

**БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ ОДНОНУКЛЕОТИДНЫХ
ВАРИАНТОВ ГЕНА F12, АССОЦИИРОВАННЫХ С
НАСЛЕДСТВЕННЫМ АНГИОТЕКОМ**

Седых А. В. ¹,
Останкова Ю. В. ¹,
Щемелев А. Н. ¹,
Тотолян А. А. ^{1,2}

¹ ФБУН «Санкт-Петербургский НИИ эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека (Санкт-Петербург, Россия).

² ФГБОУ ВО «Первый Санкт-Петербургский государственный медицинский университет имени академика И.П. Павлова» Министерства здравоохранения РФ (Санкт-Петербург, Россия).

**BIOINFORMATIC ANALYSIS OF SINGLE NUCLEOTIDE VARIANTS IN
THE F12 GENE ASSOCIATED WITH HEREDITARY ANGIOEDEMA**

Sedykh A. V. ^a,
Ostankova Yu. V. ^a,
Schemelev A. N. ^a,
Totolian A. A. ^{a, b}

^a Saint-Petersburg Pasteur Institute, Saint Petersburg, Russia.

^b First St. Petersburg State I. Pavlov Medical University, Saint Petersburg, Russia.

Резюме

Наследственный ангиоотек - генетически детерминированное заболевание, которое относится к первичным иммунодефицитам с нарушением системы комплимента. Чаще всего у пациентов заболевание характеризуется дефицитом С1-ингибитора (НАО I типа) или нарушением функциональной активности С1-ингибитора (НАО II типа). В таких случаях ставят диагноз на основании лабораторных показателей. При НАО с нормальным уровнем и активностью С1-ингибитора диагноз может быть поставлен только на основе семейного анамнеза и/или генетического анализа. У пациентов с НАО с нормальным С1-ингибитором наиболее часто встречаются мутации в гене *F12*, особенно у женщин. Однако нередко выявляются мутации, которые не имеют точного клинического значения. учитывая ограниченное количество больных НАО, невозможно экспериментально определить клиническое значение вновь обнаруживаемых полиморфных вариантов. Решением этой проблемы может стать анализ *in silico* каждого нового полиморфизма.

Целью нашей работы было оценить предсказательную возможность методов биоинформатического анализа при оценке полиморфных вариантов в гене *F12*.

Материалом служили четыре полиморфных варианта NC_000005.9:g.176831285C>G, NC_000005.9:g.176831258C>G, NC_000005.9:g.176831232G>C, NC_000005.9:g.176831232G>T с разным статусом клинической значимости. Для предсказания эффекта полиморфных вариантов на белок F12 были использованы различные веб-ресурсы, основанные на разных алгоритмах (SIFT, Polyphen-2, FATHMM-XF, MutationTaster2021, MutPred2, MUpro и I-Mutant 2, HOPE, ChimeraX).

Результаты. *In silico* анализ показал, что мутации NC_000005.9:g.176831232G>C (p.Thr328Arg) и NC_000005.9:g.176831232G>T (p.Thr328Lys) обладают патогенным эффектом, что полностью соответствует их ранее установленному клиническому статусу. В то же время, полиморфные варианты NC_000005.9:g.176831258C>G (p.Gln319His) и NC_000005.9:g.176831285C>G (p.Arg310Ser), по-видимому, не являются самостоятельной причиной заболевания, однако не исключается их возможное участие в модификации клинического фенотипа.

Биоинформатический анализ играет ключевую роль в предварительной оценке значимости вновь выявленных мутаций в гене *F12*, способствует более точному поиску патогенных вариантов. Включение биоинформатических методов в диагностику необходимо при выявлении причины заболевания у пациентов с НАО с нормальным уровнем и функциональной активностью С1-ингибитора.

Ключевые слова: первичный иммунодефицит, наследственный ангиоотек, биоинформатический анализ, *F12*, полиморфные варианты.

Abstract

Hereditary angioedema (HAE) is a genetically determined disorder classified as a primary immunodeficiency involving complement system dysfunction. In most patients, the disease is characterized by a deficiency of C1 inhibitor (type I HAE) or impaired functional activity of the C1 inhibitor (type II HAE). In such cases, the diagnosis is based on laboratory findings. In HAE with normal C1 inhibitor levels and activity, the diagnosis can only be established based on family history and/or genetic testing. Among patients with HAE with normal C1 inhibitor, mutations in the *F12* gene are most frequently observed, particularly in women. However, mutations with uncertain clinical significance are often identified. Given the limited number of HAE cases, it is not feasible to experimentally determine the clinical relevance of newly discovered polymorphic variants. A potential solution to this problem is the *in silico* analysis of each novel polymorphism.

The aim of our study was to evaluate the predictive potential of bioinformatic analysis methods in assessing polymorphic variants in the *F12* gene.

The study focused on four polymorphic variants — NC_000005.9:g.176831285C>G, NC_000005.9:g.176831258C>G, NC_000005.9:g.176831232G>C, and NC_000005.9:g.176831232G>T — with varying clinical significance statuses. To predict the effect of these polymorphic variants on the F12 protein, various web-based tools employing different algorithms were used, including SIFT, PolyPhen-2, FATHMM-XF, MutationTaster2021, MutPred2, MUpro, I-Mutant 2, HOPE, and ChimeraX.

Results. *In silico* analysis demonstrated that the mutations NC_000005.9:g.176831232G>C (p.Thr328Arg) and NC_000005.9:g.176831232G>T (p.Thr328Lys) have a pathogenic effect, which is fully consistent with their previously established clinical status. At the same time, the polymorphic variants NC_000005.9:g.176831258C>G (p.Gln319His) and NC_000005.9:g.176831285C>G (p.Arg310Ser) do not appear to be independent causes of the disease, although their potential role in modifying the clinical phenotype cannot be excluded.

Bioinformatic analysis plays a key role in the preliminary assessment of the significance of newly identified mutations in the *F12* gene and facilitates a more precise identification of pathogenic variants. The integration of bioinformatic tools into diagnostic workflows is essential for determining the cause of disease in patients with hereditary angioedema who present with normal levels and functional activity of C1 inhibitor.

Keywords: primary immunodeficiency, hereditary angioedema, bioinformatic analysis, *F12*, polymorphic variants.

1 Введение

Первичные иммунодефициты (ПИД) или врожденные ошибки иммунитета являются генетически детерминированной группой редких заболеваний, которые приводят к нарушениям иммунной системы, повышенному риску инфекций, злокачественных новообразований и аутоиммунных состояний. В настоящее время известно более 450 различных типов ПИД, которые, как правило, впервые проявляются в раннем возрасте и различаются по степени поражения компонентов иммунной системы [3]. В соответствии с современной классификацией первичные иммунодефициты разделяют на заболевания гуморального, клеточного иммунитета, с нарушениями цитокинов, с дефектами фагоцитарной системы, комбинированные иммунодефициты, аутовоспалительные заболевания, а также заболевания с нарушениями в системе комплимента и другие редкие формы [24]. К заболеваниям с дисфункцией в системе комплемента относят наследственный ангиоотек (НАО).

Наследственный ангиоотек – редкое генетически детерминированное жизнеугрожающее заболевание с аутосомно-доминантным типом наследования, проявляющееся в виде отёков кожи и слизистых/подслизистых оболочек, которые возникают из-за гиперактивации брадикинина. Клинические проявления заболевания индивидуальны для каждого пациента, но чаще всего для заболевания характерны эпизодические отеки, поражающие различные части тела, верхние дыхательные пути, конечности, половые органы и желудочно-кишечный тракт [1]. Отёки ЖКТ часто имитируют острые состояния брюшной полости, что приводит к необоснованным оперативным вмешательствам. НАО может быть потенциально опасным для жизни, так как отеки верхних дыхательных путей могут привести к асфиксии. Как правило, дебют симптомов заболевания начинается в детстве или в период полового созревания. Триггерами могут быть как физические травмы, так и психоэмоциональные потрясения. Медицинские вмешательства, например, хирургические манипуляции или стоматологическое лечение, также могут провоцировать отек [13].

На сегодняшний день НАО классифицируют по лабораторным показателям на несколько типов в зависимости от изменений уровней или функции С1-ингибитора (С1-INH) и генетической основы. НАО I типа обусловлен дефицитом С1-INH в плазме, НАО II типа характеризуется снижением функциональной активности С1-INH, при этом уровень С1-INH сохраняется в пределах нормы или повышен [2]. Оба типа НАО (I и II) связаны с мутациями в гене *SERPING1*, который кодирует ингибитор плазменной протеазы С1. Причем мутации, приводящие к НАО I типа, рассредоточены по всей протяженности гена *SERPING1*, в то время как мутации, приводящие к II типу НАО, преимущественно локализованы вокруг петли белкового реактивного центра (RCL), за единственным исключением: мутации

43 аминокислотного остатка Lys251, влияющей на функциональность после
 44 фолдинга белка. В соответствие с текущими клиническими рекомендациями
 45 НАО с нормальным уровнем и функциональным С1-INH, ранее известный как
 46 НАО III типа, предлагают классифицировать по конкретным генам, которые
 47 участвуют в патогенезе НАО, таким как *F12*, *ANGPT*, *PLG*, *KNG1*, *MYOF*,
 48 *HS3ST6* [8]. Фенотипическая картина всех форм НАО с нормальным уровнем
 49 и функциональным С1-ИНГ очень разнообразна и может отличаться не только
 50 у разных пациентов, но и у одного пациента на протяжении жизни, определить
 51 конкретный тип можно только при проведении генетического анализа [19].
 52 Однако до сих пор у большинства пациентов с нормальным уровнем и
 53 функциональным С1-ИНГ генетическая причина заболевания остается
 54 неизвестной (НАО-UNKNOWN).

55 Из-за вариабельности симптомов НАО, которые могут совпадать с
 56 симптомами других состояний, а также из-за достаточно редкой
 57 встречаемости заболевания, большинство пациентов продолжительное время
 58 живут с неверным диагнозом. Следствием этого является значительная
 59 задержка в установлении диагноза с момента появления первых симптомов
 60 [36]. Согласно новому проекту клинических рекомендаций, пациентам с
 61 симптомами НАО с нормальным уровнем и функциональным С1-ИНГ на
 62 первом этапе диагностики предлагается провести генетическое тестирование
 63 частых клинически-значимых мутаций в генах *F12* и *PLG*, при их отсутствии
 64 рекомендуется проведение полноэкзомного скрининга [4].

65 Как известно, найденные полиморфные варианты гена *F12*,
 66 кодирующего белок фактор Хагемана (Фактор XII, FXII), стали первыми
 67 биомаркерами НАО с нормальным уровнем и функциональным С1-ИНГ.
 68 Фактор XII является элементом системы крови, участвует во многих
 69 биохимических процессах, отвечая за коагуляцию и образование брадикинина
 70 [14]. Однако вклад различных полиморфных вариантов гена *F12* в развитие
 71 НАО-FXII до сих пор остается мало изучен. Не все из найденных
 72 полиморфных вариантов гена *F12* имеют одинаковое клиническое значение и
 73 влияют на развитие НАО, они могут быть как нейтральными, так и
 74 функционально значимыми, влияя на риск развития различных заболеваний
 75 [7]. Довольно низкая встречаемость заболевания препятствует накоплению
 76 достаточного объема статистически значимых данных. Кроме того, в случаях
 77 с НАО, найденные полиморфные варианты могут быть уникальными для
 78 каждого пациента, что усложняет интерпретацию их клинической значимости.
 79 Проблема усугубляется генетической гетерогенностью пациентов, что
 80 является причиной отсутствия крупных когортных исследований,
 81 объединяющих пациентов с одинаковыми мутациями для установления
 82 четкой корреляции между генотипом и фенотипом. Одним из подходов
 83 интерпретации полиморфных вариантов с неизвестным клиническим
 84 значением является биоинформационный анализ. В клинической практике
 85 биоинформатический анализ найденных мутаций может быть важным этапом

86 в ранней диагностике генетических заболеваний. Он может помочь установить
 87 причинно-следственную связь между генетическим вариантом и
 88 клиническими проявлениями болезни у конкретного пациента. Различные
 89 алгоритмы предсказания могут помочь определить мутации, связанные с
 90 повышенным риском развития определенных заболеваний, способствуя
 91 профилактическим мерам и ранней диагностике.

92 **Цель.** Целью нашей работы было оценить предсказательную
 93 возможность методов биоинформатического анализа при оценке
 94 полиморфных вариантов в гене *F12*.

95 2 Материалы и методы

96 Для биоинформатического анализа были выбраны четыре полиморфных
 97 варианта гена *F12*, обладающих разным статусом клинической значимости
 98 (табл.1).

99 Для предсказания эффекта полиморфных вариантов на белок F12 были
 100 использованы различные веб-ресурсы, основанные на разных алгоритмах.

101 *Веб-приложения для предсказания патогенности однонуклеотидных*
 102 *замен*

103 SIFT (Sorting Intolerant From Tolerant) – биоинформатический
 104 инструмент, предназначенный для прогнозирования влияния аминокислотных
 105 замен на функцию белка. Основным алгоритмом программы SIFT является
 106 множественное выравнивание белковых последовательностей, что позволяет
 107 оценить консервативность аминокислот в эволюционно родственных белках.
 108 Для каждой аминокислотной позиции в нативном белке моделируется замена
 109 на другие возможные аминокислоты, если значение замененной
 110 аминокислоты ниже или равно 0,05, такая мутация классифицируется как
 111 «вредная», если выше – как «нейтральная». Преимуществом программы
 112 является отсутствие необходимости в структуре белка, однако из-за анализа,
 113 основанного исключительно на последовательности белка, алгоритм
 114 допускает 20% ложноположительных результатов [23, 31].

115 Polyphen-2 (Polymorphism Phenotyping v2) – веб-сервер для предсказания
 116 возможного влияния полиморфного варианта на структуру и функцию белка.
 117 Прогнозирование основано на данных об аминокислотной
 118 последовательности, филогенетических и структурных признаков,
 119 характеризующих замену. В качестве основы для своей работы Polyphen-2
 120 использует базы данных белков, включая UniProt и PDB. Варианты
 121 классифицируются по трем категориям: от 0 до 0,2 баллов –
 122 доброкачественные, от 0,2 до 0,85 баллов – возможно патогенные, и от 0,85 до
 123 1 балла – с высокой вероятностью патогенные [5].

124 FATHMM- XF (Functional Analysis through Hidden Markov Models -
 125 eXtended Features) - расширенная версия инструмента прогнозирования
 126 FATHMM, которая основана на статистической модели машинного обучения
 127 (скрытых марковских моделях). Веб-ресурс включает в себя анализ
 128 некодирующих кодирующих областей генома. FATHMM – XF сопоставляет
 129 результаты исследований различных баз данных, связанных со структурной
 130 организацией белков, их экспрессией в различных тканях, эволюционными
 131 данными, геномными особенностями. Классификация вариантов делится на
 132 нейтральные/безвредные и патогенные. Результат, где р-значение выше 0,5
 133 указывает на вредное влияние мутации, ниже 0,5 – скорее имеет нейтральное
 134 значение [28].

135 MutationTaster2021 – обновленная версия программы MutationTaster,
 136 основанная на машинном обучении, которая использует алгоритм «Random
 137 Forest», вместо старой модели «Naive Bayes». Алгоритм путем
 138 множественного выравнивания анализирует изменчивость участка
 139 исследуемого гена в различных организмах, учитывает изменение в структуре
 140 белка, оценивает влияние полиморфизма на клинические проявления.
 141 Программа интегрирует информацию о вариантах из баз данных gnomAD,
 142 ClinVar, HGMD Pro, ExAC. Программа нацелена на биомедицинские
 143 исследования для выявления патогенетических мутаций, вызывающих
 144 моногенные заболевания [32].

145 MutPred2 – веб-приложение, разработанное для поиска связи между
 146 заменённой аминокислотой в мутантном белке и развитием определенных
 147 заболеваний. Программа позволяет смоделировать структурные изменения в
 148 белке. Модель патогенности MutPred2 была обучена на наборе из 53 180
 149 патогенных и 206 946 предположительно нейтральных вариантов, полученных
 150 из базы данных HGMD [33], SwissVar [22], dbSNP [26, 30].

151 Программа CADD (Combined Annotation Dependent Depletion) использует
 152 комплексный подход оценки вариантов, найденных как в кодирующей, так и
 153 некодирующей области. CADD анализирует информацию эволюционной
 154 консервативности белка, функциональные данные, эпигенетические
 155 модификации и генетический контекст варианта. Преимуществом CADD
 156 является то, что модель обучается на очень большом наборе данных из разных
 157 баз и программ. Разработчики CADD не предлагают жесткого порогового
 158 значения для определения патогенности, однако исходя из других
 159 исследований значение, превышающее 20, характерно для значимых вредных
 160 вариантов [15, 27].

161 *Веб-приложения для предсказания изменения в стабильности белка*

162 Программы MUpro и I-Mutant 2.0 предсказывают термодинамическую
 163 стабильность белка, используя различные программы машинного обучения, а
 164 также моделирование конформации мутантного белка. MUpro использует

165 машину опорных векторов (SVM) в качестве оценки значения свободной
 166 энергии Гиббса ($\Delta\Delta G$, ккал/моль) и оценки изменения стабильности белка.
 167 Программа работает только с последовательностью белка в формате FASTA,
 168 дополнительно использует информацию о структуре белка. Оценка
 169 достоверности <0 указывает на снижение стабильности белка, а оценка
 170 достоверности >0 указывает на увеличение стабильности белка при мутации
 171 [11].

172 Ресурс I-Mutant 2.0 работает в двух режимах: предсказывает
 173 стабильность на основе последовательности белка или на основе структуры
 174 белка. Положительная оценка предполагает улучшение стабильности белка,
 175 отрицательная - указывает на уменьшение стабильности белка. Значение $\Delta\Delta G$
 176 классифицируется следующим образом: значительно стабильный ($>0,5$
 177 ккал/моль), значительно нестабильный ($<-0,5$ ккал/моль) или нейтральный (от
 178 $-0,5$ до $0,5$ ккал/моль) [10].

179 *Веб-приложения для прогнозирования последствий SNP на 3D-структуру*
 180 *белка*

181 HOPE (Have Our Protein Explained) – веб-сервер, предсказывающий
 182 влияние несинонимичного полиморфного варианта на структуру и функции
 183 белка. В основе программы лежат данные из базы UniProt. Сравнение заданной
 184 белковой последовательности с диким типом приводит к получению 3D-
 185 структуры мутантного белка. HOPE не всегда выдает полные результаты из-за
 186 недостаточных данных о структуре белка [9].

187 ChimeraX – программное обеспечение для молекулярной визуализации,
 188 разработанное Центром биоинформатики Калифорнийского университета в
 189 Сан-Франциско (UCSF). Программа строит атомную структуру белков,
 190 нуклеиновых кислот, мембранных комплексов и других биомолекул с высокой
 191 степенью детализации. ChimeraX интегрируется с внешними базами данных
 192 (например, RCSB PDB, UniProt, AlphaFold), что позволяет загружать
 193 экспериментальные или предсказанные модели прямо из сети [20].

194 3 Результаты

195 Изменение структуры и функционирование мутантного белка прежде
 196 всего связано с физико-химическими свойствами заменённой аминокислоты.
 197 Каждая аминокислота обладает своими характеристиками, такими как размер,
 198 заряд, индекс гидропатии, способность к образованию химических связей, все
 199 в совокупности влияет на конформацию и стабильность белка. В таблице 2
 200 представлены характеристики аминокислот, участвующих в полиморфизме
 201 гена *F12*, исследуемых в рамках настоящей работы.

202

203 Полиморфные варианты гена *F12* были проанализированы при помощи
 204 биоинформатических предикторов HOPE и MutationTaster2021 (табл. 3).
 205 Варианты NC_000005.9:g.176831232G>T (р. Thr328Lys) и
 206 NC_000005.9:g.176831232G>C (р. Thr328Arg) затрагивают сайт
 207 гликозилирования, а также влияют на изменения в пролин богатой области
 208 белка FXII. Такие изменения могут привести к потери реакционной
 209 специфичности белка, снижению аффинности к другим белкам или
 210 кофакторам. Мутация NC_000005.9:g.176831285C>G (р. Arg310Ser) может
 211 приводить к изменениям в ионных взаимодействиях в белке FXII, что в
 212 дальнейшем может повлиять на функциональную активность белка.

213 Чтобы оценить консервативность аминокислотной последовательности
 214 использовали программу MutationTaster2021. Варианты
 215 NC_000005.9:g.176831285C>G и NC_000005.9:g.176831258C>G затрагивают
 216 более переменные позиции белка *F12*, в то время как
 217 NC_000005.9:g.176831232G>T и NC_000005.9:g.176831232G>C занимают
 218 умеренно консервативную область.

219 Все исследуемые полиморфные варианты гена *F12* показали изменение
 220 стабильности белка. Однако стоит отметить, что наибольшее значение
 221 отклонений было показано для мутаций NC_000005.9:g.176831258C>G и
 222 NC_000005.9:g.176831285C>G (табл.5).

223 При помощи веб-ресурсов PolyPhen-2 и MutationTaster2021, CADD,
 224 MutPred2 миссенс-мутации NC_000005.9:g.176831258C>G и
 225 NC_000005.9:g.176831285C>G были определены как нейтральные.
 226 Большинство программ отметило варианты NC_000005.9:g.176831232G>T и
 227 NC_000005.9:g.176831232G>C как патогенные. По прогнозам программы
 228 FATMM и SIFT все анализируемые миссенс-мутации не оказывают
 229 негативного воздействия на биологическую активность белка FXII.

230 4 Обсуждение

231 Отеки при НАО наблюдается из-за нарушения работы калликреин-
 232 кининовой системы (ККС). Основными причинами чрезмерной активации
 233 ККС при НАО являются мутации в гене *SERPING1*, а также мутации в генах,
 234 кодирующих разные компоненты ККС, что приводит к нарушению регуляции
 235 и нормального функционирования системы [29].

236 ККС — это протеолитический каскад, обеспечивающий регуляцию
 237 активности кининогена, свёртывания крови, фибринолиза. В ККС входят
 238 протеаза прекалликреина (ПК) (или фактора Флетчера), фактор свертывания
 239 XI (FXI) и XII (FXII) (или фактора Хагемана) и кофактора/субстрата
 240 высокомолекулярного кининогена (ВК). Ферментативные каскады
 241 пересекаются в патофизиологии НАО. При аутоактивации FXII происходит
 242 высвобождение FXIIa, который активирует расщепление прекалликреина
 243 плазмы (ПК) до калликреина (КК), усиливая работу ККС. Процесс активации

244 в конечном итоге приводит к расщеплению ВК, что высвобождает брадикинин
245 (БК), самый важный кинин в ККС (рис. 1) [21]. В контексте наследственного
246 ангиоотека, особенно типов с нормальным уровнем С1-ингибитора, мутации в
247 гене *F12* могут приводить к гиперактивности этого фактора и запуску
248 патогенеза НАО [12].

249 Фактор XII, кодируемый геном *F12* и состоящий из 14 экзонов и 13
250 интронов, экспрессируется в печени и циркулирует в крови как
251 одноцепочечный полипептидный зимоген с молекулярной массой 80 кДа.
252 Структурная организация белка фактора XII состоит из N-концевого домена
253 фибронектина типа II (FnII), за которым следует домен, подобный
254 эпидермальному фактору роста (EGF)¹, фибронектин типа I (FnI), домен
255 EGF², домен крингл и домен С-концевой сериновой протеазы (каталитический
256 домен) с дополнительной областью, богатой пролином [25]. В сайтах между
257 аргинином 353 (Arg³⁵³) и валином 354 (Val³⁵⁴) и между аргинином 334
258 (Arg³³⁴) и изолейцином 335 (Ile³³⁵) происходит расщепление фактора XII до
259 FXIIa [16]. Это расщепление приводит к окончательной активации белка FXII
260 и образованию двухцепочечной сериновой протеазы FXIIa, состоящей из
261 тяжелой (альфа-фактор XIIa) и легкой цепи. Высокомолекулярный кининоген
262 действует как кофактор, стабилизируя взаимодействие между калликреином и
263 фактором XII, что облегчает процесс активации [35].

264 Известно, что мутации с патогенным значением затрагивают 9 экзон гена
265 *F12*, который кодирует высокогликозилированную область белка, изменения,
266 в которой приводят к увеличению продукции активного фактора XII (фактора
267 XIIa) [14]. Все рассмотренные полиморфные варианты гена *F12* находятся в 9
268 экзоне пролин богатой области (Pro-rich). Изменения в такой области могут
269 быть триггером для запуска избыточной продукции брадикинина,
270 приводящего к образованию отеков [18].

271 По прогнозу программ MutationTaster и Polyphen-2 полиморфизм
272 NC_000005.9:g.176831232G>C (p. Thr328Arg) является патогенным. Алгоритм
273 HumVar веб-ресурса Polyphen-2 определил генетический вариант
274 NC_000005.9:g.176831232G>T (p. Thr328Lys) как повреждающий. Однако
275 анализ консервативности аминокислот (HumDiv) показал, что изменение
276 аминокислотного остатка на лизин в 328-й позиции не повлияет на функции
277 белка. Согласно базе данных популяционных частот генетических вариантов
278 gnomAD, встречаемость аминокислотной замены p.Thr328Lys в популяции
279 составляет 0.06866 и характерна для людей, имеющих симптомы НАО.
280 Примечательно, что в базах данных нет сведений о встречаемости
281 аминокислотной замены p.Thr328Arg, однако данный вариант был найден у
282 членов семьи с симптомами НАО. Несмотря на то, что мутация не затрагивает
283 консервативную область, замена треонина в 328-й позиции является
284 критичной и может привести к значительным изменениям в локальной
285 структуре белка дикого типа. Оба полиморфных варианта приводят к замене

286 полярной и нейтральной аминокислоты треонин на положительно заряженную
 287 и гидрофильную аминокислоту, что скорее всего приведет к изменению
 288 характера электростатических взаимодействий. Аргинин и лизин имеют более
 289 крупную боковую цепь и положительный заряд, предположительно это может
 290 привести к нарушению водородных связей и вызывать конформационные
 291 изменения в белке. Согласно программе NOPE, замена затрагивает
 292 неупорядоченную область белка F12, для которой характерна высокая
 293 конформационная гибкость, необходимая для взаимодействия с другими
 294 молекулами.

295 По литературным данным, структурированные домены и внутренне
 296 неупорядоченные области на уровне протеома являются двумя
 297 функциональными строительными блоками белков [34]. Неупорядоченные
 298 регионы часто участвуют в пост-трансляционных модификациях, влияя на
 299 регуляцию активности и взаимодействия белка с другими молекулами (рис.2).
 300 Мутации в таком регионе могут способствовать изменению функциональной
 301 активности белка и развитию симптомов наследственного ангиоотека.
 302 Полученные данные подтверждают предсказанную патогенную значимость
 303 вариантов NC_000005.9:g.176831232G>C (p.Thr328Arg) и
 304 NC_000005.9:g.176831232G>T (p. Thr328Lys) *in silico*.

305 Согласно веб-ресурсам MUpPro и I-Mutant Suite значительное снижение
 306 стабильности вызывает полиморфизм NC_000005.9:g.176831285C>G.
 307 Изменение большей аминокислоты по объёму аргинина (174 Å³) на меньшую
 308 серин (105 Å³) может стать причиной появления пустот в белке, потери
 309 конформации и повышения энтропийного эффекта (рис. 3). Малые по объёму
 310 аминокислоты, как правило, обладают большей вращательной способностью,
 311 что может приводить к снижению стабильности белка за счет повышения
 312 гибкости молекулы. Однако высокая встречаемость аминокислотной замены
 313 p. Arg310Ser в популяции является критерием доброкачественного статуса
 314 полиморфизма. Следует отметить, что большинство *in silico* предикторов
 315 оценило вариант как нейтральный и не влияющий на функциональность белка.

316 Рассматриваемый полиморфный вариант NC_000005.9:g.176831285C>G
 317 встречается в популяционных базах данных с частотой 0,3876%, что
 318 превышает ожидаемую частоту для заболевания наследственного ангиоотека.
 319 Исходя из полученных данных полиморфизм NC_000005.9:g.176831285C>G
 320 не вносит значимых изменений в структуру белка FXII и является
 321 нейтральным. Однако классификация патогенный/нейтральный является не
 322 точной и не всегда может указывать на проявленный эффект. Во многих
 323 случаях изменение функционирования белка зависит от дополнительного
 324 генетического контекста. Сочетание разных генетических вариантов может
 325 приводить или не приводить к функциональным последствиям в белке.
 326 Нейтральный вариант может вносить вклад в развитие заболевания в случае,
 327 если появится еще одна патогенная мутация в цис- или транс-положении этого

328 гена [6]. При НАО фенотип не коррелирует с типом патогенной мутации,
 329 различия клинических проявлений пациентов возможно зависят от сочетаний
 330 нейтральных вариантов, которые играют роль дополнительных
 331 модификаторов заболевания.

332 Локализация миссенс варианта NC_000005.9:g.176831258C>G
 333 затрагивает неупорядоченный регион, который не является строго критичным
 334 для *F12*, однако именно в этом участке полипептидной цепи найдены
 335 приведенные выше патогенные мутации p. Thr328Arg и p. Thr328Lys. Разница
 336 в свойствах аминокислот мутированного белка и белка дикого типа может
 337 повлиять на организацию изменяемой области. Однако значительная разница
 338 свойств аминокислот в случае p. Gln319His достигается только в индексе
 339 гидropатии. Боковая цепь гистидина содержит имидазольное кольцо, которое
 340 принимает положительный заряд. Такое свойство дает возможность
 341 гистидину участвовать в нехарактерных ионных взаимодействиях, которые
 342 могут дестабилизировать белковую молекулу [17] (рис.4). В то же время
 343 оценка эволюционной консервативности полиморфизма
 344 NC_000005.9:g.176831258C>G в программе MutationTaster показала, что в
 345 этой позиции мутированный остаток встречается в других гомологичных
 346 последовательностях. Скорее всего, разнообразие функциональных групп в
 347 этой точке указывает на то, что изменение одной аминокислоты не влияет на
 348 потерю функции всего неупорядоченного региона.

349 При помощи анализа *in silico* было показано, что мутации
 350 NC_000005.9:g.176831232G>C (p. Thr328Arg) и
 351 NC_000005.9:g.176831232G>T (p. Thr328Lys) являются патогенными, что
 352 полностью соответствует их ранее известным клиническим статусам.
 353 Полиморфные варианты NC_000005.9:g.176831258C>G (p. Gln319His) и
 354 NC_000005.9:g.176831285C>G (p. Arg310Ser), по всей видимости не могут
 355 вызывать заболевание самостоятельно, однако не исключено их возможное
 356 участие в формировании фенотипа заболевания опосредованно — в качестве
 357 генетических модификаторов.

358 Таким образом, биоинформатические программы позволяют быстро
 359 оценить возможную патогенность генетического варианта. Применение
 360 такого метода особенно актуально при полногеномном секвенировании, когда
 361 у пациентов находят новые варианты с неизвестной клинической
 362 значимостью. Анализ *in silico* позволяет потенциально выявлять патогенные
 363 изменения и предполагать возможные механизмы их действия на
 364 функционирование белка.

365 5 Заключение

366 Результаты проведенного биоинформатического анализа демонстрируют,
 367 что в большинстве случаев полученные данные программ совпадают между
 368 собой и согласуются с уже известными клиническими характеристиками

369 исследуемых полиморфных вариантов. Это подтверждает высокую
370 прогностическую ценность таких подходов при оценке влияния новых
371 генетических вариантов, потенциально связанных с развитием НАО.

372 Предикторы *in silico* анализируют найденные полиморфные варианты и
373 оценивают их потенциальное влияние на здоровье пациента, что особенно
374 важно для интерпретации новых вариантов.

375 **Финансирование исследования:** работа не имела финансовой поддержки.

376 **Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

ТАБЛИЦЫ

Таблица 1. Исследованные полиморфные варианты гена *F12*.

Table 1. The studied polymorphic variants of the *F12* gene.

Мутация Mutation	Частота встречаемости (GnomAD) Frequency of occurrence (GnomAD)	Clinvar	Статус клинической значимости Clinical Significance Status
NC_000005.9:g.176831285 C>G	0.3876	rs77098327	Доброкачественная
NC_000005.9:g.176831258 C>G	-	rs1763225226	Неизвестное значение
NC_000005.9:g.176831232 G>C	-	rs118204456	Патогенетическая
NC_000005.9:g.176831232 G>T	0.06866	rs118204456	Патогенетическая

Таблица 2. Характеристики аминокислот по данным IMGT Aide-mémoire.

Table 2. Characteristics of amino acids according to IMGT Aide-mémoire.

Аминокислота Acid	Молекулярная масса (Da) Mol. Weight (Da)	Количество атомов азота (N) Number of atoms	Объем (Vo) Volume	Индекс гидропатии Hydropathy index	Адонорность (Donor atoms)	Акцепторы водорода (Hydrogen acceptors)
Треонин Threonine	119	1	116,1	-0,7 (нейтральная) (neutral)	Есть (Yes)	Есть (Yes)

						es	Y		
Лизин Lysine	46	1	4	2	16	8,6	-3,9 (гидрофильная) (hydrophilic)	Есть Yes	Нет No
Аланин Alanine	9	8	3	1	88,	6	1,8 (гидрофобная) (hydrophobic)	Нет No	Нет No
Аргинин Arginine	74	1	6	2	17	3,4	-4,5 (гидрофильная) (hydrophilic)	Есть Yes	Нет No
Гистидин Histidine	55	1	0	2	15	3,2	-3,2 (нейтральная) (neutral)	Есть Yes	Есть Yes
Глутамин Glutamine	46	1	0	2	14	3,8	-3,5 (гидрофильная) (hydrophilic)	Есть Yes	Есть Yes
Серин Serine	05	1	4	1	89		-0,8 (гидрофобная) (hydrophobic)	Есть Yes	Есть Yes

Таблица 3. Анализ значимости локализации миссенс-мутаций по данным веб-ресурса MutationTaster2021.

Table 3. Analysis of the significance of localization of missense mutations according to the MutationTaster2021 web resource.

Диапазон Range	Особенности Peculiarities	Эффект Effect
NC_000005.9:g.176831285C>G		
296-349	Пролин богатая область Pro-rich region	Утрачен Lost
329-329 337-337 433-433	Наличие углеводородной связи с другими позициями, участок, отвечающий за гликозилирование Presence of a hydrocarbon bond with other positions, the site responsible for glycosylation	Может быть утрачен Might get lost
333-333	Замена аминокислоты P -> S Substitution of amino acid P -> S	Может быть утрачен Might get lost
359-359 397-397 405-405 436-436 439-439 475-475 486-486 500-500	Наличие дисульфидной связи с другими позициями Presence of disulfide bond with other positions	Может быть утрачен Might get lost

532-532		
548-548		
559-559		
569-569		
590-590		
373-614	Трипсиновый домен Trypsin domain	Может быть утрачен Might get lost
379-379	Замена аминокислоты А -> G Substitution of amino acid A -> G	Может быть утрачен Might get lost
412-412	Активный центра гистидина Histidine active site	Может быть утрачен Might get lost
461-461	ACT_SITE Charge relay system (By similarity).	Может быть утрачен Might get lost
563-563	Сериновый активный центр Serine active center	Может быть утрачен Might get lost
NC_000005.9:g.176831258C>G		
296-349	Пролин-богатая область Pro-rich	Утрачен Lost
NC_000005.9:g.176831232G>C		
296-349	Пролин-богатая область Pro-rich	Утрачен Lost

328-328 329-329 337-337 433-433	Наличие углеводородной связи с другими позициями, участок, отвечающий за гликозилирование Presence of a hydrocarbon bond with other positions, the site responsible for glycosylation	Утрачен Lost
333-333	Замена аминокислоты P -> S Substitution of amino acid P -> S	Может быть утрачен Might get lost
359-359 397-397 405-405 413-413 436-436 439-439 475-475 486-486 500-500 532-532 548-548 559-559 569-569 590-590	Наличие дисульфидной связи с другими позициями Presence of disulfide bond with other positions	Может быть утрачен Might get lost
373-614	Трипсиновый домен Trypsin domain	Может быть утрачен Might get lost

379-379	Замена аминокислоты А -> G Substitution of amino acid A-> G	Может быть утрачен Might get lost
412-412	Гистидин -активный сайт Histidine - active site	Может быть утрачен Might get lost
461-461	ACT_SITE Charge relay system (By similarity).	Может быть утрачен Might get lost
563-563	Сериновый активный центр Serine active center	Может быть утрачен Might get lost
NC_000005.9:g.176831232G>T		
296-349	Пролин богатая область Pro-rich region	Утрачен Lost
328-328 329-329 337-337 433-433	Наличие углеводородной связи с другими позициями, участок, отвечающий за гликозилирование Presence of a hydrocarbon bond with other positions, the site responsible for glycosylation	Утрачен Lost
333-333	Замена аминокислоты Р -> S Substitution of amino acid P -> S	Может быть утрачен Might get lost
359-359 397-397	Наличие дисульфидной связи с другими позициями Presence of disulfide bond with other positions	Может быть утрачен Might get lost

405-405		
413-413		
436-436		
439-439		
475-475		
486-486		
500-500		
532-532		
548-548		
559-559		
569-569		
590-590		
373-614	Трипсиновый домен Trypsin domain	Может быть утрачен Might get lost
379-379	Замена аминокислоты A-> G Substitution of amino acid A-> G	Может быть утрачен Might get lost
412-412	Гистидин -активный сайт Histidine - active site	Может быть утрачен Might get lost
461-461	ACT_SITE Charge relay system (By similarity).	Может быть утрачен Might get lost

563-563	Сериновый активный центр Serine active center	Может быть утрачен Might get lost
---------	--	--

Таблица 4. Оценка консервативности позиций выявленных мутаций по данным веб-ресурса Mutationtaster2021.

Table 4. Assessment of conservativeness of positions of identified mutations based on web resource data Mutationtaster2021.

Вид Specie	NC_000005.9:g.176831258C>G (p.Gln319His)	NC_000005.9:g.176831285C>G (p.Arg310Ser)	NC_000005.9:g.176831232G>T (p.Thr328Lys)	NC_000005.9:g.176831232G>C (p.Thr328Arg)
Human	-	-	-	-
Mutated	Не сохраняется Not saved	Не сохраняется Not saved	Не сохраняется Not saved	Не сохраняется Not saved
Proteoglycans	Сохраняется Saved	Сохраняется Saved	Не сохраняется Not saved	Не сохраняется Not saved
Mmullatta	Сохраняется Saved	Не сохраняется Not saved	Сохраняется Saved	Сохраняется Saved
Fcatus	Не сохраняется Not saved	Не сохраняется Not saved	Сохраняется Saved	Сохраняется Saved
Mmusculus	Сохраняется Saved	Не сохраняется Not saved	Сохраняется Saved	Сохраняется Saved

s	Ggallu	Нет гомологов No homologue	Нет гомологов No homologue	Нет гомологов No homologue	Нет гомологов No homologue
pes	Trubri	Нет совпадений No alignment	Нет совпадений No alignment	Нет совпадений No alignment	Нет совпадений No alignment
	Drerio	Нет совпадений No alignment	Нет совпадений No alignment	Нет совпадений No alignment	Нет совпадений No alignment
ans	Celeg	Нет гомологов No homologue	Нет гомологов No homologue	Нет гомологов No homologue	Нет гомологов No homologue
calis	Xtropi	Нет гомологов No homologue	Нет совпадений No alignment	Сохраняе тся Saved	Сохраняетс я Saved

Таблица 5. Прогноз стабильности белка при выявленных миссенс-мутациях с использованием веб-ресурсов MUpro и I-Mutant.

Table 5. Prediction of protein stability for identified missense mutations using the MUpro and I-Mutant web resources.

Ре- сурс source	Показатели Indicators	Полиморфизм Polymorphism			
		NC_000005.9:g.176831258C>G (p. Gln319His)	NC_000005.9:g.176831285C>G (p.Arg310Ser)	NC_000005.9:g.176831232G>T (p. Thr328Lys)	NC_000005.9:g.176831232G>C (p. Thr328Arg)
MUpro	ΔΔ	- 0.86549343 (decrease stability)	- 0.97052904 (decrease stability)	- 0.49383326 (decrease stability)	- 0.16351799 (decrease stability)
I-Mutant	DD	-0.95	-2.27	0.16	0.43
	RI	6	7	2	3

Таблица 6. Прогноз патогенетической значимости миссенс-мутаций по данным PolyPhen-2, MutationTaster2021, FATHMM-XF, SIFT, CADD, MutPred2, PANTHERA.

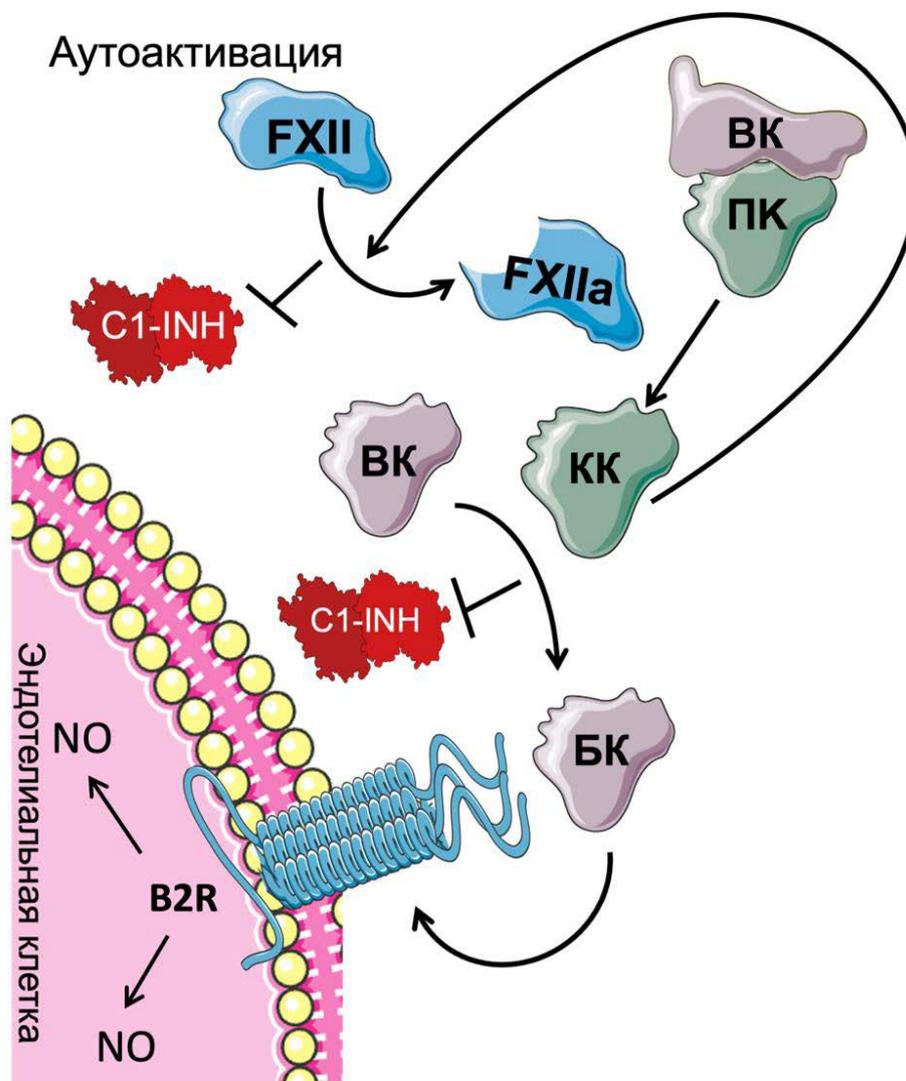
Table 6. Prediction of missense-mutations pathogenetic significance according to PolyPhen-2, MutationTaster2021, FATHMM-XF, SIFT, CADD, MutPred2, PANTHERA.

Веб-ресурс	Показатели	Миссенс-мутации			
		NC_000005.9:g.176831258 C>G (p. Gln319His)	NC_000005.9:g.176831285C>G (p. Arg310Ser)	NC_000005.9:g.176831232G>T (p. Thr328Lys)	NC_000005.9:g.176831232G>C (p. Thr328Arg)
PolyPhen-2	С татус	Нейтральная	Нейтральная	Патогенетическая/нейтральная	Патогенетическая
	S тatus	Neutral	Neutral	Pathogenetically significant / Neutral	Pathogenetically significant
	HumVar	0.152	0.001	0.910	0.973
	HumDiv	0.086	0.002	0.415	0.706
MutationTaster2021	С татус	Нейтральная	Нейтральная	Патогенетическая	Патогенетическая
	S тatus	Neutral	Neutral	Pathogenetically significant	Pathogenetically significant
	Оценка S core	24	110	78	71
FATHMM-XF		Патогенетическая	Патогенетическая	Патогенетическая	Патогенетическая

		Pathogenetically significant	Pathogenetically significant	Pathogenetically significant	Pathogenetically significant
	Оценка S core	-2.45	-2.38	-2.51	-2.53
SIFT	Статус S tatus	Нейтральная Neutral	Нейтральная Neutral	Нейтральная Neutral	Нейтральная Neutral
	Оценка S core	0.12	0.75	0.15	0.12
D CAD	Статус S tatus	Нейтральная Neutral	Нейтральная Neutral	Патогенетическая Pathogenetically significant	Патогенетическая Pathogenetically significant
	Оценка S core	1.251	9.85	22,9	22,9
MutPr ed2	Статус S tatus	Нейтральная Neutral	Нейтральная Neutral	Патогенетическая Pathogenetically significant	Патогенетическая Pathogenetically significant
	Оценка S core	0,161	0,304	0,743	0,816

РИСУНКИ

Рисунок 1. Фактор XII в калликреин-кининовой системе ККС.
Figure 1. Factor XII in the contact activation system (CAS).



Фактор XII (FXII) подвергается контактно-опосредованной активации в результате происходит высвобождение активной протеазы FXIIa. FXIIa протеолитически активирует плазменный прекалликреин (ПК), связанный с поверхностью клеток и циркулирующий в крови в виде комплекса с высокомолекулярным кининогеном (БК), в активный плазменный калликреин (КК). КК высвобождает брадикинин (БК) из ВК путем протеолиза. Свободный БК индуцирует эндотелиальную сигнализацию при связывании с рецепторами В2, увеличивая сосудистую проницаемость. Ингибитор С1-эстеразы (С1INH) является эндогенным ингибитором FXIIa и КК.

FXII - фактор XII, КК - калликреин, C1-INH - ингибитор C1-эстеразы, ПК-плазменный прекалликреин, ВК - высокомолекулярный кининоген, БК-брадикинин, B2R - рецептор брадикинина.

Factor XII (FXII) undergoes contact-mediated activation, resulting in the release of the active protease FXIIa. FXIIa proteolytically activates plasma prekallikrein (PK), which is associated with cell surfaces and circulates in the blood in complex with high-molecular-weight kininogen (HK), converting it into active plasma kallikrein (PKa). KK cleaves bradykinin (BK) from HK via proteolysis. BK induces endothelial signaling by binding to B2 receptors, thereby increasing vascular permeability. C1 esterase inhibitor (C1INH) is the endogenous inhibitor of both FXIIa and KK.

FXII - factor XII, PKa - plasma kallikrein, HK - high-molecular-weight kininogen, PK - plasma prekallikrein, C1INH - C1 esterase inhibitor, BK - bradykinin, B2R - Bradykinin receptor 2.

Рисунок 2. Визуализация структурного изменения белка FXII при миссенс-мутациях p.Thr328Arg и p. Thr328Lys.

Figure 2. Visualization of structural changes in FXII protein caused by p.Thr328Arg and p.Thr328Lys missense mutations.

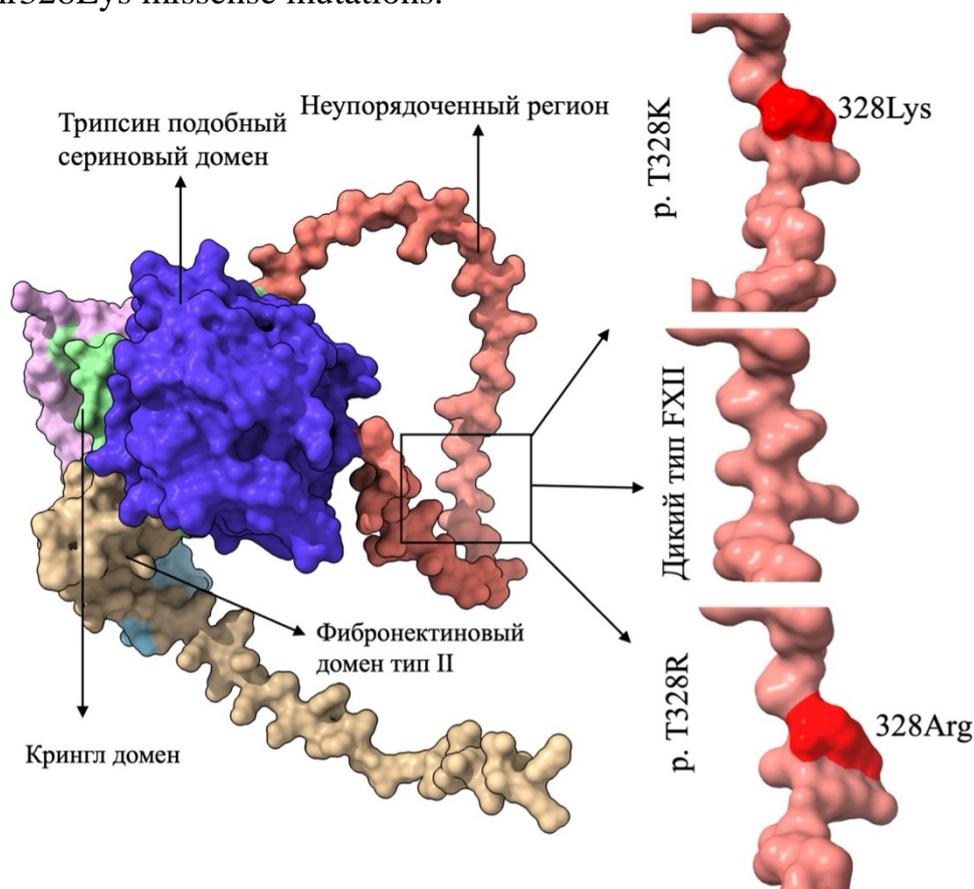
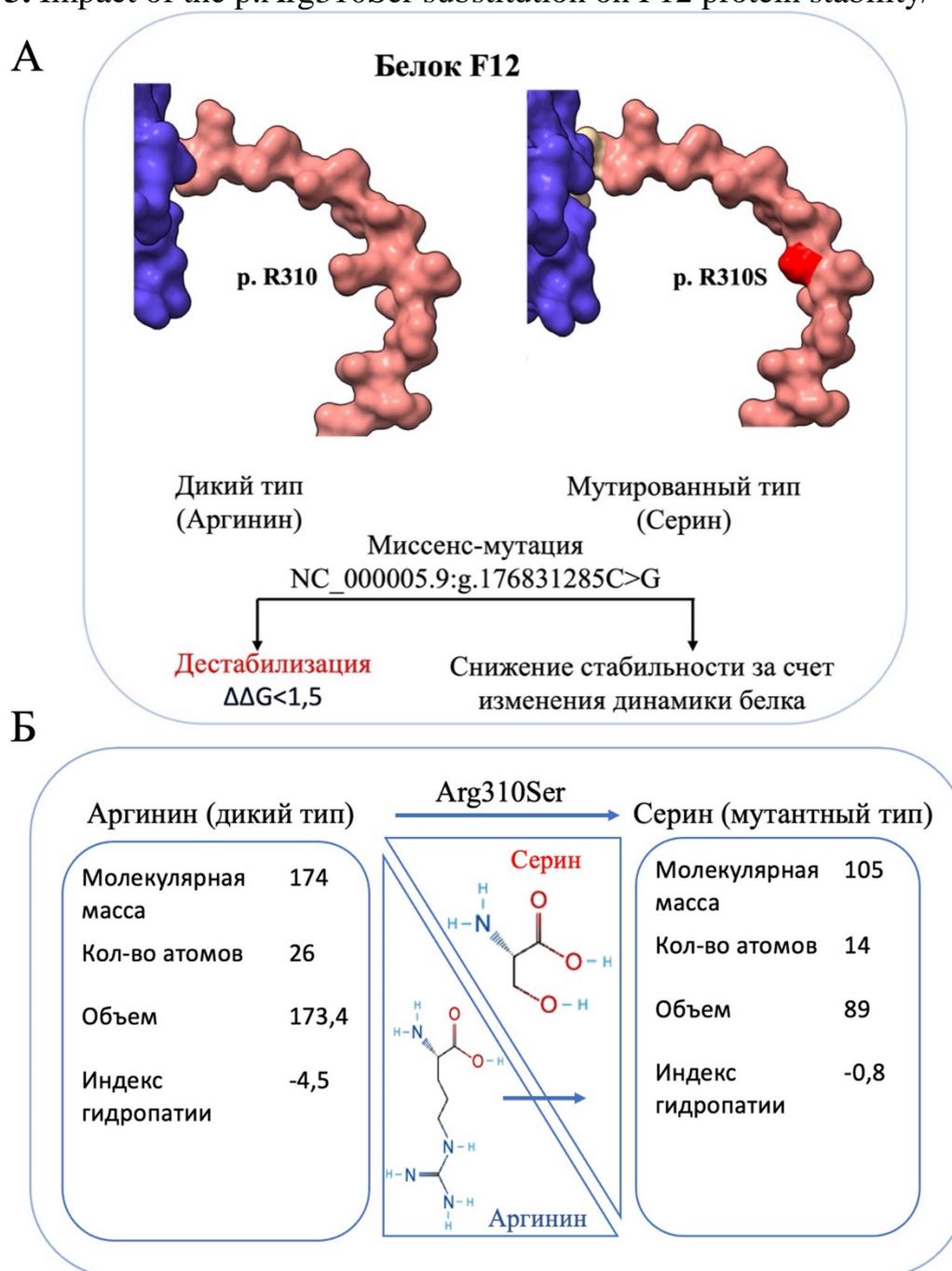


Рисунок 3. Влияние замены р. Arg310Ser на стабильность белка F12.
Figure 3. Impact of the p.Arg310Ser substitution on F12 protein stability/

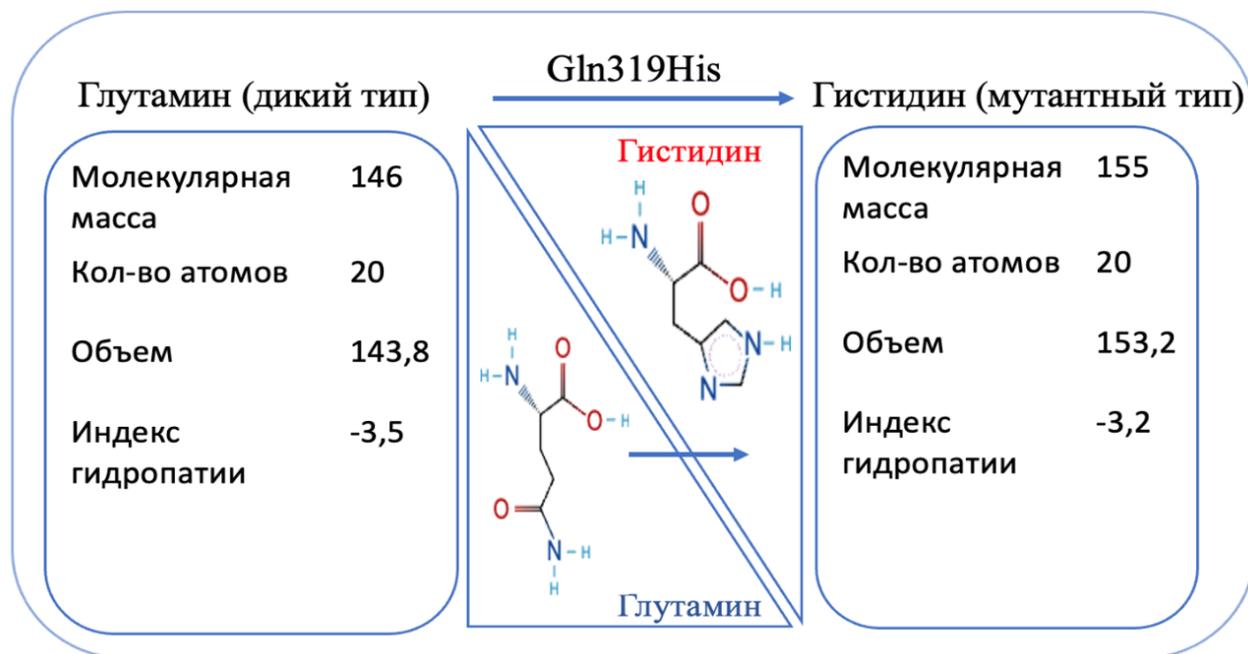


А - Прогнозирование изменений свободной энергии при аминокислотной замене р. Arg310Ser в белке F12. Б - Сравнение структурно-функциональных характеристик аргинина и серина согласно данным IMGT Aide-mémoire

A - Prediction of free energy changes upon p.Arg310Ser amino acid substitution in F12 protein. B - Comparison of structural and functional characteristics of arginine and serine according to IMGT Aide-mémoire data.

Рисунок 4. Сравнение структурно-функциональных характеристик глутамина и гистидина согласно данным IMGT Aide-mémoire.

Figure 4. Comparison of structural and functional characteristics of glutamine and histidine according to IMGT Aide-mémoire data.



ТИТУЛЬНЫЙ ЛИСТ_МЕТАДААННЫЕ

Блок 1. Информация об авторе ответственном за переписку

Останкова Юлия Владимировна – к.б.н., заведующая лабораторией иммунологии и вирусологии ВИЧ-инфекции, старший научный сотрудник лаборатории молекулярной иммунологии, ФБУН «Санкт-Петербургский НИИ эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека РФ, Санкт-Петербург, Россия адрес: 197101, Санкт-Петербург, ул. Мира, д. 14; телефоны: 8 (812) 233-20-92 /+7 (996) 929-20-91; ORCID: 0000-0003-2270-8897; e-mail: shennal@yandex.ru

Ostankova Yu. V. – PhD (Biology), Head of the Laboratory of Immunology and Virology HIV, Senior Researcher at the Laboratory of Molecular Immunology St. Petersburg Pasteur Institute, St. Petersburg, Russian Federation.
address: 197101, St. Petersburg, st. Mira, 14, St. Petersburg Pasteur Institute;
telephones: 8 (812) 233-20-92 /+7 (996) 929-20-91;
ORCID: 0000-0003-2270-8897;
e-mail: shennal@yandex.ru

Блок 2. Информация об авторах

Седых Анна Васильевна – младший научный сотрудник лаборатории иммунологии и вирусологии ВИЧ-инфекции, ФБУН «Санкт-Петербургский НИИ эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека;
Sedykh A. V. – Junior researcher at the Laboratory of Immunology and Virology of HIV Infection St. Petersburg Pasteur Institute, St. Petersburg, Russian Federation;

Щемелев Александр Николаевич – к.б.н., младший научный сотрудник лаборатории молекулярной иммунологии иммунологии и вирусологии ВИЧ-инфекции, ФБУН «Санкт-Петербургский НИИ эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека;
Schemelev Alexandr N. – PhD (Biology), Junior researcher at the Laboratory of Immunology and Virology of HIV Infection St. Petersburg Pasteur Institute, St. Petersburg, Russian Federation;

Тотолян Арег Артемович – Академик РАН, д.м.н., профессор, заведующий лабораторией молекулярной иммунологии, директор ФБУН «Санкт-Петербургский НИИ эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека, заведующий кафедрой иммунологии Первого Санкт-Петербургского медицинского университета имени академика И.П. Павлова;

Totolian Areg A. – Academician of the Russian Academy of Sciences, PhD, MD (Medicine), Professor, Head at the Laboratory of Molecular Immunology, Director of the St. Petersburg Pasteur Institute; head Department of Immunology, First St. Petersburg State Medical University named after Academician I.P. Pavlov.

Блок 3. Метаданные статьи

БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ ОДНОНУКЛЕОТИДНЫХ
ВАРИАНТОВ ГЕНА *F12*, АССОЦИИРОВАННЫХ С НАСЛЕДСТВЕННЫМ
АНГИОТЕКОМ

BIOINFORMATIC ANALYSIS OF SINGLE NUCLEOTIDE VARIANTS IN
THE F12 GENE ASSOCIATED WITH HEREDITARY ANGIOEDEMA

Сокращенное название статьи для верхнего колонтитула:

БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ ГЕНА *F12*
BIOINFORMATIC ANALYSIS OF THE *F12* GENE

Ключевые слова: первичный иммунодефицит, наследственный ангиоотек, биоинформатический анализ, *F12*, полиморфные варианты.

Keywords: primary immunodeficiency, hereditary angioedema, bioinformatic analysis, *F12*, polymorphic variants.

Оригинальные статьи.

Количество страниц текста – 10,

Количество таблиц – 6,

Количество рисунков – 4.

29.05.2025

СПИСОК ЛИТЕРАТУРЫ

Порядковый номер ссылки	Авторы, название публикации и источника, где она опубликована, выходные данные	ФИО, название публикации и источника на английском	Полный интернет-адрес (URL) цитируемой статьи или ее doi.
1	Печникова, Н. А., Останкова, Ю. В., & Тотолян, А. А. Применение биоинформатического анализа для выявления генов-кандидатов, ассоциированных с наследственным ангиоотеком //Медицинская иммунология. – 2022. – Т. 24. – №. 5. – С. 1027-1046.	Pechnikova N.A., Ostankova Y.V., Liubimova N.E., Semenov A.V., Kuznetsova R.N., Totolian A.A. Application of bioinformatic analysis to identify the clinical significance of missense mutations in the HS3ST6 gene in the development of hereditary angioedema. <i>Infektsiya i immunitet = Russian Journal of Infection and Immunity</i> , 2022, Vol. 12, no. 5, pp. 997–1005.	doi: 10.15789/1563-0625-AOB-2579.
2	Печникова, Н. А., Останкова, Ю. В., Сайтгалина, М. А., Бебяков, А. М., Денисова, А. Р., Подчерняева, Н. С., & Тотолян, А. А. Применение биоинформатического анализа для прогностической	Pechnikova N.A., Ostankova Y.V., Liubimova N.E., Semenov A.V., Kuznetsova R.N., Totolian A.A. Application of bioinformatic analysis to identify the clinical significance of missense mutations in the HS3ST6 gene in the development of hereditary angioedema. <i>Infektsiya i immunitet</i>	doi: 10.15789/1563-0625-ABA-2577

	оценки клинической значимости миссенс-мутаций гена HS3ST6 в развитии наследственного ангиоотека //Медицинская иммунология. – 2022. – Т. 25. – №. 1. – С. 135-154.	= <i>Russian Journal of Infection and Immunity</i> , 2022, Vol. 12, no. 5, pp. 997–1005.	
3	Российская Ассоциация Аллергологов и Клинических иммунологов, Национальная Ассоциация Экспертов в области Первичных Иммунодефицитов. Первичные иммунодефициты с преимущественной недостаточностью синтеза антител: клинические рекомендации. –2022		https://cr.minzdrav.gov.ru/preview-cr/267_2?utm_source=crm&utm_medium=email&utm_campaign=wim333_3_2_Gen_NAO_116002_CTE_NAO_180225&utm_content=btn_banner_top
4	Российская ассоциация аллергологов и клинических иммунологов, Союз педиатров России, Ассоциация медицинских генетиков, Национальная		https://raaci.ru/dat/pdf/KR/PI D.pdf

	Ассоциация Экспертов в области Первичных Иммунодефицитов. Клинические рекомендации. Наследственный ангионевротический отек. – 2024.		
5	Adzhubei I.A., Schmidt S., Peshkin L., Ramensky V.E., Gerasimova A., Bork P., Kondrashov A.S., Sunyaev S.R. A method and server for predicting damaging missense mutations. <i>Nat. Methods</i> , 2010, Vol. 7, no. 4, pp. 248–249.		doi: 10.1038/nmeth0410-248
6	Beer N.L., Osbak K.K., van de Bunt M., Tribble N.D., Steele A.M., Wensley K.J., Gloyn A.L. Insights into the pathogenicity of rare missense GCK variants from the identification and functional characterization of compound heterozygous and double		doi: 10.2337/dc11-2420

	mutations inherited in cis. <i>Diabetes Care</i> , 2012, Vol. 35, no. 7, pp. 1482–1484.		
7	Bork K., Wulff K., Meinke P., Wagner N., Hardt J., Witzke G. A novel mutation in the coagulation factor 12 gene in subjects with hereditary angioedema and normal C1-inhibitor. <i>Clin. Immunol.</i> , 2011, Vol. 141, no. 1, pp. 31–35.		doi: 10.1016/j.clim.2011.07.002
8	Busse P.J., Christiansen S.C., Riedl M.A., Banerji A., Bernstein J.A., Castaldo A.J., Craig T.J., Davis-Lorton M., Frank M.M., Gower R.G. US HAEA Medical Advisory Board 2020 Guidelines for the Management of Hereditary Angioedema. <i>J. Allergy Clin. Immunol. Pract.</i> , 2021, Vol. 9, no. 1, pp. 132–150.		doi: 10.1016/j.jaip.2020.08.046

9	Camilli F., Borrmann A., Gholizadeh S., te Beek T.A.H., Kuipers R.K.P., Venselaar H. The future of HOPE: what can and cannot be predicted about the molecular effects of a disease causing point mutation in a protein? <i>EMBnet.journal</i> , 2011, Vol. 17, no. 1, pp. 1–9.		doi: 10.14806/ej.17.1.212
10	Capriotti E., Fariselli P., Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. <i>Nucleic Acids Res.</i> , 2005, Vol. 33, suppl_2, pp. 306–310.		doi: 10.1093/nar/gki375
11	Cheng J., Randall A., Baldi P. Prediction of protein stability changes for single-site mutations using support vector		doi: 10.1002/prot.20810

	machines. <i>Proteins</i> , 2006, Vol. 62, no. 4, pp. 1125–1132.		
12	Cohn D.M., Renné T. Targeting factor XIIa for therapeutic interference with hereditary angioedema. <i>J. Intern. Med.</i> , 2024, Vol. 296, no. 4, pp. 311–326.		doi: 10.1111/joim.20008
13	Depetri F., Tedeschi A., Cugno M. Angioedema and emergency medicine: From pathophysiology to diagnosis and treatment. <i>Eur. J. Intern. Med.</i> , 2019, Vol. 59, pp. 8–13.		doi: 10.1016/j.ejim.2018.09.004
14	Dewald G., Bork K. A missense mutation in the factor XII gene in a family with hereditary angioedema. <i>Biochem. Biophys. Res.</i>	-	doi: 10.1016/j.bbrc.2006.03.092

	<i>Commun.</i> , 2006, Vol. 343, no. 4, pp. 1286–1289.		
15	Itan Y., Shang L., Boisson B., Ciancanelli M.J., Markle J.G., Martinez-Barricarte R., Scott E., Shah I., Stenson P.D., Gleeson J., et al. The mutation significance cutoff: gene-level thresholds for variant predictions. <i>Nat. Methods</i> , 2016, Vol. 13, no. 2, pp. 109–110.	-	doi: 10.1038/nmeth.3739
16	Lucena-Aguilar G., Sanchez- Ivanov I., Matafonov A., Sun M.F., Cheng Q., Dickeson S.K., Verhamme I.M., Gailani D. Proteolytic properties of single-chain factor XII: a mechanism for triggering contact activation. <i>Blood</i> , 2017, Vol. 129, no. 11, pp. 1527–1537.	-	doi: 10.1182/blood-2016-10-744110

17	Liao S.M., Du Q.S., Meng J.Z., Pang Z.W., Huang R.B. The multiple roles of histidine in protein interactions. <i>Chem. Cent. J.</i> , 2013, Vol. 7, Article no. 44, pp. 1–12.	-	doi: 10.1186/1752-153X-7-44
18	Liu J., Qin J., Borodovsky A., Racie T., Castoreno A., Schlegel M., Maier M.A., Zimmerman T., Fitzgerald K., Butler J., Akinc A. An investigational RNAi therapeutic targeting Factor XII (ALN-F12) for the treatment of hereditary angioedema. <i>RNA</i> , 2019, Vol. 25, no. 2, pp. 255–263.	-	doi: 10.1261/rna.068916.118
19	Manto I.A., Latysheva E.A., Sorokina L.E., Latysheva T.V. The place of scales and questionnaires in assessing the disease's severity and the long-term prophylaxis's prescribing in patients with hereditary angioedema. <i>Terapevt. Arkh.</i> ,		https://doi.org/10.26442/00403660.2021.12.201294

	2021, Vol. 93, no. 12, pp. 1498–1509		
20	Meng E.C., Goddard T.D., Pettersen E.F., Couch G.S., Pearson Z.J., Morris J.H., Ferrin T.E. UCSF ChimeraX: Tools for structure building and analysis. <i>Protein Sci.</i> , 2023, Vol. 32, no. 11, pp 23-38.		doi: 10.1002/pro.4792
21	Motta G., Juliano L., Shariat-Madar Z. Kallikrein-kinin system: insights into a multifunctional system. <i>Front. Physiol.</i> , 2023, Vol. 14, Article no. 1305981.		doi: 10.3389/fphys.2023.1305981
22	Mottaz A., David F.P.A., Veuthey A.L., Yip Y.L. Easy retrieval of single amino-acid polymorphisms and phenotype information using SwissVar.		doi: 10.1093/bioinformatics/btq028

	<i>Bioinformatics</i> , 2010, Vol. 26, no. 6, pp. 851–852.		
23	Ng P.C., Henikoff S. SIFT: Predicting amino acid changes that affect protein function. <i>Nucleic Acids Res.</i> , 2003, Vol. 31, no. 13, pp. 3812–3814.		doi: 10.1093/nar/gkg509
24	Pac M., Bernatowska E. Comprehensive activities to increase recognition of primary immunodeficiency and access to immunoglobulin replacement therapy in Poland. <i>Eur. J. Pediatr.</i> , 2016, Vol. 175, pp. 1099–1105.		doi: 10.1007/s00431-016-2746-2
25	Pathak M., Wilmann P., Awford J., Li C., Hamad B.K., Fischer P.M., Emsley J. Coagulation factor XII		doi: 10.1111/jth.12849

	protease domain crystal structure. <i>J. Thromb. Haemost.</i> , 2015, Vol. 13, no. 4, pp. 580–591.		
26	Pejaver V., Urresti J., Lugo-Martinez J., Pagel K.A., Lin G.N., Nam H.J., Mort M., Cooper D.N., Sebat J., Iakoucheva L.M., et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. <i>Nat. Commun.</i> , 2020, Vol. 11, Article no. 5918.		doi: 10.1038/s41467-020-19669-x
27	Rentzsch P., Witten D., Cooper G.M., Shendure J., Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. <i>Nucleic Acids Res.</i> , 2019, Vol. 47, D1, pp. 886–894.		doi: 10.1093/nar/gky1016

28	Rogers M.F., Shihab H.A., Mort M., Cooper D.N., Gaunt T.R., Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. <i>Bioinformatics</i> , 2018, Vol. 34, no. 3, pp. 511–513.		doi: 10.1093/bioinformatics/btx536
29	Shamanaev A., Dickeson S.K., Ivanov I., Litvak M., Sun M.F., Kumar S., Gailani D. Mechanisms involved in hereditary angioedema with normal C1-inhibitor activity. <i>Front. Physiol.</i> , 2023, Vol. 14, Article no. 1146834.		doi: 10.3389/fphys.2023.1146834
30	Sherry S.T., Ward M.H., Kholodov M., Baker J., Phan L., Smigielski E.M., Sirotkin K. dbSNP: the NCBI database		doi: 10.1093/nar/29.1.308

	of genetic variation. <i>Nucleic Acids Res.</i> , 2001, Vol. 29, no. 1, pp. 308–311.		
31	Sim N.L., Kumar P., Hu J., Henikoff S., Schneider G., Ng P.C. SIFT web server: predicting effects of amino acid substitutions on proteins. <i>Nucleic Acids Res.</i> , 2012, Vol. 40, W1, pp. 452–457.		doi: 10.1093/nar/gks539.
32	Steinhaus R., Proft S., Schuelke M., Cooper D.N., Schwarz J.M., Seelow D. MutationTaster2021. <i>Nucleic Acids Res.</i> , 2021, Vol. 49, W1, pp. 446–451.		doi: 10.1093/nar/gkab266
33	Stenson P.D., Mort M., Ball E.V., Chapman M., Evans K., Azevedo L., Hayden M.J., Heywood S., Millar D.S.,		doi: 10.1007/s00439-020-02199-3

	Phillips A.D., Cooper D.N. The Human Gene Mutation Database (HGMD®): optimizing its use in a clinical diagnostic or research setting. <i>Hum. Genet.</i> , 2020, Vol. 139, no. 10, pp. 1197–1207.		
34	Trivedi R., Nagarajaram H.A. Intrinsically disordered proteins: an overview. <i>Int. J. Mol. Sci.</i> , 2022, Vol. 23, no. 22, Article no. 14050.		doi: 10.3390/ijms232214050
35	Wu Y. Contact pathway of coagulation and inflammation. <i>Thromb. J.</i> , 2015, Vol. 13, Article no. 1, pp. 1–9.		doi: 10.1186/s12959-015-0048-y
36	Zanichelli A., Magerl M., Longhurst H., Fabien V., Maurer M. Hereditary		doi: 10.1186/1710-1492-9-29

	angioedema with C1 inhibitor deficiency: delay in diagnosis in Europe. <i>Allergy Asthma Clin. Immunol.</i> , 2013, Vol. 9, Article no. 29.		
--	---	--	--