

Рекомендации для
интерпретации данных,
полученных методом
Next Generation Sequencing
(NGS).

Москва 2016г

Авторский коллектив:

О.П. Рыжкова (ФГБНУ «МГНЦ», ryzhkova@dnalab.ru)

А.В. Поляков (ФГБНУ «МГНЦ», apol@dnalab.ru)

М.Ю. Скоблов (ФГБНУ «МГНЦ», mskoblov@gmail.com)

В.А. Степанов (НИИ Медицинской генетики, Томск,
vadim.stepanov@medgenetics.ru)

Е.Б. Прохорчук (Центр "Биоинженерия" РАН, prokhortchouk@gmail.com)

Э.К. Хуснутдинова (Институт биохимии и генетики, Уфа, elzakh@mail.ru)

А. А. Афанасьев (iBinom, a@ibinom.com)

О.Л. Кардымон (iBinom, olga@ibinom.com)

Ф. Коновалов (Геномед, fedorkonovalov@gmail.com)

А. Мазур (Геноаналитика, mazur.am@gmail.com)

Е.В. Заклязьминская (Центр хирургии им. Петровского, helenezak@gmail.com)

А.А. Костарева (Институт мол. биологии и генетики центра им. В.А. Алмазова,
Санкт-Петербург, akostareva@hotmail.com)

Оглавление

[Методология](#)

[Определение](#)

[Терминология](#)

[Номенклатура](#)

[Базы данных](#)

[Популяционные](#)

[Фенотипы](#)

[Специфичные базы данных \(локус/ болезнь/ этно/ другие\)](#)

[Кодирующая последовательность](#)

[Рекомендации по использованию баз данных](#)

[Рекомендации по использованию литературы](#)

[Компьютерные \(in silico\) предсказательные программы](#)

[Миссенс замены](#)

[Изменения в сайтах сплайсинга](#)

[Критерии для интерпретации вариантов](#)

[Патогенный \(P\)](#)

[Категория «очень сильный» \(very strong-PVS\):](#)

[Категория «сильный» \(strong-PS\):](#)

[Категория «средней тяжести» \(moderate-PM\):](#)

[Доброкачественный \(B\)](#)

[Категория «очень сильный \(независимый\)» \(stand-alone – BA\):](#)

[Категория «сильный» \(strong – BS\):](#)

[Категория «вспомогательная» \(support – BP\):](#)

[Правила комбинирования критериев для классификации вариантов.](#)

[Общие требования к отчёту и заключение о результатах](#)

[Информация о пациенте](#)

[Информация о выявленных вариантах:](#)

[Отчёт по вариантам в генах неопределённого значения, показанных к тестированию](#)

[Методология](#)

[Повторный анализ и подтверждение результатов](#)

[Особые положения](#)

[Особенности интерпретация вариантов у здоровых или людей без симптомов](#)

[Распространённые мультифакторные заболевания](#)

[Особенности интерпретации и анализа митохондриальных вариантов:](#)

[Особенности анализа соматических вариантов:](#)

[Особенности фармакогеномного анализа и интерпретации:](#)

[Напутствия и выводы настоящих руководств](#)

[для лабораторий](#)

[Для врачей](#)

[Выявлен патогенный вариант нуклеотидной последовательности](#)

[Выявлен вероятно патогенный вариант нуклеотидной последовательности](#)

Методология

В качестве основы для создания рекомендаций были взяты существующие протоколы по интерпретации результатов секвенирования нового поколения (NGS) ACMG и CAP, которые были переработаны и расширены группой ведущих специалистов в области генетики и биоинформатики РФ. Валидация рекомендаций базировалась на внутренней экспертной оценке. Комментарии, полученные от экспертов, систематизировались и обсуждались председателем и членами рабочей группы. Проект рекомендаций был рецензирован независимыми экспертами. Для окончательной редакции и контроля качества рекомендации были повторно рецензированы членами рабочей группы, которые пришли к выводу, что все замечания и комментарии были приняты во внимание, риск систематических ошибок при разработке рекомендаций сведен к минимуму.

Определение

Next Generation Sequencing — техника определения [нуклеотидной последовательности ДНК](#) и [РНК](#) для получения формального описания их [первичной структуры](#). Технология методов секвенирования нового поколения (NGS) позволяет «прочитать» одновременно сразу несколько участков [генома](#), что является главным отличием от более ранних методов секвенирования. NGS осуществляется с помощью повторяющихся циклов [удлинения цепи](#), индуцированного [полимеразой](#), или многократного [лигирования олигонуклеотидов](#). В ходе NGS могут генерироваться до сотен миллионов и миллиардов нуклеотидных последовательностей за один рабочий цикл.

Терминология

Во избежание путаницы в терминологии предложено заменить широко используемые определения “мутация” и “полиморфизм” на “вариант нуклеотидной последовательности” со следующими пятью модификациями:

- патогенный (pathogenic)
- вероятно патогенный (likely pathogenic)
- неопределенного значения (uncertain significance)
- вероятно доброкачественный (likely benign)
- доброкачественный (benign)

Рекомендуется все утверждения патогенности представлять совместно с нозологией и типом наследования, например, *c.1521_1523delCTT (p.Phe508del), патогенный, муковисцидоз, аутомно-рецессивный*.

Номенклатура

Единая, рекомендуемая мировым сообществом к использованию номенклатура <http://www.hgvs.org/mutnomen> (обязательно указание используемой версии).

Инструменты для обеспечения правильной номенклатуры HGVS при описании вариантов представлены на сайте <https://mutalyzer.nl>.

Референсная последовательность должна быть получена из:

- RefSeq базы данных Национального центра биотехнологической информации (<http://www.ncbi.nlm.nih.gov/RefSeq/>) с указанием номера версии или
- из базы данных Locus Reference Genomic (<http://www.lrg-sequence.org>).

Геномные координаты следует использовать и определять в соответствии со стандартом геномной сборки (например: GRCh37/hg19) или референсной геномной последовательности, охватывающей весь ген (в том числе 5' и 3' нетранспируемые области и промотор).

Референсный транскрипт должен представлять собой или наиболее клинически значимый или/и наиболее длинный из известных транскриптов. Референсные транскрипты, рекомендуемые мировым сообществом, могут быть идентифицированы с помощью баз данных:

- Locus Reference Genomic
- the Consensus CDS Database
- the Human Gene Mutation Database (<http://www.hgmd.cf.ac.uk>)
- ClinVar (<http://www.ncbi.nlm.nih.gov/clinvar>) или
- локус-специфичных баз данных.

Тем не менее лаборатории должны проверять клиническую значимость варианта во всех возможных транскриптах.

Рекомендуются так же использовать три особых исключения к правилам номенклатуры HGVS:

1. «X» приемлем для использования в отчетности по нонсенс вариантам в дополнение к используемым рекомендациям HGVS использовать "*" и "Ter".
2. Нумеровать экзоны в соответствии с выбранным референсным транскриптом, используемым для обозначения варианта
3. Использовать термин "патогенный" вместо выражения "влияет на функцию", т.к. клиническая интерпретация непосредственно направлена на оценку патогенности.

Базы данных

Все обнаруженные варианты необходимо классифицировать по патогенности. Оценка патогенности выявленных вариантов включает изучение медицинской и научной литературы и баз данных. Для поиска описанных ранее вариантов рекомендуется использовать следующие базы данных:

Популяционные

<p>Exome Aggregation Consortium http://exac.broadinstitute.org/</p>	<p>База данных вариантов, найденных при проведении экзомного секвенирования у 61,486 неродственных индивидуумов, являющихся участниками различных болезнь-специфичных и популяционных генетических исследований. Детские заболевания и связанные с ними лица были исключены из выборки.</p>
<p>Exome Variant Server http://evs.gs.washington.edu/EVS/</p>	<p>База данных вариантов, найденных во время экзомного секвенирования нескольких крупных когорт лиц европейского и афро-американского происхождения. Включает в себя данные о покрытии, для учета</p>

	информации об отсутствии варианта.
1000 Genomes Project http://browser.1000genomes.org/index.html	База данных вариантов, найденных во время геномного и таргетного секвенирования с низким и высоким покрытием среди 26-и популяций. Обеспечивает большее разнообразия по сравнению Exome Variant Server, но содержит данные более низкого качества, а некоторые когорты содержат данные о родственных индивидуумах.
dbSNP http://www.ncbi.nlm.nih.gov/snp	База данных коротких генетических вариантов (как правило, ≤ 50 п.о.), собранных из различных источников. Может содержать и патогенные варианты.
dbVar http://www.ncbi.nlm.nih.gov/dbvar	База данных структурных вариантов (как правило, >50 п.о.), составленная из многих источников

Фенотипы

OMIM http://www.omim.org/	База данных генов человека и генетических состояний, которая содержит репрезентативную выборку вариантов, ассоциированных с заболеваниями.
Human Gene Mutation Database http://www.hgmd.cf.ac.uk/ac/index.php	База данных аннотированных вариантов, опубликованных в литературе. Доступ к основной части контента требует оплаты.
ClinVar	База данных утверждений о клинической значимости и

http://www.ncbi.nlm.nih.gov/clinvar/	фенотипической взаимосвязи вариантов. Содержит данные низкого качества, <u>к использованию не рекомендуется.</u>
---	--

Специфичные базы данных (локус/ болезни/ этно/ другие)

Human Genome Variation Society http://www.hgvs.org/dblist/dblist.html Leiden Open Variation Database http://www.lovd.nl	Сайт общества по изучению вариаций генома человека создало список из тысяч баз данных, которые предоставляют варианты аннотации на конкретные разновидности человеческой вариации. Большая доля баз данных представлена системе Leiden Open Variation Database system.
DECIPHER https://decipher.sanger.ac.uk/	Молекулярно-цитогенетическая база данных для врачей и исследователей, связывающая геномные данные, полученных с помощью микрочипов, с фенотипом, используя геномный браузер Ensembl.

Кодирующая последовательность

NCBI Genome http://www.ncbi.nlm.nih.gov/genome	Ресурс полногеномных референсных последовательностей
RefSeqGene http://www.ncbi.nlm.nih.gov/refseq/rgsg/	Ресурс референсных последовательностей клинически релевантных генов
MitoMap http://www.mitomap.org/MITOMAP/	Пересмотренная Кембриджем референсная последовательность митохондриальной ДНК человека

Рекомендации по использованию баз данных

При использовании баз данных клинические лаборатории должны проверить следующую информацию:

1. Частота обновлений, осуществление курирования базы данных.
2. Подтвердить использование HGVS номенклатуры и определить референсные последовательности для сборки генома и транскриптов, используемые при наименовании вариантов.
3. Оценить все показатели качества, приводящиеся для оценки точности данных (может потребоваться чтение соответствующих публикаций)
4. Определить источник и независимость исследований, в которых содержится информация о варианте

Рекомендации по использованию литературы

Литература, употребляющая старую номенклатуру и классификацию должна быть использована с осторожностью. Так же настороженно следует относиться к данным, встретившимся лишь в одном литературном источнике.

Следует учитывать, что одни и те же индивидуумы, а также их родственники, могут встречаться несколько раз в различных литературных источниках в зависимости от контекста и размера выборки. Это может быть связано как с пересечением авторства и межлабораторными коллаборациями, так и с тем, что данные по пробанду и членам семьи имеются в различных клинических базах. В результате это может привести к ошибочному дублированию при подсчете интересующих случаев и ложному увеличению частоты варианта. Пересечение по авторству или учреждениям является первым признаком к потенциальному пересечению наборов данных.

Компьютерные (in silico) предсказательные программы

Если вариант не был описан в литературе ранее и не представлен ни в одной из баз данных или сведения о нем недостаточны, для решения о его значимости можно использовать результаты программ предсказаний

патогенности. Ниже представлены адреса сайтов и краткое описание наиболее используемых на данный момент программ предсказания.

Название - Вебсайт	Основа
ConSurf - http://consurftest.tau.ac.il/	Эволюционная консервативность
FATHMM - http://fathmm.biocompute.org.uk/	Эволюционная консервативность
MutationAssessor - http://mutationassessor.org/	Эволюционная консервативность
PANTHER - http://www.pantherdb.org/tools/snpScoreForm.jsp	Эволюционная консервативность
PhD-SNP - http://snps.biofold.org/phd-snp/phd-snp.html	Эволюционная консервативность
SIFT - http://sift.jcvi.org/	Эволюционная консервативность
SNPs&GO - http://snps-and-go.biocomp.unibo.it/snps-and-go/	Структура/функция белка

Миссенс замены

Название - Вебсайт	Основа
Align GVGD - http://agvgd.iarc.fr/agvgd_input.php	Структура/функция белка и эволюционная консервативность
MAPP – http://mendel.stanford.edu/SidowLab/downloads/MAPP/index.html	

MutationTaster - http://www.mutationtaster.org/	Выравнивание и измерение сходства между последовательностью варианта и последовательностью гомологичного белка
MutPred - http://mutpred.mutdb.org/	
PolyPhen-2 - http://genetics.bwh.harvard.edu/pph2/	
PROVEAN - http://provean.jcvi.org/index.php	
SIFT- http://provean.jcvi.org/index.php	
nsSNPAnalyzer - http://snpanalyzer.uthsc.edu/	
Condel - http://bg.upf.edu/fannsdb/	Объединяет SIFT, PolyPhen-2 и MutationAssessor

Изменения в сайтах сплайсинга

Название – Вебсайт	Основа
GeneSplicer - http://ccb.jhu.edu/software/genesplicer/	Модели Маркова
Human Splicing Finder - http://www.umd.be/HSF/	Основанное на положении варианта
MaxEntScan – http://genes.mit.edu/burgelab/maxent/Xmaxentscanscoreseq.html	Принцип максимальной энтропии
NetGene2 – http://www.cbs.dtu.dk/services/NetGene2/	Нейронные сети

NNSplice - http://www.fruitfly.org/seq_tools/splice.html	Нейронные сети
FSPLICE - http://www.softberry.com/berry.phtml?topic=fsplICE&group=programs&subgroup=gfind	Видо-специфичный предиктор сайтов-сплайсинга, основанный на модели весовой матрицы

Критерии для интерпретации вариантов

Варианты, описанные ранее в нескольких источниках (кроме ClinVar), являющиеся причиной развития интересующего фенотипа и подходящие под тип наследования, классифицируются как патогенные.

Для интерпретации остальных вариантов предлагается использовать два набора критериев: первый для классификации вероятно патогенных вариантов, второй для классификации вероятно доброкачественных вариантов. Каждый патогенный критерий может оцениваться несколькими категориями: очень сильный (PVS1), сильный (PS1-4); средней тяжести (PM1-6) или вспомогательный (PP1-5). Каждый доброкачественный критерий как очень сильный (независимый) (BA1), сильный (BS1-4) или вспомогательный (BP1-6). Нумерация признаков внутри категории не дает никаких усилений варианта, важность представляет только наименование, нумерация нужна для упрощения их использования.

Для каждого варианта специалист выбирает подходящие признаки, которые затем объединяет в соответствии с указанными ниже правилами, классифицируя значимость варианта по пятиуровневой системе: патогенный, вероятно патогенный, неопределенного значения, вероятно доброкачественный, доброкачественный. Если вариант не отвечает критериям, используемым любым из этих наборов (патогенных или доброкачественных), или доказательства для доброкачественности и патогенности противоречивы, то вариант по умолчанию становится неопределенного значения.

Патогенный (P)

Категория «очень сильный» (very strong-PVS):

Признаки:

PVS1 LOF-варианты - варианты, приводящие к прекращению синтеза белка (нонсенс-мутации; мутации со сдвигом рамки считывания; изменения канонических ± 1 или ± 2 нуклеотидов сайта сплайсинга; варианты, приводящие к исчезновению стоп-кодона; делеции/дупликации одного или нескольких экзонов) в генах, где данный тип варианта является известной причиной развития заболевания.

Принять во внимание:

- Существуют гены, где варианты, приводящие к прекращению синтеза белка, не являются патогенными (например, GFAP, MYH7)
- Будьте внимательны при интерпретации вариантов, расположенных близко к 3' концу гена
- Будьте внимательны со сплайс-вариантами, которые прогнозируемо приводят к пропуску экзонов, но оставляют оставшуюся часть белка нетронутой и не меняют рамку считывания.
- Будьте осторожны при наличии нескольких транскриптов.

Категория «сильный» (strong-PS):

Признаки:

PS1 вариант, приводящий к замене на ту же аминокислоту в том же положении, которая была описана ранее как патогенная для этого заболевания. Например: если описана замена G>C, приводящая к замене Val>Leu, то замена G>T, приводящая к той же аминокислотной замене, описывается признаком PS1.

Принять во внимание:

- Будьте внимательны к вариантам, патогенность которых обусловлена изменением сайта сплайсинга, а не аминокислоты (они относятся к PVS1)

PS2 *De novo* вариант: у обоих родителей пациента вариант отсутствует

Принять во внимание:

- Только подтверждения отцовства недостаточно. Донорство яйцеклеток, суррогатное материнство, ошибки при подсадке эмбриона и так далее, могут привести к неверному определению материнства.

PS3 функциональные исследования (*in vitro* или *in vivo*), подтверждают патогенный эффект варианта на ген или генный продукт.

Принять во внимание:

- Функциональные исследования должны быть валидируемыми, воспроизводимыми и проведенными лабораториями, имеющими хорошую репутацию.

PS4 Распространенность варианта у больных индивидуумов значительно выше чем у контрольной группы. Отношение шансов (OR), полученное из исследований случай-контроль > 5.0.

Принять во внимание:

- В случае очень редких вариантов, где исследования пациент-контроль не достигают показателей статистической значимости, предварительное наблюдение варианта у нескольких не связанных родством пациентов с одинаковым фенотипом и отсутствие его в контрольной группе, является вариантом средней тяжести (moderate level).

Категория «средней тяжести» (moderate-PM):

Признаки:

PM1 вариант расположен в «горячей» точке и/или важных и хорошо исследованных функциональных доменах (например, активный сайт фермента), в которых нет доброкачественных изменений.

PM2 вариант отсутствует в контрольной выборке (или встречается с крайне низкой частотой: для АД заболеваний частота не должна превышать 0,01%, для AP заболеваний - 0,5%), а также отсутствует в открытых базах данных (ESP, 1KG или EXAC)

Принять во внимание:

- Популяционные данные для инсерций/делеций могут быть плохо аннотированы в базах, данные для которых получены с помощью секвенирования следующего поколения

PM3 вариант, находящийся в транс-положении с описанным патогенным вариантом при рецессивных заболеваний

Принять во внимание:

- Требуется анализ родителей (или потомков)

PM4 Варианты, приведшие к синтезу белка с измененной длиной (инсерции/делеции в рамках считывания в неповторяющихся регионах или потеря стоп-кодона (замена на аминокислоту)).

PM5 Новые миссенс варианты в аминокислотном остатке, где другие миссенс варианты ранее были установлены как патогенные

Принять во внимание:

- остерегайтесь изменений, которые влияют не только на белковый/аминокислотный уровень, но и на сплайсинг.

Категория «вспомогательная» (supporting-PP):

Признаки:

PP1 вариант в гене, для которого точно установлена связь с болезнью, с сегрегацией с болезнью у нескольких пораженных членов семьи.

Принять во внимание:

- Может быть отнесен к категории «сильный» (PS) с накоплением сегрегационных данных

PP2 миссенс вариант в гене, который имеет низкий уровень доброкачественных миссенс изменений и в котором миссенс варианты являются обычным механизмом возникновения заболевания.

PP3 результаты не менее трех программ предсказания подтверждают патогенное воздействие на ген или генный продукт.

Принять во внимание:

Т.к. многие программы предсказаний используют одни и те же или очень схожие расчеты, результаты нескольких программ не могут считаться независимыми критериями и должны оцениваться в сочетании. PP3 может быть применен только один раз для оценки варианта.

PP4 фенотип пациента или/и семейная история высоко специфичны для заболевания с данной генетической этиологией

PP5 источники с хорошей репутацией показали патогенность варианта, но независимая оценка не проводилась.

Доброкачественный (B)

Категория «очень сильный (независимый)» (stand-alone – BA):

Признаки:

BA1 Частота аллеля >5% в ESP, 1KP или в EXAC

Категория «сильный» (strong – BS):

Признаки:

BS1 частота аллеля больше, чем ожидаемая для заболевания (для АД заболеваний частота аллеля не должна превышать 0,01%, для AP заболеваний - 0,5%)

BS2 если вариант встретился у здорового взрослого человека при заболеваниях с полной пенетрантностью, манифестирующих в детском возрасте: 1) в гомозиготном состоянии с рецессивным типом наследования; 2) в гетерозиготном состоянии с доминантным типом наследования или 3) в гемизиготном состоянии при X-сцепленном типе наследования.

BS3 функциональные исследования (*in vitro* или *in vivo*), подтверждают отсутствие патогенного эффекта на ген или генный продукт.

Принять во внимание:

- Функциональные исследования должны быть валидируемыми, воспроизводимыми и проведенными лабораториями, имеющими хорошую репутацию.

BS4 вариант в гене, для которого точно установлено отсутствие сегрегации с болезнью у пораженных членов семьи

Принять во внимание:

- Наличие фенокopies может имитировать отсутствие сегрегации у пораженных лиц.

- Наличие в семье более чем одного патогенного варианта может также привести к ложному отсутствию сегрегации наследования варианта с заболеванием у пораженных лиц

Категория «вспомогательная» (support – BP):

Признаки:

BP1 миссенс вариант в гене, где известной причиной заболевания являются мутации, приводящие к изменению длины белка.

BP2 при заболеваниях с полной пенетрантностью, если выявлена одна патогенная мутация, соответствующая фенотипу, а неописанный вариант находится с ней в транс- положении при доминантном типе наследования или в цис- положении при любом типе наследования.

BP3 делеции/инсерции с сохранением рамки считывания, если не проводилось функциональное исследование.

BP4 результаты не менее трех программ предсказания подтверждают отсутствие воздействия варианта на ген или генный продукт.

Принять во внимание:

Т.к. многие программы предсказаний используют одни и те же или очень схожие расчеты, результаты нескольких программ не могут считаться независимыми критериями и должны оцениваться в сочетании. BP3 может быть применен только один раз для оценки варианта.

BP5 источники с хорошей репутацией показали отсутствие патогенности варианта, но независимая оценка не проводилась.

BP6 Синонимичный (silent) вариант, для которого алгоритмы предсказания сплайсинга показывают отсутствие влияния на консенсусную последовательность сплайс-сайтов (отсутствует создание нового сайта-сплайсинга) и нуклеотид не является высококонсервативным в эволюции.

Правила комбинирования критериев для классификации вариантов.

Патогенный	1) 1 очень сильный (PVS1) и а) 1 и более сильный (PS1–PS4) или
-------------------	---

	<p>b) 2 и более умеренной тяжести (PM1–PM5) или c) 1 умеренной тяжести (PM1–PM5) и 1 вспомогательный (PP1–PP5) или d) 2 и более вспомогательных (PP1–PP5) или 2) 2 и более сильных (PS1-PS4) или 3) 1 сильный (PS1-PS4) и а) 3 и более средней тяжести (PM1–PM5) или b) 2 средней тяжести (PM1–PM5) и 2 и более вспомогательных (PP1–PP5) или c) 1 средней тяжести (PM1–PM5) и 4 и более вспомогательных (PP1–PP5)</p>
Вероятно патогенный	<p>1) 1 очень сильный (PVS1) и 1 умеренной тяжести (PM1-PM5) или 2) 1 сильный (PS1-PS4) и 1-2 умеренной тяжести (PM1-PM5) или 3) 1 сильный (PS1-PS4) и 2 и более вспомогательных (PP1–PP5) или 4) 3 и более средней тяжести (PM1–PM5) или 5) 2 средней тяжести (PM1–PM5) и 2 и более вспомогательных (PP1–PP5) или 6) 1 средней тяжести (PM1–PM5) и 4 и более вспомогательных (PP1–PP5)</p>
Неопределенного значения	<p>1) Вариант не описывается ни одним критерием или 2) Критерии доброкачественности и патогенности противоречат друг другу</p>
Вероятно доброкачественный	<p>1) 1 сильный (BS1-BS4) и 1 вспомогательный (BP1-BP6) или 2) 2 и более вспомогательных (BP1-BP6)</p>
Доброкачественный	<p>1) 1 очень сильный (BA1) или 2) 2 сильных (BS1-BS4)</p>

В заключении следует отражать только варианты, относящиеся к группам: патогенный, вероятно патогенный, неопределенного значения. Варианты, относящиеся к группам: доброкачественный и вероятно доброкачественный в заключении отражать не следует.

При обнаружении патогенных вариантов, хорошо характеризующих фенотип пациента и не противоречащих типу наследования (1 вариант при АД, 2 при АР и 1 при Х-сцепленном типе наследования у мальчиков), в заключении отражать только их, другие выявленные варианты в заключении отражать не следует.

При обнаружении одного патогенного и одного вероятно патогенного вариантов нуклеотидной последовательности, хорошо характеризующих фенотип пациента, при заболевании с АР типом наследования, в заключении отражать только их, другие выявленные варианты в заключении отражать не следует.

В остальных случаях отражать все выявленные варианты, относящиеся к группам вероятно патогенный и неопределенного значения.

Общие требования к отчёту и заключение о результатах

Несмотря на сложность информации отчёт должен быть краткими и ёмким.

Разделы отчёта должны включать структурированные результаты: информацию о пациенте, список выявленных патогенных, вероятно патогенных вариантов и вариантов неопределенного значения, интерпретацию, методологию, ссылки к литературе.

Информация о пациенте

Информация о пациенте должна включать: ФИО, дату рождения, пол, клинический диагноз, родственные отношения (если есть данные о других членах семьи)

Информация о выявленных вариантах:

- список патогенных, вероятно патогенных и вариантов неопределенного значения, выявленных при анализе, по номенклатуре HGVS
- название гена
- положение (позиция замены в геноме по GRCh/hg)

- зиготность
- номер экзона
- номенклатура на уровне нуклеотида и ДНК
- номенклатура на уровне белка (в случаях, где используется историческая (альтернативная) номенклатура, рекомендуется указывать оба варианта)
- номер используемой референсной последовательности (NM)
- данные о покрытии
- название заболевания
- тип наследования
- частоты варианта в базах данных
- результаты программ предсказания

Дополнительная информация:

- семейное происхождение, если известно
- при анализе конкретных вариантов, перечисление списка искомых вариантов в рамках исследования

При интерпретации результатов должна приводиться доказательная база, на основании которой классифицированы варианты в отчёте.

Стоит указывать:

- предсказательный эффект на синтезируемый белок
- встречался ли данный вариант ранее в литературе или в базе данных заболевания или контрольной группы, если да, то сколько раз (1 или более), № замены и частоту выявленного аллеля с указанием базы данных из которой получена информация
- Ссылки (при наличии), которые использовались для классификации варианта, указываются в двух местах:
 - а) рядом с описанием варианта
 - б) в конце отчёта
- рекомендации врачу для дополнительного клинического тестирования
- данные о неполной пенетрантности, фенотипической и генетической гетерогенности
- Обобщённый вывод результатов программ предсказания и консервативного эволюционного анализа

Не рекомендуется указывать частные вычислительные предсказания, а именно:

- скоры (врачи могут быть незнакомы с ограничениями предсказывающих алгоритмов)
- такие термины как «вредная» / «гибельная» мутация (высока вероятность неправильной интерпретации врачами)

Отчёт по вариантам в генах неопределённого значения, показанных к тестированию

Ген неопределённого значения - это ген без доказанной ассоциации с фенотипом пациента или с ассоциацией с другим фенотипом, отличным от фенотипа исследуемого пациента.

В случаях с генами неопределённого значения предлагаемая классификация применяться не может, по следующим причинам:

- *de novo* вариант не имеет сильной доказательной базы (в среднем человек имеет 1 *de novo* вариант на экзом и 100 *de novo* вариантов на геном)
- у здорового человека встречаются от 2 до 4 вариантов на геном, нарушающих работу гена с пагубным эффектом на синтезируемый белок

Предлагается варианты, найденные в генах неопределённого значения пометать в отчёте как “варианты в генах неопределённого значения”. Эти варианты **всегда** должны быть классифицированы как неопределённого значения.

Новые клинические случаи с редкими фенотипами и выявленными патогенными вариантами смогут позволить в дальнейшем классифицировать эти варианты по представленным рекомендациям.

Методология

Раздел методология должен включать:

- методы выделения нуклеиновых кислот (ПЦР, гибридный захват, амплификация полного генома)
- описание прибора, на котором проводилось исследование и методы анализа нуклеиновых кислот
- официальное наименование гена (при анализе конкретного гена или списка генов), заявленное в Human Genome Organization Gene Nomenclature Committee. Информация о гене может быть указана в ссылке на электронный источник (при анализе экзонов)
- номер транскрипта согласно RefSeq

- геномный референс с указанием версии
- недоступность некоторых вариантов для анализа (при недостаточной покрытости регионов при секвенировании)

В отчёте **не могут указываться** клинические рекомендации по ведению пациента.

Лаборатория может связать лечащего врача с исследовательской группой, работающей с конкретным заболеванием, в случае если вариант классифицирован как «неопределенный».

Повторный анализ и подтверждение результатов

Ввиду растущей и развивающейся базы знаний о геномных вариантах, может потребоваться повторный анализ результатов проведенного исследования. Лаборатории должны предоставить описание используемой методики по повторному анализу данных и пояснить, требуется ли дополнительная оплата за такой анализ. Поощряется разработка новых эффективных подходов для предоставления пациенту обновлённой информации. **Врачу рекомендовано** периодически запрашивать информацию у лабораторий.

Обязательно проведение подтверждения результатов анализа NGS с помощью ортогональных методов для всех вариантов сиквенса, классифицированных как патогенные, вероятно патогенные и неопределенного значения для Менделевских заболеваний. Методы могут включать, но не ограничены:

- секвенирование целевой области ДНК по Сенгеру
- рестрикционный анализ
- MLPA-анализ
- тестирование членов семьи пациента

Особые положения

Особенности интерпретация вариантов у здоровых или людей без симптомов

- для именованя варианта патогенным необходима сильная доказательная база
- вероятность обнаружения вариантов, признанных патогенными, обычно очень низкая
- предсказанная пенетрантность патогенного варианта, обнаруженного в отсутствии фенотипа или семейного анамнеза, может быть намного ниже, чем у пациента с симптомами

Распространённые мультифакторные заболевания

- определение генов, ответственных за распространённые мультифакторные заболевания основывается на популяционном подходе (GWAS), а не на семейных исследованиях
- насчитывается более 1200 аллелей с повышенным риском развития таких заболеваний и наследственных черт; большинство этих вариантов находятся за пределами генных регионов
- польза тестирования на аллели с подобным риском в клинической практике не очевидна
- настоящее руководство не рассматривает форму отчёта по аллелям для мультифакторных признаков.
- включать в отчёт эти аллели возможно в качестве побочных сведений.
- термины «патогенный» и «вероятно патогенный» не приемлемы в данном контексте, даже если ассоциация статистически достоверна
- упоминать эти варианты стоит как «аллели риска», доказательство риска может быть выражено в терминах «установленная аллель риска», «вероятная аллель риска» или «неопределённая аллель риска»

Особенности интерпретации и анализа митохондриальных вариантов:

- Номенклатура отличается от стандартной номенклатуры для ядерных генов и использует нумерацию по последовательности митохондриальной ДНК (пример: m.8992T>C) и последовательности белка.
- Действующий применимый референсный сиквенс предоставлен Revised Cambridge Reference Sequence of the Human Mitochondrial DNA (NC_012920 gi:251831106)
- Гетероплазмия или гомоплазмия должны быть включены в отчёт вместе с приблизительной оценкой уровня гетероплазмии.
- Уровень гетероплазмии в разных типах тканей может различаться в зависимости от био-образца; следовательно, гетероплазмия должна быть интерпретирована в контексте ткани
- Данные о частоте встречаемости и опубликованные исследования зачастую являются единственным доступным критерием для проверки значимости варианта

Источники для интерпретации:

- MitoMap является главным информационным источником по митохондриальным вариантам и гаплотипам.
- по частоте встречаемости (mtdb.igp.uu.se)
- вторичным структурам, сиквенсам, картированию митохондриальной транспортной ДНК (mamitrna.u-strasbg.fr)
- митохондриальных гаплогрупп (phylotree.org)
- другая информация (mtdnacommunity.org/default.aspx)

Особенности анализа соматических вариантов:

Для анализа данного типа вариантов необходимо руководствоваться рекомендациями по интерпретации соматических вариантов с перечислением ссылок на опухолевые базы данных в дополнение к базам данных для структурных вариантов. Данное руководство неприменимо для анализа соматических вариантов.

Особенности фармакогеномного анализа и интерпретации:

- фенотип проявляется только после приёма лекарственного средства
- общепринятая номенклатура фармакогеномных аллелей применяет знак (*) для обозначения аллелей, которые часто представляют гаплотип или комбинацию вариантов того же аллеля
- до сих пор применяется традиционное нумерование нуклеотида, использующее устаревшую последовательность референса
- конвертация бывшей традиционной номенклатуры к стандартизированной номенклатуре, использующей современный референс при всей необходимости, рискованна
- Интерпретация вариантов зачастую требует установления гаплотипа по комбинации обнаруженных вариантов
- Для многих генов фармакогенома (гены, кодирующие ферменты) весь фенотип берётся из диплотипа, а именно, комбинации вариантов гаплотипов обоих аллелей
- Краткая сводка по генам и клинически значимым вариантам доступна на pharmgkb.org. Аллели и номенклатура по цитохрому семейства гена P450 доступна на cypalleles.ki.se

Для классификации вариантов рекомендуется использование терминов, связанных с метаболизмом (быстрый, средний, медленный); эффективностью (резистентный, быстро реагирующий, чувствительный); или риском, может быть более уместно нежели термина патогенный.

Напутствия и выводы настоящих руководств

для лабораторий

- наилучшие результаты получаются в результате плотного взаимодействия центров здравоохранения и клинической лаборатории в процессе проведения анализа
- потребность в предоставлении центрами здравоохранения подробной клинической информации для более точной интерпретации имеет возрастающее значение; лаборатория в праве отказать в анализе, если не получает достаточную информацию о фенотипе пациента вместе с био-образцом
- лаборатория обязана оценить вариант и ген в контексте анамнеза и семейной истории пациента, физиологического осмотра и предшествующих

лабораторных тестов, и при помощи этой информации определить является ли вариант патогенным, возможно патогенным или доброкачественным.

- рекомендуется проводить трио-тестирование (мать, отец и ребёнок с заболеванием) в ходе геномного или экзомного анализа, особенно при подозрении на рецессивный или *de novo* варианты
- возможно отфильтровывание или игнорирование большинства вариантов для доминантных заболеваний, когда такие варианты наблюдаются у здоровых родственников
- рекомендуется пользоваться ресурсами специфичными для данной патологии
- Данные, полученные методом NGS должны интерпретироваться специалистом (биоинформатиком), а не врачом или пациентом.

Для врачей

- на текущий день методы анализа вариантов несовершенны и приводимые в отчетах категории вариантов не подразумевают 100% уверенности. Не рекомендуется применение этих руководств как единственного обоснования фенотипа при диагностике Менделевских заболеваний; результаты анализа следует использовать совместно с данными других объективных методов исследования
- Вариант, классифицированный как патогенный при помощи предлагаемой схемы классификации может быть использован врачом для принятия медицинских решений как базовое доказательство.
- Вариант, классифицированный как вероятно патогенный, имеет достаточное доказательство для того, чтобы врач смог использовать результаты молекулярного тестирования в принятии клинического решения в сочетании с другими доказательствами по рассматриваемому заболеванию.
- Вариант неопределённого значения не должен быть использован в принятии клинических решений.
- Вариант, классифицированный как вероятно доброкачественный, имеет достаточное доказательство для того, чтобы врач заключил, что он не является причиной заболевания у пациента в сочетании с другой информацией.
- Вариант, приведённый в отчёте как доброкачественный, имеет достаточную доказательную базу, чтобы врач заключил что он не является причиной заболевания у пациента.

Примеры заключений

Выявлен патогенный вариант нуклеотидной последовательности

ЗАКЛЮЧЕНИЕ

по результатам секвенирования ДНК (панель «Эпилепсии»)

Пациент: ФИО

Пол: М/Ж

Дата рождения: дд.мм.гг

Вид материала: Кровь (венозная)

Дата забора: дд.мм.гг

Диагноз: Симптоматическая эпилепсия.

Данные секвенирования могут быть предоставлены по запросу лечащего врача.

1. Патогенные варианты нуклеотидной последовательности, являющиеся вероятной причиной заболевания.

Ген	Положение (GRCh37/hg19)	Генотип	Экзон	Положение в кДНК	Замена АК	Частота аллеля*	Референсная последовательность	Глубина прочтения
<i>TPP1</i>	chr11:6638271G>A	A/A	6	c.622C>T	p.Arg208*	0.0173159 %	NM_000391.3	176x

*Частоты аллелей приведены по базе Exome Aggregation Consortium (выборка до 60702 человек). н/д = нет данных (не описан)

ИНТЕРПРЕТАЦИЯ

У ФИО был проведен поиск патогенных мутаций, ассоциированных с наследственными эпилепсиями, а также с другими наследственными заболеваниями со сходными фенотипическими проявлениями.

Выявлен ранее описанный вариант нуклеотидной последовательности в 6 экзоне гена *TPP1* (chr11:6638271G>A, rs119455955) в гомозиготном состоянии, приводящий к появлению сайта преждевременной терминации трансляции в 208 кодоне (p.Arg208Ter, NM_000391.3). Вариант описан как мутация, которая в гомозиготной или в компаунд-гетерозиготной форме с другими мутациями приводит к развитию нейронального цероидного липофусциноза, тип 2 (OMIM: 204500) основными симптомами которого являются атаксия, миоклонические приступы, постепенный регресс интеллектуального развития [ссылки на

источники: или 2 статьи или база данных, в которой указано сколько раз встретился вариант (обязательно больше 2)]. Частота варианта в контрольной выборке ExAC составляет 0.0173%.

Результат требует проверки альтернативными методами (секвенирование по Сенгеру) и тщательного сопоставления с клиническими признаками.

Других значимых изменений, соответствующих критериям поиска, не обнаружено.

Результаты данного исследования могут быть правильно интерпретированы только врачом-генетиком.

ОПИСАНИЕ ИССЛЕДОВАНИЯ (стандартное, не зависит от типа выявленных вариантов)

Анализ ДНК пациента проведен на секвенаторе нового поколения (название) методом (пример: парно-концевое чтение (2x151 п.о.)) со средним покрытием (пример: не менее 70-100x). Для пробоподготовки была использована методика (пример: селективный захват участков ДНК, относящихся к кодирующим областям (количество генов) генов с известным клиническим значением). Список исследованных генов: (при секвенировании клинического или полного экзома можно ограничиться ссылкой на ресурс, в котором представлены эти данные).

Для названия выявленных вариантов использовалась номенклатура (пример: представленная на сайте <http://www.hgvs.org/mutnomen> версия 2.1511)

Обработка данных секвенирования проведена с использованием автоматизированного алгоритма (название используемой программы для биоинформатического анализа и/или ее описание, пример: алгоритм включает выравнивание прочтений на референсную последовательность генома человека (GRCh37/hg19), постпроцессинг выравнивания, выявление вариантов и фильтрацию вариантов по качеству, а также аннотацию выявленных вариантов по всем известным транскриптам каждого гена из базы RefSeq с применением ряда методов предсказания патогенности замен (SIFT, PolyPhen2-HDIV, PolyPhen2-HVAR, MutationTaster, LRT), а также методов расчета эволюционной консервативности позиций (PhyloP, PhastCons)).

Для оценки популяционных частот выявленных вариантов использованы выборки проектов «1000 геномов», ESP6500 и Exome Aggregation Consortium. Для оценки клинической релевантности выявленных вариантов использованы база данных OMIM, специализированные базы данных по отдельным заболеваниям (при наличии) и литературные данные. В заключение включены

только варианты, имеющие возможное отношение к клиническим проявлениям у пациента. Полиморфизмы, классифицированные по различным критериям как нейтральные, в заключение не включены.

Ограничения методики (например: метод не позволяет выявлять инсерции и делеции длиной более 10 п.о., мутации в интронных областях (за исключением канонических сайтов сплайсинга), вариации длины повторов (в том числе экспансии триплетов), а также мутации в генах, у которых в геноме существует близкий по последовательности паралог (псевдоген). Метод не предназначен для определения цис-, транс-положения пар гетерозиготных мутаций, а также для оценки уровня метилирования, выявления хромосомных перестроек, полиплоидии, выявления мутаций в состоянии мозаицизма).

СВЕДЕНИЯ О КАЧЕСТВЕ ИССЛЕДОВАНИЯ

Всего прочтений	16051259	Всего выявлено вариантов	42310
Длина прочтений	2x151 п.о.	Вариантов после фильтрации по базовым критериям патогенности и оценки по клиническим критериям	1
Прочитано нуклеотидов	4.42 млрд.		
Среднее покрытие	147.8x	Количество фрагментов с покрытием менее X10	43

Данные по покрытию генов (указать какие участки нуклеотидной последовательности генов не были проанализированы или были проанализированы частично. Список должен включать название гена, № экзона и название референсной последовательности (NM). При секвенировании клинического или полного экзона можно ограничиться информацией о покрытии в графическом виде с указанием (в %) количества фрагментов нуклеотидной последовательности с покрытием менее X10 и информацией о том, как получить сведения о покрытии конкретной области).

ССЫЛКИ НА ИСПОЛЬЗОВАННЫЕ БАЗЫ ДАННЫХ И ЛИТЕРАТУРУ

1. Ссылка на статью
2. Ссылка на статью
3. <http://www.omim.org/>
4. <http://www.ncbi.nlm.nih.gov/snp/>
5. <http://evs.gs.washington.edu/EVS/>
6. <http://exac.broadinstitute.org/>

Дата

Подпись врача лабораторного генетика

Подпись биоинформатика

Выявлен вероятно патогенный вариант нуклеотидной последовательности

ЗАКЛЮЧЕНИЕ

по результатам секвенирования ДНК (клиническое секвенирование экзома)

Пациент: ФИО

Пол: М/Ж

Дата рождения: дд.мм.гг

Вид материала: кровь (венозная)

Дата забора: дд.мм.гг

Диагноз: Криптогенная фокальная эпилепсия. Аутичное поведение. Задержка психоречевого развития.

Данные секвенирования могут быть предоставлены по запросу лечащего врача.

1. Патогенные варианты нуклеотидной последовательности, являющиеся вероятной причиной заболевания.

Ген	Положение (GRCh37/hg19)	Генотип	Экзон	Положение в кДНК	Замена АК	Частота аллеля *	Референсная последовательность	Глубина прочтения
-----	-------------------------	---------	-------	------------------	-----------	------------------	--------------------------------	-------------------

Не выявлено

2. Вероятно патогенные варианты нуклеотидной последовательности, являющиеся возможной причиной заболевания.

Ген	Положение (GRCh37/hg19)	Генотип	Экзон	Положение в кДНК	Замена АК	Частота аллеля*	Референсная последовательность	Глубина прочтения
<i>KANSL1</i>	chr17:44248700CA G>C	N/del	2	c.808_809delCT	p.Leu270fs	0.003295 5%	NM_015443.3	25x

3. Варианты нуклеотидной последовательности неопределенного значения, имеющие возможное отношение к фенотипу.

Ген	Положение (GRCh37/hg19)	Генотип	Экзон	Положение в кДНК	Замена АК	Частота аллеля*	Референсная последовательность	Глубина прочтения
<i>RPGRIP1L</i>	chr16:53686809G >GA	N/ins	15	c.1789_1790insT	p.Thr597fs	н/д	NM_015272.2	220x
<i>CPS1</i>	chr2:211502472T >G	T/G	22	c.2734T>G	p.Phe912Val	0.001649 1%	NM_001875.4	199x
<i>LAMB1</i>	chr7:107569594T >C	T/C	31	c.4802A>G	p.Glu1601Gly	н/д	NM_002291.2	149x

*Частоты аллелей приведены по базе Exome Aggregation Consortium (выборка до 60702 человек). н/д = нет данных (не описан)

ИНТЕРПРЕТАЦИЯ

У ФИО был проведен поиск патогенных мутаций, ассоциированных с наследственными формами эпилепсии, эпилептической энцефалопатией, аутизмом, синдромальными и несиндромальными формами умственной отсталости, а также с другими наследственными заболеваниями со сходными фенотипическими проявлениями.

Выявлен ранее не описанный у больных вариант нуклеотидной последовательности во 2 экзоне гена *KANSL1* (chr17:44248700CAG>C, rs551541795) в гетерозиготном состоянии, приводящий к сдвигу рамки считывания, начиная с 270 кодона (p.Leu270fs, NM_015443.3). Мутации, приводящие к сдвигу рамки считывания, в гене *KANSL1* в гетерозиготном состоянии описаны у пациентов с синдромом Кулена - де Ври (OMIM: 610443). Частота выявленного варианта нуклеотидной последовательности в контрольной выборке ExAC составляет 0.0033% (4 мутантных аллеля на 121376 хромосом; в гомозиготном состоянии не зарегистрирован). Выявленный вариант, нарушает синтез полноразмерного белка, однако, в силу ненулевой частоты встречаемости в контрольной выборке его следует расценивать как

вероятно патогенный вариант, который может иметь отношение к фенотипу пациента в случае получения дополнительных подтверждающих данных.

Результат требует проверки альтернативными методами (секвенирование по Сенгеру), тщательного сопоставления с клиническими признаками и анализа происхождения выявленного варианта (унаследованный или *de novo*).

Выявлен ранее не описанный вариант нуклеотидной последовательности в 15 экзоне гена *RPGRIP1L* (chr16:53686809G>GA) в гетерозиготном состоянии, приводящий к сдвигу рамки считывания, начиная с 597 кодона (p.Thr597fs, NM_015272.2). Мутации в гене *RPGRIP1L* в гомозиготном и компаунд-гетерозиготном состоянии описаны у пациентов с синдромом Жубер, тип 7 (OMIM: 611560), синдромом Меккеля, тип 5 (OMIM: 611561) и синдромом COACH (OMIM: 216360). Выявленный вариант нуклеотидной последовательности не зарегистрирован в контрольных выборках "1000 геномов", ESP6500 и ExAC. Поскольку данный вариант нуклеотидной последовательности нарушает синтез полноразмерного белка, его следует расценивать как патогенный, однако, применительно к данному случаю (в силу невыявления второй мутации в гене) - как вариант с неопределенной клинической значимостью, который, тем не менее, может иметь отношение к фенотипу пациента в случае получения дополнительных подтверждающих данных.

Результат требует проверки альтернативными методами (секвенирование по Сенгеру) и тщательного сопоставления с клиническими признаками.

Выявлен ранее не описанный у больных вариант нуклеотидной последовательности в 22 экзоне гена *CPS1* (chr2:211502472T>G, rs753496027) в гетерозиготном состоянии, приводящий к замене аминокислоты в 912 позиции белка (p.Phe912Val, NM_001875.4). Мутации в гене *CPS1* в гомозиготном и компаунд-гетерозиготном состоянии описаны, в частности, у пациентов с недостаточностью карбамоилфосфат-синтетазы 1 (OMIM: 237300). Частота выявленного варианта нуклеотидной последовательности в контрольной выборке ExAC составляет 0.0016%. Алгоритмы предсказания патогенности расценивают данный вариант как вероятно патогенный (SIFT, Polyphen2_HDIV, Polyphen2_HVAR, MutationTaster, PROVEAN, LRT). По совокупности сведений, выявленный вариант нуклеотидной последовательности следует расценивать как вариант с неопределенной клинической значимостью, который, тем не менее, может иметь отношение к фенотипу пациента в случае получения дополнительных подтверждающих данных.

Результат требует проверки альтернативными методами (секвенирование по Сенгеру) и тщательного сопоставления с клиническими признаками.

Выявлен ранее не описанный вариант нуклеотидной последовательности в 31 экзоне гена *LAMB1* (chr7:107569594T>C) в гетерозиготном состоянии, приводящий к замене аминокислоты в 1601 позиции белка (p.Glu1601Gly, NM_002291.2). Мутации в гене *LAMB1* в гомозиготном состоянии описаны у пациентов с лиссэнцефалией, тип 5 (OMIM: 615191). Выявленный вариант нуклеотидной последовательности не зарегистрирован в контрольных выборках "1000 геномов", ESP6500 и ExAC. Алгоритмы предсказания патогенности расценивают данный вариант как вероятно патогенный (SIFT, Polyphen2_HDIV, Polyphen2_HVAR, MutationTaster, PROVEAN). По совокупности сведений, выявленный вариант нуклеотидной последовательности следует расценивать как вариант с неопределенной клинической значимостью, который, тем не менее, может иметь отношение к фенотипу пациента в случае получения дополнительных подтверждающих данных.

Результат требует проверки альтернативными методами (секвенирование по Сенгеру) и тщательного сопоставления с клиническими признаками.

Других значимых изменений, соответствующих критериям поиска, не обнаружено.

Результаты данного исследования могут быть правильно интерпретированы только врачом-генетиком.

ОПИСАНИЕ ИССЛЕДОВАНИЯ (стандартное, не зависит от типа выявленных вариантов)