



# ЕСТЕСТВЕННЫЕ НАУКИ: АКТУАЛЬНЫЕ ПРОБЛЕМЫ И ИННОВАЦИИ

---

---

## СТАТЬИ

УДК 616.831-005:577.21  
ББК 28.707.3

### БИОИНФОРМАЦИОННЫЙ ПОИСК И ПОСТРОЕНИЕ БАЗЫ ДАННЫХ ГЕНОВ, АССОЦИИРОВАННЫХ С ВЫСОКИМ РИСКОМ РАЗВИТИЯ ЦЕРЕБРОВАСКУЛЯРНОЙ ПАТОЛОГИИ

**Дергилева Александра Яковлевна**

Студент кафедры биоинженерии и биоинформатики,  
Волгоградский государственный университет  
saschenkadergileva@gmail.com  
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

**Кишеневский Александр Михайлович**

Аспирант кафедры биоинженерии и биоинформатики,  
Волгоградский государственный университет  
kish-sanek@mail.ru  
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

**Аннотация.** В работе рассматриваются основы патогенеза цереброваскулярных заболеваний, факторы риска, генетические аспекты формирования сосудистых заболеваний головного мозга и их зависимость от пенетрантности генетических полиморфизмов. Проведен биоинформационный поиск с использованием открытых ресурсов PubMed, SNP, Gene, KEGG, а также полнотекстовых версий статей в PubMedCentral и Научной электронной библиотеке Российского индекса цитирования (eLibrary). В процессе исследования был проведен анализ порядка 350 генов, задействованных в процессах, обуславливающих возникновение ЦВЗ. Эти гены были связаны с контролем таких процессов, как вязкость и микроциркуляция крови; функция тромбоцитов и сосудистого эндотелия; макрофагальные и иммунные реакции; чувствительность нейронов к гипоксии; защитная функция нейроглии. Количество точечных полиморфизмов в каждом из найденных генов колебалось в пределах от 500 до 3 000.

Была создана модель для построения базы данных генетических полиморфизмов, ассоциированных с высоким риском развития цереброваскулярной патологии, которая позволяет составить полноценное представление о структуре генетических полиморфизмов, ассоциированных с высоким риском развития цереброваскулярной патологии. База данных имеет практическое применение, поскольку может быть использована для виртуального скрининга и последующего моделирования потенциальной эффективности SNP-полиморфизмов в качестве диагностических тест-систем по выявлению лиц, предрасположенных к сосудистым заболеваниям головного мозга.

**Ключевые слова:** ишемия мозга, инсульт, цереброваскулярная патология, генетический полиморфизм, базы данных.

В настоящее время цереброваскулярные заболевания (далее – ЦВЗ) занимают лидирующие позиции по причинам смертности и инвалидности населения РФ, уступая лишь болезням сердца. Среди заболеваний нервной системы, приводящих к стойкой утрате трудоспособности, ЦВЗ занимают 1-е место. Данные регистров показывают, что в России происходит около 450 000 инсультов ежегодно, летальность при которых достигает 35 % [5].

Это указывает на сосудистые заболевания головного мозга как на важную медико-социальную проблему, которая требует тщательного изучения, разработки новых методов диагностики и лечения, профилактики данного вида заболеваний.

В настоящий момент «золотым стандартом» диагностики сосудистых заболеваний головного мозга является компьютерная томография и магнитно-резонансная томография, которые позволяют достоверно зафиксировать признаки инсульта в первые часы от начала заболевания с точностью от 78 % до 100 % (в зависимости от вида нарушения мозгового кровообращения и индивидуальных особенностей пациента) [1].

Однако при весьма широком арсенале методов инструментальной и лабораторной диагностики выявление факторов риска и наличия хронических заболеваний головного мозга, а также профилактика таких грозных осложнений, как острое нарушение мозгового кровообращения, остаются востребованной и перспективной задачей в поддержании здоровья населения.

Факторами риска ЦВЗ являются: артериальная гипертензия, стеноз сосудов головного мозга, коагулопатии с развитием тромбозов и эмболии (в том числе кардиального генеза), дислипидемии с развитием атеросклероза со-

судистой стенки, заболевания сердца, сахарный диабет, применение оральных контрацептивов, курение, злоупотребление алкоголем, малоподвижный образ жизни [2; 3; 5; 6]. Все эти факторы могут быть исключены, а их влияние нивелировано при наличии своевременных и последовательных мер профилактики.

Ключевую роль в патогенезе данного ряда заболеваний играют глутаматная эксайтотоксичность, местные воспалительные реакции под воздействием провоспалительных цитокинов и ряда других медиаторов, оксидантный стресс, явления кальций-индуцированного апоптоза, недостаточная активация холинергической и ГАМК-ергической систем, гипоперфузия ишемизированного участка мозговой ткани, геморрагическое пропитывание тканей [4; 7–9].

При наличии довольно сложного процесса патогенеза ЦВЗ остается актуальной задача диагностики и последующей профилактики данной группы заболеваний на ранних этапах. Одним из достоверных методов оценки является анализ генетических полиморфизмов, ответственных за развитие факторов риска и патологических изменений, приводящих к повышенному риску развития цереброваскулярной патологии.

**Цель работы.** Составление базы данных генетических полиморфизмов цереброваскулярной патологии для последующего выявления зависимости наличия определенных групп полиморфизмов к пенетрантности генетической предрасположенности и развитию сосудистых заболеваний головного мозга.

**Материалы и методы исследования.** Для создания структурированной и удобной в использовании базы данных (далее – БД) был проведен биоинформационный поиск с использованием открытых ресурсов PubMed, SNP, Gene,

KEGG, а также полнотекстовых версий статей в PubMedCentral и Научной электронной библиотеке Российского индекса цитирования (eLibrary). При работе над проектированием базы данных использовались программы Microsoft Office Excel. Макет БД был получен с помощью программы Microsoft Visio. При проектировании поля и таблицы БД формировались с учетом следующих критериев: структурированность, полнота, взаимосвязанность данных.

Поиск генов, имеющих отношение к проблеме, и кодируемых ими белков осуществлялся в следующей последовательности:

- выбор семантической группировки (cerebral ischemia, cerebral atherosclerosis, hypertensive encephalopathy, cerebral thrombosis, stroke с учетом всех возможных синонимических конструкций – свыше 20);

- формирование выборки генов в GenBank и преобразование ее в выборку кодируемых белков;

- переход к ProteinDataBank и функциональное аннотирование выбранных полимеров;

- анализ и описание процессов, происходящих при возникновении полиморфизмов и нарушениях в структуре кодируемых белков, которые имеют доказанную связь с развитием ЦВЗ.

**Результаты и их обсуждение.** С учетом критериев был составлен следующий список полей и таблиц (см. рис. 1):

- ProcessesTable (таблица процессов).
- ProcessId (идентификатор процесса в спроектированной базе данных).

- ProcessName (полное название процесса).
- GenesTable (таблица генов).
- GeneId (идентификатор гена в спроектированной базе данных).
- GeneBankId (идентификатор гена в GeneBank).
- ShortName (общепринятое сокращение).
- FullName (полное наименование гена).
- PolymorphismsLink (конкретные полиморфизмы данного гена).
- EncodedProteinId (кодируемый данным геном белок).
- ProteinsTable (таблица белков).
- ProteinId (идентификатор белка в спроектированной базе данных).
- PDBId (идентификатор белка в ProteinDataBank).
- ShortName (общепринятое сокращение).
- FullName (полное наименование белка).
- ProteinFunktionTable (таблица функций белка).
- FunctionId (идентификатор функции в спроектированной базе данных).
- FunctionName (наименование функции, краткое описание).

Таблица ProcessesTable предназначена для записи процессов, которые имеют непосредственное отношение к ЦВЗ. Таблица GenesTable включает в себя список генов, обладающих полиморфизмами. Поле PolymorphismsLink содержит сведения об известных полиморфизмах данного гена в трех вариантах конкретизации: все известные по-

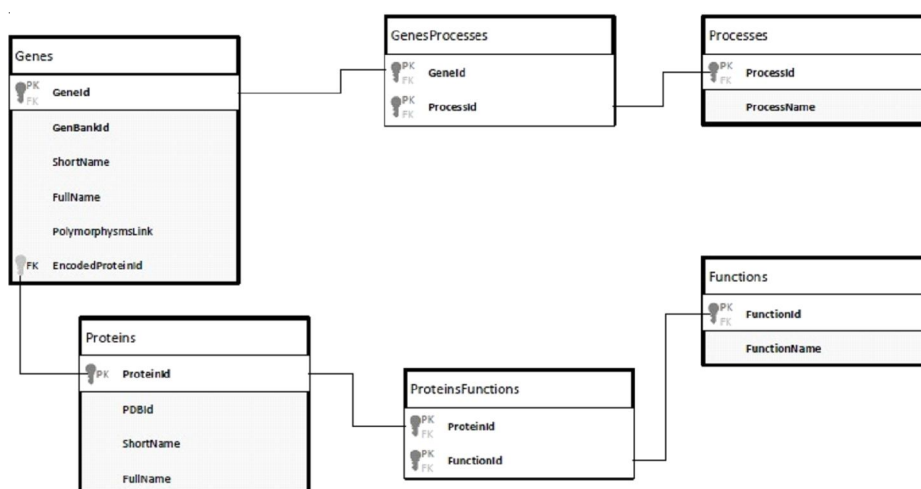


Рис. 1. Модель для построения базы данных генетических полиморфизмов, ассоциированных с высоким риском развития цереброваскулярной патологии

лиморфизмы; аннотированные полиморфизмы; аннотированные в отношении причастности к цереброваскулярной патологии. Таблица ProteinsTable несет информацию о кодируемых белках. Таблица ProteinFunctionTable описывает функции полиморфного белка и дает возможность составить представление о роли приобретаемых полиморфизмов в возникновении ЦВЗ. Остальные поля в приведенных таблицах несут общую информацию о генах, белках и функциях. Поля «ProcessId», «GeneId», «EncodedProteinId», «ProteinId», «FunctionId» являются идентификаторами процессов, генов, белков и функций в спроектированной нами базе данных. Они индивидуальны для каждого пункта и взаимосвязаны между собой.

В процессе исследования был проведен анализ порядка 350 генов, задействованных в процессах, обуславливающих возникновение ЦВЗ. Эти гены были связаны с контролем таких процессов, как вязкость и микроциркуляция крови; функция тромбоцитов и сосудистого эндотелия; макрофагальные и иммунные реакции; чувствительность нейронов к гипоксии; защитная функция нейроглии.

Количество точечных полиморфизмов в каждом из найденных генов колебалось в пределах от 500 до 3 000. Однако большинство из этих полиморфизмов не были аннотированы. В созданной БД по каждому гену приведен список полиморфизмов, полностью описанных в PubMedCentral и доказательно соотношенных с факторами риска конкретных заболеваний человека (от 5 до 46 на каждый ген). С помощью кросс-поиска с использованием семантической группировки ЦВЗ встречается полностью, получено в среднем от 2 до 5 полиморфизмов в каждом гене.

Данные о полиморфизме генов человека, которые наиболее часто используются при исследовании больных ЦВЗ, представлены на рисунках 2 и 3.

Эти гены контролируют ряд ключевых процессов, связанных с регуляцией реологии крови, тонуса сосудов и функции сосудистого эндотелия, обмена липидов и внутриклеточного переноса сигналов.

*ace* – ген, кодирующий ангиотензин, приводит к превращению фермента, который преобразует ангиотензин I в ангиотензин II. Потенцирует высвобождение норадреналина за счет прямого действия на постганглионарные симпатические нервные волокна.

*apoe* – ген, кодирующий аполипопротеин E. Белок оказывает антиатеросклеротическое действие. Один из важнейших апобелков, участвующий независимо в обмене липидов в крови, с одной стороны, и в обмене холестерина в мозге (и в некоторых других органах) – с другой.

*edn1* – кодирование белка эндотелин 1, который является сосудосуживающим пептидом из 21 аминокислоты, играющим ключевую роль в гомеостазе кровеносных сосудов. Основной механизм действия эндотелина 1 заключается в высвобождении кальция, что вызывает следующие процессы: 1) стимуляцию всех фаз гемостаза, начиная с агрегации тромбоцитов и заканчивая образованием красного тромба; 2) сокращение и рост гладких мышц сосудов, приводящее к утолщению стенки сосудов и уменьшению их диаметра – вазоконстрикции.

*f2* – ген, кодирующий протромбин. Напрямую отвечает за образование тромба.

*hmox1* – представляет собой ген, кодирующий фермент гемооксигеназу 1. Продук-

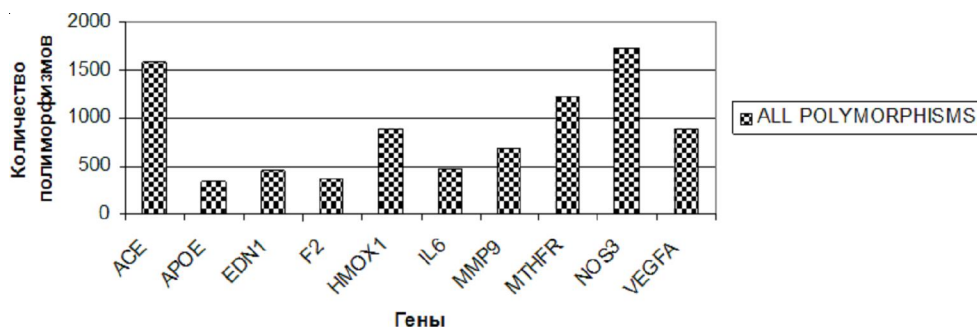


Рис. 2. Общее количество обнаруженных полиморфизмов для 10 генов, наиболее часто встречающихся в исследованиях цереброваскулярной патологии человека

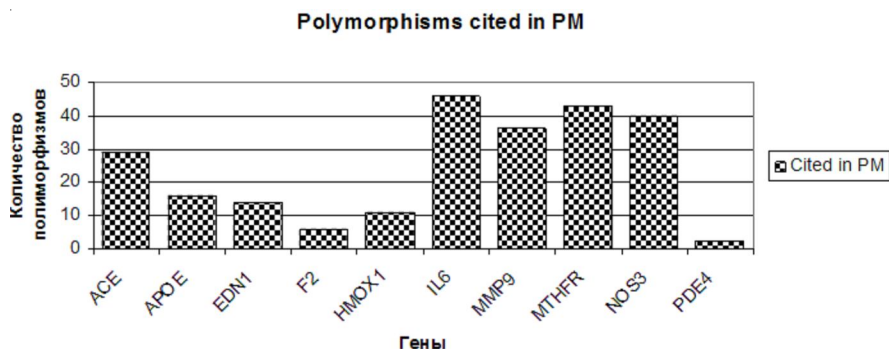


Рис. 3. Количество аннотированных полиморфизмов для 10 генов, наиболее часто встречающихся в исследованиях цереброваскулярной патологии человека

ты окисления гема обладают противовоспалительными свойствами за счет подавления активности интерлейкина-1.

*il6* – кодирует интерлейкин 6, цитокин с широким спектром биологических функций. (стимулирует иммунный ответ, стимулирует пролиферацию и дифференцировку В- и Т-клеток, стимулирует лейкоцитопозез).

*mmp9* – ген, кодирующий ММР9, разрушающую коллагены IV и V типов.

*mthfr* – ген, кодирующий метилен-тетрагидрофолат-редуктазу. Фермент необратимо преобразует 5,10-метилентетрагидрофолат в 5-метилтетрагидрофолат.

*nos3* – ген, кодирующий кальцийзависимую NO-синтазу, которая регулирует сосудистый тонус и ингибирует агрегацию тромбоцитов.

*pde4* – ген, кодирующий фосфодиэстеразу, фермент, относящийся к семейству металлофосфогидролаз, регулирующий внеклеточное соотношение циклических аденозинмонофосфата и гуанозинмонофосфата за счет их расщепления до 5'-монофосфатного нуклеотида.

**Заключение.** Созданная БД позволяет составить полноценное представление о структуре генетических полиморфизмов, ассоциированных с высоким риском развития цереброваскулярной патологии. БД имеет практическое применение, поскольку может быть использована для виртуального скрининга и последующего моделирования потенциальной эффективности SNP-полиморфизмов в качестве диагностических тест-систем по выявлению лиц, предрасположенных к сосудистым заболеваниям головного мозга.

### СПИСОК ЛИТЕРАТУРЫ

1. Верещагин, Н. В. Принципы диагностики и лечения больных с острыми ишемическими нарушениями мозгового кровообращения / Н. В. Верещагин, З. А. Суслина, М. А. Пирадов // Атмосфера. Нервные болезни. – 2002. – № 1. – С. 8–11.
2. Гайфуллина, Р. Ф. Роль генетического полиморфизма в патогенезе цереброваскулярных заболеваний / Р. Ф. Гайфуллина, М. Н. Катина, Ф. Ф. Ризванова // Казанский медицинский журнал. – Т. 93, № 4. – С. 663–666.
3. Гусев, Е. И. Нейропротективная терапия ишемического инсульта / Е. И. Гусев, В. И. Скворцова // Атмосфера. Нервные болезни. – 2002. – № 1. – С. 3–7.
4. Мушкамбаров, Н. Н. Молекулярная биология : учеб. пособие для студентов мед. вузов / Н. Н. Мушкамбаров, С. Л. Кузнецов. – М. : МИА, 2010. – 326 с.
5. Опыт применения препарата церепро (холина альфоцерат) при лечении амбулаторных больных с хроническими прогрессирующими сосудистыми заболеваниями головного мозга / Е. В. Костенко [и др.] // Журнал неврологии и психиатрии им. С. С. Корсакова. – 2012. – Т. 112, № 3. – С. 24–26.
6. Петров В. И. Вариабельность повреждения нейронов и соотношение активности изоформ нитроксидсинтазы при ишемии-реперфузии головного мозга у крыс / В. И. Петров [и др.]. // Бюллетень экспериментальной биологии и медицины. – 2007. – № 11. – С. 495–500.
7. Пономарев, Э. А. Морфологические параметры нейропротекции при ишемии-реперфузии головного мозга у крыс / Э. А. Пономарев, В. В. Новочадов, Н. Н. Стрепетов // Вестник Волгоградского государственного медицинского университета. – 2010. – № 1. – С. 91–93.

8. Постнова, М. В. Морфофункциональные особенности ядер переднего гипоталамуса в обеспечении стресс-реактивности организма / М. В. Постнова, В. В. Новочадов, М. Б. Потанин // *Фундаментальные исследования*. – 2013. – № 10–2. – С. 366–370.

9. Contestabile A. Neuronal-glia interactions define the role of nitric oxide in neural functional processes / A. Contestabile, B. Monti, E. Polazzi // *Curr. Neuropharmacol.* – 2012. – Vol. 10, № 4. – P. 303–310.

## BIOINFORMATIC SCREENING AND DATABASE CREATION FOR GENES ASSOCIATED TO HIGH RISK OF CEREBROVASCULAR PATHOLOGY

**Dergileva Aleksandra Yakovlevna**

Student, Bioengineering and Bioinformatics Department,  
Volgograd State University  
saschenkadergileva@gmail.com  
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

**Kishenevsky Aleksandr Mikhailovich**

Postgraduate Student, Bioengineering and Bioinformatics Department,  
Volgograd State University  
kish-sanek@mail.ru  
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

**Abstract.** The paper deals with the pathogenesis of cerebrovascular disease, risk factors and genetic aspects of the formation of vascular diseases of the brain and their dependence on the penetrance genetic polymorphisms. A bio-information search was carried out with the use of public resources PubMed, SNP, Gene, KEGG and all articles presented in PubMedCentral and Scientific Electronic Library of Russian citation index (eLibrary). About 350 genes involved in the processes determining the occurrence of CVD were analyzed. Those genes were related to such control processes as viscosity and microcirculation; platelet function and vascular endothelium; macrophage and immune responses; sensitivity of neurons to hypoxia; the protective function of the neuroglia. The number of point polymorphisms found in each of the genes ranged from 500 to 3000.

A model for building a database of genetic polymorphisms associated with a high risk of cerebrovascular pathology was created. The model let present a full picture of the structure of genetic polymorphisms associated with a high risk of cerebrovascular disease. The database has a practical application because it can be used for virtual screening and subsequent modeling of the potential effectiveness of SNP-polymorphisms as diagnostic test systems to identify individuals predisposed to cerebrovascular diseases.

**Key words:** cerebral ischemia, stroke, hypertensive encephalopathy, cerebrovascular pathology, gene polymorphism, database.