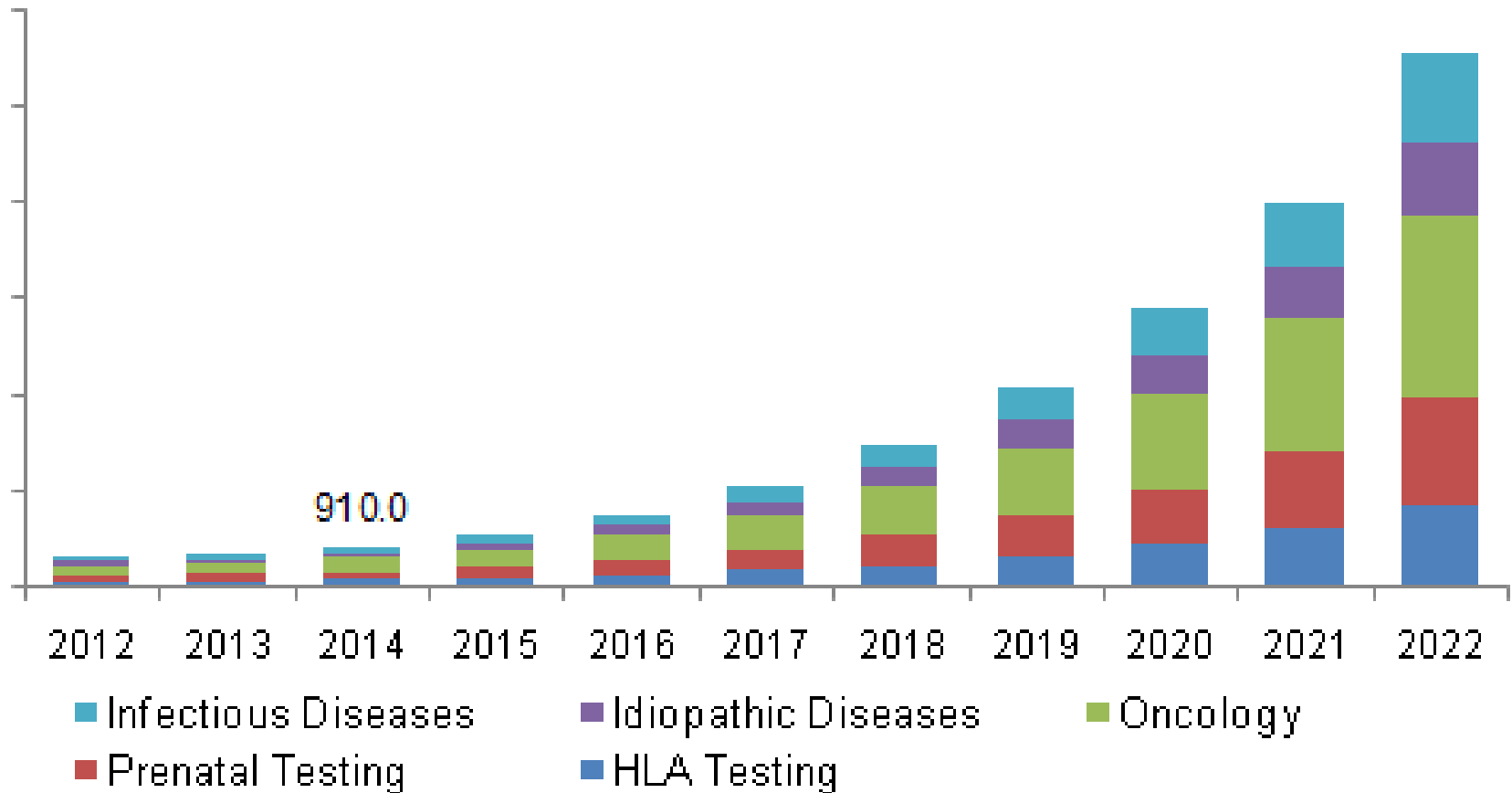




# **Применение NGS-технологий в медицинской генетике. Эффективность. Ограничения**

Скоблов Михаил Юрьевич  
заведующий лабораторией функциональной геномики  
ФГБНУ "Медико-генетический научный центр"

# U.S. next generation sequencing market, by application, 2012-2022 (USD Million)







Медико-генетический  
научный центр имени  
академика Н.П. Бочкова

115522, Москва,  
ул. Москворечье, д. 1

Запись на приём  
+7 (495) 111-03-03  
Пн-Пт: с 9:00 до 17:00

Eng

ГОТОВНОСТЬ АНАЛИЗОВ

О ЦЕНТРЕ

НАУЧНАЯ ДЕЯТЕЛЬНОСТЬ

ОБРАЗОВАНИЕ

МЕДИЦИНСКИЕ УСЛУГИ

СПЕЦИАЛИСТАМ

ПРЕСС-ЦЕНТР



Федеральное государственное  
бюджетное научное учреждение  
«Медико-генетический научный центр  
имени академика Н.П. Бочкова»

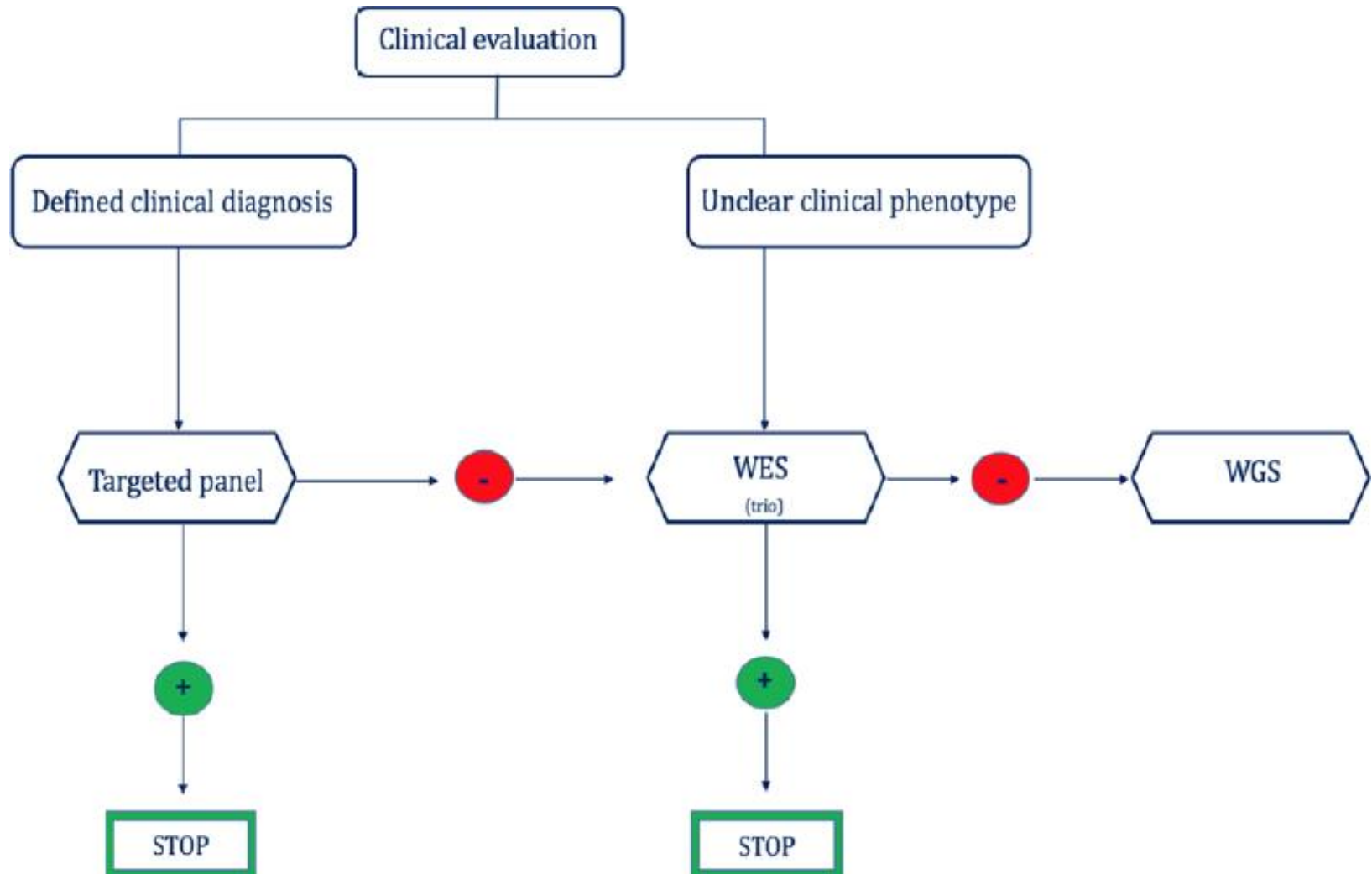
Федеральное государственное бюджетное научное учреждение  
«Медико-генетический научный центр имени академика Н.П. Бочкова»  
является ведущей медико-генетической организацией в России.

Учредителем ФГБНУ «МГНЦ» является Министерство науки и высшего  
образования Российской Федерации. Научная деятельность ФГБНУ  
«МГНЦ» осуществляется под научно-методическим руководством  
Российской академии наук (РАН).

## 13 тысяч молекулярных анализов в 2019 году:

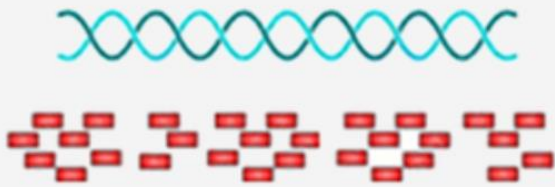
- 20 геномов
- 700 экзоменов
- 2000 панелей
- 10 тысяч секвенирований по Сэнгеру

# Применение NGS в медицинской генетике



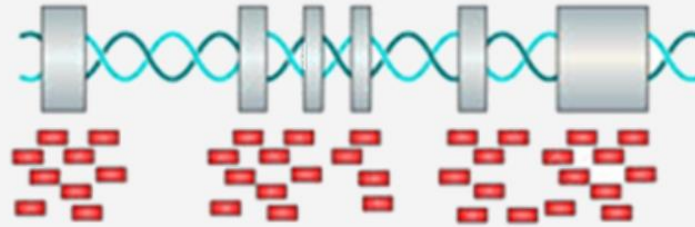
# Применение NGS в медицинской генетике

## Whole genome sequencing



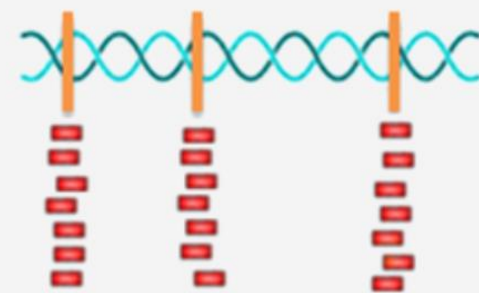
- Sequencing region : whole genome
- Sequencing Depth: >30X
- Covers everything – can identify all kinds of variants including SNPs, INDELs and SV.

## Whole exome sequencing



- Sequencing region: whole exome
- Sequencing Depth : >50X ~ 100X
- Identify all kinds of variants including SNPs, INDELs and SV in coding region.
- Cost effective

## Targeted sequencing



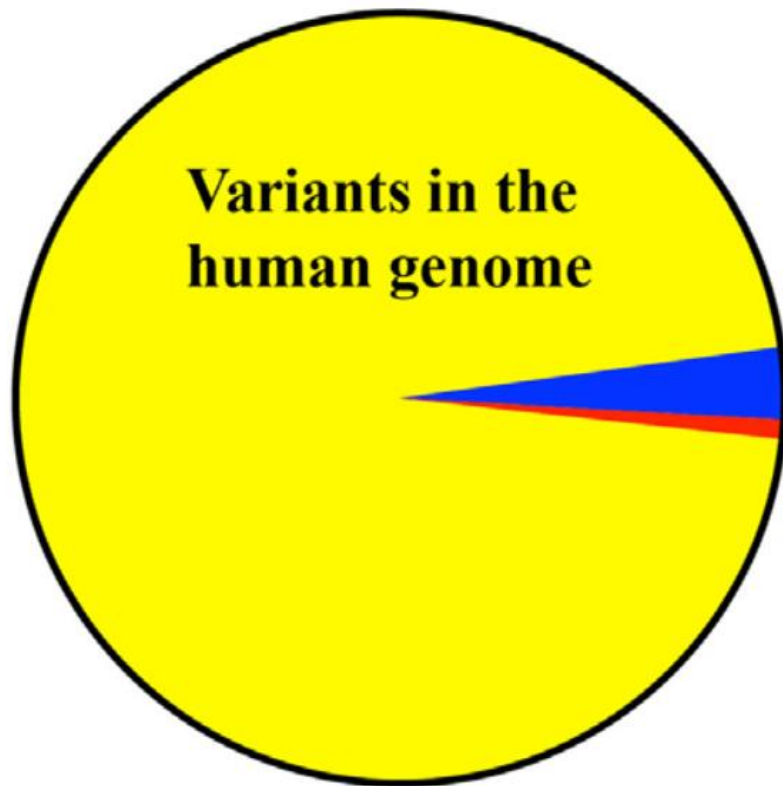
- Sequencing region: specific regions (could be customized)
- Sequencing Depth : >500X
- Identify all kinds of variants including SNPs, INDELs in specific regions
- Most Cost effective



# WES vs WGS

	Полный экзом	Полный геном
Область исследование	95% код. части генома (1-2% от всего генома)	98% всего генома
% не обнаруженных вариантов из HGMD	2,1%	0,3%
CNV	Возможна по анализу покрытия. Необходимо подтверждение другим методом	Разрешение точных границ перестроек выше ХМА высокого разрешения.
Некодирующая область	-	+
мтДНК	-	С прочтением 2000X
Детекция экспансии тринуклеотидных повторов	-	Возможна в некоторых случаях
Качество/покрытие	Хорошее/более x100	Хорошее/более x30
Реанализ данных	Делают но не на старых	Более долговечно

# Количество данных получаемых при NGS



## WGS:

Variants throughout the genome  
Structural variations  
Copy number variations



## WES:

Variants in coding regions and splice sites  
Copy number variations

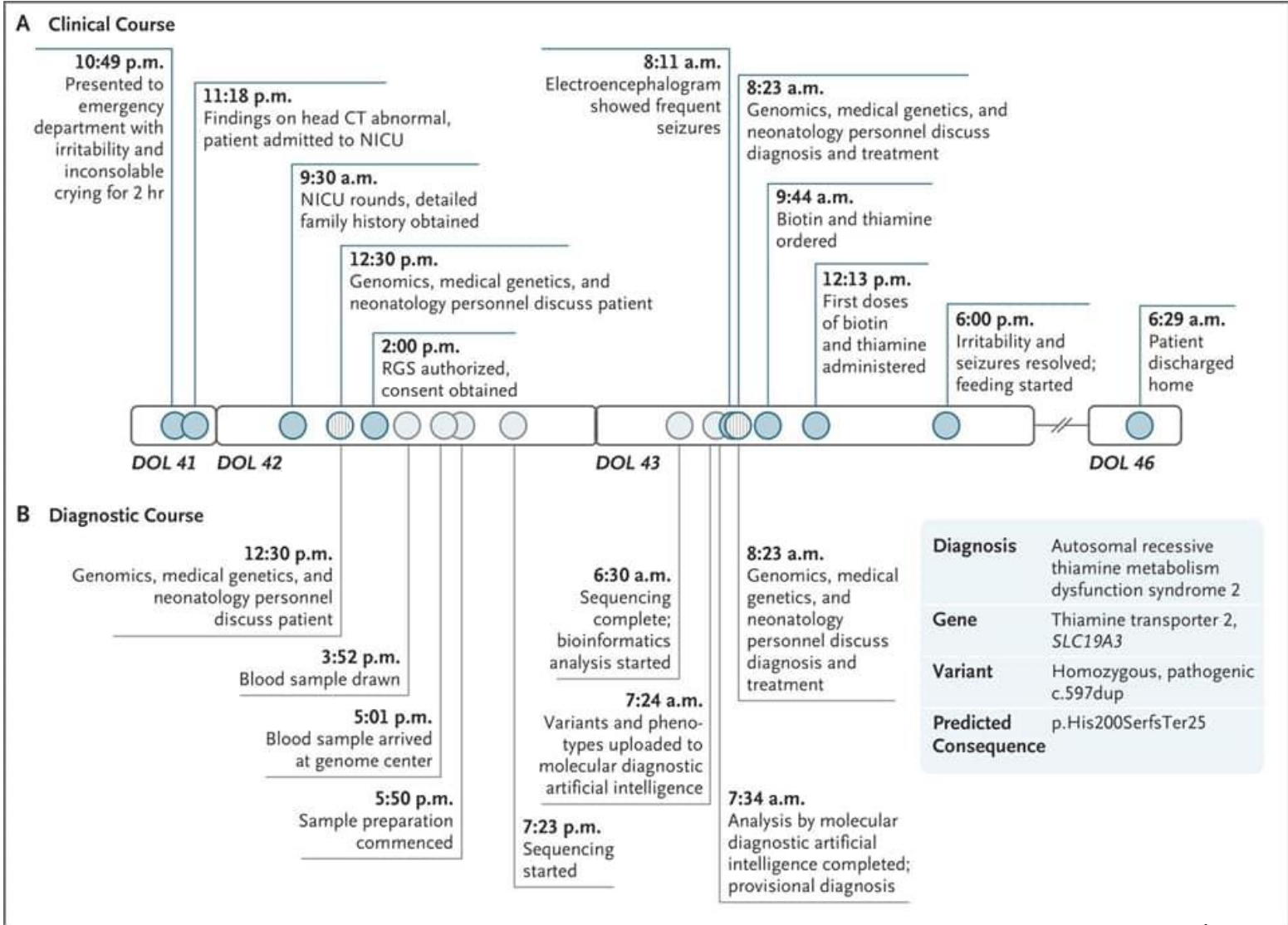


## TGP:

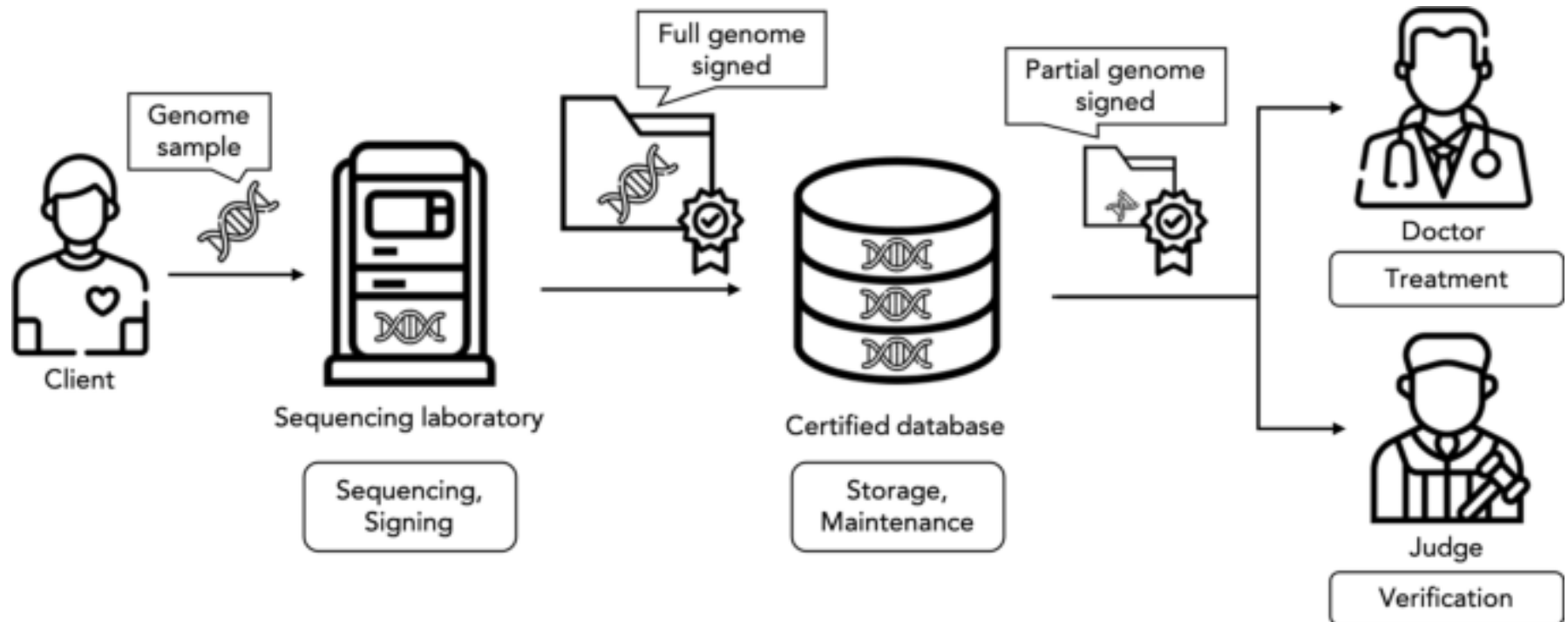
Variants in pre-selected genes  
Deletions



# Диагностика метаболизма тиамина на основе быстрого секвенирования была сделана в течение 13 часов



# Какова эффективность NGS диагностики?



Diagnostic yield for commonly ordered gene panels (n ≥ 10 samples) at the University of Minnesota from 2012 to 2017.

Panel	Total cases	Number of genes <sup>a</sup>	Diagnostic findings	Possible diagnostic findings	Negative findings	Diagnostic yield (%) (Diagnostic + Possible diagnostic)
Phenylketonuria	52	1	49	1	2	96
Fanconi anemia	39	1–18	30	4	5	87
Epidermolysis bullosa	27	1–13	20	3	4	85
Retinal dystrophy panel	69	32–315	40	15	14	80
Adrenoleukodystrophy	18	1	13	1	4	78
Albinism	42	1–24	20	12	10	76
Congenital hyperinsulinism	12	7–14	3	5	4	67
Craniosynostosis	13	1–20	5	2	6	54
Hearing loss (all subpanels combined)	173	2–149	50	37	86	50
Achondroplasia	10	1	5	0	5	50
Congenital myopathy	35	1–29	9	8	18	49
Alport syndrome	48	3	20	3	25	48
Stickler Syndrome	13	1–6	5	1	7	46
Ataxia/Hereditary Spastic Paraparesis	23	28–101	3	7	13	43
Hereditary spastic paraparesis	26	1–74	7	4	15	42
Limb girdle muscular dystrophy	24	1–36	6	4	14	42
Ataxia	25	21–67	3	7	15	40
Carnitine acetyltransferase deficiency	10	1	1	3	6	40
Cystic fibrosis	10	1	3	1	6	40
Noonan syndrome	40	5–22	15	0	25	38
Hereditary hemorrhagic telangiectasia	11	1–4	4	0	7	36
Periodic paralysis syndromes	11	2–5	3	1	7	36
Polycystic kidney disease	11	1–9	3	1	7	36
Charcot Marie Tooth	117	1–58	32	9	76	35
Glycogen storage disease	17	1–25	5	0	12	29
Complex neurologic	45	6–266	4	8	33	27
-----						
Marfan syndrome	38	1–3	2	3	33	13
Li Fraumeni syndrome	25	1–3	3	0	22	12
Macrocephaly/Overgrowth	10	3–18	1	0	9	10
Hereditary breast/ovarian cancer	126	2–18	11	1	114	10
Connective tissue disorder	48	2–29	2	2	44	8
Developmental eye panel	12	14–31	1	0	11	8
Renal coloboma syndrome	13	1	1	0	12	8
Ehlers Danlos syndrome	111	1–16	5	2	104	6
Dystonia	17	1–18	1	0	16	6
Motor neuron disease	19	5–85	1	0	18	5
Myoclonus dystonia	20	1–3	1	0	19	5

# Whole Genome Sequencing Increases Molecular Diagnostic Yield Compared with Current Diagnostic Testing for Inherited Retinal Disease

Jamie M Ellingford<sup>1</sup>, Stephanie Barton<sup>2</sup>, Sanjeev Bhaskar<sup>2</sup>, Simon G Williams<sup>2</sup>,

Panagiotis I Sergouniotis<sup>3</sup>, Jamie G Martin<sup>2</sup>, William G Newman<sup>1</sup>, Paul N Bischoff  
> Neurology. 2021 Mar 30;96(13):e1770-e1782. doi: 10.1212/WNL.00000000000011655.

Stuart Bayliss<sup>2</sup>, Simon C Ramsdell  
Epub 2021 Feb 10.

Affiliations + expand

## Diagnostic Yield of Whole Genome Sequencing After Nondiagnostic Exome Sequencing or Gene Panel in Developmental and Epileptic Encephalopathies

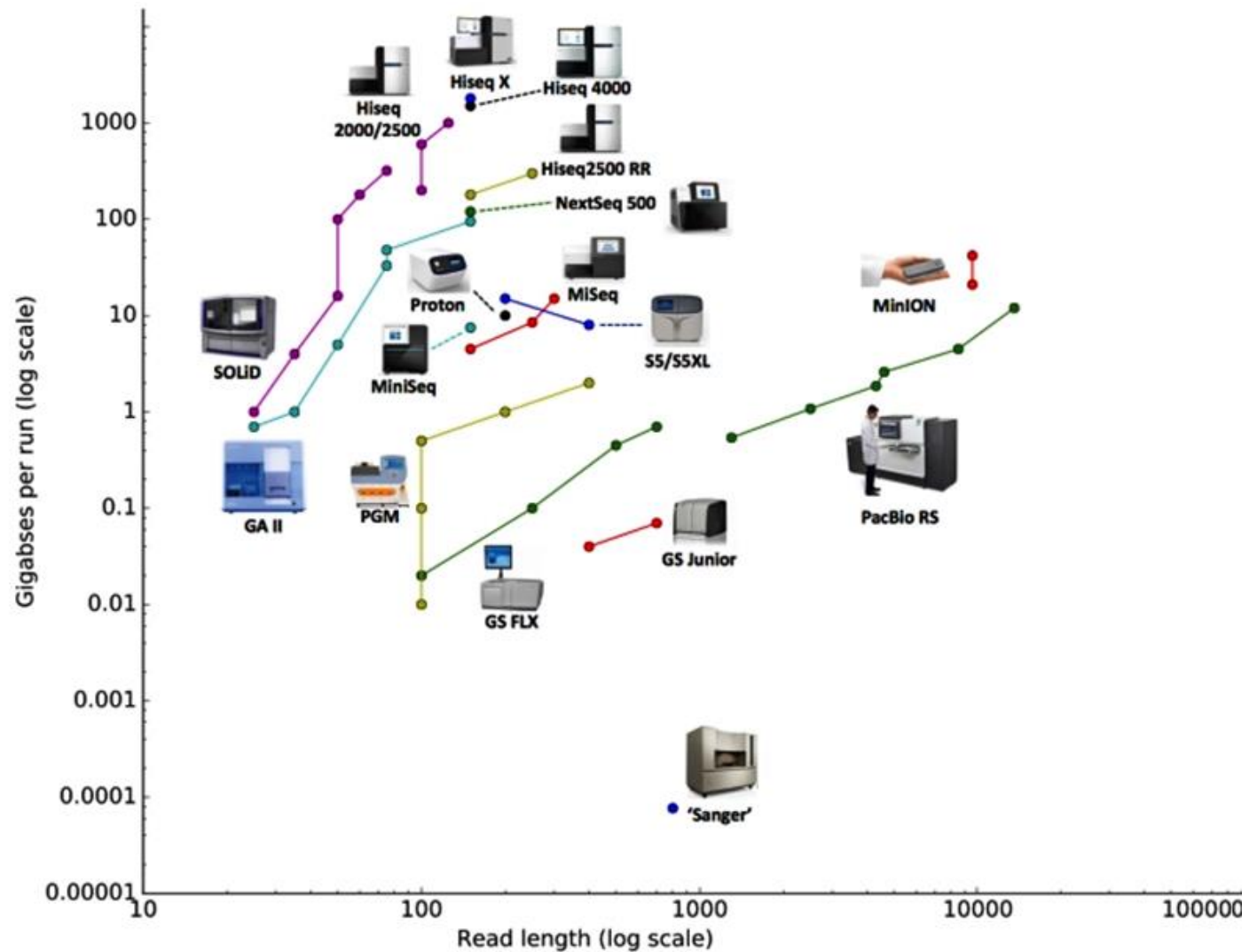
Elizabeth Emma Palmer<sup>1</sup>, Rani Sachdev<sup>2</sup>, Rebecca Macintosh<sup>2</sup>, Uirá Souto Melo<sup>2</sup>,  
Stefan Mundlos<sup>2</sup>, Sarah Righetti<sup>2</sup>, Tejaswi Kandula<sup>2</sup>, Andre E Minoche<sup>2</sup>, Clare Puttick<sup>2</sup>,  
Velimir Gayevskiy<sup>2</sup>, Luke Hesson<sup>2</sup>, Senel Idrisoglu<sup>2</sup>, Cheryl Shoubridge<sup>2</sup>,  
Monica Hong Ngoc Thai<sup>2</sup>, Ryan L Davis<sup>2</sup>, Alexander P Drew<sup>2</sup>, Hugo Sampaio<sup>2</sup>,  
Peter Ian Andrews<sup>2</sup>, John Lawson<sup>2</sup>, Michael Cardamone<sup>2</sup>, David Mowat<sup>2</sup>, Alison Colley<sup>2</sup>,  
Sarah Kummerfeld<sup>2</sup>, Marcel E Dinger<sup>2</sup>, Mark J Cowley<sup>2</sup>, Tony Roscioli<sup>2</sup>, Ann Bye<sup>2</sup>, Edwin Kirk<sup>2</sup>

- Прирост в эффективности диагностики после WES от 5 до 15%
- В основном из-за более высокого качества данных (равномерное покрытие в отличие от экзема) и различных CNV/инсерций мобильных элементов.

**Почему эффективность ДНК-диагностики с применением NGS технологий сегодня около 40%?**

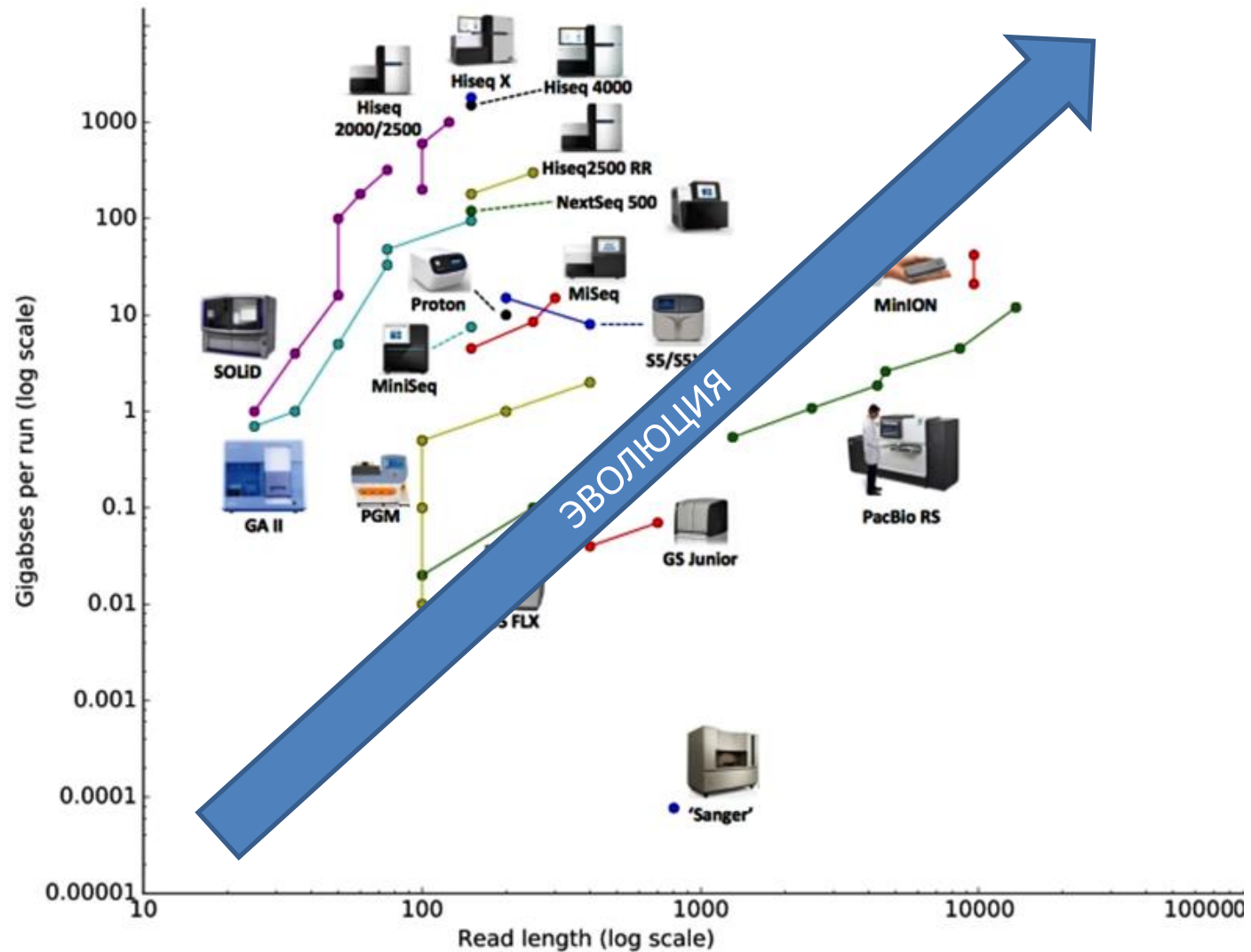


# 1. Несовершенство приборов



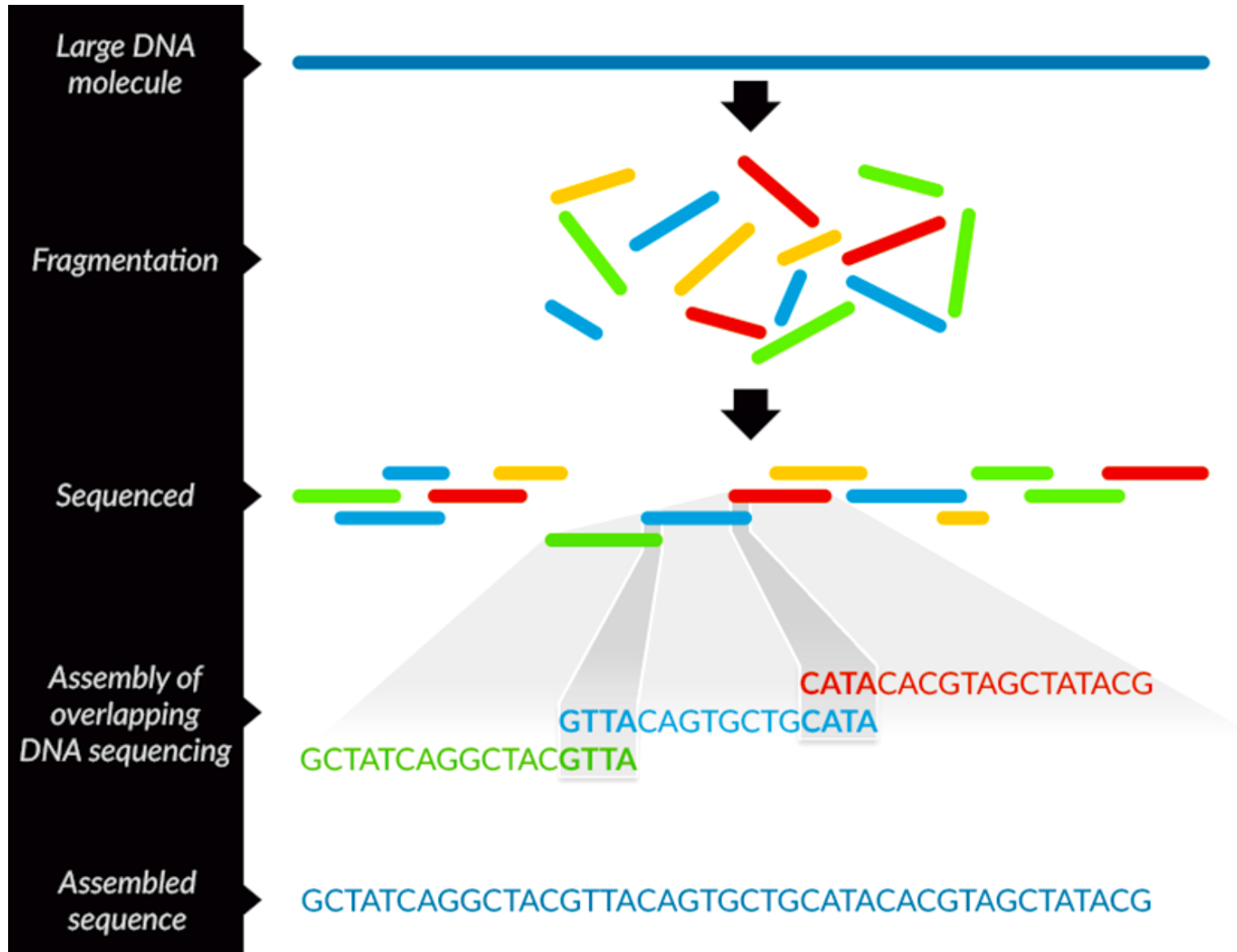


# 1. Несовершенство приборов

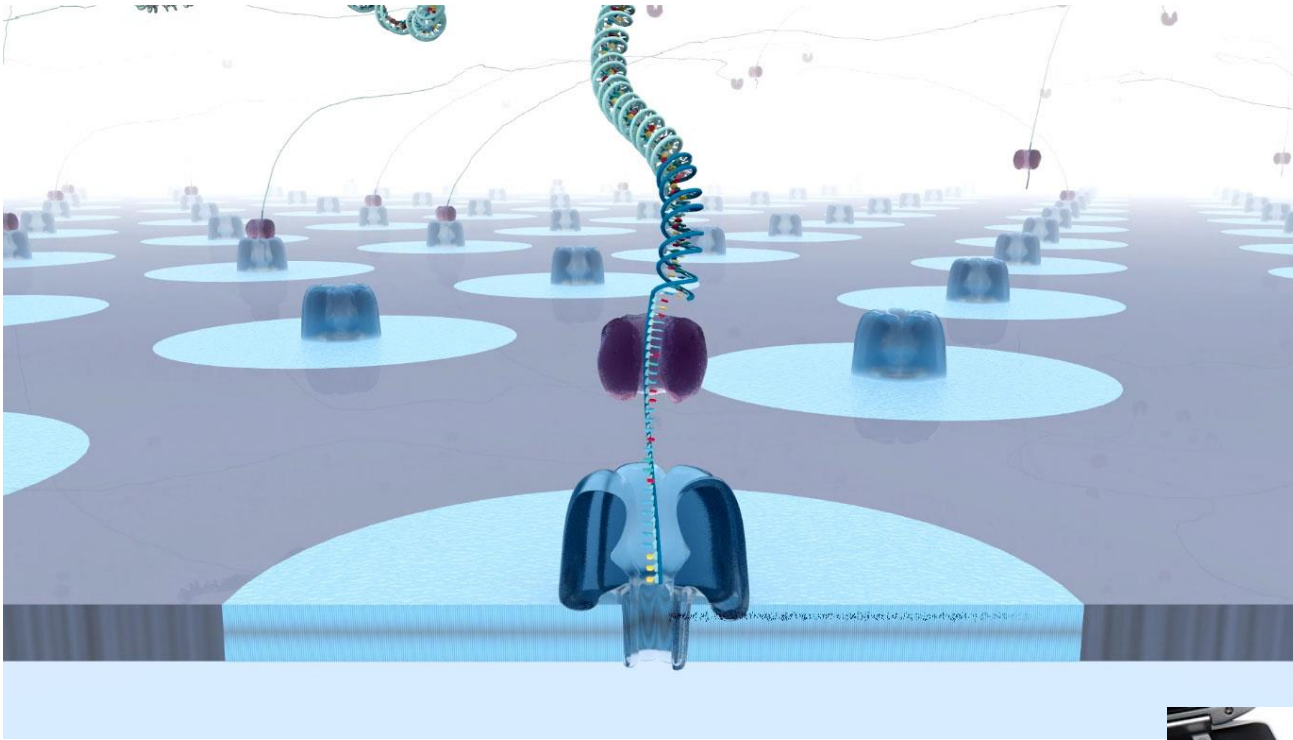




## 2. Несовершенство технологий



## 2. Несовершенство технологий



# 3. Несовершенство биоинформатического анализа NGS-данных

## Understanding the limitations of next generation sequencing informatics, an approach to clinical pipeline validation using artificial data sets

Robert Daber\*, Shrey Sukhadia, Jennifer J.D. Morrisette

*Center for Personalized Diagnostics, University of Pennsylvania School of Medicine, Philadelphia, PA*

The advantages of massively parallel sequencing are quickly being realized through the adoption of comprehensive genomic panels across the spectrum of genetic testing. Despite such widespread utilization of next generation sequencing (NGS), a major bottleneck in the implementation and capitalization of this technology remains in the data processing steps, or bioinformatics. Here we describe our approach to defining the limitations of each step in the data processing pipeline by utilizing artificial amplicon data sets to simulate a wide spectrum of genomic alterations. Through this process, we identified limitations of insertion, deletion (indel), and single nucleotide variant (SNV) detection using standard approaches and described novel strategies to improve overall somatic mutation detection. Using these artificial data sets, we were able to demonstrate that NGS assays can have robust mutation detection if the data can be processed in a way that does not lead to large genomic alterations landing in the unmapped data (i.e., trash). By using these pipeline modifications and a new variant caller, AbsoluteVar, we have been able to validate SNV mutation detection to 100% sensitivity and specificity with an allele frequency as low 4% and detection of indels as large as 90 bp. Clinical validation of NGS relies on the ability for mutation detection across a wide array of genetic anomalies, and the utility of artificial data sets demon-



**Каждая задача решается применением уникальной комбинацией скриптов/программ/алгоритмов....**



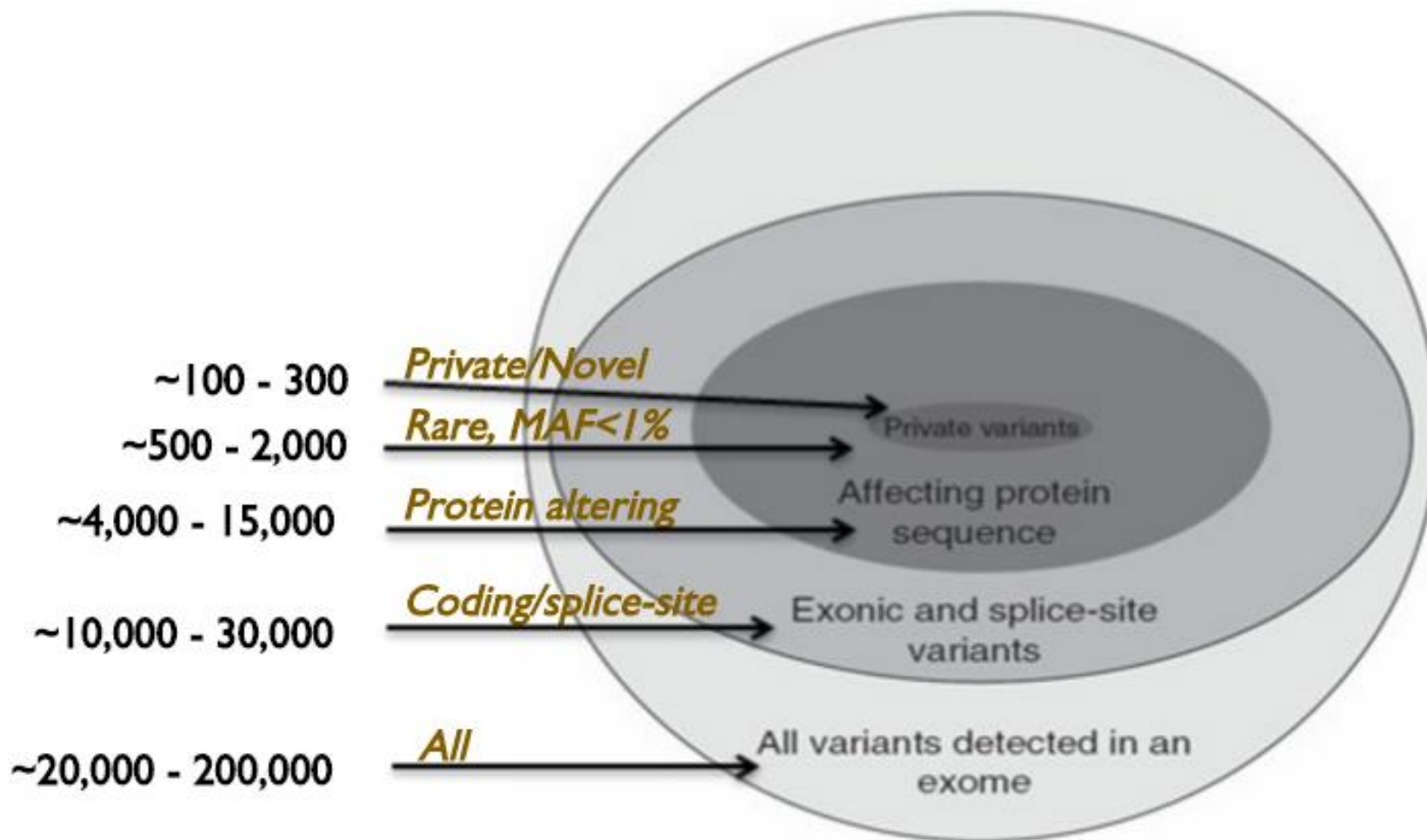
## 4. Сложность объекта – генома человека

- около 5-10 миллионов SNP (в зависимости от популяции);
- 25 000–50 000 редких вариантов (собственные мутации или с частотой меньше 0,5%);
- около 75 de novo мутаций, не обнаруженных в геномах родителей;
- 3-7 новых CNV приблизительно размером в 500 т.п.н. ДНК;
- 200 000–500 000 инделов (длинной от 1 до 50 п.о.) (в зависимости от численности населения);
- 500-1000 делеций размером от 1 до 45 т.п.н., затрагивающих около 200 генов;
- около 150 инделов, захватывающих рамку считывания;
- около 200-250 сдвигов рамки считывания;
- 10 000–12 000 синонимичных SNP;
- 8000–11000 несинонимичных SNP в 4000–5000 генах;
- 175-500 редких несинонимичных вариантов;
- 1 de novo несинонимичная мутация;
- около 100 преждевременных стоп-кодонов;
- 40-50 вариантов, нарушающих сайт сплайсинга;
- 250–300 генов с вероятными вариантами потери функции;
- около 25 генов могут быть полностью инактивированы.

# 4. Сложность объекта – генома человека

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y			
	Genom	Cl	Start	R	A	UCSC	GnomAD	Func.refGene	Gene.refGene	GeneDetail.refGeneWithVer	ExonicFunc.refGene	AAChange.refGeneWith	CLN	dbnp1	AF-exom	AF_popmax	AF2-genomes	AF_popmax3	regsn_fpu	regsn	regu	ID	POSITIO	ARIAT	REQU			
488	ref	4	8021150	C	G	chr4:8021150	4-8021150-C-G	splicing	ALB1M2	NM_001130085.2:exon12:c.1270-779G>C	-	-	-	-	-	-	0,0002	0,0003	0,613013699	B	off	rs12500190	8021151	C/T	0,3004			
489	alt	4	8021150	C	G	chr4:8021150	4-8021150-C-G	splicing	ALB1M2	NM_001130085.2:exon12:c.1270-779G>C	-	-	-	-	-	-	0,0002	0,0003	0,613013699	B	off	rs12500190	8021151	C/T	0,3004			
498	ref	17	56386119	T	C	chr17:56386119	17-56386119-T-C	exonic,splicing	TSP0AP1,TSP0AP	NM_004758.4:exon20:c.3699+1754A>G,N	nonsynonymous SNV	TSP0AP1:NM_004418.3:exon21	-	-	-	-	3,19E-05	0,0006	-	-	-	rs12680689	56386118	A/G	0,3021			
499	alt	17	56386119	T	C	chr17:56386119	17-56386119-T-C	exonic,splicing	TSP0AP1,TSP0AP	NM_004758.4:exon20:c.3699+1754A>G,N	nonsynonymous SNV	TSP0AP1:NM_004418.3:exon21	-	-	-	-	3,19E-05	0,0006	-	-	-	rs12680689	56386118	A/G	0,3021			
2846	ref	11	14369770	G	C	chr11:14369770	11-14369770-G-C	intronic	RRA52	-	-	-	-	-	-	-	-	3,19E-05	-	-	-	-	rs11023194	14369769	T/C	0,3026		
2847	alt	11	14369770	G	C	chr11:14369770	11-14369770-G-C	intronic	RRA52	-	-	-	-	-	-	-	-	3,19E-05	-	-	-	-	rs11023194	14369769	T/C	0,3026		
3054	ref	21	41451760	C	T	chr21:41451760	21-41451760-C-T	splicing	DSCAM	NM_001271534.3:exon25:c.4420+319G>A	-	-	-	-	-	-	3,19E-05	-	-	-	-	rs2297262	41451761	G/A	0,3029			
3055	alt	21	41451760	C	T	chr21:41451760	21-41451760-C-T	splicing	DSCAM	NM_001271534.3:exon25:c.4420+319G>A	-	-	-	-	-	-	3,19E-05	-	-	-	-	rs2297262	41451761	G/A	0,3029			
3430	ref	14	104135962	T	G	chr14:10413596	14-104135962-T-G	splicing	KLC1	NM_001301017.1:exon6:c.885+27T>G,N	-	-	-	-	-	-	0,0003	0,0011	0,232876712	B	off	rs12896171	104135966	A/G	0,3049			
3431	alt	14	104135962	T	G	chr14:10413596	14-104135962-T-G	splicing	KLC1	NM_001301017.1:exon6:c.885+27T>G,N	-	-	-	-	-	-	0,0003	0,0011	0,232876712	B	off	rs12896171	104135966	A/G	0,3049			
3444	ref	12	31271924	G	C	chr12:31271924	12-31271924-G-C	ncRNA_exonic	OVO52	-	-	-	-	-	-	-	-	0,0005	0,0008	-	-	-	rs78263854	31271923	C/T	0,3072		
3445	alt	12	31271924	G	C	chr12:31271924	12-31271924-G-C	ncRNA_exonic	OVO52	-	-	-	-	-	-	-	-	0,0005	0,0008	-	-	-	rs78263854	31271923	C/T	0,3072		
3702	ref	13	31317879	T	G	chr13:31317879	13-31317879-T-G	splicing	ALOX5AP	NM_001204406.1:exon3:c.242-318T>G,N	-	-	-	-	-	-	3,19E-05	0,0001	-	-	-	rs4254165	31317878	A/G	0,3086			
3703	alt	13	31317879	T	G	chr13:31317879	13-31317879-T-G	splicing	ALOX5AP	NM_001204406.1:exon3:c.242-318T>G,N	-	-	-	-	-	-	3,19E-05	0,0001	-	-	-	rs4254165	31317878	A/G	0,3086			
3734	ref	8	10481532	C	G	chr8:10481532	8-10481532-C-G	splicing	RP1L1	NM_178857.6:exon2-UTR5	-	-	-	-	-	-	6,37E-05	0,0024	-	-	-	rs10090876	10481531	C/A	0,3097			
3735	alt	8	10481532	C	G	chr8:10481532	8-10481532-C-G	splicing	RP1L1	NM_178857.6:exon2-UTR5	-	-	-	-	-	-	6,37E-05	0,0024	-	-	-	rs10090876	10481531	C/A	0,3097			
3800	ref	2	69330120	G	T	chr2:69330120	2-69330120-G-T	splicing	ANTXR1	NM_018153.3:exon10:c.802+48G>T,N	-	-	-	-	-	-	3,18E-05	6,48E-05	0,20890411	B	off	rs6710260	69330118	A/G	0,3118			
3801	alt	2	69330120	G	T	chr2:69330120	2-69330120-G-T	splicing	ANTXR1	NM_018153.3:exon10:c.802+48G>T,N	-	-	-	-	-	-	3,18E-05	6,48E-05	0,20890411	B	off	rs6710260	69330118	A/G	0,3118			
920	ref	8	87932356	A	G	chr8:87932356	8-87932356-A-G	intronic	CNBD1	-	-	-	-	-	-	-	-	3,19E-05	0,0001	-	-	-	rs6981203	87932352	A/G	0,3134		
921	alt	8	87932356	A	G	chr8:87932356	8-87932356-A-G	intronic	CNBD1	-	-	-	-	-	-	-	-	3,19E-05	0,0001	-	-	-	rs6981203	87932352	A/G	0,3134		
9374	ref	10	98792426	A	C	chr10:98792426	10-98792426-A-C	splicing	SUT1	NM_003061.3:exon23:c.2438+1802T>G,N	-	-	-	-	-	-	9,56E-05	0,0019	-	-	-	rs7086186	98792425	A/G	0,3148			
9375	alt	10	98792426	A	C	chr10:98792426	10-98792426-A-C	splicing	SUT1	NM_003061.3:exon23:c.2438+1802T>G,N	-	-	-	-	-	-	9,56E-05	0,0019	-	-	-	rs7086186	98792425	A/G	0,3148			
5380	ref	4	2698523	C	G	chr4:2698523	4-2698523-C-G	splicing	FAM193A	NM_001366316.1:exon17:c.3105+1667C>	-	-	-	-	-	-	3,19E-05	6,48E-05	0,239726027	B	off	rs4690086	2698524	G/A	0,3169			
5381	alt	4	2698523	C	G	chr4:2698523	4-2698523-C-G	splicing	FAM193A	NM_001366316.1:exon17:c.3105+1667C>	-	-	-	-	-	-	3,19E-05	6,48E-05	0,239726027	B	off	rs4690086	2698524	G/A	0,3169			
7964	ref	19	40389740	C	T	chr19:40389740	19-40389740-C-T	exonic	FCGBP	-	synonymous SNV	FCGBP:NM_003890.2:exon18:c.G	-	-	-	-	3,92E-05	0,0001	-	-	-	rs141158749	40389741	A/G	0,3169			
7965	alt	19	40389740	C	T	chr19:40389740	19-40389740-C-T	exonic	FCGBP	-	synonymous SNV	FCGBP:NM_003890.2:exon18:c.G	-	-	-	-	3,92E-05	0,0001	-	-	-	rs141158749	40389741	A/G	0,3169			
8010	ref	15	42378412	T	C	chr15:42378412	15-42378412-T-C	exonic,splicing	PLA2G4D,PLA2G4I	NM_178034.4:exon2:c.118+1414A>G,N	nonsynonymous SNV	PLA2G4D:NM_178034.4:exon2	-	-	-	rs55146761	0,0062	0,0136	0,0039	0,05	-	rs761088353	42378409	A/C	0,3222			
8011	alt	15	42378412	T	C	chr15:42378412	15-42378412-T-C	exonic,splicing	PLA2G4D,PLA2G4I	NM_178034.4:exon2:c.118+1414A>G,N	nonsynonymous SNV	PLA2G4D:NM_178034.4:exon2	-	-	-	rs55146761	0,0062	0,0136	0,0039	0,05	-	rs761088353	42378409	A/C	0,3222			
8784	ref	8	144403486	C	T	chr8:144403486	8-144403486-C-T	exonic	TOP1MT	-	nonsynonymous SNV	TOP1MT:NM_001238447.1:exon	-	-	-	-	rs77591811	8,02E-06	3,28E-05	9,57E-05	0,0002	-	rs11544482	144403485	G/C	0,3250		
8785	alt	8	144403486	C	T	chr8:144403486	8-144403486-C-T	exonic	TOP1MT	-	nonsynonymous SNV	TOP1MT:NM_001238447.1:exon	-	-	-	-	rs77591811	8,02E-06	3,28E-05	9,57E-05	0,0002	-	rs11544482	144403485	G/C	0,3250		
11102	ref	21	45454780	C	T	chr21:45454780	21-45454780-C-T	intronic	TRAPPC10	-	nonsynonymous SNV	TOP1MT:NM_001238447.1:exon	-	-	-	-	rs10132934	-	-	-	-	-	rs2238712	45454778	A/G	0,3295		
11103	alt	21	45454780	C	T	chr21:45454780	21-45454780-C-T	intronic	TRAPPC10	-	nonsynonymous SNV	TOP1MT:NM_001238447.1:exon	-	-	-	-	rs10132934	-	-	-	-	-	rs2238712	45454778	A/G	0,3295		
11104	ref	16	84228198	G	C	chr16:84228198	16-84228198-G-C	ncRNA_exonic,sp	LOC654780,ADAD	NM_001145400.2:exon2:c.559+10G>C,N	-	-	-	-	-	-	rs20198554	2,73E-05	0,0002	3,19E-05	0,0012	-	rs11864937	84228199	G/T	0,3297		
11105	alt	16	84228198	G	C	chr16:84228198	16-84228198-G-C	ncRNA_exonic,sp	LOC654780,ADAD	NM_001145400.2:exon2:c.559+10G>C,N	-	-	-	-	-	-	rs20198554	2,73E-05	0,0002	3,19E-05	0,0012	-	rs11864937	84228199	G/T	0,3297		
11446	ref	16	84228197	G	T	chr16:84228197	16-84228197-G-T	ncRNA_exonic,sp	LOC654780,ADAD	NM_001145400.2:exon2:c.559+9G>T,N	-	-	-	-	-	-	rs7610938	4,56E-06	7,50E-05	3,19E-05	0,0001	-	rs11864937	84228199	G/T	0,3297		
11447	alt	16	84228197	G	T	chr16:84228197	16-84228197-G-T	ncRNA_exonic,sp	LOC654780,ADAD	NM_001145400.2:exon2:c.559+9G>T,N	-	-	-	-	-	-	rs7610938	4,56E-06	7,50E-05	3,19E-05	0,0001	-	rs11864937	84228199	G/T	0,3297		
12102	ref	11	117988536	G	T	chr11:11798853	11-117988536-G-T	splicing	TMFR554	NM_001173551.1:exon12:c.1296+367G>T	-	-	-	-	-	-	rs10204101	-	-	-	-	-	rs671111	117988539	G/A	0,3303		
12103	alt	11	117988536	G	T	chr11:11798853	11-117988536-G-T	splicing	TMFR554	NM_001173551.1:exon12:c.1296+367G>T	-	-	-	-	-	-	rs10204101	-	-	-	-	-	rs671111	117988539	G/A	0,3303		
12252	ref	11	82920904	A	G	chr11:82920904	11-82920904-A-G	splicing	ANKRD42	NM_001300976.2:exon4:c.331-438A>G,N	-	-	-	-	-	-	3,18E-05	-	-	-	-	-	rs613175	82920900	C/G	0,3314		
12253	alt	11	82920904	A	G	chr11:82920904	11-82920904-A-G	splicing	ANKRD42	NM_001300976.2:exon4:c.331-438A>G,N	-	-	-	-	-	-	3,18E-05	-	-	-	-	-	rs613175	82920900	C/G	0,3314		
12780	ref	X	30993058	A	G	chrX:30993058	X-30993058-A-G	intergenic	TAB3,FTHL17	dist=85687,dist=96292	-	-	-	-	-	-	9,19E-05	0,0003	-	-	-	-	-	rs5927653	30993057	T/C	0,3319	
12781	alt	X	30993058	A	G	chrX:30993058	X-30993058-A-G	intergenic	TAB3,FTHL17	dist=85687,dist=96292	-	-	-	-	-	-	9,19E-05	0,0003	-	-	-	-	-	rs5927653	30993057	T/C	0,3319	
13032	ref	16	87877249	C	T	chr16:87877249	16-87877249-C-T	intronic	SLC7A5	-	-	-	-	-	-	-	-	6,39E-05	0,0012	-	-	-	-	-	rs58852522	87877246	T/C	0,3418
13033	alt	16	87877249	C	T	chr16:87877249	16-87877249-C-T	intronic	SLC7A5	-	-	-	-	-	-	-	-	6,39E-05	0,0012	-	-	-	-	-	rs58852522	87877246	T/C	0,3418
13089	ref	11	60102508	T	G	chr11:60102508	11-60102508-T-G	exonic	MS446E	-	nonsynonymous SNV	MS446E:NM_139248.2:exon1:c.T	-	-	-	-	3,19E-05	0,0012	-	-	-	-	-	rs2304933	60102507	G/T	0,3456	
13098	alt	11	60102508	T	G	chr11:60102508	11-60102508-T-G	exonic	MS446E	-	nonsynonymous SNV	MS446E:NM_139248.2:exon1:c.T	-	-	-	-	3,19E-05	0,0012	-	-	-	-	-	rs2304933	60102507	G/T	0,3456	
13156	ref	11	60102508	T	G	chr11:60102508	11-60102508-T-G	exonic	MS446E	-	nonsynonymous SNV	MS446E:NM_139248.2:exon1:c.T	-	-	-	-	3,19E-05	0,0012	-	-	-	-	-	rs2304933	60102507	G/T	0,3456	
13157	alt	11	601025																									

# Интерпретация NGS данных



Gilissen C. et al. *Eur. J. Hum. Genet.* (2012) 20:490-497.



## 5. Сложность объекта – заболевания человека

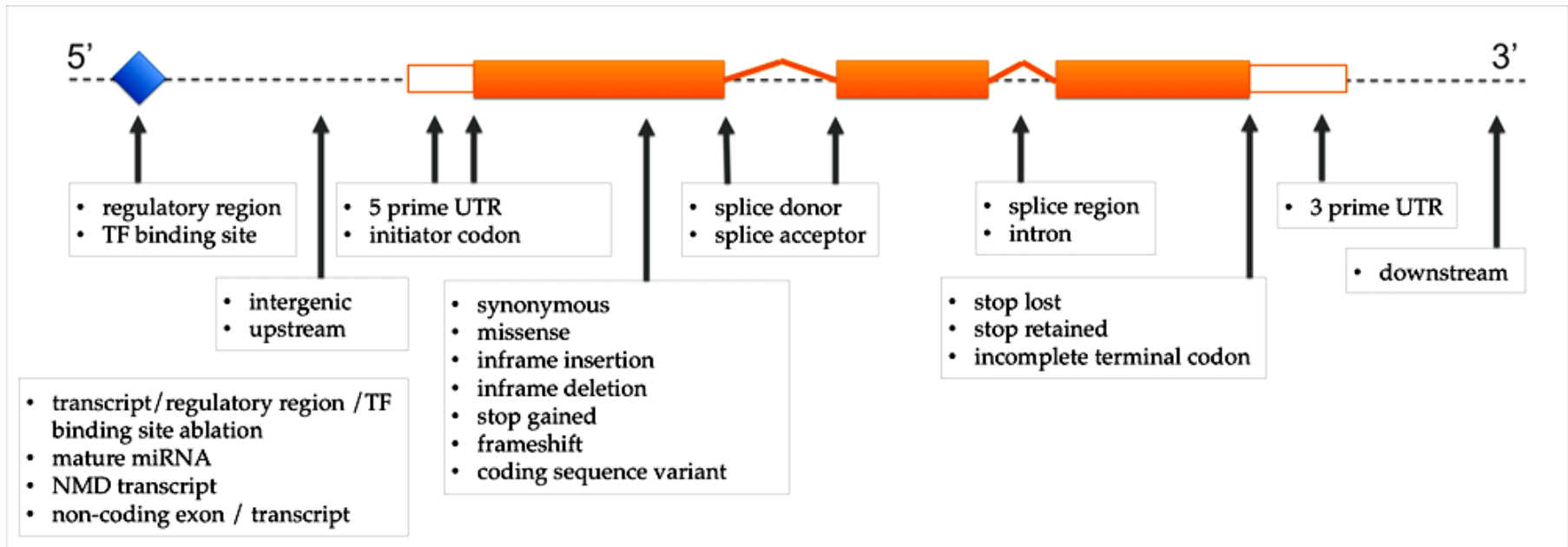
	Количество
Описание фенотипа с известными молекулярными основами	6'111
Описание фенотипа с неизвестными молекулярными основами	1'527
Фенотипы с предполагаемым менделевским наследованием	1'756

## 5. Сложность объекта – заболевания человека

	Количество
Описание фенотипа с известными молекулярными основами	6'111
Описание фенотипа с неизвестными молекулярными основами	1'527
Фенотипы с предполагаемым менделевским наследованием	1'756



## 6. Огромное количество всевозможных механизмов патогенеза заолеваний



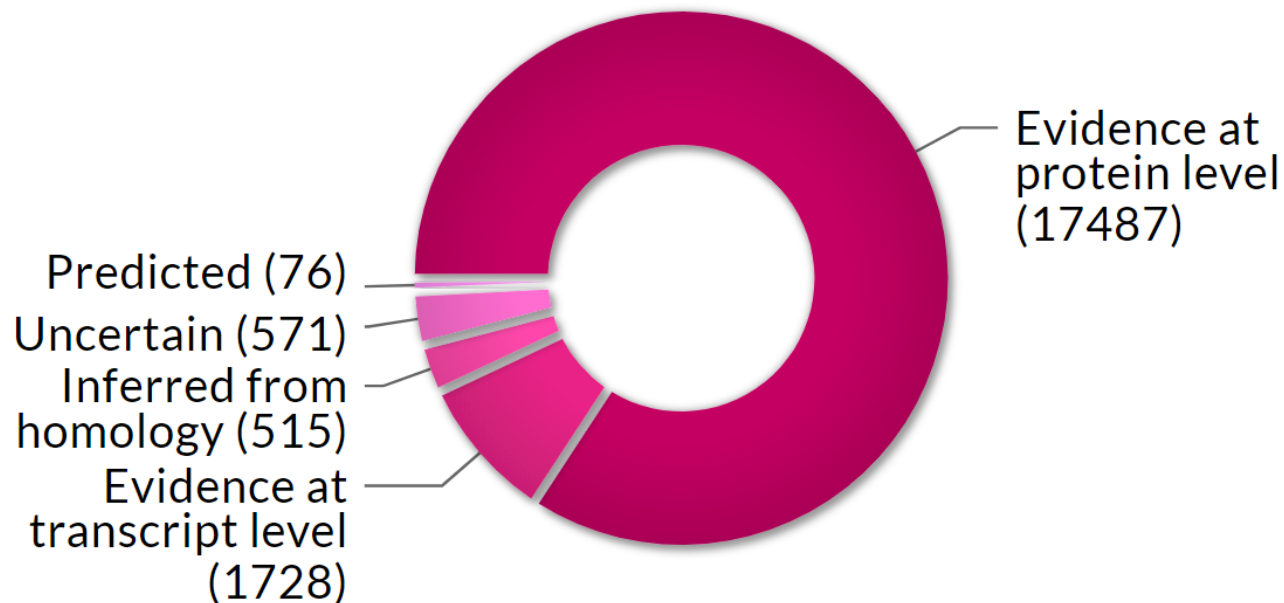
# До сих пор непонятно сколь генов и какие они

	RefSeq	Ensembl	Gencode (v. 29)
<b>Total No of Genes</b>	<b>6118</b>	<b>57365</b>	<b>58721</b>
Protein-coding genes	20216	20418	19940
Long non-coding RNA genes	18533	15014	16066
Small non-coding RNA genes		4871	7577
Pseudogenes	16435	15195	14505

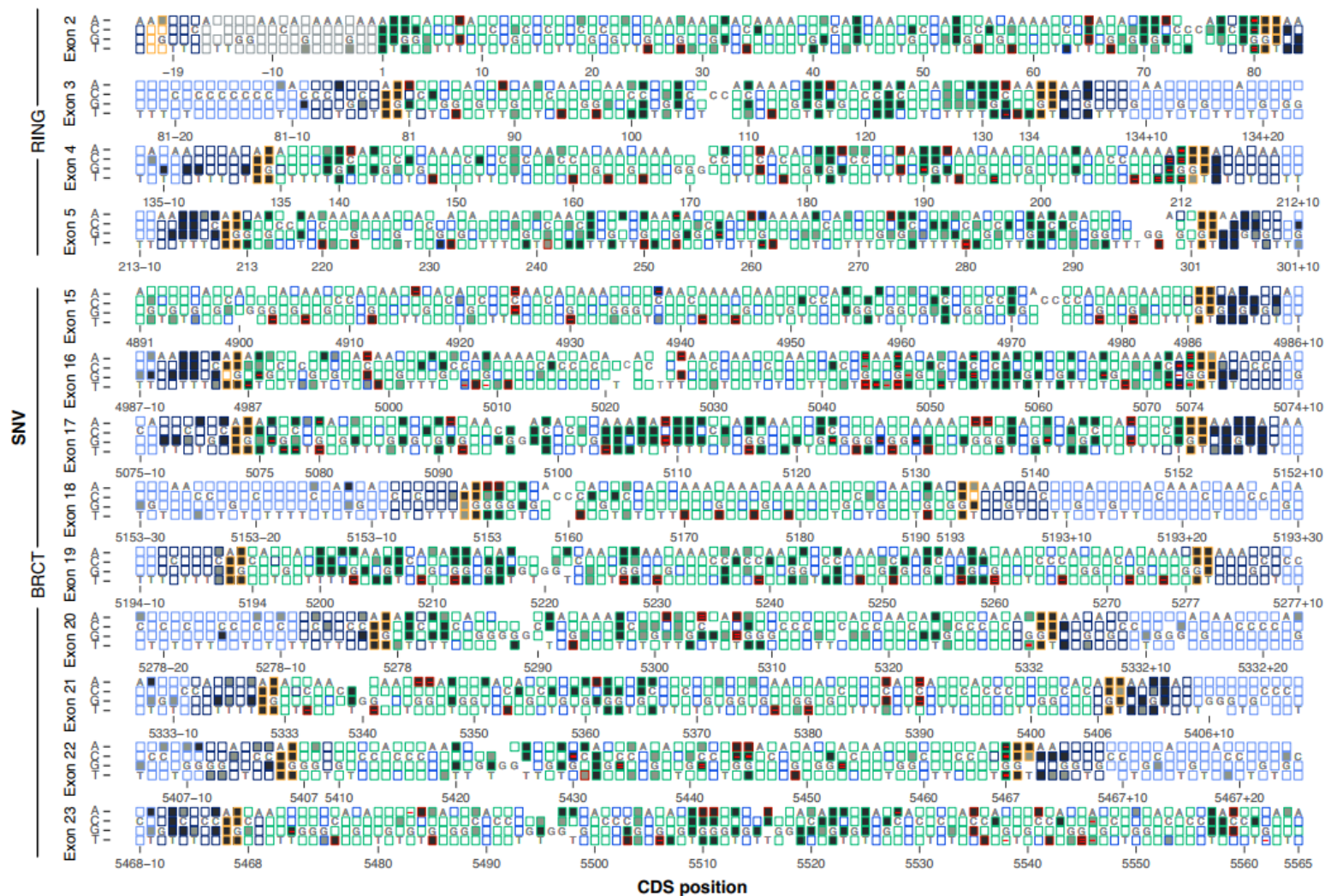
# До сих пор не понятно кто что кодирует и в чем тут смысл

neXtprot

Protein existence in neXtProt



# Патогенность вариантов известна только для небольшого их числа (меньше 1%)

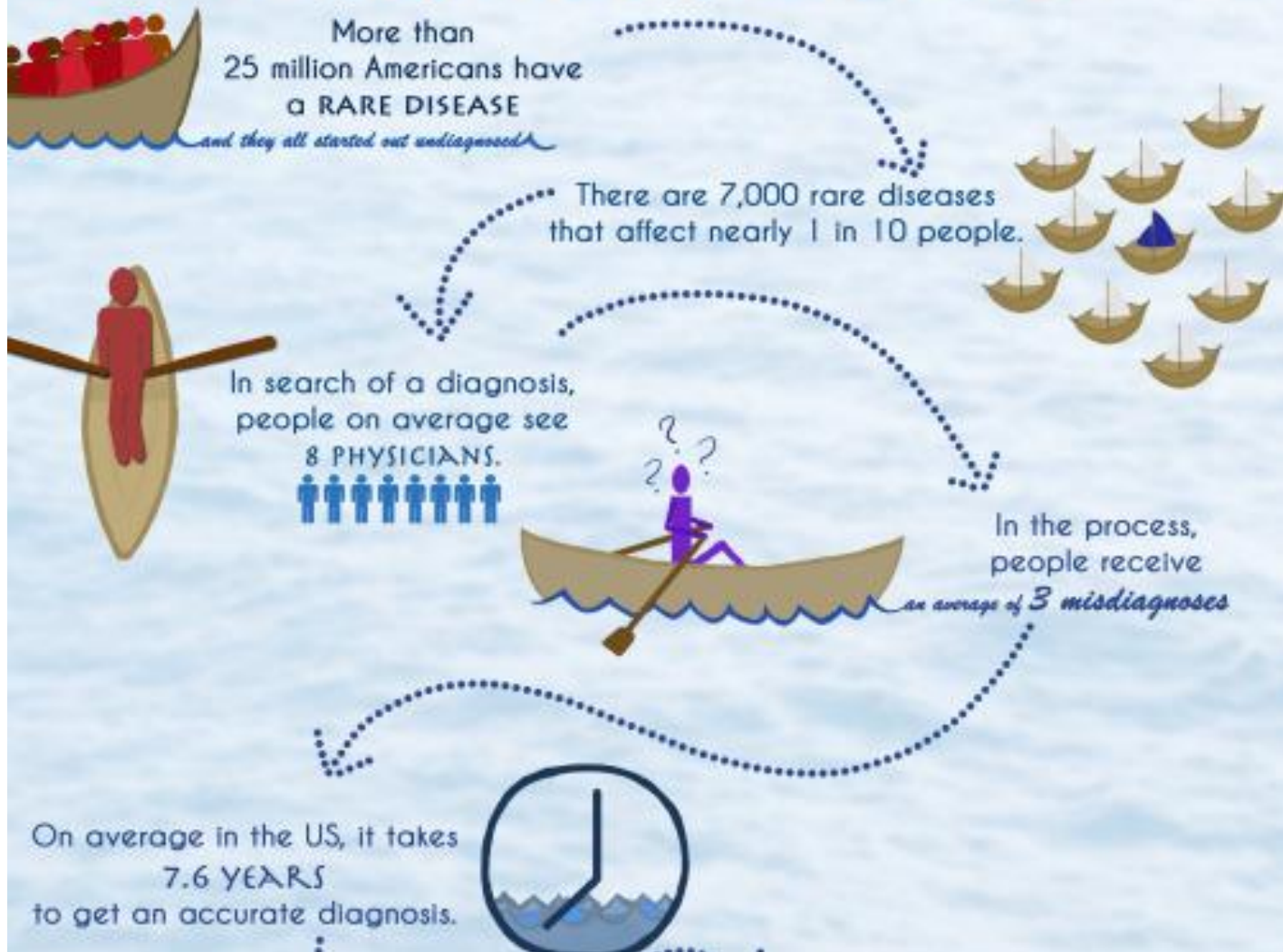


# Вы сейчас находитесь здесь:

- Нет прибора позволяющего получать идеальные данные
- Нет алгоритмов позволяющих идеально процессировать данные
- Слишком много функционально-значимых мест в геноме, и слишком мало мы о них знаем
- Очень много заболеваний и мало мы знаем о их механизмах



# A DIAGNOSTIC ODYSSEY



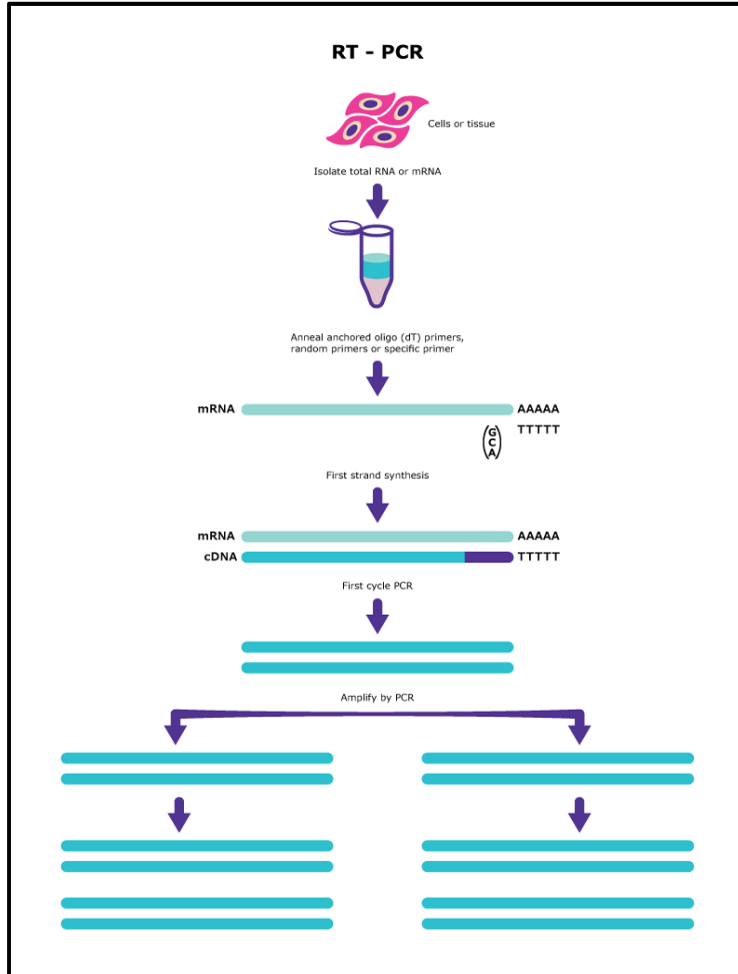
# Альтернативные подходы



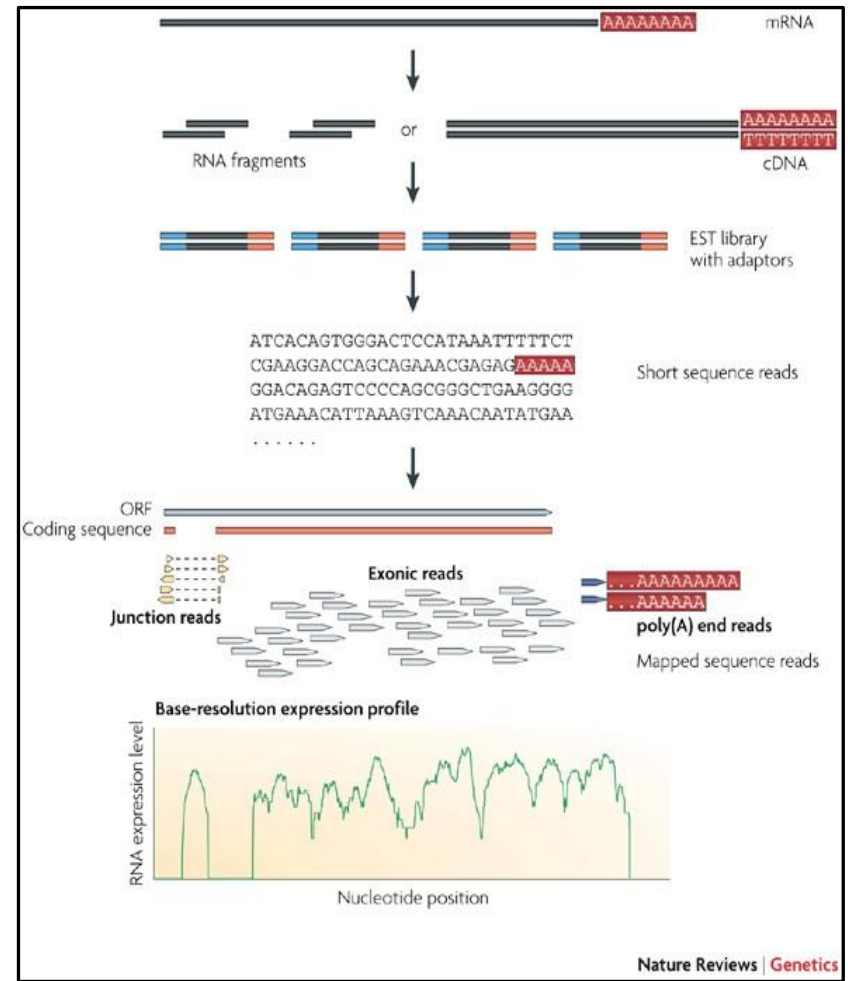
- Десятки разных алгоритмов предсказывающих .....
- Разные экспериментальные подходы позволяющие идентифицировать перестройки, химерные транскрипты, lncRNA, miRNA, конформационные изменения ДНК и т.п.
- РНК-анализ

# РНК-анализ

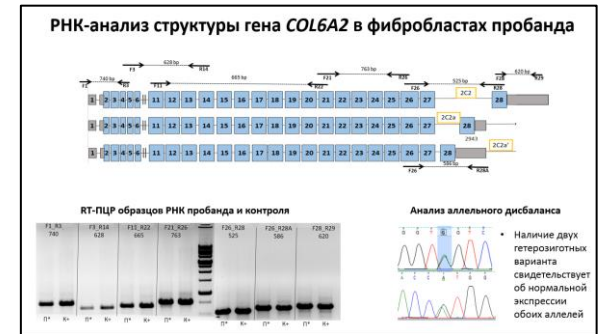
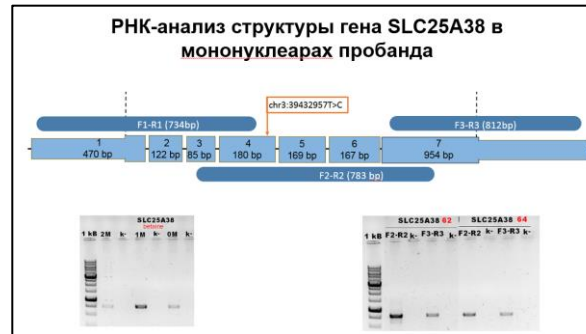
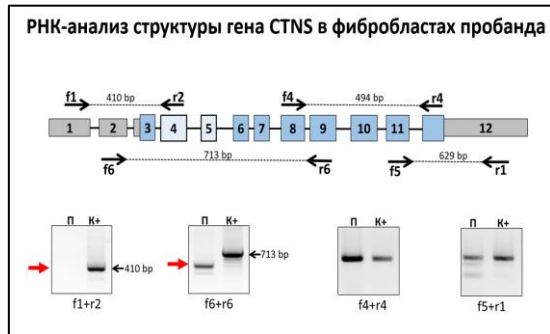
- Локальный



- Глобальный



# Преимущества использования РНК анализа



- Обнаружение нарушений структуры мРНК за счёт глубоко-интронных вариантов
- Поиск ранее не найденных вариантов
- Анализ изменения экспрессии
- Обнаружение ошибок ДНК-диагностики
- Позволяет с высокой степенью уверенности исключить ген из списка генов-кандидатов для конкретного генетического заболевания.



Черная пятница: 178 евро

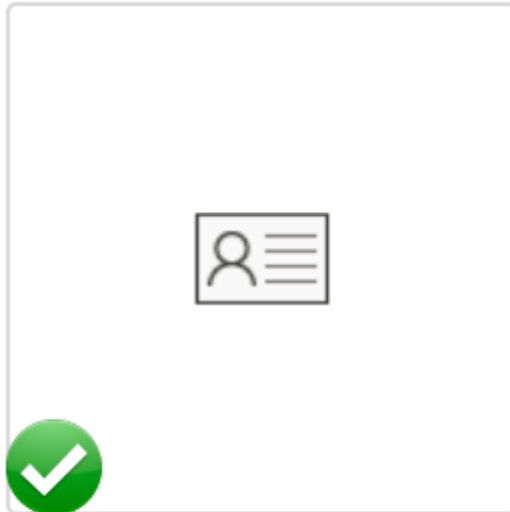
# Сравнение результатов секвенирования индивидуальных геномов человека

Individual	Ploidy	Technology	Av Depth	Total SNPs [M]
Venter	2n	Sanger	7.5 ×	3.21
Watson	2n	Roche 454	7.4 ×	3.32
Chinese (YH)	2n	Illumina	36.0 ×	3.07
African (NA18507)*	2n	Illumina	40.6 ×	3.61
African (NA18507)*	2n	AB SOLiD	17.9 ×	3.86
Korean (SJK)	2n	Illumina	28.9 ×	3.43
Korean (AK1)	2n	Illumina	27.8 ×	3.45
Khoisan (KB1)	2n	Roche 454	10.2 ×	4.05
D. Tutu (ABT)	2n	AB SOLiD	30.0 ×	3.62
Lupski	2n	AB SOLiD	29.6 ×	3.42

Gonzaga-Jauregui C, Lupski JR, Gibbs RA.  
Human genome sequencing in health and disease.  
Annu Rev Med. 2012;63:35-61.



# 3'415'465 персональных вариаций



56001801066408A.snp.vcf

860 Mb



56001801066408A-6584-Pharmacogenetics Report.pdf



Wellness-and-Longevity-12269  
18-56001801066408A\_snp\_vcf\_  
gz-18Dec21.pdf



# Genetic Report

Confidential Report Number

1226918

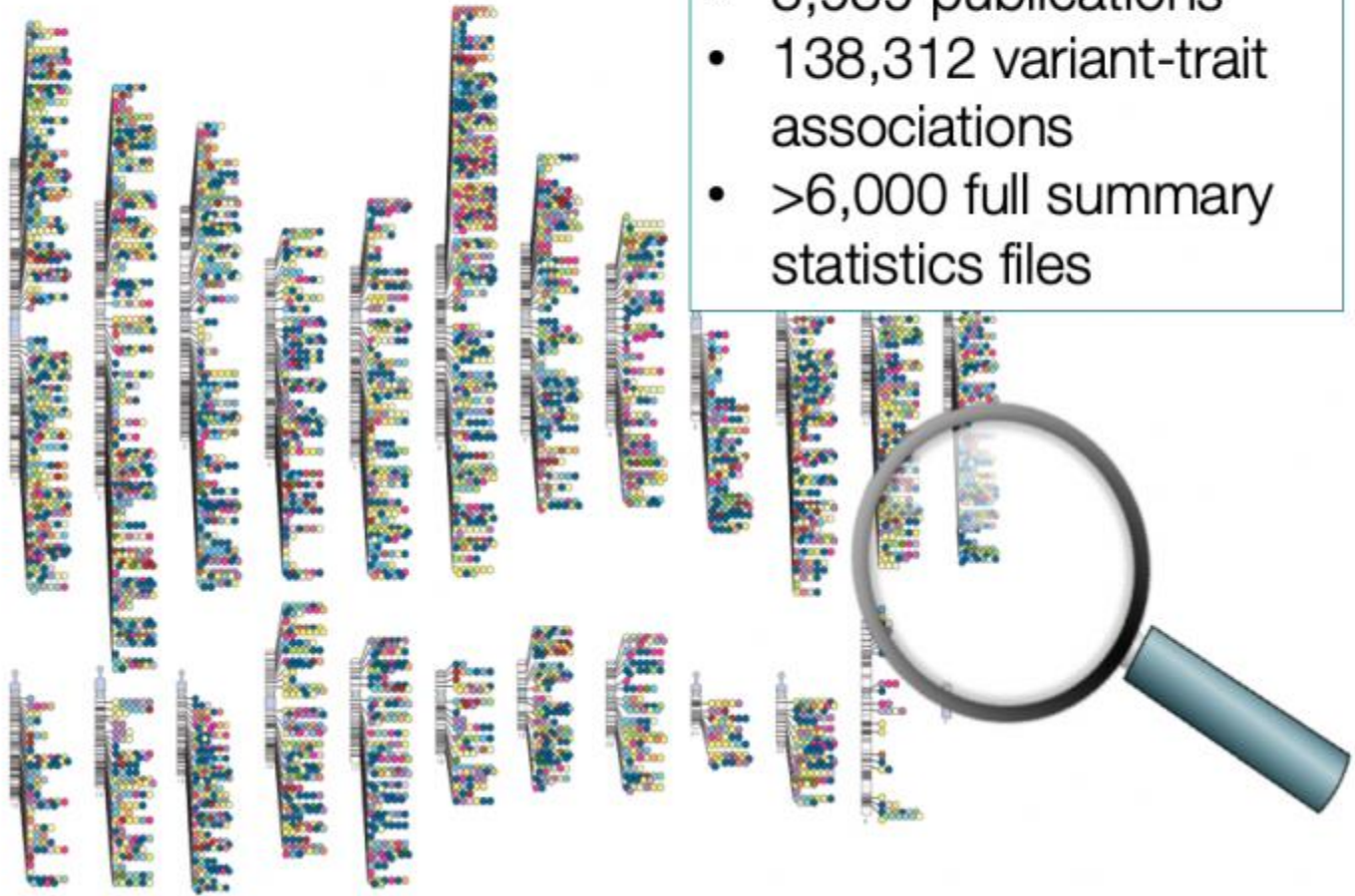
## Wellness & Longevity App

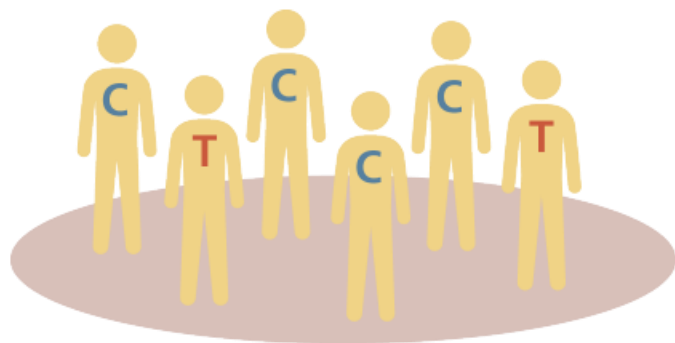
- The Report analyzes a large amount of genomic data, associating genetic variants found in the **genomic files** with variants known from the scientific literature. While this Report does not require FDA/EMA approval, we do want to point out that it has not been approved by the FDA/EMA for such use.
- We do not independently judge the validity or accuracy of such published scientific information.
- Because scientific and medical information changes over time, your risk assessment and genetically tailored prevention for one or more of the medications contained within this report may also change over time.
- Therefore, this report may not be 100% accurate (e.g., new research could mean different results) and may not predict actual results or outcomes.

# GWAS Catalog

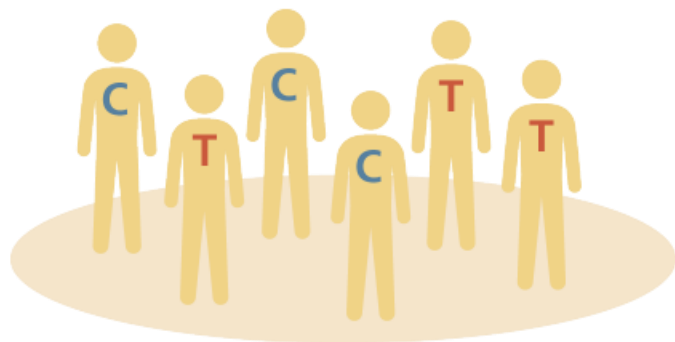
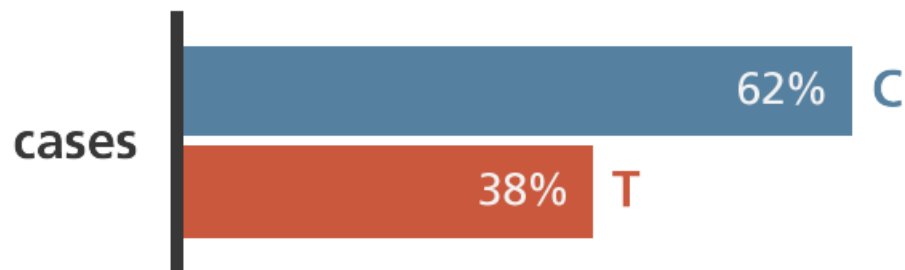
As of May 2019

- 3,989 publications
- 138,312 variant-trait associations
- >6,000 full summary statistics files





**cases (n=1,000)**  
people with heart disease



**controls (n=1,000)**  
people without heart disease



# «Genetic Report» на 163 страницах

## Your Genetic Testing Data



**Variant ID** column lists the exact position that was tested for within the gene listed in column #2. This variant ID can be thought of as the exact "GPS coordinate" within the gene.

**Gene** column refers to the gene that's being tested for in that row.

**No Risk** column refers to the letter of the genetic code that is usually not associated with having an increased or decreased risk of the disease, condition, or trait listed in column #6.

**Risk** column refers to the letter of the genetic code that is likely to be associated with having either an increased or decreased risk of the disease, condition, or trait listed in column #6.

**Your Genetic Makeup** column refers to the exact letters of YOUR genetic makeup at that position (column #1) in that gene (column #2). Single most genes come in pairs, there are usually two letters at each position. If only one letter is listed, this means you only have one copy of that gene (which is perfectly normal for some genes).

The letters of the genetic code are G, A, T, and C. You may also see an I (Insertion) or D (Deletion). Two dashed lines "--" means that variant's data did not pass quality control and therefore the data was excluded from your analysis.

**Condition / Trait Assessed** column lists what exactly is being analyzed at that specific position within the gene listed in column #2.

**Reference(s)** column refers to the scientific research studies that found that the specific position (column #1) within the gene (column #2) is associated with the specific disease, condition, or trait listed in column #6. You can find these papers by either searching pubmed.com or Google for the reference listed (one at a time) along with the Gene and the name of the condition.

Variant ID	Gene	No Risk	Risk	Your Genetic Makeup	Condition / Trait Assessed	Reference(s)
40652	AMPD1	G	A	GA	Surmountable Exercise-induced Fatigue	Morisaki (1992), Rico-Sanz (2003), Lucia (2005), Ruiz (2009), Eynon (2013)
44	ACTN3	C	T	CC	Athletic Predisposition	North (1999), Suminaga (2000), Yang (2003), MacArthur (2004), Niemi (2005), Lucia (2006), MacArthur (2007), Moran (2007), Roth (2007), Santiago (2008), Eynon (2009), Shang (2010), Berman (2010), Gentil (2011), Hagberg (2011), Chiu (2011), Ciężczyk (2011), Puthucherry (2011), Eynon (2011), Shang (2012), Pimenta (2012), Zilberman-Schapira (2012), Mehlman (2012), Kikuchi (2012), Eynon (2012), Eynon (2013), Tucker (2013), Ahmetov (2013), Guth (2013), Grealis (2013), Pimenta (2013), Seto (2013), Maffulli (2013), Massidda (2014), Mikami (2014), Morucci (2014), Garatachea (2014), Kim (2014), Tringali (2014), Kikuchi (2015), Orysiak (2015), Ben-Zaken (2015), Ahmetov (2015), Deschamps (2015), Head (2015), Kikuchi (2015), Coelho (2016), Sarzynski (2016), Pasqua (2016), Garton (2016), Papadimitriou (2016), Baumert (2016), Lee (2016), Del Coso (2017), Dionísio (2017), Galeandro (2017), Li (2017), Yang (2017)
70740	MTCYB	G	A	?	Insurmountable Exercise-induced Fatigue	Bouzidi (1993), Dumoulin (1996), Andreu (1999), Katirji (2013)
73822	MTTG	T	C	?	Insurmountable Exercise-induced Fatigue	Nishigaki (2002), Kitzman (2008), Bhatia (2015)

- 3'450 описанных вариаций

# «Genetic Report» на 163 страницах

Lactose Intolerance

Increased risk of becoming lactose intolerant  
as an adult

Asthma

RISK DETECTED

Resistance to HIV Infection

No Resistance Detected

# «Genetic Report» на 163 страницах

CONDITION NAME	RESULTS	MAIN MESSAGE
Ethanol	✓	No variants detected
Heroin	✓	No variants detected
Metformin	✓	No variants detected
Methadone	⚠	We found a variant related to your reaction to Methadone
Methotrexate	✗	We found a variant related to your reaction to Methotrexate
Mirtazapine	⚠	We found a variant related to your reaction to Mirtazapine
Morphine	⚠	We found a variant related to your reaction to Morphine



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