

# Интерпретация NGS данных

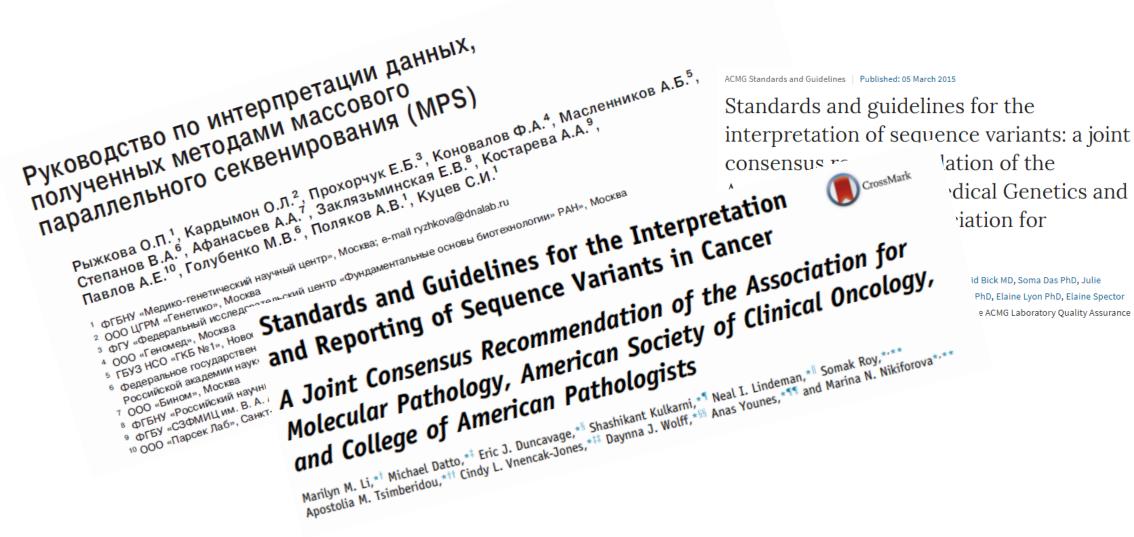
Спарбер П.А.

# Проблема интерпретации вариантов нуклеотидной последовательности

Как определить патогенность ранее неописанных вариантов?

- Анализ сегрегации выявленного варианта в семье
- Популяционные данные
- Биоинформатические подходы
- Экспериментальное подтверждение

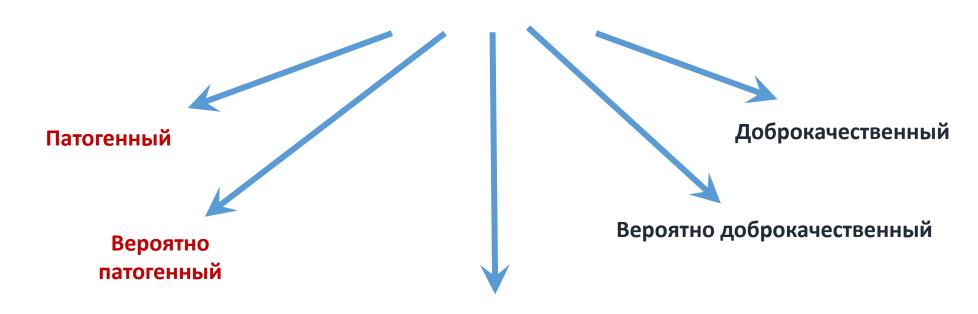
# Руководства для интерпретации вариантов



# Руководства по интерпретации вариантов нуклеотидной последовательности

- Отказываемся от понятий «полиморфизм» и «мутация», как вводящие в заблуждения
- Вместо этого используем термин «вариант нуклеотидной последовательности»
- 5 градаций: доброкачественный, вероятно-доброкачественный, вариант неопределенного значения, вероятно-патогенный, патогенный
- Введена система критериев более объективная оценка

# Найденный при секвенировании вариант нуклеотидной последовательности



Вариант неопределённого значения

50%

# Все ли так просто?



### Кпитепии патогенности и

## Да пошло оно всё!

Критерий «очень сильный»

PVS1: LOF-варианты — к прекращению синтеза белк ции со сдвигом рамки считы ческих (±1 или ±2) нуклеоти рианты, приводящие к изме кодоне; делеции/дупликации экзонов), если данный тип в следовательности гена являет болевания.

Критерий «вспомоги ВР1: миссенс-вари Р что только варианты н ту, к приводящие к изменен го за чиной заболевания.

опис **BP2:** если выявлен заме ветствующий фенотип заме тностью, а неописан

С в транс-положении пр генн ния заболевания или в спла гом типе наследовани: рию PVS1.

PS2: de novo вариант, от телей пациента.

PP4: фенотип пациента і соко специфичны для забоз венной этиологией

PP5: источники с хорошеи репутацией указ патогенность варианта, но независимая оценка водилась.

Патогенный вар



»горячей» точке и/или ах функциональных доій сайт фермента), в коенные изменения.

контрольной выборке ой частотой): для ауто-

Р3: делеции/инсерции с сохранением рамки счиния, если не проводилось функциональное иссление.

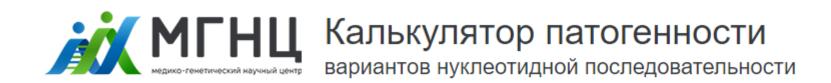
Р4: результаты не менее трех программ предсказанатогенности in silico подтверждают отсутствие вознаия варианта на ген или генный продукт. При этом ет учитывать, что многие программы используют и те же или очень схожие алгоритмы, поэтому реаты использования нескольких программ не могут пъся независимыми и должны оцениваться в сочеи. Как и РР3, критерий ВР4 используется для оценприанта один раз, представляя собой комбинацию пътатов всех программ предсказания патогенности.

Р5: источники с хорошей репутацией сообщили об ствии патогенности варианта, но независимая ка не проводилась.

Р6: синонимичный (silent) вариант с отсутствием ния на консенсусную последовательность ис-сайтов по данным алгоритмов предсказания исинга (отсутствует новые сайты сплайсинга) и еотид не является эволюционно высоко консерваым.

# Калькулятор патогенности вариантов нуклеотидной последовательности

http://calc.generesearch.ru



Данный инструмент создан с целью облегчения использования критериев, описанных в российском Руководстве по интерпретации данных, полученных методами массового параллельного секвенирования (MPS), и автоматической оценки патогенности/доброкачественности вариантов нуклеотидной последовательности в зависимости от выбранных критериев.

# 1. Вариант в гене АРС

•	Ппобанл — у	МОПП	пипа 12 чет Циагпоз. с	ПОЛИПОЗ РЗК ТОЛСТОЙ КИІ <u>Gavert (2002) Hum Mutat 19, 664</u>	IJКИ
•	c.1312+1G>A	DM	Adenomatous polyposis coli	Xiong (2015) Science 347: 1254806 [Additional report]  Zhang (2016) Gene 577: 187 [Functional characterisation]	APC:
	c.1312+1G>C	$\mathbf{DM}$	Adenomatous polyposis coli	Kerr (2013) J Mol Diagn 15, 31	
•	c.1312+2T>C	DM	Adenomatous polyposis coli	Friedl (2005) Hered Cancer Clin Pract 3, 95  Xiong (2015) Science 347: 1254806 [Additional report]	ы у
	c.1312+2T>G	DM?	Adenomatous polyposis coli	<u>Gundesen (2005) Hum Genet 117 300</u> <u>Xiong (2015) Science 347: 1254806</u> [Additional report]	)30M
	c.1312+3A>C	DM	Adenomatous polyposis coli, attenuated	Nielsen (2007) Clin Genet 71, 427	
•	c.1312+3A>G	DM	Adenomatous polyposis coli	Olschwang (1993) Am J Hum Genet 52, 273  Aretz (2004) Hum Mutat 24: 370 [Functional characterisation]  Kerr (2013) J Mol Diagn 15: 31 [Additional report]  1 more reference(s)	
•	c.1312+5G>A	DM	Adenomatous polyposis coli	Aretz (2004) Hum Mutat 24, 370  Schwarzová (2013) Fam Cancer 12: 35 [Functional characterisation]	,
•	c.1312+5G>C	DM?	Adenomatous polyposis coli	Kerr (2013) J Mol Diagn 15, 31	те с
	c.1312+5G>T	DM	Adenomatous polyposis coli	Varesco (1994) Hum Genet 93, 281  Grandval (2014) Hum Mutat 35: 532 [Functional characterisation]	

# 1. Вариант в гене АРС

- Пробанд женщина 15 лет. Диагноз: полипы, рак толстой кишки
- Выявлен неописанный вариант в гене *APC:* NM\_000038:c.1312+1G>T
- Гетерозиготные патогенные варианты в этом гене описаны у пациентов с доминантным семейным аденоматозным полипозом (ОМІМ #175100)
- Анализ сегрегации не проводился
- Данный вариант отсутствует в контрольных выборках
- Ген не толерантен к вариантам с потерей функции
- В HGMD описано множество патогенных вариантов в этом сайте с функциональным анализом
- Программы предсказания сплайсинга поломка донорного сайта сплайсинга
- Ваша оценка?
- Ответ: 1)Патогенный вариант (PM2, PVS1,PP3)
  - 2) Патогенный вариант (PM2,PVS1, PS3,PP3)

# 2. Вариант в гене *ТР53*

- Пробанд женщина, возраст? Диагноз: первично-множественные злокачественные заболевания
- Секвенированием по Сенгеру выявлен гетерозиготный вариант в гене *TP53:* NM\_001126114: c.322\_327del6 (p.Gly108\_Phe109del)
- Гетерозиготные патогенные варианты в гене *TP53* описаны у больных с синдромом Ли-Фраумени
- Вариант неописанный, отсутствует в контрольных выборках
- Делеция 2-х аминокислот в ДНК-связывающем домене
- Ваша оценка?
- Ответ: Вероятно патогенный вариант (РМ1, РМ2, РМ4)

# 2. Вариант в гене ТР53

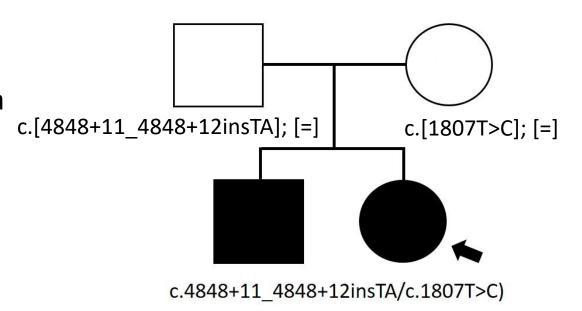
- Что нужно делать дальше?
  - о Функциональный анализ
  - о Дополнительные клинические обследования
  - Анализ сегрегации
  - Полногеномное секвенирование

# 3. Вариант в гене АВСА4

- Пробанд женщина возраст (?). Диагноз: прогрессирующая макулодистрофия сетчатки
- При секвенировании панели генов выявлено 2 варианта в гене ABCA4: c.1807T>C (p.Tyr603His) и c.4848+11\_4848+12insTA
- Гомозиготные или компаунд гетерозиготные варианты в гене АВСА4 описаны у пациентов с болезнью Штаргардта
- Первый вариант описан множество раз как патогенный.
- Второй неописанный, отсутствует в контрольных выборках

# 3. Вариант в гене *АВСА4*. Анализ сегрегации

- Варианты находятся в транс-положении
- Программы предсказания не влияет на сплайснг
- Ваша оценка?



Ответ: Вариант неопределенного значения (РМ2,РМ3)

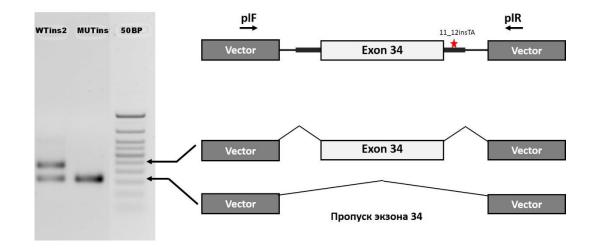
# 3. Вариант в гене АВСА4

- Что нужно делать дальше?
  - о Функциональный анализ
  - о Дополнительные клинические обследования
  - о Исследование бабушек и дедушек
  - Ничего не сделаешь 🕾

# Результаты функционального анализа варианта c.4848+11\_4848+12insTA

- В результате пропуска 34 экзона на белковом уровне происходит делеция 25 аминокислот. p.(Gly1592\_Lys1616del)
- Экзон 34 кодирует часть второго внеклеточного домена (ECD2)
- Описаны несколько патогенных миссенс вариантов в ECD2, изменяющих его конформацию

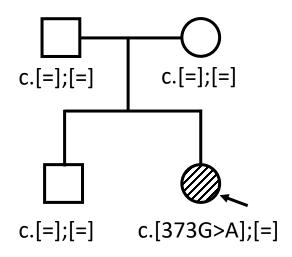
Ваша оценка?



Ответ: Вероятно патогенный вариант (РМ2, РМ4, РМ3)

# 4. Вариант в гене *GABRB2*

- Пробанд девочка 4 года. Диагноз: криптогенная фокальная эпилепсия. Задержка моторного развития.
- WGS квадро выявлен de novo вариант NM\_021911.2(*GABRB2*\_v001):c.373G>A p.(Asp125Asn) в гене *GABRB2*.
- Гетерозиготные патогенные варианты в этом гене описаны у пациентов с ранней эпилептической энцефалопатией 2 (ОМІМ #617829)
- В базе данных HGMD описан как патогенный
- Данный вариант отсутствует в контрольных выборках
- Ваша оценка?
- Ответ: Вероятно патогенный вариант (РМ2, PS2, PP5)



# Данные из статьи

- 197 больных с эпилепсией
- Всем провели WGS или WES
- Имеется только клиническое описание

#### **ARTICLE**

### High Rate of Recurrent *De Novo* Mutations in Developmental and Epileptic Encephalopathies

Fadi F. Hamdan,¹ Candace T. Myers,² Patrick Cossette,³,⁴ Philippe Lemay,¹ Dan Spiegelman,⁵ Alexandre Dionne Laporte,⁵ Christina Nassif,¹ Ousmane Diallo,⁵ Jean Monlong,⁶,⁷ Maxime Cadieux-Dion,³,ጾ,⁰ Sylvia Dobrzeniecka,³ Caroline Meloche,³ Kyle Retterer,¹⁰ Megan T. Cho,¹⁰ Jill A. Rosenfeld,¹¹ Weimin Bi,¹¹,¹² Christine Massicotte,¹ Marguerite Miguet,¹ Ledia Brunga,¹³ Brigid M. Regan,¹⁴ Kelly Mo,¹⁴ Cory Tam,¹⁴ Amy Schneider,¹⁵ Georgie Hollingsworth,¹⁵ Deciphering Developmental Disorders Study,¹⁶ David R. FitzPatrick,¹ˀ Alan Donaldson,¹³ Natalie Canham,¹⁰ Edward Blair,²⁰ Bronwyn Kerr,²¹ Andrew E. Fry,²² Rhys H. Thomas,²³ Joss Shelagh,²⁴

(Author list continued on next page)



• Клинические данные: Пробанд, мальчик 4,5 лет, имеет грубое отставание психомоторного развития, дебютировавшее с судорог в 6 мес. Сестра пробанда — 6 лет, не разговаривает. Родители здоровы.

#### • WES:

оквам ва,	гь влияние на клиниче	вскую ка	ртину болезни			
Ген	Референская последовательность	Экзон/ интрон	Геномная координата (GRCh37)	Нуклеотидная замена	Аминокислотн ая замена	Оценка, согласно рекомендациям АСМО
SCN1A	NM_001165963.1	19	166872146	c.3521C>G (aemeposuaoma)	p.T1174S	Патогенная

- Анализ сегрегации не проводился
- Гетерозиготные патогенные варианты в гене *SCN1A* описаны у пациентов с семейными фибрильными судорогами, GEFS+, синдромом Драве (РЭЭ 6 типа)
- Описан в HGMD несколько раз как ассоциированный с гемиплегической мигренью, миоклонической эпилепсией, синдромом Драве.
  - 1. Gargus (2007) Pediatr Neurol 37: 407 PubMed: 18021921
  - 2. Cestèle (2013) Epilepsia 54: 927 PubMed: 23398611
  - 3. Escayg (2001) Am J Hum Genet 68: 866 PubMed: 11254445
  - 4. Frosk (2013) J Child Neurol 28: 389 PubMed: 22550089
  - 5. Lal (2016) PLoS One 11: e0150426 PubMed: 26990884
  - 6. Rilstone (2012) Epilepsia 53: 1421 PubMed: 22780858
  - 7. Yordanova (2011) Neurosci Lett 494: 180 PubMed: 21396429
- Ваша оценка?

Thr1174Ser c.3521C>G p.T1174S





RESEARCH ARTICLE

Evaluation of Presumably Disease Causing SCN1A Variants in a Cohort of Common Epilepsy Syndromes

Dennis Lal<sup>1,2,3,4</sup> \*, Eva M. Reinthaler<sup>5</sup>, Borislav Dejanovic<sup>6</sup>, Patrick May<sup>7</sup>,

- В данной работе замена была охарактеризована как фактор риска развития эпилепсии с небольшим вкладом, исходя из **несколько** более высокой частоты в группе больных (OR=0.32).
- В работах с относительно большими выборками (165 пациентов, 2001 год) встречался также в контрольных образцах!

• Что мы забыли?

Ваша оценка?

Population Frequencies							
Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency			
European (Finnish)	87	25122	2	0.003463			
European (non- Finnish)	362	128886	1	0.002809			
▶ Ashkenazi Jewish	8	10362	0	0.0007721			
▶ Other	5	7220	0	0.0006925			
▶ African	9	24968	0	0.0003605			
▶ South Asian	7	30606	0	0.0002287			
▶ Latino	4	35426	0	0.0001129			
▶ East Asian	0	19942	0	0.000			
Total	482	282532	3	0.001706			

Ответ: Доброкачественный вариант (BS1, BS2)

# 6. Вариант в гене *CFTR*

Population Frequencies @

•	Пробанд маль
	респираторные
	вязкой мокрот
	«барабанные г

• Проведен поис автоматическо

• Выявлен вариа

• Вариант затраг

• Описан как пат

•	•	•				
Population	Allele Count	Allele Number		Number of Homozygotes	Allele Frequency	нуситы,
European (non- Finnish)	1598	129034	1		0.01238	ением густой,
▶ Other	49	7214	0		0.006792	пу
Ashkenazi Jewish	58	10368	0		0.005594	
▶ Latino	135	35426	0		0.003811	0.50
▶ African	65	24958	0		0.002604	ΟΓΟ
▶ European (Finnish)	61	25074	0		0.002433	
▶ South Asian	61	30608	0		0.001993	ной форме
▶ East Asian	0	19948	0		0.000	τονι φορινίε
Female	939	129374	0		0.007258	лка
Male	1088	153256	1		0.007099	
Total	2027	282630	1		0.007172	

Ваша оценка?

Ответ: Патогенный вариант

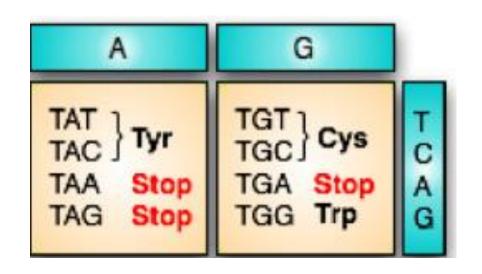
- Пробанд мальчик 4 лет с ранней эпилептической энцефалопатией
- Проведен WES и выявлен вариант NM\_001165963.2:c.4285-4A>G в гене *SCN1A*.
- Отсутствует в контрольных выборках
- Вариант описан как патогенный в базе данных HGMD
- Анализ сегрегации невозможен отец недоступен

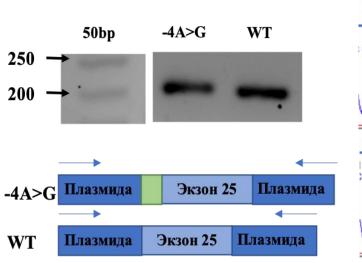


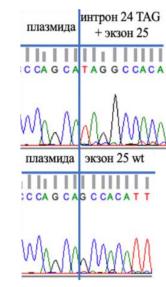
Ваша оценка?

Ответ: Вариант неопределенного значения (РМ2,РР5)

- В лаборатории функциональной геномики проведен функциональный анализ
- Наблюдается удлинение экзона на 3 буквы TAG
- Ваша оценка?
- Ответ: Вероятно патогенный вариант (РМ2, PP5, PS3)

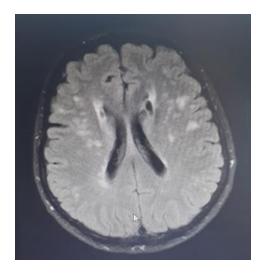


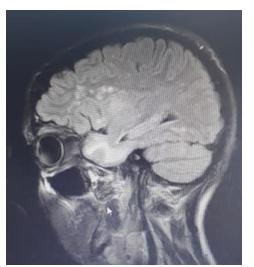




# 8. Вариант в гене *NOTCH3*

- Пробанд женщина 39 лет. Диагноз: Демиелинизирующее заболевание нервной системы неуточненное. Лейкоэнцефалопатия. Возможный рассеянный склероз, инсультоподобный дебют с преходящей пирамидной недостаточностью справа, гемигипестезией справа, дизартрией? Синдром, имитирующий рассеянный склероз? Системный васкулит? Синдром CADASIL? MELAS? Рефлекторный тетрапарез. Чувствительные нарушения в анамнезе.
- Проведено секвенирование гена *NOTCH3*





# 8. Вариант в гене *NOTCH3*

- Родители недоступны для анализа сегрегации
- У отца микроинсульты в 50. умер в 59

#### Заключение:

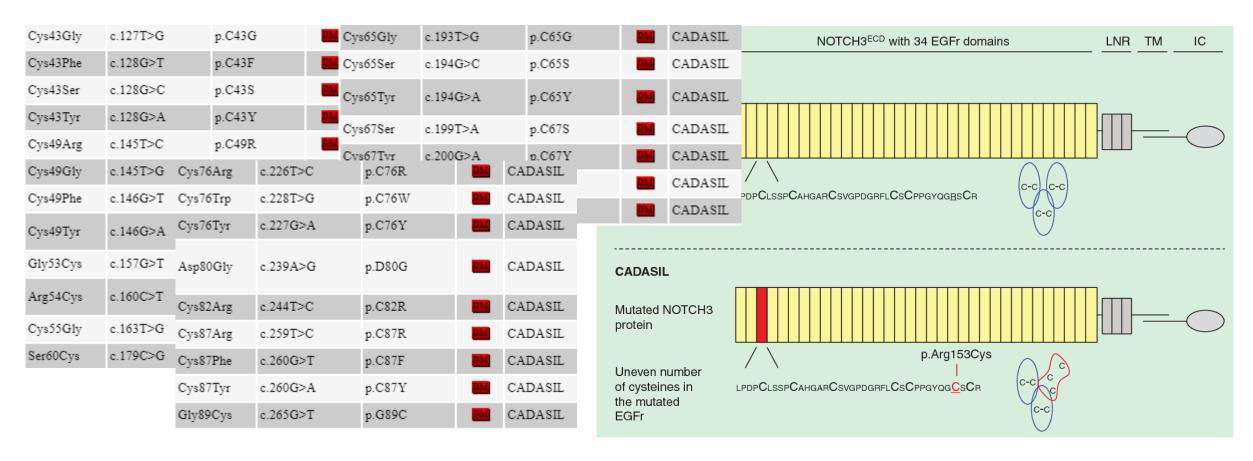
Методом прямого автоматического секвенирования проведен частичный анализ гена NOTCH3 (CADASIL синдром, MN\_000435). Исследованы 2, 3, 4, 5, 6 экзоны данного гена и прилежащие интронные области. В интроне 3 гена обнаружена не описанная ранее замена с.341-1G>C в гетерозиготном состоянии. Анализ биоинформатическими методами (Human Splicing Finder) показал высокую вероятность патогенности данной замены (нарушение сайта сплайсинга). Рекомендуется консультация лечащего врача.

# Популяционные данные

- Ген относительно толерантен к вариантам с потерей функции
- 22 варианта с потерей функции описано у здоровых людей 50-60-70 лет

Variant ID	▲ Source	Consequence	Annotation Flags	Allele Count	Allele Number
19-15276597-C-T	E	c.5667+1G>A	<ul><li>splice donor</li></ul>	1	247348
19-15276713-CG-C	E	p.Arg1851ValfsTer60	frameshift	1	251166
19-15276816-ATG-A	E	p.Thr1816llefsTer3	frameshift	1	251428
19-15278168-G-A	E	p.Gln1752Ter	<ul><li>stop gained</li></ul>	1	249776
19-15278221-ACC-A	E	c.5200-1_5200delGG	<ul><li>splice acceptor</li></ul>	1	244354
19-15278224-T-A	E	c.5200-2A>T	<ul><li>splice acceptor</li></ul>	1	243908
19-15280950-C-A	E	p.Glu1716Ter	<ul><li>stop gained</li></ul>	1	234766
19-15281235-A-AG	E	p.Leu1674ProfsTer5	<ul><li>frameshift</li></ul>	1	249550
19-15281495-C-CAGTGGGTAG	C E	p.Arg1627ProfsTer68	frameshift	1	196594
19-15285132-AGCCGCAC-A	E	p.Glu1492AspfsTer84	frameshift	4	158142
19-15288380-G-C	E	p.Tyr1453Ter	<ul><li>stop gained</li></ul>	1	235552
19-15288406-C-CG	E	p.Ala1445ArgfsTer118	frameshift	1	220096
19-15288465-CA-C	E	p.Trp1425GlyfsTer153	frameshift	1	160712
19-15290000-AG-A	E	p.Leu1185TrpfsTer87	frameshift	1	250252
19-15290039-G-C	E	p.Ser1172Ter	<ul><li>stop gained</li></ul>	1	249862
19-15290262-C-A	E	p.Glu1125Ter	<ul><li>stop gained</li></ul>	1	251264
19-15290305-A-T	E	p.Cys1110Ter	<ul><li>stop gained</li></ul>	1	251040
19-15291784-GC-G	E	p.Gly994AlafsTer278	frameshift	1	209140
19-15291974-C-T	E	c.2793-1G>A	<ul><li>splice acceptor</li></ul>	1	238108
19-15291975-T-C	G	c.2793-2A>G	<ul><li>splice acceptor</li></ul>	1	31390

### Патогенез CADASIL



• Ваша оценка?

# Вариант c.341-2A>G

Neurology<sup>®</sup>

- RT-PCR показал, что замена приводит к делеции 21 нуклеотидов
- Splice site mutation causing a seven amino acid *Notch3* in-frame deletion in CADASIL A. Joutel, H. Chabriat, K. Vahedi, et al.

• На белковом уровне делеция 7 аминокислот включая 6 остаток цистеина во втором EGF домене

Neurology 2000;54;1874 DOI 10.1212/WNL.54.9.1874

• Ваша оценка?

Ответ: 1)Вероятно патогенный вариант (РМ2, РМ4, РМ1)

2)Патогенный вариант (РМ2, РМ4, PS3, РМ1)

# Take home messages

- 1. Даже если вариант описан как патогенный в HGMD, всегда необходимо читать статьи
- 2. Анализ сегрегации очень важен и может перевести статус VUS как в вероятно патогенный, так и в вероятно доброкачественный.
- 3. Анализ сегрегации необходимо делать всегда, когда это клинически целесообразно
- 4. Для грамотной интерпретации необходимо знать этиопатогенез заболевания
- 5. Часто лишь функциональный анализ может поставить окончательную точку в вопросе значимости выявленного варианта
- 6. Не всегда существует единственный верный вариант интерпретации
- 7. Если для «вашего» заболевания существует адаптированный экспертами вариант руководства лучше пользоваться им
- 8. Из каждого правило будет исключения