ИЭ в S-белке осталась идентичной ИЭ уханьского штамма (см. табл. 3), поэтому вакцина «Спутник V» будет обладать достаточной (хотя и пониженной) эффективностью как против штамма «δ», так и против штамма «ο», и она по-прежнему может быть рекомендована для иммунизации против COVID-19. Критическим фактором вакцинации является длительность ее эффекта. Что касается пептидных вакцин, то при сильной ограниченности их по составу ИЭ эффективность их против разных вариантов SARS-CoV-2 будет сильно варьировать в зависимости от того, приходится ли мутации в S-белке на те сайты, которые представлены в самой вакцине.

Другой возникающий вопрос: будет ли штамм «ο» последним в череде волн пандемии COVID-19 и завершится ли она в 2022 г. с трансформацией SARS-CoV-2 в сезонный штамм по аналогии с вирусами гриппа? Если сравнивать потенциал

изменчивости вирусов гриппа и коронавирусов, то у последних он несравненно выше и не исчерпан, прежде всего, из-за большего размера генома и S-белка, не ограниченного в размерах и способах мутирования (транзиции, трансверсии, делеции и вставки), отсутствия сезонной привязанности. У вирусов же гриппа отмечается сезонное активирование, ограничение в размерах гемагглютинина и консервативность его остова (у H1 гемагглютинина он сохраняется на протяжении, по крайней мере, уже более ста лет, и существуют строгие ограничения в его генетическом коде).

Если сопоставить мутации S-белка штаммов SARS-CoV-2, возникших на протяжении двух лет пандемии, то нетрудно заметить отсутствие какой-либо тенденции или ограничений по их числу и составу. У штамма «ε» (B.1.429) их всего 6, у штамма C.1.2 их 20, а рекордсменом предстает штамм «ο» – более 30 мутаций, причем по составу у последнего они отличаются явным преобладанием мутирования трансверсиями в 5 позициях аспарагина в лизин (обретение вирусом новых положительно заряженных аминокислот в S-белке (лизина и аргинина) – верный маркер возрастания его контагиозности), а у C.1.2 чаще всего (в 6 позициях) мутировал пролин. Вовлечение в мутационный процесс множества сайтов в S-белке не только изменяет его иммуногенность, но и позволяет разнообразить круг его рецепторов в клетке (и тем самым увеличить вероятность заражения, т.е. контагиозность), взаимодействуя с которыми, вирус инфицирует организм.

Пандемия COVID-19 расширила круг хозяев коронавирусов, включив в него и человека. Присущая штаммам C.1.2 и омикрону столь высокая частота повторяемости мутирования в S-белке одних и тех же аминокислот (например, пролина и аспарагина), не входящих в первую десятку наиболее часто встречающихся в нем аминокислот и определяющих вторичную структуру белков (обе аминокислоты локализируются в местах изгиба белковой молекулы), естественно ставят вопрос об ее происхождении. Многократная повторяемость мутаций, затрагивающих одни и те же аминокислоты (не только аспарагин или пролин), склоняет к мысли, что они возникли не в результате последовательного переселения коронавируса от одного носителя к другому, а, вероятнее всего, из-за длительного пребывания штамма в одном носителе – одна и та же среда обитания с однотипным мутагенным воздействием, т.е. при длительном пребывании коронавируса в организме носителя могут последовательно рождаться, как и при ВИЧ инфекции, новые мутанты. Примечательно, что первое описание омикрона связано с выделением его именно у пациента с иммунным дефицитом, обеспечившим вирусу длительное беспрепятственное воспроизведение потомства. Организм при ослабленном ВИЧ-инфекцией иммунитете, по сообщениям ученых из ЮАР, предрасполагает

к многоступенчатым мутациям в геноме коронавируса. Согласно опубликованному ими отчету, из-за перерыва в лечении ВИЧ-инфекции у заразившейся COVID-19 женщины коронавирус «δ» на протяжении 9 месяцев претерпел 21 мутацию (<https://nauka.tass.ru/>). Не исключено возникновение новых штаммов коронавирусов при затяжном течении COVID-19 и при длительном бессимптомном вирусоносительстве. С учетом множества источников возникновения мутантов SARS-CoV-2 ожидания по блокированию моноштаммспецифичной вакциной пандемии COVID-19 становятся призрачными, и как в случае пандемии ВИЧ-инфекции, вакцинацию в борьбе против COVID-19, вероятно, потеснит фармакотерапия.

Пестрая мозаика мутаций по количеству и составу в S-белке, наряду с его крупными размерами (у большого корабля – большое плавание!), не дают оснований полагать, что штаммом «ο» иссякает потенциал изменения SARS-CoV-2 и с его распространением пандемия завершится. Уже выявленный новый подтип омикрона штамм BA.2 содержит в S-белке меньше мутаций и отличается их составом, но обладает большей контагиозностью. Поэтому более вероятными представляются ожидания увядания пандемии COVID-19 на протяжении последующих двух лет, т.е. когда большая часть человечества «успеет» переболеть COVID-19 с варьирующей степенью тяжести, если не будут найдены препараты, специфически блокирующие размножение коронавирусов, или новые иммуномодуляторы, памятуя, что человек не создан для бесконечной ревакцинации против одного и того же патогена через каждые 4–6 месяцев. А биоинформационным ориентиром для прогнозирования свойств (контагиозности (трансмиссивности), патогенности и иммуногенности) новых штаммов SARS-CoV-2 могли бы послужить особенности первичной структуры S-белка [13] и состава его ИЭ. Используя их, штамм «ο» можно охарактеризовать, как обладающий высокой трансмиссивностью, умеренной патогенностью и измененной иммуногенностью. Особенности аминокислотного, ди- и трипептидного составов S-белка SARS-CoV-2, как и его мутантов, не свидетельствуют в пользу предположения, что прививка способна обеспечить долговременный иммунитет [13].

Другое подтверждение предположения о маловероятности формирования долговременного иммунитета вследствие иммунизации вакцинами на основе S-белка исходит из сравнения особенностей аминокислотного и дипептидного состава белков субъединичных вакцин, которые разработаны, как и вакцины против SARS-CoV-2, на основе только одного белка вируса (вакцины против гепатитов В и Е и папилломавируса человека) и обеспечивают долговременную защиту. Сопоставление именно аминокислотного и дипептидного состава белков этих вакцин обусловлено тем, что контакты в МНС-ИЭ-Т-клеточный рецептор,

как было отмечено ранее, реализуются через короткие фрагменты длиной в 1–3 аминокислоты, и в них возможно доминирование некоторых аминокислот. Представленные в таблице 4 данные подтверждают это предположение: среди доминирующих аминокислот в белках, использованных в вакцинах против гепатита В и Е и папилломавируса человека, выступает пролин. Высокой частоте его встречаемости сопутствует также высокая частота аминокислот, составляющих основу внутренне дезорганизованных областей (intrinsically disordered regions) в белках: глицина, аспарагина, глутамина, треонина и серина. Дополняющим биомаркером субъединичных вакцин, индуцирующих долговременный иммунитет, могут служить особенности дипептидного состава использующихся в них вирусных белков – доминирование дипептидов РХ и ХР (Х – любая аминокислота).

В таблицах 5 и 6 приведены составы дипептидов соответственно для S-белка SARS-CoV-2 и суммарного белка, охватывающего 4 белка (поверхностные белки вируса гепатита В и папилломавируса человека и иммуногенный белок вируса гепатита Е) и обозначенного как Σ -белок. Их различия особенно очевидны по составу отсутствующих и наиболее часто встречающихся дипептидов, особенно пролин-содержащих. Эту общую особенность аминокислотного и дипептидного состава, характерную для столь разных белков у разных по своей природе вирусов, вряд ли можно счесть за случайность. Она свойственна также поверхностному белку Е1 вируса краснухи, разные виды вакцин против которой вызывают длительный иммунитет. Вирус краснухи, как и вирусы кори, паротита, полиомиелита и бешенства, в числе тех, что вызывают, как и разработанные на их основе цельновирионные вакцины, длительный иммунитет. Общая особенность их поверхностных белков – повышенная частота встречаемости дипептидов РХ и ХР [13, табл. 2 и 3]. Повторяемость у разных вирусов доминирования в первичной структуре вирусных белков пролина с аминокислотами, составляющими основу внутренне дезорганизованных областей, служит основанием рассматривать их кандидатами в биомаркеры потенциала вирусного белка

формировать долговременную иммунную память. Дополнительной аргументацией этому может служить персистенция на протяжении, по крайней мере, 2 лет высоких титров антител к гликопротеину вируса Эбола после однократной инъекции вакцины rVSV-ZEBOV[22]. Отличительная особенность структурных белков разных вариантов вируса Эбола (гликопротеина, нуклеопротеина и VP40) – высокое и доминирующее содержание в них пролина. Но следует сразу оговорить возможность существования и других (еще не распознанных, например, у прокариот) биомаркеров долговременного иммунитета как проявление функциональной вырожденности. В широком аспекте функциональная вырожденность определяется как способность элементов, которые структурно различны, выполнять одинаковые функции либо определять одно и то же состояние. Функциональная вырожденность пронизывает все уровни биологической организации и составляет основу функционирования и врожденной, и адаптивной ИС.

Как известно, пролин избегает вторичные структуры в белках и сосредоточен преимущественно в их внутренне дезорганизованных областях, что предполагает линейную, а не пространственную структуру ИЭ, индуцирующих долговременную иммунную память. С ИЭ, индуцирующими долговременную иммунную память, по-видимому, связано и иммунодоминирование среди ИЭ.

По общей статистике белков, по рангу частоты встречаемости пролин среди 20 аминокислот занимает примерно 12-ю позицию, а в S-белке SARS-CoV-2 – даже 13-ю позицию (см. табл. 2), что следует рассматривать как аргумент против ожидания долговременного иммунитета к SARS-CoV-2 при использовании вакцин на основе его S-белка с соответствующим прогнозом: для поддержания иммунитета против SARS-CoV-2 необходимы повторные ревакцинации. Среди других белков SARS-CoV-2 лишь в белке N в ряду частот встречаемости аминокислот пролин занимает более высокую (8-ю) позицию.

Если внутренне дезорганизованные области в белках являются вероятными кандидатами в ИЭ,

Таблица 4. Аминокислотный состав белков субъединичных вакцин
Table 4. Amino acid composition of subunit vaccine proteins

Возбудитель The pathogen	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T	(L)
HBV	3	14	6	11	3	52	14	46	17	14	21	6	18	26	34	9	17	17	41	31	(400 а.к.)
HEV	0	8	3	3	0	21	8	13	3	10	11	0	0	4	9	2	1	3	12	3	(114 а.к.)
HPV16	34	19	11	28	21	37	13	44	26	36	30	25	7	27	35	11	30	21	33	43	(531 а.к.)
HPV18	27	28	16	35	14	47	16	51	25	43	31	29	8	23	35	11	22	26	44	37	(568 а.к.)
Rubella	9	24	17	16	21	47	24	34	11	40	55	16	15	13	46	3	12	17	22	39	(481 а.к.)

Примечание: HBV – вирус гепатита В, HEV – вирус гепатита Е, иммуногенный белок; HPV16 и HPV18 – папилломавирус человека типов 16 и 18, главный капсидный белок L1; Rubella – вирус краснухи, гликопротеин Е1, (L) – длина белка.

Note: HBV – Hepatitis B virus, HEV – Hepatitis E virus, small immunogenic protein; HPV16 и HPV18 Human papillomavirus type 16 and 18 major capsid protein L1 Rubella virus, E1 envelope glycoprotein; (L) protein length.

Problem-Solving Article

Таблица 5. Частота встречаемости дипептидов в Σ -белке
Table 5. Occurrence of dipeptides in Σ -protein

	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
K	3	5	1	3	4	9	1	4	2	3	3	3	0	9	4	1	2	4	2	0
R	8	5	5	2	3	7	0	5	2	9	4	0	1	5	3	0	1	3	2	3
H	1	1	0	0	0	5	0	3	2	2	3	2	3	1	2	0	3	2	5	1
D	1	0	3	4	1	8	2	8	3	4	3	6	3	3	3	2	3	2	5	13
E	2	0	2	6	2	1	1	2	1	4	2	4	1	0	0	0	4	2	2	2
P	4	6	2	11	2	20	3	23	8	6	13	7	0	3	9	1	6	5	15	13
C	5	1	0	0	1	7	5	7	3	2	5	1	2	1	1	0	0	3	1	6
L	6	8	3	12	6	15	7	16	5	11	3	7	4	9	13	0	2	11	8	7
I	4	0	3	0	0	5	7	13	4	1	1	3	1	11	5	0	1	2	5	4
V	0	6	1	9	3	13	2	4	2	6	4	6	3	4	10	2	6	4	9	9
A	4	5	0	4	1	9	4	3	5	6	9	2	0	3	12	3	6	5	7	5
Y	3	5	5	4	1	3	0	5	8	2	1	1	3	4	4	1	4	2	2	2
W	0	1	1	1	2	1	0	5	0	1	5	1	1	1	3	2	5	0	2	1
F	2	3	1	0	0	8	4	11	6	7	3	4	3	6	7	0	4	3	5	3
G	2	5	6	8	4	7	2	12	3	9	8	2	2	6	8	3	4	4	9	9
M	0	0	0	3	1	1	2	1	0	5	1	0	1	3	4	1	2	2	5	1
N	11	4	1	1	0	9	1	4	2	8	2	1	0	4	2	2	5	3	5	5
Q	2	3	0	3	0	5	0	11	1	2	8	5	2	3	7	4	1	0	8	2
S	3	7	2	4	5	16	4	10	8	5	9	1	3	1	10	6	4	5	12	15
T	3	4	0	2	2	8	6	7	6	10	6	4	0	3	6	1	7	5	21	13

Таблица 6. Частота встречаемости дипептидов в S-белке SARS-CoV-2 (Wuhan-Hu-1)
Table 6. Occurrence of dipeptides in SARS-CoV-2 (Wuhan-Hu-1) S-protein

	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T
K	2	4	1	1	2	2	3	5	3	5	1	3	1	3	4	1	8	4	5	3
R	2	1	0	5	2	0	0	5	1	5	5	0	0	5	3	0	1	1	4	2
H	1	1	0	1	0	0	0	1	0	3	3	1	1	1	1	0	0	0	1	2
D	4	2	0	2	3	5	3	9	7	4	3	2	0	3	3	0	2	1	8	1
E	3	1	1	2	1	3	3	4	5	6	1	1	0	3	3	1	3	2	4	1
P	1	3	1	4	1	3	2	6	4	4	4	1	1	8	3	0	1	4	2	5
C	1	0	1	2	1	1	4	3	0	5	3	2	0	1	4	1	2	0	4	5
L	5	1	5	7	3	11	3	10	8	6	4	4	0	3	8	2	6	9	5	8
I	4	4	2	4	1	3	4	4	3	4	12	4	1	1	6	1	3	2	5	8
V	4	4	0	5	4	2	2	12	5	7	6	8	0	8	3	1	8	2	8	8
A	1	2	1	6	4	4	0	8	7	4	6	5	2	0	8	1	3	5	8	4
Y	2	3	2	1	4	1	0	3	3	4	2	3	0	4	4	0	4	5	5	4
W	0	1	0	0	0	1	0	1	1	0	0	1	0	2	0	1	2	0	0	2
F	6	4	1	2	3	4	3	7	3	6	4	1	0	1	7	0	10	5	4	6
G	7	2	0	6	1	1	3	3	7	12	7	4	3	6	3	0	3	2	4	8
M	0	0	0	1	1	0	0	1	1	1	1	1	0	1	0	0	0	1	3	2
N	4	2	1	4	3	1	5	10	5	5	4	3	0	9	8	0	7	4	8	5
Q	3	1	0	5	3	6	1	5	6	2	2	4	0	3	2	1	2	2	6	8
S	6	1	0	0	4	3	4	5	3	10	8	4	1	10	5	1	13	4	8	9
T	5	5	1	4	7	7	0	6	4	4	3	2	2	5	7	2	10	9	7	6

необязательно вызывающими формирование к ним долговременной памяти, то потенциальным фактором, препятствующим формированию долговременной иммунной памяти к ним, может быть присутствие в их последовательностях иммуносупрессивных доменов. Наделен ими и белок S коронавируса

SARS-CoV-2, на основе которого разработаны вакцины против COVID-19. В качестве примера приведем последовательности только двух обнаруженных нами в белке S кандидатов в иммуносупрессивные домены: 815RSFIEDLLFNK825 и 237RFQTLALHR246. Они гомологичны иммуносупрессивным доменам,

Таблица 7. Аминокислотный состав белков SARS-CoV-2
Table 7. Amino acid composition of SARS-CoV-2 proteins

Белки	K	R	H	D	E	P	C	L	I	V	A	Y	W	F	G	M	N	Q	S	T	(L)
S	61	42	17	62	48	58	40	108	76	97	79	54	12	77	82	14	88	62	99	97	(1273ак)
N	31	29	4	24	12	28	0	27	14	8	37	11	5	13	43	7	22	35	37	32	(419ак)
E	2	3	0	1	2	2	3	14	3	13	4	4	0	5	1	1	5	0	8	4	(75ак)
M	7	14	5	6	7	5	4	35	20	12	19	9	7	11	14	4	11	4	15	13	(222ак)
ORF3a	11	6	8	13	11	12	7	30	21	25	13	17	6	14	14	4	8	9	22	24	(275ак)
ORF6	4	1	1	4	5	1	0	8	10	3	1	2	1	3	0	3	4	3	4	3	(61ак)
ORF7a	7	5	3	2	8	6	6	15	8	8	9	5	0	10	4	1	2	5	7	10	(121ак)
ORF8	5	4	4	7	6	7	7	10	10	12	5	7	1	8	5	1	2	6	9	5	(121ак)
ORF10	0	2	0	1	0	1	1	4	3	4	2	3	0	4	1	2	5	1	2	2	(38ак)

встречающимся в ретровирусах животных и человека, в синцитинах человеческой плаценты и в белках высокопатогенных вирусов (вирусы гриппа, Эбола, Ласса, Марбурга и др.). К сожалению, и на сегодняшний день неясны механизмы действия иммуносупрессивных доменов в составе белков, не установлено, как они действуют на компоненты ИС. Сами структуры иммуносупрессивных доменов могут заметно дивергировать, и для их функционализации важны особенности фланкирующих их последовательностей. Единство же гомологии иммуносупрессивных доменов разного происхождения проявляется в наличии гидрофобного ядра и фланкирования их с обоих концов остатками аргинина [23]. С учетом сильной вырожденности последовательностей иммуносупрессивных доменов их представленность в белке S, вероятно, не исчерпывается приведенными выше кандидатами в иммуносупрессивные домены.

В аспекте выявленного в белках SARS-CoV-2 множества ИЭ, рестриктированных по МНС II, возможно следующее объяснение феномена гибридного иммунитета, проявившегося широко в связи с развернутой кампанией иммунизации населения, включая и переболевших ранее COVID-19. Оказалось, что у последних после иммунизации обнаруживали более сильный гуморальный иммунный ответ (обеспечивавший защиту от новых штаммов SARS-CoV-2), чем те, которые не были ранее инфицированы SARS-CoV-2 и были иммунизированы [15]. Известно, что формирование долговременной иммунной памяти и антител с широким профилем специфичности является длительным процессом и связан с возникновением мутаций в геноме В-клеток в герминальных центрах с образованием генов иммуноглобулинов к разным ИЭ. Наличие их множества в вирусных белках необходимо для выработки разных клонов В-клеток памяти с формированием генов, кодирующих иммуноглобулины с аномально длинными переменными областями Н-цепей. Такие иммуноглобулины характерны при ВИЧ-инфекции [16–18], выявляются они и при перенесенной гриппозной инфекции [19]. Важность роли временного

фактора в проявлении гибридного иммунитета против SARS-CoV-2 (чем позднее проведена вакцинация и больше времени между введениями первой и второй дозы, тем сильнее эффект) подчеркивают его сходство с феноменами образования антител с широкой нейтрализующей способностью при ВИЧ-инфекции и/или повторном гриппе. Примечательно, что возникают такие широко нейтрализующие антитела через месяцы и годы после инфекции и в контексте повторной встречи с компонентами инфекционного агента. Большая реакция гуморального иммунитета перенесшего ранее COVID-19 на иммунизацию двумя дозами вакцины против SARS-CoV-2, по существу, является результатом трехкратной (и длительной по времени) встречей переболевшего с SARS-CoV-2, и в основе проявления ее как феномена гибридного иммунитета, по-видимому, заложен тот же механизм, что и при образовании антител с широкой нейтрализующей способностью при других инфекциях.

Оборотная сторона наличия множества ИЭ в вирусных белках может проявляться в нескольких аспектах. Во-первых, ими обеспечивается образование широкого спектра антител к вирусу: как протективных, так и нейтральных. Последние не обладают способностью блокировать сам патоген, но связываются с ним, вызывая антителозависимое усиление инфицирования, фагоцитоза клеток и связывания комплемента и спасая вирус от полного уничтожения. Ускользнувший от ИС вирус порождает новое потомство мутантов, из которых селекционируются наиболее жизнеспособные. Не удивителен в этом аспекте сценарий развития на протяжении двух лет пандемии COVID-19 с ее чередой волн заражения возникающими новыми штаммами SARS-CoV-2 с меняющимся пандемическим потенциалом (по уровням заболеваемости и летальности) и калейдоскопической картиной распространения COVID-19 уже на фоне все большего охвата населения в мире вакцинацией. Во-вторых, при обладании коронавирусами самым крупным поверхностным белком и высокой плотностью в нем ИЭ инфекция ими, как и вакцинации против них (особенно многократные и частые)

сопряжены с перегрузкой ИС с различными последствиями.

Возможности ИС не беспредельны, и поражение новыми штаммами SARS-CoV-2, и многократная вакцинация против них способны привести к ее супрессии со многими осложнениями. Истощение пула наивных лимфоцитов (особенно у пожилых лиц) лишает ИС надзорной функции, вследствие чего нарушается элиминация из организма патологически измененных клеток и возрастает частота онкогенных поражений, а в итоге – увеличение смертности населения. Резистентность огромного контингента населения к вирусу может проявляться в бессимптомном его носительстве или бессимптомной инфекции благодаря вирусостатическому эффекту, реализуемому врожденной ИС

и/или резидентными Т-клетками памяти адаптивной ИС, т.е. при обследовании этой части населения антител к SARS-CoV-2 не выявить. Будет ли многократная вакцинация в этом случае лучшей защитой от вируса? Окажется ли благом посягательство на естественный иммунитет?

Подводя итог рассмотрения возможностей развиваемого метода прогнозировать иммуногенность белков, хотелось бы заметить, что в таблице 3 представлено также распределение Т-клеточных ИЭ в других белках, в частности, в гистоне H4, участвующем в формировании нуклеосом в хроматине, являющемся одним из наиболее древних и консервативных (от дрожжей до человека) белков в эволюции и обладающем низкой иммуногенностью. Для него, как и ожидалось, характерна низкая плотность и малая протяженность ИЭ. Специфическая картина распределения Т-клеточных ИЭ характерна для инсулина, проявляясь наличием протяженных блоков неИЭ. Богата Т-клеточными ИЭ первичная структура тиреопероксидазы человека, к которой часто выявляются антитела при аутоиммунном поражении щитовидной железы. Оба последних белка нередко являются мишенями возникающих в организме аутоантител. Обилие ИЭ, рестриктированных по МНС II, отмечается и для N-белка SARS-CoV-2, гемагглютинаина вируса гриппа H1N1, поверхностных белков E1 и E2 вируса гепатита С, для гемагглютинаина и белка слияния вируса кори и других вирусов. В совокупности же приведенные примеры подтверждают информативность разработанного метода распознавания ИЭ.

Сфера применения вакцин в последнее десятилетие расширилась внедрением их для лечения онкологических заболеваний, которые характеризуются наличием специфических опухолевых антигенов (неоантигенов). Последние выбираются как мишени при иммунотерапии с использованием индивидуальных раковых вакцин [20]. Получение индивидуальной раковой вакцины, как и других вакцин, – процесс многостадийный, и первый из них – анализ белков опухоли *in silico* на присутствие в них раковых антигенов. С разработкой

передовых методов секвенирования обнаружение мутаций в белках опухоли уже не представляет трудностей. Проблема – идентификация в белке областей, соответствующих ИЭ, миновавших селекцию на уровне центральных механизмов толерантности и способных индуцировать иммунный ответ. Новый метод может быть использован для обнаружения в белках опухоли неоантигенов, как узнаваемых МНС, так и видимых Т-клетками.

Полезен он и для предсказания потенциальной иммуногенности белковых препаратов человеческого происхождения, используемых широко в терапии при различных патологиях. В числе их гормоны, интерфероны, моноклональные антитела, факторы роста, очищенные фракции крови и др. При длительном их использовании, несмотря на человеческую природу их происхождения, у пациентов к ним возникает иммунная реакция, зависящая от Т-клеток и блокирующая их биоактивность. Экспериментально показано, что устранение иммунной реакции возможно путем деиммунизации и индукции толерантности. Первая достигается, например, путем производства белкового лекарственного препарата, не содержащего последовательности Т-клеточного ИЭ. Толерантность же достигается путем введения белкового лекарственного препарата вместе с ИЭ Т-регуляторных клеток, которые реализуют супрессию иммунной реакции [21].

Наконец, нельзя обойти вниманием возможность практического приложения разработанного метода при аутоиммунных нарушениях. Часто они возникают из-за не единичного, а множественного нарушения толерантности, затрагивающего разные белки пораженного органа. Секвенирование протеома на уровне одной клетки расширяет возможности идентификации в белках новых ИЭ и позволяет более полно постигнуть патогенез развивающегося аутоиммунного поражения органа.

В завершение хотелось бы подчеркнуть, что пробелы в наших знаниях о механизмах функционирования ИС и сегодня, как во времена Пастера, но на другом научном и технологическом уровне, низводят поиски вакцин до ловли удачи.

Секвенирован геном человека, расшифровка генома любого вируса сегодня не составляет проблему, разработано множество вариантов платформ для вакцин, но это не приблизило нас к полному пониманию, как формируется эффективный специфичный иммунный ответ. За короткий срок на протяжении пандемии COVID-19 удалось создать несколько вакцин, не способных, однако, противостоять разным штаммам SARS-CoV-2, обеспечивая долговременный протективный эффект. Пятая волна пандемии COVID-19, порождаемая штаммом «о», опрокинула все надежды на формирование коллективного иммунитета в странах с предшествующим высоким уровнем иммунизации моноштаммспецифичной вакциной. Амплитуды третьей, четвертой и пятой волн заражений коррелирует с контагиозностью (уханьский < дельта

< омикрон) вызвавших их штаммов. Особенность пятой волны COVID-19 проявится во всем мире ее наибольшим масштабом, по сравнению с предыдущими волнами, обусловленным высокой контагиозностью самого штамма «О», срабатыванием иммунного импринтинга («первородного греха») у ранее переболевших и вакцинированных и зимним сезоном.

В России распространение волн заражения COVID-19 имеет свои особенности. Спад третьей волны заражений обозначился в самом начале развертывания вакцинации. Четвертая волна пошла на спад, когда вакцинацией не было охвачено даже половины населения страны, и ее большая амплитуда объяснима большей контагиозностью циркулировавших штаммов и благоприятствующим им осенне–зимним сезоном. Следовательно, трудно полагать, что уровень вакцинации населения является определяющим фактором в распространении COVID-19. Это утверждение особенно справедливо для третьей волны заражений.

Препятствием для создания эффективных вакцин против COVID-19 на основе белка S коронавируса SARS-CoV-2, обеспечивающих долговременный иммунитет, служат особенности самого белка S: его крупные размеры, содержание в нем большого числа последовательностей, гомологичных белкам человека, отсутствие маркеров долговременного иммунитета, наличие иммуносупрессивных доменов, множество механизмов его мутирования и быстрота возникновения мутаций, быстрое выведение из циркуляции антител к нему. Примечательно, что индуцирующие лишь непродолжительный иммунитет белки S сезонных коронавирусов человека HCoV-NL63, HCoV-229E, HCoV-NKU1 и HCoV-OC43, принадлежащих разным линиям подсемейства *Coronavirinae*, и белок S пандемического коронавируса SARS-CoV-2 при наличии существенных различий их первичных структур имеют плотности распределения

пролина ≤ 46 остатков пролина на 1000 аминокислот, т.е. они сходны по отсутствию в них биомаркеров долговременного иммунитета. Совокупность этих особенностей не внушает оптимизма в отношении окончания пандемии в 2022 г. и расширения поисков вакцин против COVID-19, нацеленных на использование новых векторов, или полиштаммспецифичных вакцин на основе белка S. Не оказались спасительными мРНК-вакцины, перспективными представляются попытки создать комбинированную вакцину против вирусов гриппа и коронавирусов. Для преодоления пандемии COVID-19 (памятуя и о возможном новом «явлении MERS-CoV-2 народу») необходимы вакцины с широкой специфичностью и длительным защитным эффектом, что предполагает развитие новых подходов в вакцинологии и пересмотр границ ее возможностей исходя из более глубоких знаний об ИС.

Трудности в предсказании иммуногенных свойств белков инфекционных агентов еще не преодолены из-за сложности механизмов формирования Т- и В-клеточной иммунной памяти (двух ее путей: независимого и зависимого от герминальных центров), соучастия в них функционально и транскрипционно различных клеточных популяций и разных информационных потоков с обязательной стадией образования комплексов МНС-ИЭ-Т-клеточный рецептор, и новый метод распознавания ИЭ и выявленные биомаркеры индуцирования долговременной иммунной памяти могут быть использованы как иммуноинформативные инструменты для прогнозирования свойств вакцин (против разных инфекционных патогенов) на стадии их дизайна. Они оказались эффективными и для распознавания в белках кандидатов В-клеточных ИЭ. С привлечением данных по составу рецепторов Т- и В-клеток и репертуару антител они могут способствовать раскрытию новых феноменов ИС, отражающих ее многоуровневую сложность.

Литература

- Joglekar AV, Li G. T cell antigen discovery. *Nature Methods*. 2020. Doi:10.1038/s41592-020-0867-z
- Peters B, Nielsen M, Sette A. T Cell Epitope Predictions. *Annu. Rev. Immunol.* 2020. 38:123–45. Doi 10.1146/annurev-immunol-082119-124838
- Loan Ping Eng, Tin Wee Tan, Joo Chuan Tong, Söllner J. Building MHC Class II Epitope Predictor Using Machine Learning Approaches. In: Peng Zhou and Jian Huang (eds.), *Computational Peptidology, Methods in Molecular Biology*, vol. 1268, DOI 10.1007/978-1-4939-2285-7_4.
- Söllner J. Computational Peptide Vaccinology. In: Peng Zhou and Jian Huang (eds.), *Computational Peptidology, Methods in Molecular Biology*, vol. 1268, DOI: 10.1007/978-1-4939-2285-7_13
- Пригожин И. Р., Стенгерс И. Порядок из хаоса. М.: Прогресс 1986:432.
- Харченко Е. П. Иммуноэпитопный континуум родства белков и полуреактивность и аутореактивность антител. *Медицинская иммунология*. 2015. Т.17, № 4. С. 335–346.
- Харченко Е. П., Дитятев А. Э. Предсказание пептидных антигенных структур, узнаваемых основным комплексом гистосовместимости. Докл. АН СССР 1990. Т.318(4):1013–1016.
- Харченко Е. П. Распознавание эпитопов, связывающихся с разными DR-типами главного комплекса гистосовместимости класса II. *Иммунология*. 1995. Т.16(6):21–23.
- De Groot AS, Moise L, Terry F, Gutierrez AH, et al. Better epitope discovery, precision immune engineering, and accelerated vaccine design using immunoinformatics tools. *Front. Immunol.* 2020 11:442. doi: 10.3389/fimmu.2020.00442
- Meyers LM, Gutiérrez AH, Boyle CM, et al. Highly conserved, non-human-like, and cross-reactive SARS-CoV-2 T cell epitopes for COVID-19 vaccine. design and validation. *npj Vaccines*. 2021. Vol. 6:71. doi 10.1038/s41541-021-00331-6
- Grifoni A, Sidney J., Zhang Y, et al. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2 Cell Host & Microbe. 2020. Vol.27, P671–680, doi 10.1016/j.chom.2020.03.002
- Halperin G, Shoenfeld Y. SARS-CoV-2, the autoimmune virus. *Autoimmunity Reviews*. 2020. Vol. 19. 102695. Doi:10.1016/j.autrev.2020.102695
- Харченко Е. П. Общие особенности коронавирусной пандемии и пандемий гриппа и поверхностных белков их возбудителей. *Параллели. Эпидемиология и Вакцинопрофилактика*. 2021;20(4):4–18. doi:10.31631/2073-3046-2021-20-4-4-18
- Харченко Е. П. Вакцины против COVID-19: сравнительная оценка рисков аденовирусных векторов. *Эпидемиология и Вакцинопрофилактика*. 2020;19(5):4–17. <https://doi.org/10.31631/2073-3046-2020-19-5-4-17>
- Callaway E. COVID-19 super-immunity: one of the pandemic's great puzzles. *Nature*. 2021. doi.org/10.1038/d41586-021-02795-x.

16. Burton DR. Advancing an HIV vaccine advancing vaccinology. *Nat Rev Immunol.* 2019;19(2):77–78. doi: 10.1038/s41577-018-0103-6
17. Sok D, Burton DR. Recent progress in broadly neutralizing antibodies to HIV. *Nat Immunol.* 2018;19(11):1179–1188. doi: 10.1038/s41590-018-0235-7.
18. Andrabi R, Bhiman JN, Burton DR. Strategies for a multi-stage neutralizing antibody-based HIV vaccine. *Curr Opin Immunol.* 2018;53:143–151. doi: 10.1016/j.coi.2018.04.025
19. Bajic G, van der Poel CE, Kuraoka M, et al. Autoreactivity profiles of influenza hemagglutinin broadly neutralizing antibodies. *Sci Rep.* 2019;9(1):3492. doi: 10.1038/s41598-019-40175-8
20. Blass E, Ott PA. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. *Nat Rev Clin Oncol.* 2021 Apr;18(4):215–229. doi:10.1038/s41571-020-00460-2
21. Jawa V, Terry F, Gokemeijer J, et al. T-Cell Dependent Immunogenicity of Protein Therapeutics Pre-clinical Assessment and Mitigation—Updated Consensus and Review 2020. *Front. Immunol.* 11:1301. doi: 10.3389/fimmu.2020.01301
22. Huttner A, Siegrist C.A. Durability of single-dose rVSV-ZEBOV vaccine responses: what do we know? *Expert Rev Vaccines.* 2018, 17:1105-1110. doi:10.1080/14760584.2018.1546582.
23. Киселев О. И. Беременность, иммуносупрессия, грипп и плацентарная экспрессия эндогенных ретровирусов. Санкт-Петербург. Изд-во Росток. 2014:316.

Reference

1. Joglekar AV, Li G. T cell antigen discovery. *Nature Methods.* 2020. Doi:10.1038/s41592-020-0867-z
2. Peters B, Nielsen M, Sette A. T Cell Epitope Predictions. *Annu. Rev. Immunol.* 2020. 38:123–45. Doi 10.1146/annurev-immunol-082119-124838
3. Loan Ping Eng, Tin Wee Tan, Joo Chuan Tong, Söllner J. Building MHC Class II Epitope Predictor Using Machine Learning Approaches. In: Peng Zhou and Jian Huang (eds.), *Computational Peptidology, Methods in Molecular Biology*, vol. 1268, DOI 10.1007/978-1-4939-2285-7_4,
4. Söllner J. Computational Peptide Vaccinology. In: Peng Zhou and Jian Huang (eds.), *Computational Peptidology, Methods in Molecular Biology*, vol. 1268, DOI: 10.1007/978-1-4939-2285-7_13
5. Prigozhin I.R., Stengers I. Order from chaos. *Progress* 1986. s432.
6. Kharchenko EP. Immune epitope continuum of the protein relationships, poly- and autoreactivity of antibodies. *Medical Immunology.* 2015;17(4):335–346 (In Russ.). doi:10.15789/1563-0625-2015-4-335-346.
7. Kharchenko E.P., Dityatev A.E. Prediction of peptide antigenic structures recognized by the main histocompatibility complex. *Dokl. USSR ACADEMY OF SCIENCES* 1990. T.318 N 4 C.1013-1016.
8. Kharchenko E.P. Recognition of epitopes binding to different DR types of the main histocompatibility complex of class II. *Immunology.* 1995. T.16 N 6 S.21-23.
9. De Groot AS, Moise L, Terry F, Gutierrez AH, et al. Better epitope discovery, precision immune engineering, and accelerated vaccine design using immunoinformatics tools. *Front. Immunol.* 2020 11:442. doi: 10.3389/fimmu.2020.00442
10. Meyers LM, Gutiérrez AH, Boyle CM, et al. Highly conserved, non-human-like, and cross-reactive SARS-CoV-2 T cell epitopes for COVID-19 vaccine. design and validation. *npj Vaccines.* 2021. Vol. 6:71. doi 10.1038/s41541-021-00331-6
11. Grifoni A, Sidney J., Zhang Y, et al. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2 Cell Host & Microbe. 2020. Vol.27, P.671–680, doi 10.1016/j.chom.2020.03.002
12. Halperin G, Shoenfeld Y. SARS-CoV-2, the autoimmune virus. *Autoimmunity Reviews.* 2020. Vol. 19. 102695. Doi:10.1016/j.autrev.2020.102695
13. Kharchenko EP. Common Features of Coronavirus and Influenza Pandemics and Surface Proteins of their Pathogens. *Parallels. Epidemiology and Vaccinal Prevention.* 2021;20(4): 4–18 (In Russ.). [https://doi: 10.31631/2073-3046-2021-20-4-4-18](https://doi.org/10.31631/2073-3046-2021-20-4-4-18)
14. Kharchenko EP. Vaccines against Covid-19: the Comparative Estimates of Risks in Adenovirus Vectors. *Epidemiology and Vaccinal Prevention.* 2020; 19 (5): 4–17 (In Russ.). doi: 10.31631/2073-3046-2020-19-5-4-17.
15. Callaway E. COVID-19 super-immunity: one of the pandemic's great puzzles. *Nature.* 2021. doi.org/10.1038/d41586-021-02795-x.
16. Burton DR. Advancing an HIV vaccine advancing vaccinology. *Nat Rev Immunol.* 2019;19(2):77–78. doi: 10.1038/s41577-018-0103-6
17. Sok D, Burton DR. Recent progress in broadly neutralizing antibodies to HIV. *Nat Immunol.* 2018;19(11):1179–1188. doi: 10.1038/s41590-018-0235-7.
18. Andrabi R, Bhiman JN, Burton DR. Strategies for a multi-stage neutralizing antibody-based HIV vaccine. *Curr Opin Immunol.* 2018;53:143–151. doi: 10.1016/j.coi.2018.04.025
19. Bajic G, van der Poel CE, Kuraoka M, et al. Autoreactivity profiles of influenza hemagglutinin broadly neutralizing antibodies. *Sci Rep.* 2019;9(1):3492. doi: 10.1038/s41598-019-40175-8
20. Blass E, Ott PA. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. *Nat Rev Clin Oncol.* 2021 Apr;18(4):215–229. doi:10.1038/s41571-020-00460-2
21. Jawa V, Terry F, Gokemeijer J, et al. T-Cell Dependent Immunogenicity of Protein Therapeutics Pre-clinical Assessment and Mitigation—Updated Consensus and Review 2020. *Front. Immunol.* 11:1301. doi: 10.3389/fimmu.2020.01301
22. Huttner A, Siegrist C.A. Durability of single-dose rVSV-ZEBOV vaccine responses: what do we know? *Expert Rev Vaccines.* 2018, 17:1105-1110. doi:10.1080/14760584.2018.1546582.
23. Kiselev O.I. Pregnancy, immunosuppression, influenza and the placental expression of endogenous retroviruses. 2014. St. Petesburg. Publisher «Rostok»:316 (In Russ.).

Об авторе

- **Евгений Петрович Харченко** – д. б. н., ведущий научный сотрудник Института эволюционной физиологии и биохимии им. И. М. Сеченова, РАН. 194223, Россия, Санкт-Петербург, пр. Тореза, 44. +7 (904) 338-22-80, neuro.children@mail.ru

Поступила 10.01.2022. Принята к печати: 19.02.2022.

Контент доступен под лицензией CC BY 4.0.

About the Author

- **Eugene P. Kharchenko** – Dr. Sci. (Biol.), leader researcher of I. Sechenov Institute of Evolutionary Physiology and Biochemistry, Russian Academy Sciences. 194223, Russian Federation, St. Petersburg, Toreza pr., 44. +7 (904) 338-22-80, neuro.children@mail.ru

Received: 10.01.2022. Accepted: 19.02.2022.

Creative Commons Attribution CC BY 4.0.