


Перевод на английский язык <https://vavilov.elpub.ru/jour>

## Генетический контроль N-гликозилирования белков плазмы крови человека

С.Ж. Шарапов<sup>1</sup>, А.Н. Тимошук<sup>1</sup>, Ю.С. Аульченко<sup>1, 2</sup> 

<sup>1</sup> Институт перспективных исследований проблем искусственного интеллекта и интеллектуальных систем  
Московского государственного университета им. М.В. Ломоносова, Москва, Россия

<sup>2</sup> Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия  
 [yurii@bionet.nsc.ru](mailto:yurii@bionet.nsc.ru)

**Аннотация.** Гликозилирование является важной модификацией белков, которая влияет как на их физико-химические свойства, так и на выполняемые ими биологические функции. Масштабные популяционные исследования показали, что уровни различных N-гликанов белков плазмы крови ассоциированы с риском развития ряда мультифакторных заболеваний человека. Найденные ассоциации стали основанием для рассмотрения N-гликанов в качестве потенциального источника биомаркеров и терапевтических мишеней. Биохимические пути N-гликозилирования хорошо изучены, однако понимание механизмов общей и тканеспецифической регуляции этих биохимических реакций *in vivo* весьма ограничено. Это затрудняет как интерпретацию наблюдаемых ассоциаций уровней N-гликанов с заболеваниями человека, так и разработку биомаркеров и молекулярных мишеней на их основе. Прогресс в области технологий анализа N-гликозилирования белков позволил к началу 2010-х годов проводить исследования регуляции N-гликозилирования с помощью методов генетического анализа, в том числе полногеномного исследования генетических ассоциаций. Применение этих методов дает возможность находить новые, ранее неизвестные регуляторы N-гликозилирования и расширяет представление о роли N-гликанов в контроле мультифакторных заболеваний и комплексных признаков человека. В данном обзоре мы рассматриваем современное состояние исследований генетического контроля популяционной изменчивости уровней N-гликозилирования белков плазмы крови человека. Описаны современные физико-химические методы измерения N-гликомного профиля, приведены базы данных, содержащие гены, вовлеченные в биосинтез N-гликанов. Систематизированы результаты исследований вклада средовых и генетических факторов в популяционную изменчивость N-гликанов, а также результаты картирования геномных локусов N-гликанов методом полногеномного исследования ассоциаций. Представлены результаты последующих функциональных исследований *in vitro* и *in silico*, позволивших предложить новые гены-кандидаты, регулирующие N-гликозилирование белков. В заключение кратко показан текущий прогресс в области гликогеномики человека и описаны возможные пути дальнейших исследований N-гликома.

Ключевые слова: гликом; гликаны; N-гликозилирование; генетика; ПГИА.


**Для цитирования:** Шарапов С.Ж., Тимошук А.Н., Аульченко Ю.С. Генетический контроль N-гликозилирования белков плазмы крови человека. *Вавиловский журнал генетики и селекции*. 2023;27(3):224-239. DOI 10.18699/VJGB-23-29

## Genetic control of N-glycosylation of human blood plasma proteins

S.Zh. Sharapov<sup>1</sup>, A.N. Timoshchuk<sup>1</sup>, Y.S. Aulchenko<sup>1, 2</sup> 

<sup>1</sup> MSU Institute for Artificial Intelligence, Lomonosov Moscow State University, Moscow, Russia

<sup>2</sup> Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

 [yurii@bionet.nsc.ru](mailto:yurii@bionet.nsc.ru)

**Abstract.** Glycosylation is an important protein modification, which influences the physical and chemical properties as well as biological function of these proteins. Large-scale population studies have shown that the levels of various plasma protein N-glycans are associated with many multifactorial human diseases. Observed associations between protein glycosylation levels and human diseases have led to the conclusion that N-glycans can be considered a potential source of biomarkers and therapeutic targets. Although biochemical pathways of glycosylation are well studied, the understanding of the mechanisms underlying general and tissue-specific regulation of these biochemical reactions *in vivo* is limited. This complicates both the interpretation of the observed associations between protein glycosylation levels and human diseases, and the development of glycan-based biomarkers and therapeutics. By the beginning of the 2010s, high-throughput methods of N-glycome profiling had become available, allowing research into the genetic control of N-glycosylation using quantitative genetics methods, including genome-wide association studies (GWAS). Application of these methods has made it possible to find previously unknown regulators of N-glycosylation and expanded the understanding of the role of N-glycans in the control of multifactorial diseases and human complex traits. The present review considers the current knowledge of the genetic control of variability in the levels of N-glycosylation of plasma proteins in human populations. It briefly describes the most popular physical-chemical methods of N-glycome profiling and the databases that contain genes involved in the biosynthesis of N-glycans. It also reviews the results of

studies of environmental and genetic factors contributing to the variability of N-glycans as well as the mapping results of the genomic loci of N-glycans by GWAS. The results of functional *in vitro* and *in silico* studies are described. The review summarizes the current progress in human glycogenomics and suggests possible directions for further research. Key words: glycome; glycans; N-glycosylation; genetics; GWAS.

**For citation:** Sharapov S.Zh., Timoshchuk A.N., Aulchenko Y.S. Genetic control of N-glycosylation of human blood plasma proteins. *Vavilovskii Zhurnal Genetiki i Selektcii = Vavilov Journal of Genetics and Breeding*. 2023;27(3):224-239. DOI 10.18699/VJGB-23-29

## Гликомика – раздел гликобиологии

Гликаны – полисахариды или олигосахариды, полимеры, состоящие из моносахаридных звеньев, соединенных гликозидными связями. Гликаны могут ковалентно присоединяться к белкам и липидам путем создания гликозидной связи, образуя гликопротеины и гликолипиды соответственно. Гликозилирование является одной из самых распространенных (Craveur et al., 2014) пост- и ко-трансляционных модификаций белков. Известно, что в природе 20 % всех видов белков гликозилированы (Khougu et al., 2011). При этом более 40 % (по массе) всех белков плазмы крови человека N-гликозилированы (Clerc et al., 2016). Гликозилирование влияет как на физико-химические свойства белков – растворимость, пространственную конфигурацию, фолдинг и т. п. (Varki, 1993; Ohtsubo, Marth, 2006; Skropeta, 2009), так и на их биологическую функцию. Клетки всех без исключения многоклеточных организмов содержат гликоконъюгаты – гликопротеины и гликолипиды с ковалентно присоединенными гликанами (Varki, Kornfeld, 2017).

Гликопротеины и гликолипиды, расположенные на поверхности клеточных мембран, участвуют в процессах клеточных взаимодействий – между клетками, клетками и внеклеточным матриксом и между клетками и макромолекулами, а также в процессах взаимодействий между организмами (хозяин-паразит, симбионт-симбионт и т. п.) (Ohtsubo, Marth, 2006; Skropeta, 2009; Lauc et al., 2016; Poole et al., 2018), тем самым играя огромную роль в развитии и функционировании многоклеточных организмов (Gagneux et al., 2015–2017).

Множество исследований химической структуры гликанов и их метаболизма были проведены в первой половине XX в. Однако в то время гликаны рассматривались в основном как структурные элементы и источники энергии живых систем. Бурное развитие химических, физических и молекулярно-биологических методов исследования гликанов привело к появлению нового раздела молекулярной биологии – гликобиологии. Эта область включает исследования химических и физических свойств гликанов, энзимологии процессов синтеза и деградации гликанов, их эволюции, механизмов распознавания гликанов белками, роли гликанов в функционировании биологических систем и в развитии заболеваний человека, а также разработку методов лечения, профилактики, диагностики и предсказания заболеваний (Varki et al., 2017). В настоящее время гликобиология – быстро развивающаяся наука, имеющая большое значение для многих смежных областей знания, включая биомедицину и биотехнологии (Nikolac Perkovic et al., 2014; Varki, Kornfeld, 2017).

По аналогии с геномикой, транскриптомикой, протеомикой и метаболомикой, гликомика представляет собой

систематическое изучение гликома – совокупности всех гликанов и их концентраций в конкретном образце – клеточной культуре, ткани, органе или целом организме. Разнообразие возможных гликоконъюгатов поражает воображение. Хотя число мономеров, входящих в структуру гликанов, относительно невелико, мономеры могут образовывать различные гликозидные связи, порождая огромное число возможных гликанов. Разнообразие гликоконъюгатов также увеличивается за счет допустимого наличия в белке не одного, а нескольких сайтов гликозилирования.

В начале XXI в. были разработаны высокопроизводительные физико-химические методы, позволившие проводить масштабные когортные исследования ассоциаций N-гликома с заболеваниями и признаками человека. На сегодняшний день показана ассоциация N-гликанов со множеством мультифакторных заболеваний человека (Gudelj et al., 2018a; Dotz, Wuhler, 2019; Reily et al., 2019), включая диабет 2-го типа и моногенные формы диабета (Thanabalasingham et al., 2013; Keser et al., 2017), ревматоидный артрит (Gudelj et al., 2018b), болезнь Паркинсона (Russell et al., 2017), воспалительное заболевание кишечника (Trbojević Akmačić et al., 2015; Clerc et al., 2018), сердечно-сосудистые (Connelly et al., 2016; Wang et al., 2016) и онкологические (Fuster, Esko, 2005; Saldova et al., 2014; Mehta et al., 2015; Taniguchi, Kizuka, 2015) заболевания.

Результаты обсервационных исследований ассоциации гликозилирования белков с заболеваниями не дают ответа на вопрос о причинно-следственных связях между ними и тем более о молекулярно-биологических механизмах, лежащих в основе данных связей. С 1983 г. проводятся исследования функциональных последствий изменений в гликозилировании белков (Anthony et al., 2012; Cobb, 2020). Гликозилирование иммуноглобулинов класса G (IgG) является наиболее изученным с этой точки зрения ввиду важности данного гликопротеина в адаптивном иммунитете. Молекула IgG имеет консервативный сайт гликозилирования Asn297, находящийся в консервативном домене CN2 тяжелой цепи. Этот домен играет важную роль в связывании с Fcγ-рецепторами, что в свою очередь влияет на эффекторную функцию IgG. Показано, что снижение уровня фукозилирования IgG усиливает антителозависимую клеточную цитотоксичность (Peirp et al., 2008). Дальнейшие исследования с использованием методов кристаллографии показали, что отсутствие фукозилирования IgG вызывает увеличение аффинности между Fc доменом IgG и рецептором FcγR3A, а наличие фукозы в составе гликана IgG ведет к появлению стерических затруднений при взаимодействии (Mizushima et al., 2011).

Изучение процессов гликозилирования является фундаментально важным для решения задач диагностики, пред-

сказания, профилактики и лечения заболеваний человека. В 2012 г. Национальная академия наук США представила доклад о необходимости полномасштабного изучения гликома, так как гликаны напрямую вовлечены в патогенез практически всех известных заболеваний ([http://www.nap.edu/catalog.php?record\\_id=13446](http://www.nap.edu/catalog.php?record_id=13446)).

### Строение и разнообразие гликанов

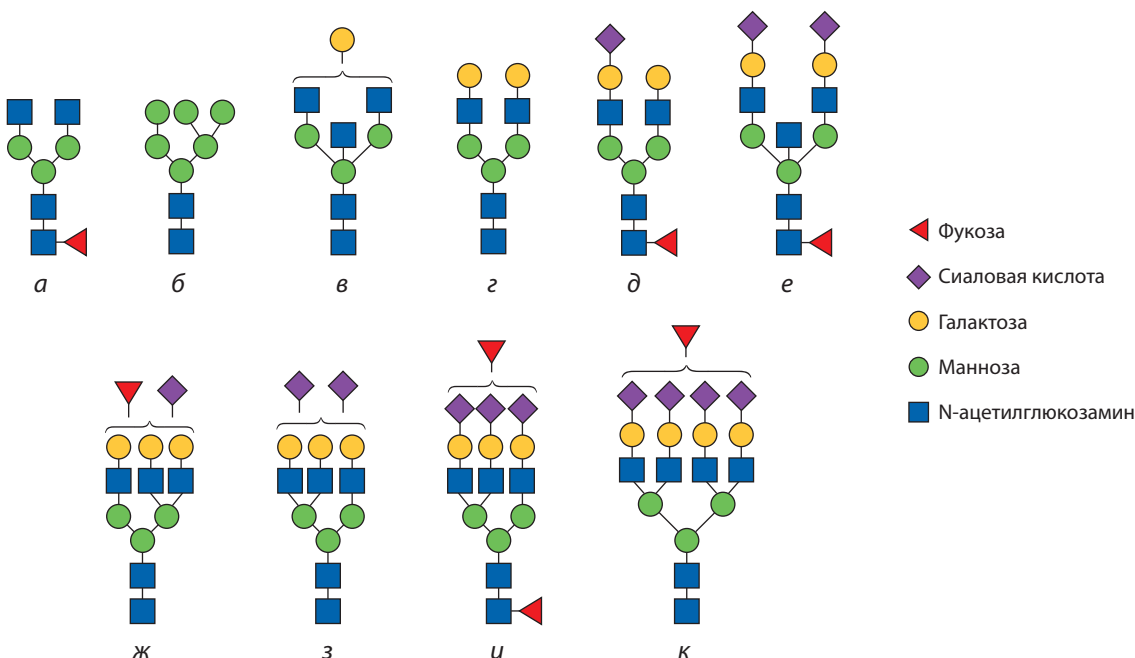
Углеводы – одна из основных групп макромолекул, выделяемых в биологии наряду с белками, липидами и нуклеиновыми кислотами. Благодаря способности к полимеризации и наличию высокого числа хиральных атомов моносахариды могут формировать огромное количество различных стерео- и региоизомеров. По степени полимеризации различают четыре основные группы углеводов: моносахариды (глюкоза, фруктоза, галактоза, т. д.); дисахариды (молекулы, состоящие из двух моносахаридов, связанных гликозидной связью, например, сахароза, лактоза, мальтоза); полисахариды с повторяющимися звеньями, образующие как линейные, так и разветвленные соединения (О-антигены бактерий, амилоза, целлюлоза, хитин), и гликаны – сложные олигосахариды с неповторяющимися звеньями, свободные или связанные с белками или липидами (гликопротеины, протеоглики и гликолипиды).

В свою очередь, гликаны, связанные с белками, подразделяются на N-, O- и C-гликаны. N-гликаны создают гликозидную связь с атомом азота (N), находящимся в составе аминокислоты аспарагина; O-гликаны – гликозидную связь с гидроксилом аминокислот серина или треонина; C-гликаны – гликозидную связь с атомом углерода, входящим в состав аминокислоты триптофана. C-гли-

козилирование встречается редко по сравнению с N- и O-гликозилированием (Chauhan et al., 2013).

Важным отличием N-гликозилирования от O-гликозилирования является то, что N-гликозидная связь образуется только с аспарагином, входящим в состав мотива Asn-X-Ser/Thr, где в качестве «X» может выступать любая аминокислота, за исключением пролина, в то время как для O-гликанов такого мотива не известно. Более того, существует фермент PNGase F, специфично расщепляющий N-гликозидную связь между гликаном и белком, в результате чего N-гликаны высвобождаются в раствор для дальнейшего анализа (Vilaj et al., 2021). В отличие от N-гликозилирования, сайт O-гликозилирования не имеет консенсусной последовательности, и существующие методы выделения O-гликанов (бета-элиминирование) обладают меньшей специфичностью по сравнению с таковым для N-гликанов (Mulloy et al., 2017). Отчасти поэтому на сегодняшний день наиболее хорошо разработаны технологии и протоколы определения структуры и высокопроизводительного анализа уровней N-гликанов, которым и посвящен настоящий обзор.

Наиболее распространенные мономеры в составе N-гликанов – моносахариды манноза, фукоза, галактоза, N-ацетилглюкозамин (GlcNAc) и сиаловая кислота (рис. 1). В структуре N-гликанов всегда присутствует остов (Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-Asn-X-Ser/Thr), к которому с помощью гликозидных связей присоединяются остальные мономеры. Структура гликанов может быть разветвленной и включать в себя одну или несколько ветвей – антенн. К каждой из антенн могут присоединяться мономеры – галактоза, сиаловая кислота или фукоза. Фукоза также может присоединять-



**Рис. 1.** Примеры структур N-гликанов: б – гликан с большим числом остатков маннозы, остальные гликаны имеют комплексную структуру; а, в, г – биантеннарные гликаны; ж–и – триантеннарные гликаны; к – тетраантеннарный гликан.

У структур а, д, е и и присутствует фукозилирование остова N-гликана; структуры ж, и и к антеннарно фукозилированы (фукоза присоединена к антенне). В структурах в и е присутствует бисектный (bisecting) остаток N-ацетилглюкозамина (GlcNAc). В структурах в–к и д–к присутствуют остатки галактозы и сиаловой кислоты соответственно.

ся непосредственно к остову. Отрицательно заряженная сиаловая кислота – единственный мономер в составе N-гликанов, который несет заряд. Гликаны, не несущие в своем составе сиаловой кислоты, являются нейтральными. В составе N-гликана сиаловая кислота всегда связана с остатком галактозы.

Одна из наиболее популярных в настоящее время так называемая оксфордская номенклатура гликанов (Harvey et al., 2009) следует правилам:

1. Символ “F” в самом начале обозначения говорит о наличии фукозы, присоединенной к остову.
2. Далее идет последовательность AN, где N – число антенн (ветвей) в структуре гликана.
3. Затем, в случае наличия рассечения сахарного остова, добавляется символ “B” (bisecting).
4. Далее, в случае фукозилирования антеннарных ветвей, добавляется символ “F”.
5. Если в структуре гликана присутствует галактоза, присоединенная к одной или нескольким антеннам, то далее идет последовательность G[n1,n2,...]N, где N – число галактоз в составе гликана, а n1 – атом углерода в составе галактозы, с которым образована гликозидная связь.
6. Если в структуре гликана присутствует сиаловая кислота, присоединенная к одной или нескольким антеннам, то далее идет последовательность S[n1,n2,...]N, где N – число остатков сиаловой кислоты в составе гликана, а n1 – атом углерода в составе сиаловой кислоты, с которым образована гликозидная связь.

Например, “FA2” обозначает, что в структуре гликана присутствуют остаток фукозы, присоединенный к остову, а также две антенны. “A3BG3S1” говорит о том, что в структуре гликана имеются три антенны, рассечение сахарного остова, галактозилирование трех антенн и сиалирование одной антенны.

### Гены, вовлеченные в биологические пути N-гликозилирования

В отличие от мРНК и белков, закодированных в последовательности геномной ДНК и синтезирующихся в результате матричных процессов, структура гликанов не закодирована в геноме напрямую. Биосинтез гликанов представляет собой разветвленную сеть биохимических реакций (Lombard et al., 2014). Конечная структура гликана определяется взаимодействием множества молекул и факторов – субстратов и их доступностью, активностью ферментов биосинтеза и деградации гликанов, их локализацией и конкуренцией за субстрат, и белками-транспортерами (Kukuruzinska, Lennon, 1998; Nairn et al., 2008, 2012; Moremen et al., 2012). Показано также, что структура и разнообразие N-гликанов, присутствующих в определенных клетках или тканях, отчасти регулируются на уровне транскрипции генов, кодирующих белки, которые участвуют в синтезе и деградации гликанов (Nairn et al., 2008, 2012; Moremen et al., 2012).

Биосинтез гликанов происходит в эндоплазматическом ретикулеуме и аппарате Гольджи. На сегодняшний день в базе данных KEGG содержится информация о более чем 300 ферментах, участвующих в процессах синтеза и деградации гликанов (Kanehisa et al., 2017). Одними из основных ферментов, непосредственно участвующих

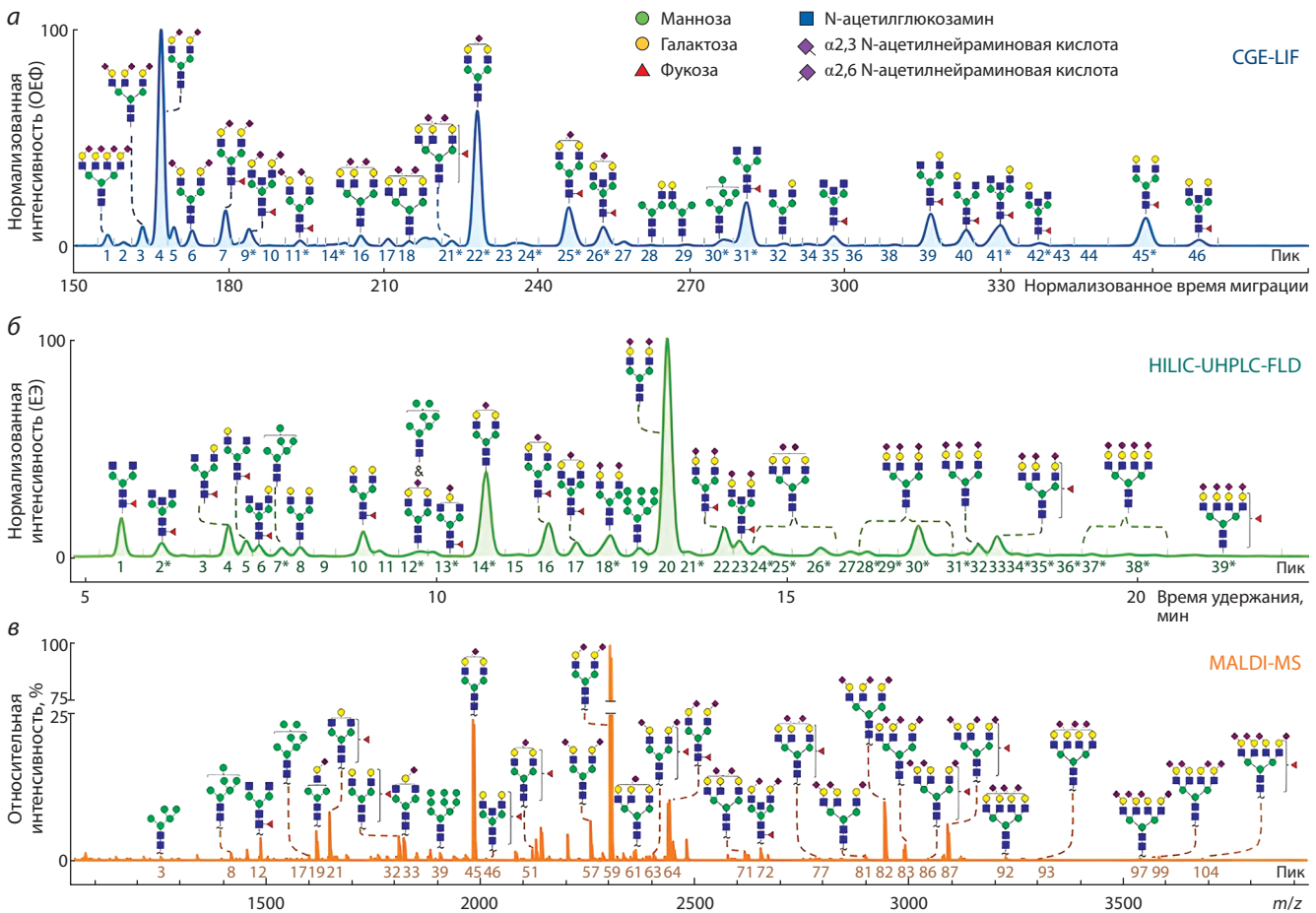
в биосинтезе N-гликанов, являются гликозилтрансферазы – ферменты, которые переносят активированные моносахариды к растущему гликану. В специализированной базе данных CAZy (Lombard et al., 2014) содержатся аннотация и классификация более 200 гликозилтрансфераз, по крайней мере 40 из которых относятся к пути N-гликозилирования белков. Кроме того, К.С. Егорова с коллегами разработали базу данных CSDB\_GT, содержащую экспериментально подтвержденные активности CAZy (Egorova et al., 2021). Этапы биосинтеза N-гликанов и соответствующие гены гликозилтрансфераз подробно описаны в обзорах (Saito, Ishii, 2002; Mohanty et al., 2020).

### Физико-химические методы высокопроизводительного секвенирования N-гликанов

Бурное развитие гликобиологии в совокупности с огромным успехом популяционных эпидемиологических исследований ускорило разработку методов высокопроизводительного измерения гликома белков плазмы крови. За последние десять лет создано несколько методов высокопроизводительного профилирования N-гликома (Huffman et al., 2014): высокоэффективная и сверхвысокоэффективная жидкостная хроматография (ВЭЖХ и СВЭЖХ, high and ultra-high performance liquid chromatography, HPLC и UHPLC), мультиплексный капиллярный гель-электрофорез с флуоресцентной детекцией (multiplex capillary gel electrophoresis with laser induced fluorescence detection, xCGE-LIF), жидкостная хроматография и тандемная масс-спектрометрия (liquid chromatography electrospray mass spectrometry, LC-MS), матрично-активированная лазерная десорбция/ионизация и тандемная масс-спектрометрия (matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, MALDI-TOF-MS). Репрезентативные профили N-гликозилирования белков плазмы крови человека, измеренные тремя различными методами, представлены на рис. 2.

В обзорной статье (de Naan et al., 2022) проведено подробное сравнение пяти наиболее популярных методов измерения N-гликома белков плазмы крови человека. Несмотря на разнообразие технологий, которые используются в данных методах, все они состоят из нескольких ключевых этапов – подготовки образца (клеточной культуры, ткани, органа или организма), выделения N-гликанов (например, путем отщепления от гликоконъюгатов), разделения N-гликанов и измерения их концентрации (абсолютной или относительной) (Huffman et al., 2014).

Применение того или иного метода для анализа гликанов имеет свои достоинства и недостатки. Методами UHPLC и xCGE-LIF анализируют гликаны, отщепленные от белка, в то время как методами MALDI-TOF-MS и LC-MS, основанными на масс-спектрометрии, можно анализировать конкретные гликопептиды, в состав которых входит участок белка с ковалентно соединенным гликаном. По сравнению с методами СВЖХ и xCGE-LIF, методы MALDI-TOF-MS и LC-MS идентифицируют гликаны различной молекулярной массы, но не могут разделять стереоизомеры гликанов. Методы СВЭЖХ и xCGE-LIF обладают достаточно хорошей разрешающей способностью и позволяют получать более точные количественные



**Рис. 2.** Профили N-гликозилирования белков плазмы крови, полученные с применением методов СВЭЖХ, xCGE-LIF и MALDI-MS.

N-гликомные профили, полученные методами: а – xCGE-LIF (Reiding et al., 2019); б – СВЭЖХ (Reiding et al., 2019; Zaytseva et al., 2020); в – MALDI-MS после дифференциальной этерификации сиаловой кислоты (Vreeker et al., 2018; Reiding et al., 2019; Zaytseva et al., 2020). Рисунок адаптирован из статьи (de Haan et al., 2022).

оценки. Данные методы имеют высокую производительность и меньшую начальную себестоимость по сравнению с методами, основанными на масс-спектрометрии.

Среди перечисленных методов высокопроизводительного измерения N-гликома белков плазмы крови наибольшее распространение получил метод СВЭЖХ (Актаџић et al., 2015) благодаря его относительной дешевизне, улучшенному разрешению (по сравнению с ВЭЖХ) и высокой производительности. На октябрь 2022 г. в мировом масштабе исследован гликом плазмы крови человека примерно 200000 образцов. Из этого объема около 80 % изучено с применением технологии СВЭЖХ (Г. Лауц, личное сообщение).

### Наследуемость уровней N-гликозилирования белков плазмы крови человека

Развитие методов высокопроизводительного измерения гликома и генетического анализа позволило к началу 2010-х гг. провести первые исследования генетического контроля гликозилирования на материале когортных исследований. N-гликом плазмы крови стал основным объектом исследования по ряду причин: во-первых, по сравнению с другими тканями человека плазма крови – более

доступный объект исследования; во-вторых, как было сказано ранее, на сегодняшний день наиболее хорошо разработаны технологии определения структуры и измерения уровней N-гликанов. Самыми представленными гликопротеинами плазмы крови человека являются иммуноглобулины G, A, M, фибриноген, трансферрин, гаптоглобин и др. (Clerc et al., 2016). Главным источником гликопротеинов плазмы крови человека служат клетки печени и клетки, продуцирующие антитела (Uhlén et al., 2015; Clerc et al., 2016).

В работе (Кнезевић et al., 2009) изучены популяционная изменчивость гликанов плазмы крови человека, их наследуемость (доля дисперсии признака, обусловленная генетическими различиями), а также влияние различных факторов среды на уровень гликанов. В данной работе уровни гликанов измеряли с использованием технологии ВЭЖХ. Авторы сделали несколько важных выводов. Во-первых, была установлена высокая популяционная изменчивость уровней гликозилирования. Во-вторых, обнаружен достоверный эффект пола и возраста человека на уровень различных гликанов, что говорит о важности учета данных параметров. В-третьих, наследуемость уровней гликанов варьировалась (средний коэффициент наследо-

вания  $h^2 = 34.7\%$ , стандартное отклонение –  $15.5\%$ , что говорит о том, что гликаны находятся под контролем как генетических, так и средовых факторов.

В работе (Zaytseva et al., 2020) исследователи оценили наследуемость 39 N-гликаных признаков, измеренных технологией СВЭЖХ. Показано, что для 24 из 39 признаков наследуемость составляет более  $50\%$  (средний коэффициент наследования  $h^2 = 48.0\%$ , стандартное отклонение –  $17.7\%$ ), что подтверждает гипотезу о существенном влиянии на гликом плазмы крови как средовых, так и генетических факторов. Авторы отметили самую высокую наследуемость ( $> 50\%$ ) у биантеннарных гликанов с фукозилированием остова и пониженным уровнем сиалирования антеннарных цепей, находящихся в составе иммуноглобулинов, преимущественно IgG, который является самым представленным гликопротеином среди всех белков плазмы крови человека (Clerc et al., 2016; Uhlén et al., 2019). Среднюю и высокую наследуемость ( $30\text{--}62\%$ ) показали би- и триантеннарные гликаны с высоким уровнем сиалирования антеннарных цепей. О.О. Zaytseva с коллегами (2020) полагают, что в этом случае высокая наследуемость может объясняться как наличием данных структур в составе большого числа гликопротеинов (трансферрин, гемопексин, альфа-1-анти-трипсин, альфа-1-кислый гликопротеин), что приводит к погрешностям в оценивании генетических факторов для каждого из них по отдельности, так и нахождением этих гликанов в составе преимущественно гликопротеинов, продуцируемых клетками печени, гликозилирование которых подвержено значительному влиянию окружающей среды ввиду принадлежности к белкам острой фазы воспаления (Jain et al., 2011).

Несмотря на то что изучение наследуемости дает возможность оценить, какая доля изменчивости признака находится под контролем генома, такие исследования не позволяют найти конкретные участки генома, влияющие на проявление признака. Поиск таких участков возможен с применением методов картирования генов количественных признаков, в частности метода полногеномного анализа ассоциаций.

### **Полногеномные исследования ассоциаций уровней N-гликанов, связанных с белками плазмы крови**

Самый широко используемый метод картирования локусов комплексных признаков и заболеваний человека – полногеномное исследование ассоциации (ПГИА). Этот метод предполагает проведение анализа ассоциаций между большим числом (от сотен тысяч до десятков миллионов) генетических маркеров, распределенных по всему геному, и исследуемым признаком. При этом, как правило, анализируются большие (от нескольких тысяч до миллионов) выборки особей или индивидов. Наличие таких данных позволяет тестировать практически весь геном на присутствие ассоциаций с исследуемым признаком и, таким образом, найти новые, ранее неизвестные ассоциации между локусами и признаками. Чаще всего в ПГИА используется материал не одной, а нескольких выборок. Результаты исследования каждой из выборок объединяются с применением методов полногеномного

метаанализа (Winkler et al., 2014), что увеличивает общий размер выборки и тем самым повышает статистическую мощность анализа ассоциаций.

Наличие ассоциации геномного локуса с изучаемым признаком не дает ответа на вопрос о молекулярно-биологических механизмах, лежащих в основе полученной ассоциации. В выявленных локусах могут располагаться от одного до десятков генов либо генов в локусе может не быть вообще (рис. 3) (Visscher et al., 2012, 2017). При этом существует множество причин возникновения ассоциации: наличие в локусе кодирующих замен, влияющих на структуру и функционирование продукта гена (белка или РНК); замен, влияющих на специфичность связывания транскрипционных факторов с регуляторными областями. Само число функциональных вариантов может варьироваться от одного до многих (Yang et al., 2012).

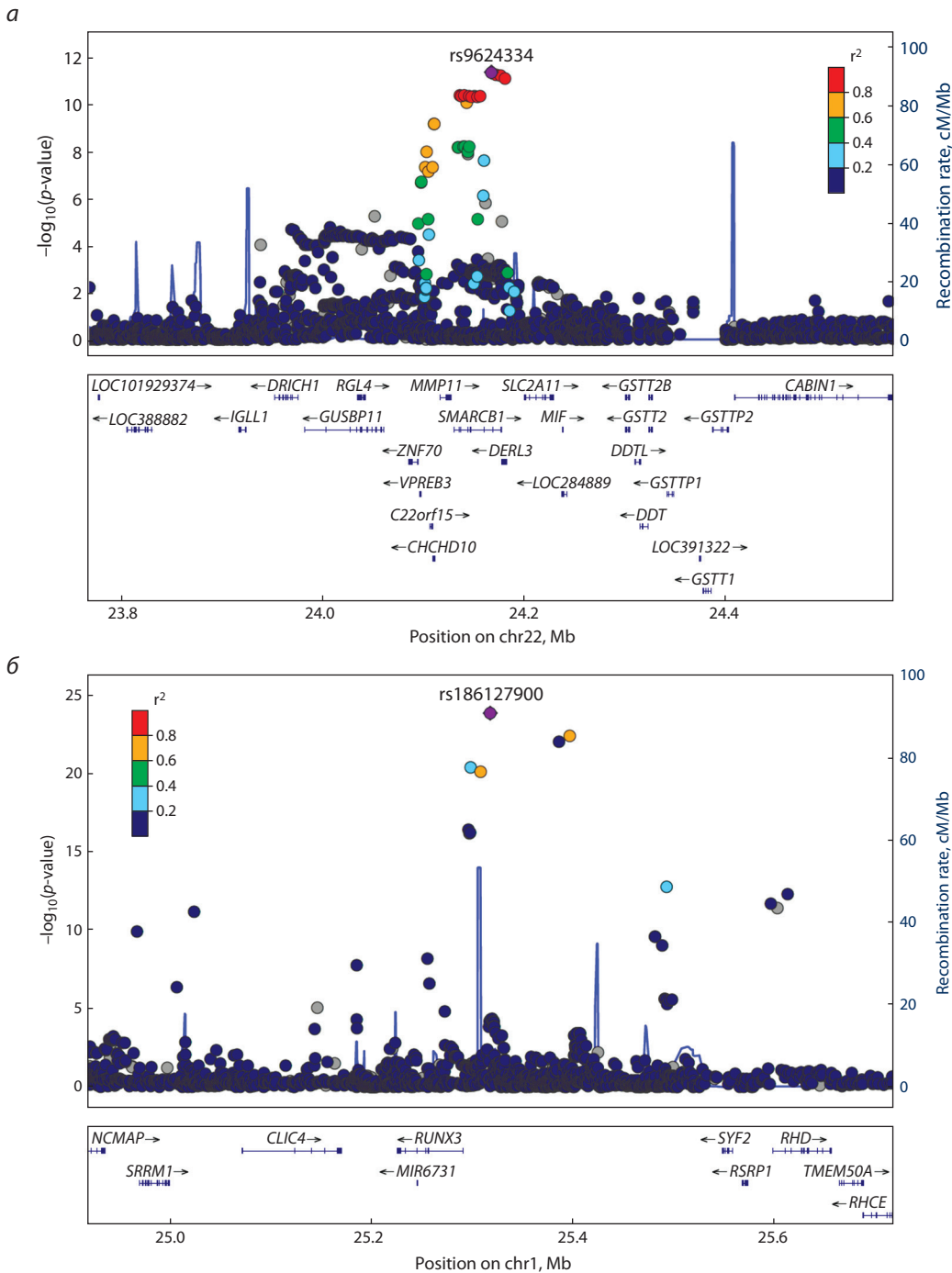
Определение функциональных генов в найденных локусах и механизмов их влияния на исследуемый признак – важная задача функциональных исследований, проводимых с применением методов молекулярной и клеточной биологии. При этом число возможных гипотез, требующих тестирования, растет (теоретически) в геометрической прогрессии в зависимости от числа возможных молекулярных механизмов ассоциации. Принимая во внимание сложность, дороговизну и трудоемкость методов молекулярной и клеточной биологии, крайне важным является проведение первичной биоинформатической приоритизации гипотез о механизмах возникновения ассоциации. К настоящему времени разработано большое число методов функциональной аннотации *in silico* (Yang et al., 2012; Bulik-Sullivan et al., 2015; Pers et al., 2015; McLaren et al., 2016; Staley et al., 2016; Zhu et al., 2016; Pasaniuc, Price, 2017; Hemani et al., 2018). Применение данных методов позволяет приоритизировать гипотезы о механизме ассоциации, тем самым повышая эффективность дальнейших молекулярно-биологических исследований.

К настоящему времени в качестве объектов для исследования генетического контроля гликозилирования белков методом ПГИА использованы: общий N-гликом белков плазмы крови человека (чему посвящен наш обзор) (Lauc et al., 2010a, b; Huffman et al., 2011; Sharapov et al., 2019, 2020), N-гликом иммуноглобулина G – наиболее представленного N-гликопротеина плазмы крови (Lauc et al., 2013; Shen et al., 2017; Wahl et al., 2018; Klarić et al., 2020; Shadrina et al., 2021), а также N-гликом трансферрина (Landini et al., 2022), секретируемого печенью.

На сегодняшний день опубликованы результаты пяти ПГИА общего N-гликома белков плазмы крови человека (Lauc et al., 2010a, b; Huffman et al., 2011; Sharapov et al., 2019, 2020).

### **ПГИА общего N-гликома белков плазмы крови человека**

Первые ПГИА уровней N-гликозилирования белков человека проведены в 2010–2011 гг. (Lauc et al., 2010a, b; Huffman et al., 2011). В этих работах для анализа уровней гликозилирования использованы метод ВЭЖХ и генетические данные с невысокой (по сегодняшним меркам) плотностью покрытия генома маркерами – до 2.5 млн ОНП на геном. В ПГИА 2010–2011 гг. было найдено шесть ло-



**Рис. 3.** Примеры графиков региональной ассоциации, визуализирующих ассоциацию между признаком и генетическими маркерами, расположенными в локусе.

По оси ординат отложен отрицательный десятичный  $p$ -value генетической ассоциации, по оси абсцисс – геномные координаты генетического маркера (ОНП). Сигнал ассоциации может быть расположен в кодирующей части нескольких генов (а), может не включать гены (б).

кусков (*FUT8*, *HNF1A*, *FUT3/FUT5/FUT6*, *MGAT5*, *B3GAT1*, *SLC9A9*), ассоциированных с N-гликозилированием белков плазмы крови человека. Отметим, что ни в одном из этих исследований не использованы независимые образцы для подтверждения результатов.

В исследовании, опубликованном в 2019 г. (Sharapov et al., 2019), были взяты данные выборки британского

близнецового исследования – TwinsUK Registry (Spector, Williams, 2006; Moayyeri et al., 2013). Для 2763 участников исследования были доступны данные полногеномного генотипирования и первичные данные СВЭЖХ N-гликома белков плазмы крови. Число ОНП было увеличено с 2.5 до 8.5 млн в результате проведенной импутации с использованием данных проекта «1000 геномов» и со-

Локусы, ассоциация которых (за исключением *KREMEN1*) подтверждена на независимых выборках в исследовании (Sharapov et al., 2020)

ОНП	Позиция	Ген	ЭФФ/РЕФ	EAF	Признак	BETA	SE	<i>p</i>	<i>N</i>
rs186127900	1:25318225	<i>RUNX3</i>	G/T	0.99	FA2G2S2	1.24	0.19	1.16E-10	1245
rs1257220	2:135015347	<i>MGAT5</i>	A/G	0.26	G4total, A4total	0.22	0.02	6.11E-20	4343
rs4839604	3:142960273	<i>SLC9A9</i>	C/T	0.80	FBS2/FS2	-0.20	0.03	3.87E-11	3592
rs17775791	3:186722362	<i>ST6GAL1</i>	C/T	0.26	FG1S1/(FG1+FG1S1)	-0.49	0.02	8.60E-97	4343
rs3115663	6:31601843	<i>PRRC2A</i>	C/T	0.18	M9	-0.15	0.03	1.63E-07	4343
rs6421315	7:50355207	<i>IKZF1</i>	C/G	0.37	A2[6]BG1n	-0.23	0.02	1.19E-27	4802
rs13297246	9:33128617	<i>B4GALT1</i>	A/G	0.17	FA2G2n	0.31	0.03	1.28E-24	4051
rs3967200	11:126232385	<i>ST3GAL4</i>	C/T	0.86	G4S3/G4S4	0.63	0.03	1.20E-106	4802
rs7928758	11:134265967	<i>B3GAT1</i>	G/T	0.15	A4G4S3	-0.36	0.03	6.43E-27	3592
rs735396	12:121438844	<i>HNFB1A</i>	C/T	0.35	G3Fa/G3total	-0.21	0.02	4.91E-20	4343
rs11621121	14:65822493	<i>FUT8</i>	C/T	0.42	FG3/G3total	-0.31	0.02	8.94E-45	4187
rs35590487	14:105989599	<i>IGH, TMEM121</i>	C/T	0.75	FA2[3]G1n	-0.20	0.03	1.38E-09	2469
rs3760776	19:5839746	<i>FUT6</i>	A/G	0.09	G3Fa/G3total	-0.48	0.05	3.85E-23	2469
rs9624334	22:24166256	<i>SMARCB1, DERL3, CHCHD10</i>	C/G	0.17	FA2[6]BG1n	-0.31	0.03	7.15E-26	4051
rs909674	22:39859169	<i>MGAT3</i>	A/C	0.70	FBn	-0.22	0.02	1.88E-20	4343
rs140053014	22:29550678	<i>KREMEN1</i>	Ins/Del	0.98	FA2[3]G1n	-0.68	0.23	0.0027	459

Примечание. ОНП – ОНП, показавший наименьшее *p*-value; ЭФФ/РЕФ – эффекторный, референсный аллель; EAF – частота эффекторного аллеля; BETA – оценка эффекта эффекторного аллеля на признак; SE – его стандартная ошибка; *p* – *p*-value ассоциации ОНП с признаком.

ответствующего контроля качества. Была подтверждена ассоциация 5 из 6 локусов, найденных ранее (всех за исключением *SLC9A9*), а также выявлено 10 новых локусов.

По результатам четырех исследований (Lauc et al., 2010a, b; Huffman et al., 2011; Sharapov et al., 2019) в общей сложности найдены ассоциации с 16 локусами, 15 из них были в дальнейшем подтверждены в работе (Sharapov et al., 2020) (см. таблицу), в которой использована крупнейшая (на момент проводимого исследования) коллекция геномных и гликомных данных для 4802 образцов из четырех выборок – EPIC-Potsdam, PainOmics, SOCCS и SABRE, детальное описание которых представлено в дополнительных примечаниях к статье (Sharapov et al., 2020). Участники данных исследований были генотипированы с использованием следующих ДНК чипов: EPIC-Potsdam (Human660W, 560 000 ОНП, HumanCoreExome, 410 000 ОНП, InfiniumOmniExpressExome, 850 000 ОНП), PainOmics (Illumina HumanCore BeadChip, 720 000 ОНП, Illumina GSA, 300 000 ОНП), SOCCS (HumanHap300/ HumanHap240S, 510 000 ОНП), SABRE (Illumina Human Core Bead Chip, 330 000 ОНП).

Когортное исследование EPIC-Potsdam включает 27 000 участников в возрастном диапазоне 35–65 лет. Участники были набраны случайным образом из популяции в г. Подстам (Германия), в период с 1994 по 1998 г. (Boeing et al., 1999). PainOmics (Allegri et al., 2016) – исследование типа «случай-контроль», направленное на поиск потенциальных биомаркеров дорсалгии и терапевтических мишеней для его лечения. В период с 2014 по 2016 г. была собрана выборка из 3400 участников, проживающих в Италии, Бельгии, Англии и Хорватии.

Шотландское исследование SOCCS (Theodoratou et al., 2016; Vučković et al., 2016) является исследованием «случай-контроль», направленным на изучение факторов риска колоректального рака. Были собраны данные о 2000 пациентах с колоректальным раком и 2100 контрольных испытуемых. SABRE – это популяционное исследование, начатое в 1988 г. (Tillin et al., 2012). Всего были собраны данные о 4800 участниках в возрасте от 40 до 69 лет, проживающих в Западном Лондоне (Великобритания).

Для приоритизации новых генов-регуляторов гликозилирования белков в подтвержденных локусах и для выдвижения гипотез о потенциальных механизмах действия этих локусов в работе (Sharapov et al., 2019) был применен ансамбль методов количественной генетики и биоинформатики. Эти методы и подходы описаны ниже.

1. Приоритизация генов на основании результатов анализов колокализации с eQTL. Методы колокализации, в частности использованный авторами метод SMR/HEIDI (Zhu et al., 2016), позволяют идентифицировать гены, изменение экспрессии (на уровне мРНК) которых опосредует ассоциацию между ОНП и изучаемым признаком.
2. Приоритизация генов, основанная на определении возможных функциональных последствий нуклеотидных замен, находящихся в высоком неравновесии по сцеплению с ОНП, ассоциированными с N-гликомными признаками. С помощью методов VEP (McLaren et al., 2016), FATHMM-XF (Rogers et al., 2018), FATHMM-InDel (Ferlino et al., 2017) выбирались ОНП, замены в которых изменяли первичную аминокислотную последовательность белка и/или классифицировались как



патогенные. Гены, последовательность которых была затронута такими заменами, приоритизировались как гены-кандидаты.

3. Приоритизация генов на основании их вовлеченности в различные биологические пути. Метод DEPICT (Pers et al., 2015) приоритизирует гены и биологические пути на базе результатов анализа обогащения (перепредставленности генов того или иного биологического пути в ассоциированных локусах), который в свою очередь проводится на основе заранее рассчитанной вероятности принадлежности того или иного локуса к конкретной генной сети и/или биологическому пути.

Если для того или иного локуса применение описанных выше методов не привело к приоритизации какого-то гена, выбирали ген, ближайший к ОНП с самой достоверной ассоциацией в локусе.

### Гены-кандидаты, вовлеченные в процессы N-гликозилирования белков плазмы крови

В результате проведенных *in silico* анализов в работе по изучению общего N-гликома плазмы крови (Sharapov et al., 2019) для 15 локусов были приоритизированы 20 генов-кандидатов (рис. 4).

В этом разделе приведено подробное описание этих генов и гипотезы об их возможной роли в регуляции N-гликозилирования белков плазмы крови человека.

В 8 из 15 локусов (*MGAT5*, *MGAT3*, *FUT3/FUT5/FUT6*, *FUT8*, *ST6GAL1*, *ST3GAL4*, *B4GALT1*, *B3GAT1*) в качестве генов-кандидатов содержатся гены, кодирующие ферменты – гликозилтрансферазы, участвующие в биосинтезе N-гликанов.

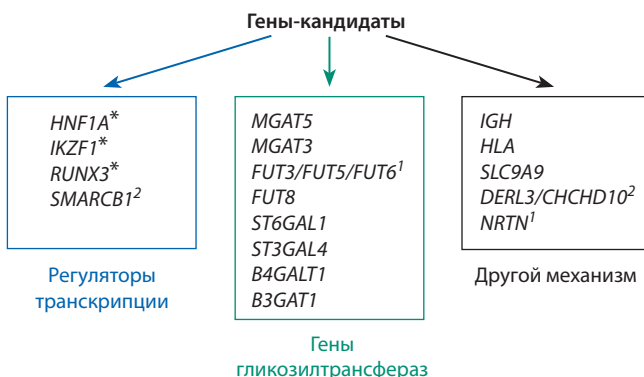
В локусе на 2-й хромосоме, 125 м. п. н., геном-кандидатом является *MGAT5*, кодирующий фермент GnT-V – альфа-1,6-маннозилгликопротеин 6-бета-N-ацетилглюкозаминилтрансферазу. Данный фермент переносит остаток N-ацетилглюкозамина к маннозе в составе N-гликана, при этом образуется три- или тетраантеннарный N-гликан. Локус с геном *MGAT5* показал ассоциацию с гликомными признаками, отражающими уровень три- и тетраантеннарных гликанов (Sharapov et al., 2019).

В локусе на 22-й хромосоме, 39 м. п. н., ген-кандидат – *MGAT3*, кодирующий N-ацетилглюкозаминилтрансферазу GnT-III – бета-1,4-маннозилгликопротеин 4-бета-N-ацетилглюкозаминилтрансферазу. Этот фермент переносит остаток N-ацетилглюкозамина к маннозе в составе N-гликана в определенную позицию, при этом образуется рассеяние остова. Показано наличие плейотропного эффекта данного локуса как на уровни N-гликанов, так и на уровень экспрессии гена *MGAT3* в CD19+ клетках (B-лимфоцитах) (Sharapov et al., 2019; Klarić et al., 2020).

В локусе на 14-й хромосоме, 66 м. п. н., ген-кандидат – *FUT8*. Этот ген кодирует фермент Fuc-TVIII – альфа-(1, 6)-фукозилтрансферазу. Данный фермент переносит остаток фукозы к N-ацетилглюкозамину, находящемуся в остове N-гликана, и тем самым отвечает за фукозилирование остова N-гликанов. Следует отметить, что локусы *FUT8* и *MGAT3* показали ассоциацию с признаками FBS2/(FS2+FBS2) и FBS2/FS2, отражающими представленность рассеяния остова у биантеннарных гликанов с фукозилированием остова (Sharapov et al., 2019), что подтверждает известный факт интерференции активностей ферментов Fuc-TVIII и GnT-III (Brockhausen, Schachter, 1997).

В локусе на 19-й хромосоме, 5,8 м. п. н., генами-кандидатами являются *FUT6*, *FUT5*, *FUT3*, *NRTN*. Ген *NRTN* кодирует нейротропический фактор, регулирующий выживание и функционирование нейронов. Гены *FUT6* и *FUT3/FUT5* кодируют ферменты Fuc-TVI и Fuc-TIII – фукозилтрансферазы 6 и 3 соответственно, переносящие остаток фукозы с ГДФ-фукозы к N-ацетилглюкозамину путем образования альфа-1,3(4)-гликозидной связи. Эти ферменты ответственны за антеннарное фукозилирование N-гликанов. В работе (Sharapov et al., 2019) продемонстрировано, что этот локус ассоциирован с антеннарным фукозилированием три- и тетраантеннарных гликанов. Нужно отметить, что в гене *FUT6* расположен ОНП rs17855739. Этот ОНП кодирует замену G>A (частота аллеля A ~12 % в популяциях человека, согласно базе данных TopMED), которая приводит к замещению отрицательно заряженной глутаминовой кислоты на положительно заряженный лизин в 247-й позиции (замена p.Glu247Lys). Эта замена расположена в каталитическом домене фермента Fuc-TVI и приводит к инактивации фермента. Таким образом, этот вариант может иметь функциональное влияние на гликозилирование белков плазмы крови человека. Нужно отметить, что гены *FUT3*, *FUT5* и *FUT6* произошли от общего предкового гена в результате двух дупликаций (Duruu et al., 2002). При этом экспрессия гена *FUT5* на уровне транскрипции и трансляции в организме человека гораздо слабее по сравнению с *FUT3* и *FUT6* (Taniguchi et al., 2014).

В локусе на 3-й хромосоме, 186 м. п. н., ген-кандидат – *ST6GAL1*. Ген *ST6GAL1* кодирует альфа-2,6-сиалилтрансферазу 1. Этот фермент катализирует образование альфа-2,6-гликозидной связи между сиаловой кислотой и N-ацетилглюкозамином, связанным с галактозой, в составе N-гликана. В работе (Sharapov et al., 2019) локус *ST6GAL1* ассоциирован с уровнями моно- и дисиалилированных N-гликанов и их предшественников.



**Рис. 4.** Предложенные в работе (Sharapov et al., 2019) гены-кандидаты, регулирующие уровни гликозилирования белков плазмы крови человека.

Звездочкой отмечены гены, роль которых в регуляции гликозилирования доказана в результате экспериментов. Верхний индекс обозначает приоритизацию генов-кандидатов внутри одного локуса.

В локусе на 11-й хромосоме, 126 м.п.н., геном-кандидатом является *ST3GAL4*. Ген *ST3GAL4* кодирует фермент альфа-2,3-сиалилтрансферазу, переносящую остаток сиаловой кислоты. Данный локус показал ассоциацию с уровнями различных сиалилированных N-гликанов (Sharapov et al., 2019).

В локусе на 9-й хромосоме, 33 м.п.н., ген-кандидат – *B4GALTI*. Ген *B4GALTI* кодирует фермент галактозилтрансферазу, присоединяющий галактозу к различным субстратам, в том числе и к N-ацетилглюкозамину. Локус *B4GALTI* был ассоциирован с уровнями галактозилированных биантеннарных N-гликанов и их предшественников (Sharapov et al., 2019). Известно, что ряд мутаций в гене *B4GALTI* приводит к врожденному заболеванию гликозилирования (Staretz-Chacham et al., 2020).

В локусе на 11-й хромосоме, 134 м.п.н., геном-кандидатом является *B3GAT1*, кодирующий фермент галактозилгалактозилксилосилпротеин-3-бета-глюкуронозилтрансферазу 1-го типа. Этот фермент катализирует перенос глюкуроновой кислоты в ходе биосинтеза HNK-1 эпитопа. Этот эпитоп экспрессируется на лимфоцитах, однако его присутствие на белках плазмы крови до определенного момента не было известно. Ассоциация данного локуса с уровнями N-гликанов белков плазмы крови впервые приведена в работе (Huffman et al., 2011). В исследовании (Sharapov et al., 2019) обнаружено присутствие глюкуроновой кислоты в N-гликоме плазмы крови, что может объяснять ассоциацию этого локуса.

В семи других локусах гены-кандидаты не были генами гликозилтрансфераз. В локусе на 22-й хромосоме, 39 м.п.н., были приоритизированы три гена – *SMARCB1*, *DERL3*, *CHCHD10*. В кодирующей последовательности гена *SMARCB1* расположен самый сильный сигнал ассоциации в этом локусе. Ген *SMARCB1* кодирует белок – часть комплекса hSWI/SNF, ремоделера хроматина. Продукт гена *SMARCB1* играет важную роль в ингибировании канцерогенеза, пролиферации и дифференциации клеток (Pottier et al., 2007).

Ген *DERL3* кодирует фермент, участвующий в деградации люминальных гликопротеинов с некорректной третичной структурой, происходящей в эндоплазматическом ретикулуме (Oda et al., 2006). В данном локусе находится патогенный вариант rs3177243, расположенный в кодирующей последовательности гена *DERL3*.

Ген *CHCHD10* кодирует митохондриальный белок, встречающийся в фибриллах крист митохондрий. Показано, что генетическая ассоциация этого локуса с уровнями N-гликанов белков может быть опосредована влиянием нуклеотидных замен на экспрессию гена *CHCHD10* в клетках крови (Sharapov et al., 2019). Ранее непосредственное участие митохондриальных белков в процессах гликозилирования не было известно, однако в 2017 г. опубликована работа, показавшая роль фрагментации митохондрий и числа контактов эндоплазматического ретикулума с митохондриями в представленности сиалилированных гликанов на поверхности глиобластомных клеток, что, в свою очередь, влияло на узнавание лимфоцитами клеток глиобластомы (Martinvalet, 2018).

Локус на 14-й хромосоме, 105 м.п.н., содержит кластер генов *IGH*, кодирующих тяжелые цепи иммуноглобули-

нов. Иммуноглобулин G (IgG) – наиболее представленный N-гликопротеин плазмы крови человека (Clerc et al., 2016), а его конститутивный сайт N-гликозилирования расположен в тяжелой цепи.

В локусе на 3-й хромосоме, 142 м.п.н., ген-кандидат – *SLC9A9*. Ген *SLC9A9* кодирует Na<sup>+</sup>/H насос, который, предположительно, регулирует уровень pH в аппарате Гольджи (АГ). Гликозилирование белков происходит в АГ, и, согласно имеющимся данным, этот процесс чувствителен к изменению pH (Kellokumpu, 2019). Процессы, происходящие в АГ, влияют на формирование гетеродимерных комплексов, осуществляющих гликозилирование (Hassinen et al., 2011). В работе (Rivinoja et al., 2009) продемонстрировано, что увеличение pH в АГ может нарушить терминальное N-гликозилирование (включая сиалилирование) из-за неверной локализации гликозилтрансфераз. Согласно этой гипотезе, в работе (Huffman et al., 2011) локус *SLC9A9* показал ассоциацию с уровнем тетрасиалилированных N-гликанов, а в работе (Sharapov et al., 2019) – с уровнем сиалилированных N-гликанов.

В локусе на 12-й хромосоме, 121 м.п.н., ген-кандидат – *HNF1A*. Подробное функциональное исследование этого локуса, проведенное G. Lauc с коллегами (Lauc et al., 2010a), выявило, что ген *HNF1A*, кодирующий фактор транскрипции гепатоцитов, регулирует экспрессию большинства генов, кодирующих фукозилтрансферазы, – *FUT3*, *FUT5*, *FUT6*, *FUT8*, *FUT10*, *FUT11* – в клеточной линии HepG2, полученной из клеток печени. В том же исследовании продемонстрировано, что *HNF1A* регулирует экспрессию генов, кодирующих ключевые ферменты для синтеза ГДФ-фукозы – субстрата для фукозилтрансфераз, что в совокупности говорит о ключевой роли гена *HNF1A* в процессах фукозилирования гликанов.

Для локуса на 7-й хромосоме в 50 м.п.н. ген-кандидат – *IKZF1*. Ранее в работе (Lauc et al., 2013) была найдена ассоциация данного локуса с гликозилированием IgG, в которой авторы предположили, что ген *IKZF1* является кандидатным для данного локуса. Ген *IKZF1* кодирует ДНК-связывающий белок Ikaros, – регулятор транскрипции, контролирующий процесс ремоделирования хроматина. Интересно, что локус *IKZF1* ассоциирован с уровнями N-гликанов белков плазмы крови с фукозилированием остова, с которыми был ассоциирован локус *FUT8* (Sharapov et al., 2019). *IKZF1* рассматривается как важный регулятор дифференциации лимфоцитов (Sellars et al., 2009; Marke et al., 2018).

Поскольку клетки, секретирующие IgG, – производные лимфоцитов, ген *IKZF1* был выбран в качестве кандидата в этом локусе. Была выдвинута гипотеза о его роли в регуляции фукозилирования остова N-гликанов IgG путем регуляции экспрессии гена *FUT8* (Sharapov et al., 2019). Более того, в исследовании (Klarić et al., 2020) экспериментальным путем установлено, что нокаун гена *IKZF1* в IgG секретирующих клетках *MATAT6* приводит к более чем трехкратному увеличению экспрессии *FUT8* и увеличению уровня фукозилирования секретируемого IgG.

В локусе на 1-й хромосоме, 25 м.п.н., ген-кандидат – *RUNX3*. Этот ген кодирует Runt-домен-содержащий бе-



крови (рис. 5). Эта сеть визуализирует ассоциацию между найденными локусами и уровнями N-гликанов белков плазмы крови.

Для построения сети проведена классификация гликомных признаков. Признаки были классифицированы на четыре группы согласно ткани, секретирующей N-гликопротеины в плазму крови. Первая группа – признаки, отражающие уровни N-гликанов иммуноглобулинов (IgA, IgG, IgD, IgE, IgM), секретируемых клетками лимфоцитарного ряда – В-лимфоцитами, плазмобластами и плазмочитами. Вторая группа – признаки, отражающие уровни N-гликанов белков (трансферрин, гаптоглобин и т. д.), секретируемых в основном гепатоцитами – клетками печени. Третья группа – признаки, отражающие уровни N-гликанов белков, секретируемых как В-лимфоцитами и их потомками, так и гепатоцитами. Четвертая группа – неклассифицированные признаки. Классификация была проведена на основе данных представленности гликопротеинов, опубликованных в работе (Clerc et al., 2016), в которой исследователи оценили вклад каждого из N-гликопротеинов в N-гликом плазмы крови человека.

Локусы и ассоциированные признаки в этой сети могут быть визуально разделены на две частично перекрывающиеся подсети. Первая подсеть образована локусами *ST3GAL4*, *B3GAT1*, *MGAT5*, *HNF1A*, *FUT3/FUT6*, *FUT8* и *SLC9A9*. Эта подсеть ассоциирована с уровнями N-гликанов, связанных с N-гликопротеинами, секретируемыми в кровотоке клетками печени. Большинство из этих признаков отражает уровни N-гликанов с тремя или четырьмя антеннами, не встречающимися в составе иммуноглобулинов. В данную сеть входит локус *HNF1A* – транскрипционного фактора гепатоцитов. При этом локус *HNF1A* показал ассоциацию с теми же признаками, что и локус *FUT3/FUT6*, это соответствует ранее доказанной роли гена *HNF1A* в регуляции уровня экспрессии фукозилтрансфераз (Lauc et al., 2010a). Полученные результаты позволяют выдвинуть гипотезу о том, что эта подсеть генов регулирует процессы гликозилирования в гепатоцитах. Роль генов-кандидатов, входящих в эту сеть, следует проверять на клетках печени, например на гепатоцитах, или на клетках, близких к ним, например на клеточной линии HepG2.

Вторая подсеть образована локусами *FUT8*, *FUT6/FUT3*, *SLC9A9*, *IKZF1*, *MGAT3*, *RUNX3*, *SMARCB1/DERL3/CHCHD10*, *B4GALT1*, *ST6GAL1* и *IGH/TMEM121*. Эти локусы ассоциированы с уровнями N-гликанов, связанных с иммуноглобулинами, секретируемыми в кровотоке клетками лимфоцитарного ряда. Более того, в исследованиях ПГИА уровней N-гликанов иммуноглобулина G было показано, что данные локусы ассоциированы с N-гликозилированием IgG (Shen et al., 2017; Klarić et al., 2020). Основываясь на том, что IgG является наиболее представленным гликопротеином плазмы крови, можно предположить: гены-кандидаты из данной сети регулируют процессы N-гликозилирования в В-лимфоцитах и их потомках. Роль генов-кандидатов, входящих в данную сеть, скорее всего, следует проверять на материале клеток, секретирующих антитела, или клеток, близких к ним.

В работе (Klarić et al., 2020) установлена роль транскрипционного фактора *IKZF1* в регуляции экспрессии

гена *FUT8* в лимфоидной линии *GM12878*. Нокдаун *IKZF1* приводил к повышению уровня фукозилированных белков, что доказывает роль фактора транскрипции *IKZF1* в регуляции фукозилирования белков в результате *in vitro* эксперимента.

В общей сложности с помощью метода полногеномного исследования ассоциаций были найдены 16 локусов, ассоциация 15 из которых подтверждена на независимых выборках. Для этих локусов было проведено *in silico* исследование и предложено 20 генов-кандидатов. В результате *in vitro* экспериментов доказаны роль фактора транскрипции *IKZF1* в регуляции фукозилирования белков и роль гена *HNF1A* в регуляции уровня экспрессии фукозилтрансфераз. Роль транскрипционного фактора *RUNX3* в регуляции N-гликозилирования подтверждена в результате таргетного редактирования генома (с использованием системы CRISPR-dCas9) в клеточных линиях VPR-dCas9 и KRAB-dCas9 HEK-293F, секретирующих IgG в среду. Сравнение профиля N-гликозилирования IgG с контрольной немодифицированной клеточной линией продемонстрировало, что увеличение экспрессии гена *RUNX3* приводит к значительному уменьшению галактозилированных структур с последующим увеличением агалактозилированных структур (Mijakovac et al., 2022).

## Заключение

Результаты полногеномных исследований ассоциации уровней N-гликанов белков плазмы крови доказывают факт о том, что N-гликозилирование белков плазмы крови человека – это сложный процесс, который находится под контролем генов, принадлежащих разным биологическим путям и экспрессирующихся в разных тканях. Гены-кандидаты, предложенные в результате полномасштабного *in silico* исследования (Sharapov et al., 2019) подтвержденных локусов, позволяют выдвинуть функциональные гипотезы о механизмах влияния найденных локусов на N-гликозилирование белков плазмы крови. Эти гипотезы будут востребованы при планировании молекулярно-генетических *in vitro* и *in vivo* исследований гликома и изучения роли гликома в патогенезе социально и экономически важных заболеваний человека. Результаты проведенных экспериментов *in vitro* подкрепляют научную состоятельность функциональных гипотез о генах-кандидатах, выдвинутых с помощью метода ПГИА.

Существует несколько направлений развития популяционной гликогеномики человека. Будут проведены более мощные ПГИА уровней N-гликанов с применением выборок большего размера. Станут доступны новые данные функциональной геномики, применимые для изучения процессов N-гликозилирования, совместный анализ которых с результатами ПГИА позволит определить большее число локусов и потенциальных регуляторов N-гликозилирования. В настоящее время применение метода ПГИА для изучения регуляции гликозилирования ограничено анализом общего N-гликома плазмы крови, N-гликома IgG и трансферрина. Развитие технологий профилирования N-гликома расширит набор белков, для которых будут изучены индивидуальные профили N-гликозилирования. С другой стороны, скорее всего, появятся высокопроизводительные технологии профилирования N-гликома дру-

гих тканей человека. Все это позволит составить более полную картину регуляции N-гликозилирования белков человека, что, в свою очередь, даст возможность установить роль гликозилирования в патогенезе гликом-ассоциированных заболеваний и ускорить процессы разработки методов прогнозирования, профилактики, диагностики и лечения данных заболеваний.

## Список литературы / References

- Akmačić I.T., Ugrina I., Štambuk J., Gudelj I., Vučković F., Lauc G., Pučić-Baković M. High-throughput glycomics: optimization of sample preparation. *Biochemistry (Mosc.)*. 2015;80(7):934-942. DOI 10.1134/S0006297915070123.
- Allegri M., De Gregori M., Minella C.E., Klersy C., Wang W., Sim M., Gieger C., Manz J., Pemberton I.K., MacDougall J., Williams F.M., Van Zundert J., Buyse K., Lauc G., Gudelj I., Primorac D., Skelin A., Aulchenko Y.S., Karssen L.C., Kapural L., Rauck R., Fanelli G., PainOMICS Group "Omics" biomarkers associated with chronic low back pain: protocol of a retrospective longitudinal study. *BMJ Open*. 2016;6(10):e012070. DOI 10.1136/bmjopen-2016-012070.
- Anthony R.M., Wermeling F., Ravetch J.V. Novel roles for the IgG Fc glycan. *Ann. N. Y. Acad. Sci.* 2012;1253(1):170-180. DOI 10.1111/j.1749-6632.2011.06305.x.
- Boeing H., Korfmann A., Bergmann M.M. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. *Ann. Nutr. Metab.* 1999;43(4):205-215. DOI 10.1159/000012787.
- Brockhausen I., Schachter H. Glycosyltransferases involved in N- and O-glycan biosynthesis. In: Gabius H.-J., Gabius S. (Eds.). *Glycosciences*. Weinheim: Chapman & Hall, 1997;79-113. DOI 10.1002/9783527614738.ch5.
- Bulik-Sullivan B., Finucane H.K., Anttila V., Gusev A., Day F.R., Loh P.-R., ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Duncan L., Perry J.R.B., Patterson N., Robinson E.B., Daly M.J., Price A.L., Neale B.M. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* 2015;47(11):1236-1241. DOI 10.1038/ng.3406.
- Chauhan J.S., Rao A., Raghava G.P.S. In silico platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences. *PLoS One*. 2013;8(6):e67008. DOI 10.1371/journal.pone.0067008.
- Clerc F., Novokmet M., Dotz V., Reiding K.R., de Haan N., Kammeijer G.S.M., Dalebout H., Bladergroen M.R., Vukovic F., Rapp E., IBD-BIOM Consortium, Targan S.R., Barron G., Manetti N., Lattiano A., McGovern D.P.B., Anness V., Lauc G., Wuhler M. Plasma N-glycan signatures are associated with features of inflammatory bowel diseases. *Gastroenterology*. 2018;155(3):829-843. DOI 10.1053/j.gastro.2018.05.030.
- Clerc F., Reiding K.R., Jansen B.C., Kammeijer G.S.M., Bondt A., Wuhler M. Human plasma protein N-glycosylation. *Glycoconj. J.* 2016;33(3):309-343. DOI 10.1007/s10719-015-9626-2.
- Cobb B.A. The history of IgG glycosylation and where we are now. *Glycobiology*. 2020;30(4):202-213. DOI 10.1093/glycob/cwz065.
- Connelly M.A., Gruppen E.G., Otvos J.D., Dullaart R.P.F. Inflammatory glycoproteins in cardiometabolic disorders, autoimmune diseases and cancer. *Clin. Chim. Acta*. 2016;459:177-186. DOI 10.1016/j.cca.2016.06.012.
- Craveur P., Rebehmed J., de Brevern A.G. PTM-SD: a database of structurally resolved and annotated posttranslational modifications in proteins. *Database (Oxford)*. 2014;2014:bau041. DOI 10.1093/database/bau041.
- de Haan N., Pučić-Baković M., Novokmet M., Falck D., Lageveen-Kammeijer G., Razdorov G., Vučković F., Trbojević-Akmačić I., Gornik O., Hanić M., Wuhler M., Lauc G., The Human Glycome Project. Developments and perspectives in high-throughput protein glycomics: enabling the analysis of thousands of samples. *Glycobiology*. 2022;32(8):651-663. DOI 10.1093/glycob/cwac026.
- Dotz V., Wuhler M. N-glycome signatures in human plasma: associations with physiology and major diseases. *FEBS Lett.* 2019;593(21):2966-2976. DOI 10.1002/1873-3468.13598.
- Dupuy F., Germot A., Marena M., Oriol R., Blancher A., Julien R., Maftah A.  $\alpha$ 1,4-fucosyltransferase activity: A significant function in the primate lineage has appeared twice independently. *Mol. Biol. Evol.* 2002;19(6):815-824. DOI 10.1093/oxfordjournals.molbev.a004138.
- Egorova K.S., Smirnova N.S., Toukach P.V. CSDB\_GT, a curated glycosyltransferase database with close-to-full coverage on three most studied nonanimal species. *Glycobiology*. 2021;31(5):524-529. DOI 10.1093/glycob/cwaa107.
- Ferlaino M., Rogers M.F., Shihab H.A., Mort M., Cooper D.N., Gaunt T.R., Campbell C. An integrative approach to predicting the functional effects of small indels in non-coding regions of the human genome. *BMC Bioinformatics*. 2017;18(1):442. DOI 10.1186/s12859-017-1862-y.
- Fuster M.M., Esko J.D. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer*. 2005;5(7):526-542. DOI 10.1038/nrc1649.
- Gagneux P., Aebi M., Varki A. Evolution of glycan diversity. In: Varki A., Cummings R.D., Esko J.D., Stanley P., Hart G.W., Aebi M., Darvill A.G., Kinoshita T., Packer N.H., Prestegard J.H., Schnaar R.L., Seeberger P.H. (Eds.). 3rd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2015-2017. Chapter 20. PMID 28876829. DOI 10.1101/glycobiology.3e.020.
- Gudelj I., Lauc G., Pezer M. Immunoglobulin G glycosylation in aging and diseases. *Cell. Immunol.* 2018a;333:65-79. DOI 10.1016/j.cellimm.2018.07.009.
- Gudelj I., Salo P.P., Trbojević-Akmačić I., Albers M., Primorac D., Perola M., Lauc G. Low galactosylation of IgG associates with higher risk for future diagnosis of rheumatoid arthritis during 10 years of follow-up. *Biochim. Biophys. Acta Mol. Basis Dis.* 2018b;1864(6 Pt. A):2034-2039. DOI 10.1016/j.bbadis.2018.03.018.
- Harvey D.J., Merry A.H., Royle L., Campbell M.P., Dwek R.A., Rudd P.M. Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *Proteomics*. 2009;9(15):3796-3801. DOI 10.1002/pmic.200900096.
- Hassinen A., Pujol F.M., Kokkonen N., Pieters C., Kihlström M., Korhonen K., Kellokumpu S. Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. *J. Biol. Chem.* 2011;286(44):38329-38340. DOI 10.1074/jbc.M111.277681.
- Hemani G., Zheng J., Elsworth B., Wade K.H., Haberland V., Baird D., Laurin C., Burgess S., Bowden J., Langdon R., Tan V.Y., Yarmolinsky J., Shihab H.A., Timpson N.J., Evans D.M., Relton C., Martin R.M., Davey Smith G., Gaunt T.R., Haycock P.C. The MR-Base platform supports systematic causal inference across the human phenotype. *Elife*. 2018;7:e34408. DOI 10.7554/elifesciences.34408.
- Huffman J.E., Knezevic A., Vitart V., Kattla J., Adamczyk B., Novokmet M., Igl W., Pucic M., Zgaga L., Johannson Å., Redzic I., Gornik O., Zemunik T., Polasek O., Kolcic I., Pehlic M., Koeleman C.A.M., Campbell S., Wild S.H., Hastie N.D., Campbell H., Gyllenstein U., Wuhler M., Wilson J.F., Hayward C., Rudan I., Rudd P.M., Wright A.F., Lauc G. Polymorphisms in *B3GAT1*, *SLC9A9* and *MGAT5* are associated with variation within the human plasma N-glycome of 3533 European adults. *Hum. Mol. Genet.* 2011;20(24):5000-5011. DOI 10.1093/hmg/ddr414.
- Huffman J.E., Pučić-Baković M., Klarić L., Hennig R., Selman M.H.J., Vučković F., Novokmet M., Krišić J., Borowiak M., Muth T., Polašek O., Razdorov G., Gornik O., Plomp R., Theodoratou E., Wright A.F., Rudan I., Hayward C., Campbell H., Deelder A.M., Reichl U., Aulchenko Y.S., Rapp E., Wuhler M., Lauc G. Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. *Mol. Cell. Proteomics*. 2014;13(6):1598-1610. DOI 10.1074/mcp.M113.037465.

- Jain S., Gautam V., Naseem S. Acute-phase proteins: As diagnostic tool. *J. Pharm. Bioallied. Sci.* 2011;3(1):118-127. DOI 10.4103/0975-7406.76489.
- Kanehisa M., Furumichi M., Tanabe M., Sato Y., Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017;45(D1):D353-D361. DOI 10.1093/nar/gkw1092.
- Kellokumpu S. Golgi pH, ion and redox homeostasis: How much do they really matter? *Front. Cell Dev. Biol.* 2019;7:93. DOI 10.3389/fcell.2019.00093.
- Kennedy A.E., Ozbek U., Dorak M.T. What has GWAS done for HLA and disease associations? *Int. J. Immunogenet.* 2017;44(5):195-211. DOI 10.1111/iji.12332.
- Keser T., Gornik I., Vučković F., Selak N., Pavić T., Lukić E., Gudelj I., Gašparović H., Biočina B., Tilin T., Wennerström A., Männistö S., Salomaa V., Havulinna A., Wang W., Wilson J.F., Charutvedhi N., Perola M., Campbell H., Lauc G., Gornik O. Increased plasma N-glycome complexity is associated with higher risk of type 2 diabetes. *Diabetologia.* 2017;60(12):2352-2360. DOI 10.1007/s00125-017-4426-9.
- Khoury G.A., Baliban R.C., Floudas C.A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* 2011;1:90. DOI 10.1038/srep00090.
- Klarić L., Tsepilov Y.A., Stanton C.M., Mangino M., Sikka T.T., Esko T., Pakhomov E., ... Wilson J.F., Zoldoš V., Vitart V., Spector T., Aulchenko Y.S., Lauc G., Hayward C. Glycosylation of immunoglobulin G is regulated by a large network of genes pleiotropic with inflammatory diseases. *Sci. Adv.* 2020;6(8):eaax0301. DOI 10.1126/sciadv.aax0301.
- Knezević A., Polasek O., Gornik O., Rudan I., Campbell H., Hayward C., Wright A., Kolcic I., O'Donoghue N., Bones J., Rudd P.M., Lauc G. Variability, heritability and environmental determinants of human plasma N-glycome. *J. Proteome Res.* 2009;8(2):694-701. DOI 10.1021/pr800737u.
- Kukuruzinska M.A., Lennon K. Protein N-glycosylation: molecular genetics and functional significance. *Crit. Rev. Oral Biol. Med.* 1998; 9(4):415-448. DOI 10.1177/10454411980090040301.
- Landini A., Trbojević-Akmačić I., Navarro P., Tsepilov Y.A., Shara-pov S.Z., Vučković F., Polašek O., Hayward C., Petrović T., Vilaj M., Aulchenko Y.S., Lauc G., Wilson J.F., Klarić L. Genetic regulation of post-translational modification of two distinct proteins. *Nat. Commun.* 2022;13(1):1586. DOI 10.1038/s41467-022-29189-5.
- Lauc G., Essafi A., Huffman J.E., Hayward C., Knežević A., Kattla J.J., Polašek O., ... Gyllensten U., Wilson J.F., Wright A.F., Hastie N.D., Campbell H., Rudd P.M., Rudan I. Genomics meets glycomics-the first GWAS study of human N-glycome identifies HNF1 $\alpha$  as a master regulator of plasma protein fucosylation. *PLoS Genet.* 2010a; 6(12):e1001256. DOI 10.1371/journal.pgen.1001256.
- Lauc G., Huffman J.E., Pučić M., Zgaga L., Adamczyk B., Mužinić A., Novokmet M., ... Wuhrer M., Wright A.F., Rudd P.M., Hayward C., Aulchenko Y., Campbell H., Rudan I. Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet.* 2013; 9(1):e1003225. DOI 10.1371/journal.pgen.1003225.
- Lauc G., Pezer M., Rudan I., Campbell H. Mechanisms of disease: The human N-glycome. *Biochim. Biophys. Acta.* 2016;1860(8):1574-1582. DOI 10.1016/j.bbagen.2015.10.016.
- Lauc G., Rudan I., Campbell H., Rudd P.M. Complex genetic regulation of protein glycosylation. *Mol. Biosyst.* 2010b;6(2):329-335. DOI 10.1039/b910377e.
- Lombard V., Golaconda Ramulu H., Drula E., Coutinho P.M., Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014;42(D1):D490-D495. DOI 10.1093/nar/gkt1178.
- Marke R., van Leeuwen F.N., Scheijen B. The many faces of IKZF1 in B-cell precursor acute lymphoblastic leukemia. *Haematologica.* 2018;103(4):565-574. DOI 10.3324/haematol.2017.185603.
- Martinvalet D. The role of the mitochondria and the endoplasmic reticulum contact sites in the development of the immune responses. *Cell Death Dis.* 2018;9(3):336. DOI 10.1038/s41419-017-0237-7.
- McLaren W., Gil L., Hunt S.E., Riat H.S., Ritchie G.R.S., Thormann A., Flicek P., Cunningham F. The ensembl variant effect predictor. *Genome Biol.* 2016;17(1):122. DOI 10.1186/s13059-016-0974-4.
- Mehta A., Herrera H., Block T. Glycosylation and liver cancer. *Adv. Cancer Res.* 2015;126:257-279. DOI 10.1016/bs.acr.2014.11.005.
- Mijakovac A., Miškec K., Krištić J., Vičić Bočkor V., Tadić V., Bošković M., Lauc G., Zoldoš V., Vojta A. A transient expression system with stably integrated CRISPR-dCas9 fusions for regulation of genes involved in immunoglobulin G glycosylation. *CRISPR J.* 2022;5(2):237-253. DOI 10.1089/crispr.2021.0089.
- Mizushima T., Yagi H., Takemoto E., Shibata-Koyama M., Isoda Y., Iida S., Masuda K., Satoh M., Kato K. Structural basis for improved efficacy of therapeutic antibodies on defucosylation of their Fc glycoforms. *Genes Cells.* 2011;16(11):1071-1080. DOI 10.1111/j.1365-2443.2011.01552.x.
- Moayyeri A., Hammond C.J., Hart D.J., Spector T.D. The UK adult twin registry (TwinsUK resource). *Twin Res. Hum. Genet.* 2013; 16(1):144-149. DOI 10.1017/thg.2012.89.
- Mohanty S., Chaudhary B., Zoetewey D. Structural insight into the mechanism of N-linked glycosylation by oligosaccharyltransferase. *Biomolecules.* 2020;10(4):624. DOI 10.3390/biom10040624.
- Moremen K.W., Tiemeyer M., Nairn A.V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* 2012;13(7):448-462. DOI 10.1038/nrm3383.
- Mulloy B., Dell A., Stanley P., H. Prestegard J. Structural analysis of glycans. In: Varki A., Cummings R.D., Esko J.D., Stanley P., Hart G.W., Aebi M., Darvill A.G., Kinoshita T., Packer N.H., Prestegard J.H., Schnaar R.L., Seeberger P.H. (Eds.). *Essentials of Glycobiology*. [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2015. Chapter 50, 2017. DOI 10.1101/glycobiology.3e.050.
- Nairn A.V., Aoki K., dela Rosa M., Porterfield M., Lim J.-M., Kulik M., Pierce J.M., Wells L., Dalton S., Tiemeyer M., Moremen K.W. Regulation of glycan structures in murine embryonic stem cells. *J. Biol. Chem.* 2012;287(45):37835-37856. DOI 10.1074/jbc.M112.405233.
- Nairn A.V., York W.S., Harris K., Hall E.M., Pierce J.M., Moremen K.W. Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J. Biol. Chem.* 2008;283(25):17298-17313. DOI 10.1074/jbc.M801964200.
- Nikolac Perkovic M., Pucic Bakovic M., Kristic J., Novokmet M., Huffman J.E., Vitart V., Hayward C., Rudan I., Wilson J.F., Campbell H., Polasek O., Lauc G., Pivac N. The association between galactosylation of immunoglobulin G and body mass index. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* 2014;48:20-25. DOI 10.1016/j.pnpbp.2013.08.014.
- Oda Y., Okada T., Yoshida H., Kaufman R.J., Nagata K., Mori K. Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation. *J. Cell Biol.* 2006;172(3):383-393. DOI 10.1083/jcb.200507057.
- Ohtsubo K., Marth J.D. Glycosylation in cellular mechanisms of health and disease. *Cell.* 2006;126(5):855-867. DOI 10.1016/j.cell.2006.08.019.
- Pasaniuc B., Price A.L. Dissecting the genetics of complex traits using summary association statistics. *Nat. Rev. Genet.* 2017;18(2):117-127. DOI 10.1038/nrg.2016.142.
- Peipp M., Lammerts van Bueren J.J., Schneider-Merck T., Bleeker W.W.K., Dechant M., Beyer T., Repp R., van Berkel P.H.C., Vink T., van de Winkel J.G.J., Parren P.W.H.I., Valerius T. Antibody fucosylation differentially impacts cytotoxicity mediated by NK and PMN effector cells. *Blood.* 2008;112(6):2390-2399. DOI 10.1182/blood-2008-03-144600.
- Pers T.H., Karjalainen J.M., Chan Y., Westra H.-J., Wood A.R., Yang J., Lui J.C., Vedantam S., Gustafsson S., Esko T., Frayling T., Speliotes E.K., Boehnke M., Raychaudhuri S., Fehrmann R.S.N.,

- Hirschhorn J.N., Franke L. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* 2015;6(1):5890. DOI 10.1038/ncomms6890.
- Poole J., Day C.J., von Itzstein M., Paton J.C., Jennings M.P. Glycointeractions in bacterial pathogenesis. *Nat. Rev. Microbiol.* 2018; 16(7):440-452. DOI 10.1038/s41579-018-0007-2.
- Pottier N., Cheok M.H., Yang W., Assem M., Tracey L., Obenaus J.C., Panetta J.C., Relling M.V., Evans W.E. Expression of SMARCB1 modulates steroid sensitivity in human lymphoblastoid cells: identification of a promoter SNP that alters PARP1 binding and SMARCB1 expression. *Hum. Mol. Genet.* 2007;16(19):2261-2271. DOI 10.1093/hmg/ddm178.
- Reiding K.R., Bondt A., Hennig R., Gardner R.A., O'Flaherty R., Trbojević-Akmačić I., Shubhakar A., Hazes J.M.W., Reichl U., Fernandes D.L., Pučić-Baković M., Rapp E., Spencer D.I.R., Dolhain R.J.E.M., Rudd P.M., Lauc G., Wuhrer M. High-throughput serum N-glycomics: method comparison and application to study rheumatoid arthritis and pregnancy-associated changes. *Mol. Cell. Proteomics.* 2019;18(1):3-15. DOI 10.1074/mcp.RA117.000454.
- Reily C., Stewart T.J., Renfrow M.B., Novak J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* 2019;15(6):346-366. DOI 10.1038/s41581-019-0129-4.
- Rivinoja A., Hassinen A., Kokkonen N., Kaupilla A., Kellokumpu S. Elevated Golgi pH impairs terminal N-glycosylation by inducing mislocalization of Golgi glycosyltransferases. *J. Cell. Physiol.* 2009; 220(1):144-154. DOI 10.1002/jcp.21744.
- Rogers M., Shihab H.A., Mort M., Cooper D., Gaunt T.R., Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics.* 2018;34(3):511-513. DOI 10.1093/bioinformatics/btx536.
- Russell A.C., Šimurina M., Garcia M.T., Novokmet M., Wang Y., Rudan I., Campbell H., Lauc G., Thomas M.G., Wang W. The N-glycosylation of immunoglobulin G as a novel biomarker of Parkinson's disease. *Glycobiology.* 2017;27(5):501-510. DOI 10.1093/glycob/cwx022.
- Saito M., Ishii A. T3Gal-V (GM3 synthase, SAT-I). In: Handbook of Glycosyltransferases and Related Genes. Tokyo: Springer, 2002; 289-294. DOI 10.1007/978-4-431-67877-9\_39.
- Saldova R., Asadi Shehni A., Haakensen V.D., Steinfeld I., Hilliard M., Kifer I., Helland A., Yakhini Z., Børresen-Dale A.-L., Rudd P.M. Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC. *J. Proteome Res.* 2014;13(5):2314-2327. DOI 10.1021/pr401092y.
- Sellers M., Reina-San-Martin B., Kastner P., Chan S. Ikaros controls isotype selection during immunoglobulin class switch recombination. *J. Exp. Med.* 2009;206(5):1073-1087. DOI 10.1084/jem.2008.2311.
- Shadrina A.S., Zlobin A.S., Zaytseva O.O., Klarić L., Sharapov S.Z., Pakhomov E., Perola M., Esko T., Hayward C., Wilson J.F., Lauc G., Aulchenko Y.S., Tsepilov Y.A. Multivariate genome-wide analysis of immunoglobulin G N-glycosylation identifies new loci pleiotropic with immune function. *Hum. Mol. Genet.* 2021;30(13):1259-1270. DOI 10.1093/hmg/ddab072.
- Sharapov S.Z., Shadrina A.S., Tsepilov Y.A., Elgaeva E.E., Tiys E.S., Feoktistova S.G., Zaytseva O.O., ... Dagostino C., Gieger C., Allegri M., Williams F., Schulze M.B., Lauc G., Aulchenko Y.S. Replication of fifteen loci involved in human plasma protein N-glycosylation in 4,802 samples from four cohorts. *Glycobiology.* 2020; 31(2):82-88. DOI 10.1093/glycob/cwaa053.
- Sharapov S.Z., Tsepilov Y.A., Klarić L., Mangino M., Thareja G., Shadrina A.S., Simurina M., ... Louis E., Georges M., Suhre K., Spector T., Williams F.M.K., Lauc G., Aulchenko Y.S. Defining the genetic control of human blood plasma N-glycome using genome-wide association study. *Hum. Mol. Genet.* 2019;28(12):2062-2077. DOI 10.1093/hmg/ddz054.
- Shen X., Klarić L., Sharapov S., Mangino M., Ning Z., Wu D., Trbojević-Akmačić I., Pučić-Baković M., Rudan I., Polašek O., Hayward C., Spector T.D., Wilson J.F., Lauc G., Aulchenko Y.S. Multivariate discovery and replication of five novel loci associated with Immunoglobulin G N-glycosylation. *Nat. Commun.* 2017;8(1):447. DOI 10.1038/s41467-017-00453-3.
- Skropeta D. The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.* 2009;17(7):2645-2653. DOI 10.1016/j.bmc.2009.02.037.
- Spector T.D., Williams F.M.K. The UK adult twin registry (TwinsUK). *Twin Res. Hum. Genet.* 2006;9(6):899-906. DOI 10.1375/183242706779462462.
- Staley J.R., Blackshaw J., Kamat M.A., Ellis S., Surendran P., Sun B.B., Paul D.S., Freitag D., Burgess S., Danesh J., Young R., Butterworth A.S. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics.* 2016;32(20):3207-3209. DOI 10.1093/bioinformatics/btw373.
- Staretz-Chacham O., Noyman I., Wormser O., Abu Quider A., Hazan G., Morag I., Hadar N., Raymond K., Birk O.S., Ferreira C.R., Koifman A. B4GALT1-congenital disorders of glycosylation: Expansion of the phenotypic and molecular spectrum and review of the literature. *Clin. Genet.* 2020;97(6):920-926. DOI 10.1111/cge.13735.
- Taniguchi N., Honke K., Fukuda M., Narimatsu H., Yamaguchi Y., Angata T. (Eds.). Handbook of Glycosyltransferases and Related Genes. Tokyo: Springer, 2014. DOI 10.1007/978-4-431-54240-7.
- Taniguchi N., Kizuka Y. Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. *Adv. Cancer Res.* 2015;126:11-51. DOI 10.1016/bs.acr.2014.11.001.
- Thanabalasingham G., Huffman J.E., Kattla J.J., Novokmet M., Rudan I., Gloyd A.L., Hayward C., ... Hastie N.D., Campbell H., McCarthy M.I., Rudd P.M., Owen K.R., Lauc G., Wright A.F. Mutations in HNF1A result in marked alterations of plasma glycan profile. *Diabetes.* 2013;62(4):1329-1337. DOI 10.2337/db12-0880.
- Theodoratou E., Thaçi K., Agakov F., Timofeeva M.N., Štambuk J., Pučić-Baković M., Vučković F., Orchard P., Agakova A., Din F.V.N., Brown E., Rudd P.M., Farrington S.M., Dunlop M.G., Campbell H., Lauc G. Glycosylation of plasma IgG in colorectal cancer prognosis. *Sci. Rep.* 2016;6:28098. DOI 10.1038/srep28098.
- Tillin T., Forouhi N.G., McKeigue P.M., Chaturvedi N., SABRE Study Group. Southall and Brent REvisited: Cohort profile of SABRE, a UK population-based comparison of cardiovascular disease and diabetes in people of European, Indian Asian and African Caribbean origins. *Int. J. Epidemiol.* 2012;41(1):33-42. DOI 10.1093/ije/dyq175.
- Trbojević Akmačić I., Ventham N.T., Theodoratou E., Vučković F., Kennedy N.A., Krištić J., Nimmo E.R., Kalla R., Drummond H., Štambuk J., Dunlop M.G., Novokmet M., Aulchenko Y., Gornik O., Campbell H., Pučić Baković M., Satsangi J., Lauc G. Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin G glycome. *Inflamm. Bowel Dis.* 2015;21(6):1237-1247. DOI 10.1097/mib.0000000000000372.
- Uhlén M., Fagerberg L., Hallström B.M., Lindskog C., Oksvold P., Mardinoglu A., Sivertsson Å., ... Forsberg M., Persson L., Johansson F., Zwahlen M., von Heijne G., Nielsen J., Pontén F. Proteomics. Tissue-based map of the human proteome. *Science.* 2015; 347(6220):1260419. DOI 10.1126/science.1260419.
- Uhlén M., Karlsson M.J., Hober A., Svensson A.-S., Scheffel J., Kottol D., Zhong W., ... Voldborg B.G., Tegel H., Hober S., Forsström B., Schwenk J.M., Fagerberg L., Sivertsson Å. The human secretome. *Sci. Signal.* 2019;12(609):eaaz0274. DOI 10.1126/scisignal.aaz0274.
- Varki A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology.* 1993;3(2):97-130. DOI 10.1093/glycob/3.2.97.
- Varki A., Cummings R.D., Esko J.D., Stanley P., Hart G.W., Aebi M., Darvill A.G., Kinoshita T., Packer N.H., Prestegard J.H., Schnaar R.L., Seeberger P.H. (Eds.). Essentials of Glycobiology. [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2017.
- Varki A., Kornfeld S. Historical background and overview. In: Varki A., Cummings R.D., Esko J.D., Stanley P., Hart G.W., Aebi M., Dar-

- vill A.G., Kinoshita T., Packer N.H., Prestegard J.H., Schnaar R.L., Seeberger P.H. (Eds.). *Essentials of Glycobiology*. [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2017.
- Vilaj M., Lauc G., Trbojević-Akmačić I. Evaluation of different PNGase F enzymes in immunoglobulin G and total plasma N-glycans analysis. *Glycobiology*. 2021;31(1):2-7. DOI 10.1093/glycob/cwaa047.
- Visscher P.M., Brown M.A., McCarthy M.I., Yang J. Five years of GWAS discovery. *Am. J. Hum. Genet.* 2012;90(1):7-24. DOI 10.1016/j.ajhg.2011.11.029.
- Visscher P.M., Wray N.R., Zhang Q., Sklar P., McCarthy M.I., Brown M.A., Yang J. 10 years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* 2017;101(1):5-22. DOI 10.1016/j.ajhg.2017.06.005.
- Vreeker G.C.M., Nicolardi S., Bladergroen M.R., van der Plas C.J., Mesker W.E., Tollenaar R.A.E.M., van der Burgt Y.E.M., Wuhrer M. Automated plasma glycomics with linkage-specific sialic acid esterification and ultrahigh resolution MS. *Anal. Chem.* 2018;90(20):11955-11961. DOI 10.1021/acs.analchem.8b02391.
- Vučković F., Theodoratou E., Thači K., Timofeeva M., Vojta A., Štambuk J., Pučić-Baković M., Rudd P.M., Đerek L., Servis D., Wennerström A., Farrington S.M., Perola M., Aulchenko Y., Dunlop M.G., Campbell H., Lauc G. IgG glycome in colorectal cancer. *Clin. Cancer Res.* 2016;22(12):3078-3086. DOI 10.1158/1078-0432.CCR-15-1867.
- Wahl A., van den Akker E., Klaric L., Štambuk J., Benedetti E., Plomp R., Razdorov G., Trbojević-Akmačić I., Deelen J., van Heemst D., Slagboom P.E., Vučković F., Grallert H., Krumsiek J., Strauch K., Peters A., Meitinger T., Hayward C., Wuhrer M., Beekman M., Lauc G., Gieger C. Genome-wide association study on immunoglobulin G glycosylation patterns. *Front. Immunol.* 2018;9:277. DOI 10.3389/fimmu.2018.00277.
- Wang Y., Klarić L., Yu X., Thači K., Dong J., Novokmet M., Wilson J., Polasek O., Liu Y., Krištić J., Ge S., Pučić-Baković M., Wu L., Zhou Y., Ugrina I., Song M., Zhang J., Guo X., Zeng Q., Rudan I., Campbell H., Aulchenko Y., Lauc G., Wang W. The association between glycosylation of immunoglobulin G and hypertension. *Medicine (Baltimore)*. 2016;95(17):e3379. DOI 10.1097/md.0000000000003379.
- Winkler T.W., Day F.R., Croteau-Chonka D.C., Wood A.R., Locke A.E., Mägi R., Ferreira T., Fall T., Graff M., Justice A.E., Luan J., Gustafsson S., Randall J.C., Vedantam S., Workalemahu T., Kilpeläinen T.O., Scherag A., Esko T., Kutalik Z., Heid I.M., Loos R.J.F., Genetic Investigation of Anthropometric Traits (GIANT) Consortium. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* 2014;9(5):1192-1212. DOI 10.1038/nprot.2014.071.
- Yang J., Ferreira T., Morris A.P., Medland S.E., Genetic Investigation of Anthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium, Madden P.A.F., Heath A.C., Martin N.G., Montgomery G.W., Weedon M.N., Loos R.J., Frayling T.M., McCarthy M.I., Hirschhorn J.N., Goddard M.E., Visscher P.M. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* 2012;44(4):369-375. S1-3. DOI 10.1038/ng.2213.
- Zaytseva O.O., Freidin M.B., Keser T., Štambuk J., Ugrina I., Šimurina M., Vilaj M., Štambuk T., Trbojević-Akmačić I., Pučić-Baković M., Lauc G., Williams F.M.K., Novokmet M. Heritability of human plasma N-glycome. *J. Proteome Res.* 2020;19(1):85-91. DOI 10.1021/acs.jproteome.9b00348.
- Zhu Z., Zhang F., Hu H., Bakshi A., Robinson M.R., Powell J.E., Montgomery G.W., Goddard M.E., Wray N.R., Visscher P.M., Yang J. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* 2016;48(5):481-487. DOI 10.1038/ng.3538.

#### ORCID ID

S.Zh. Sharapov [orcid.org/0000-0003-0279-4900](https://orcid.org/0000-0003-0279-4900)  
A.N. Timoshchuk [orcid.org/0000-0001-8529-4498](https://orcid.org/0000-0001-8529-4498)  
Y.S. Aulchenko [orcid.org/0000-0002-7899-1575](https://orcid.org/0000-0002-7899-1575)

**Благодарности.** Исследование осуществлено в рамках Программы исследований Института искусственного интеллекта при МГУ. Работа Ю.А. частично поддержана бюджетным проектом № FWNR-2022-0020.

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

Поступила в редакцию 17.11.2022. После доработки 20.01.2023. Принята к публикации 23.01.2023.