Платформы для секвенирования. Приборы, наборы и софт

Александр Лавров

Медико-генетический научный центр
им. Н.П. Бочкова

Этапы высокопроизводительного секвенирования

1. Подготовка ДНК библиотеки
2. Амплификация ДНК библиотеки
3. Секвенирование ДНК библиотеки
4. Анализ данных

Платформа = 1 + 2+3 + 4

Приборы

Ion Torrent next-generation sequencing systems

Ion GeneStudio S5 System

Ion Torrent Genexus System

iSeq 100

MiniSeq

MiSeq Series ⊕

NextSeq 550 Series ⊕ NextSeq 1000 & 2000

NextSeq 550 Series ⊕

NextSeq 1000 & 2000

NovaSeq 6000

Sequencers ⊕

Sequencers ⊕

Sequencers ⊕

Sequencers ⊕

DNBSEQ-T7	DNBSEQ-G400	DNBSEQ-G50	DNBSEQ-G400 FAST




	iSeq 100 System	MiniSeq System	MiSeq Series ⊕	NextSeq Series ⊕
Popular Applications & Methods	Key Application	Key Application	Key Application	Key Application
Large Whole-Genome Sequencing (human, plant, animal)				●
Small Whole-Genome Sequencing (microbe, virus)	●	●	●	●
Exome Sequencing				●
Targeted Gene Sequencing (amplicon, gene panel)	●	●	●	●
Maximum Reads Per Run	4 million	25 million	25 million †	400 million
Maximum Read Length	2 × 150 bp	2 × 150 bp	2 × 300 bp	2 × 150 bp




	NextSeq Series ⊕	HiSeq 4000 System	HiSeq X Series‡	NovaSeq 6000 System
Popular Applications & Methods	Key Application	Key Application	Key Application	Key Application
Large Whole-Genome Sequencing (human, plant, animal)	●	●	●	●
Small Whole-Genome Sequencing (microbe, virus)	●	●		●
Exome Sequencing	●	●		●
Targeted Gene Sequencing (amplicon, gene panel)	●	●		●
Maximum Reads Per Run	400 million	5 billion	6 billion	20 billion
Maximum Read Length	2 × 150 bp	2 × 150 bp	2 × 150 bp	2 x 250**

Ion GeneStudio S5 System

Ion GeneStudio S5 Plus System

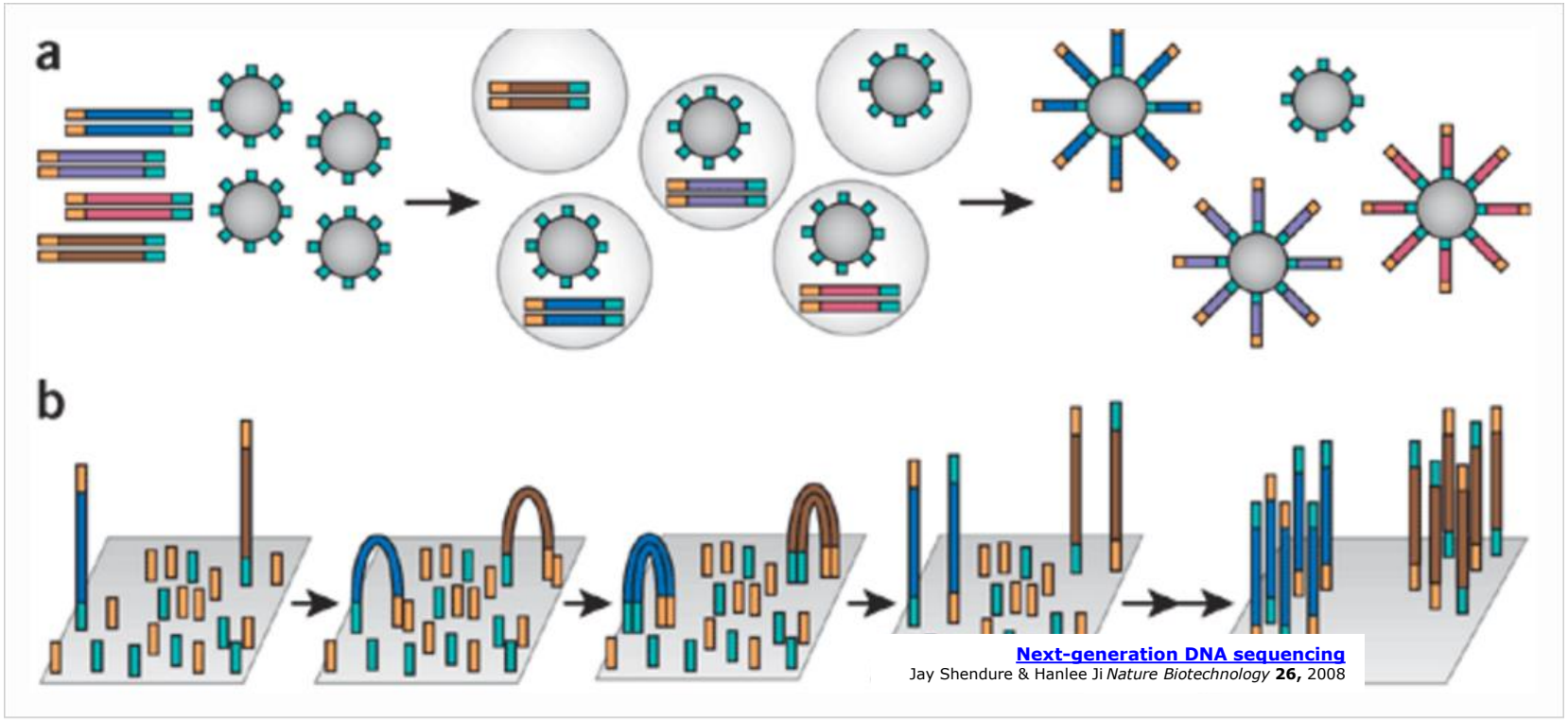
Ion GeneStudio S5 Prime System

Chip type	Number of reads	Read length (output*)	Ion GeneStudio™ S5 System	Ion GeneStudio™ S5 Plus System	Ion GeneStudio™ S5 Prime System
			Turnaround time (sequencing run** plus analysis time)		
Ion 510 Chip	2–3 million	200 bp (0.3–0.5 Gb)	4.5 hr	3 hr	3 hr
		400 bp (0.6–1 Gb)	10.5 hr	5 hr	5 hr
Ion 520 Chip	4–6 million	200 bp (0.6–1 Gb)	7.5 hr	3.5 hr	3 hr
		400 bp (1.2–2 Gb)	12 hr	5.5 hr	5.5 hr
	3–4 million	600 bp (0.5–1.5 Gb)	12 hr	5.5 hr	5.5 hr
Ion 530 Chip	15–20 million	200 bp (3–4 Gb)	10.5 hr	5 hr	4 hr
		400 bp (6–8 Gb)	21.5 hr	8 hr	6.5 hr
	9–12 million	600 bp (1.5–4.5 Gb)	21 hr	8 hr	7 hr
Ion 540 Chip	60–80 million	200 bp (10–15 Gb)	19 hr	10 hr	6.5 hr
		200 bp (20–30 Gb) 2 runs in 1 day	NA	20 hr	10 hr†
Ion 550 Chip	100–130 million	200 bp (20–25 Gb)	NA	11.5 hr	8.5 hr
		200 bp (40–50 Gb) 2 runs in 1 day	NA	NA	12 hr†



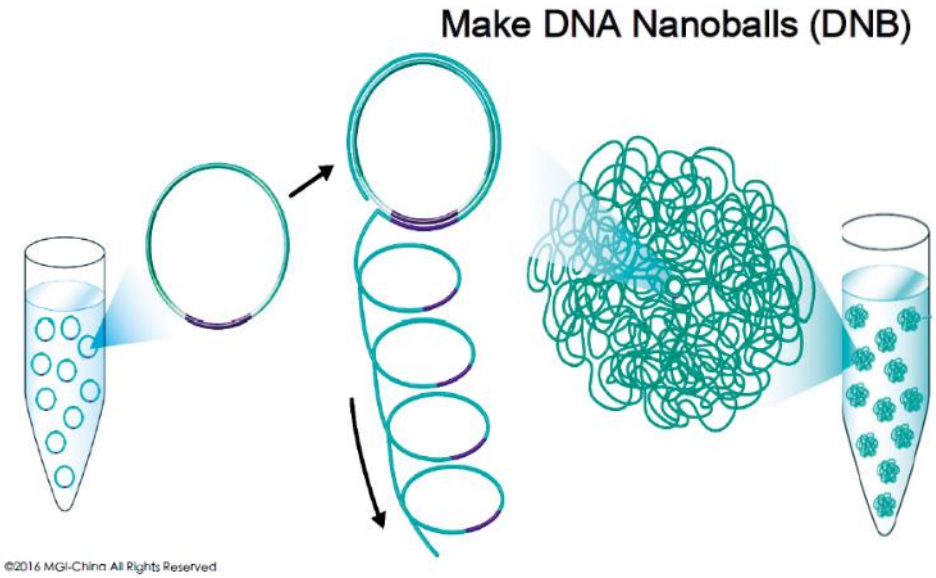


DNBSEQ-T7	DNBSEQ-G400 （previously MGISEQ-2000）	DNBSEQ-G50 （previously MGISEQ-200）	BGISEQ-500	BGISEQ-50
Ultra-high Throughput	Adaptive	Effective	Reliable	Fast
Whole Genome Sequencing,Deep Exome Sequencing,Transcriptome Sequencing,and Targeted Panel Projects.	WGS, WES, Transcriptome sequencing and more	Targeted DNA, RNA, Microbial sequencing	Targeted DNA, RNA, Epigenetics and clinical applications	Pathogen Rapid tests, NIPT, PGS and CNV tests
FC	FCL & FCS	FCS	FCL	FCS
——	4 lane & 2 lane	1 lane	2 lane	2 lane
Ultra-high Throughput	High Throughput	Medium Throughput	High Throughput	Low Throughput
6Tb	1440Gb	60Gb	520Gb	225Gb
5000M	1500-1800M	280-300M	1300M	375M



Наборы.

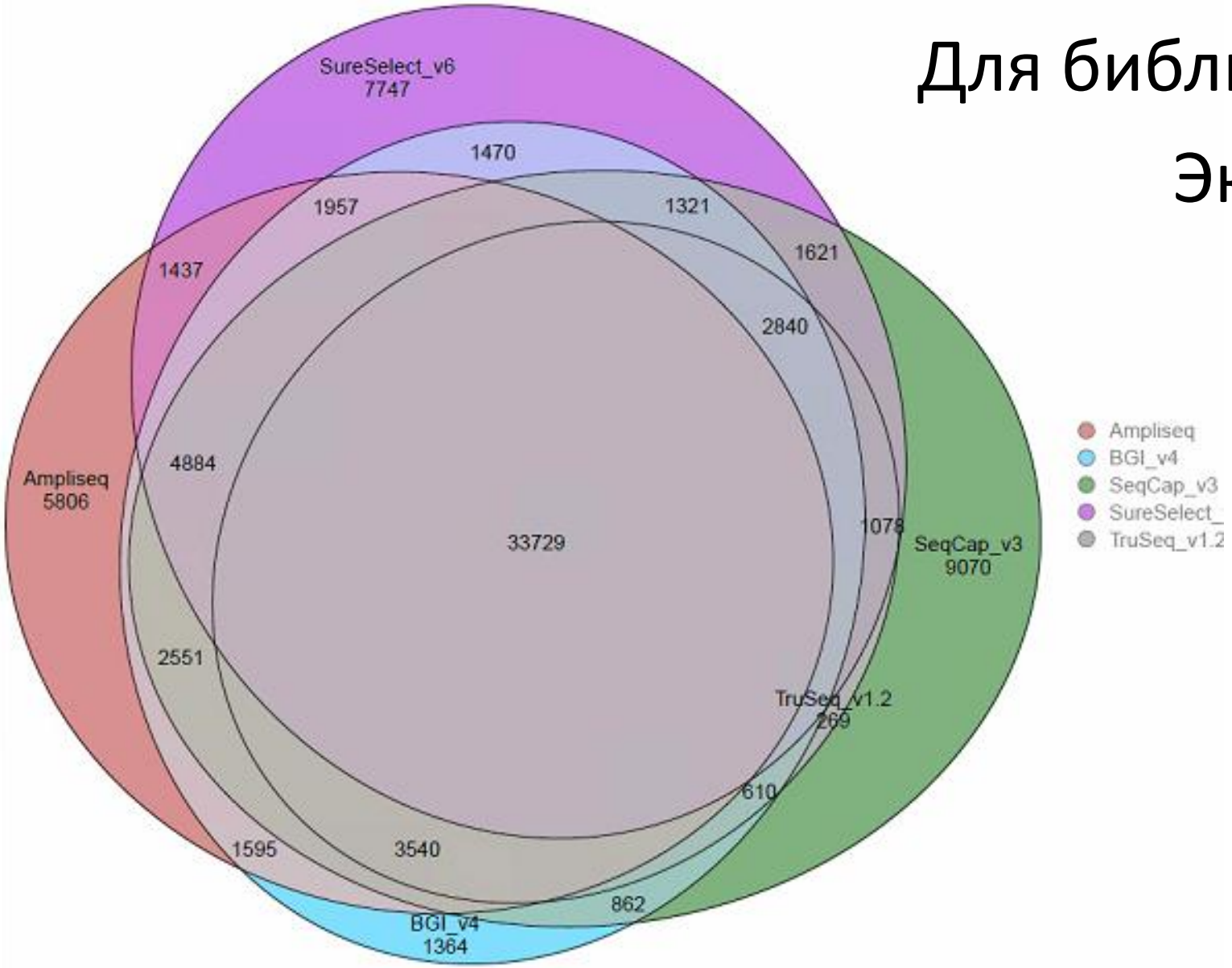
Для сиквенса



Наборы.

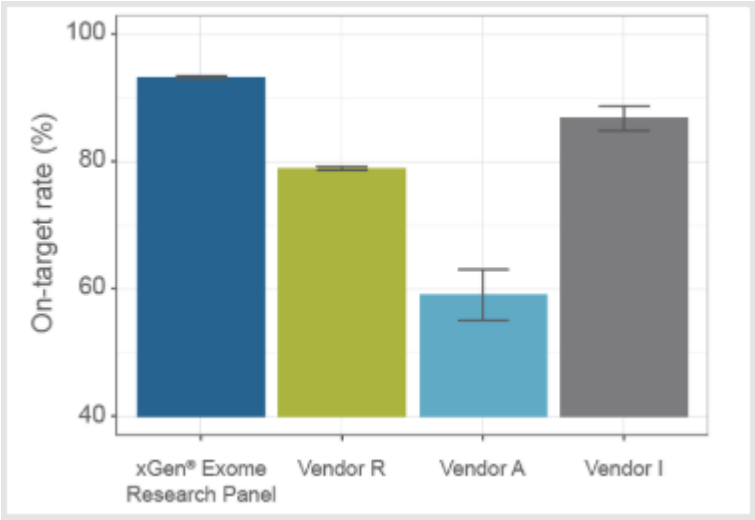
Для библиотек.

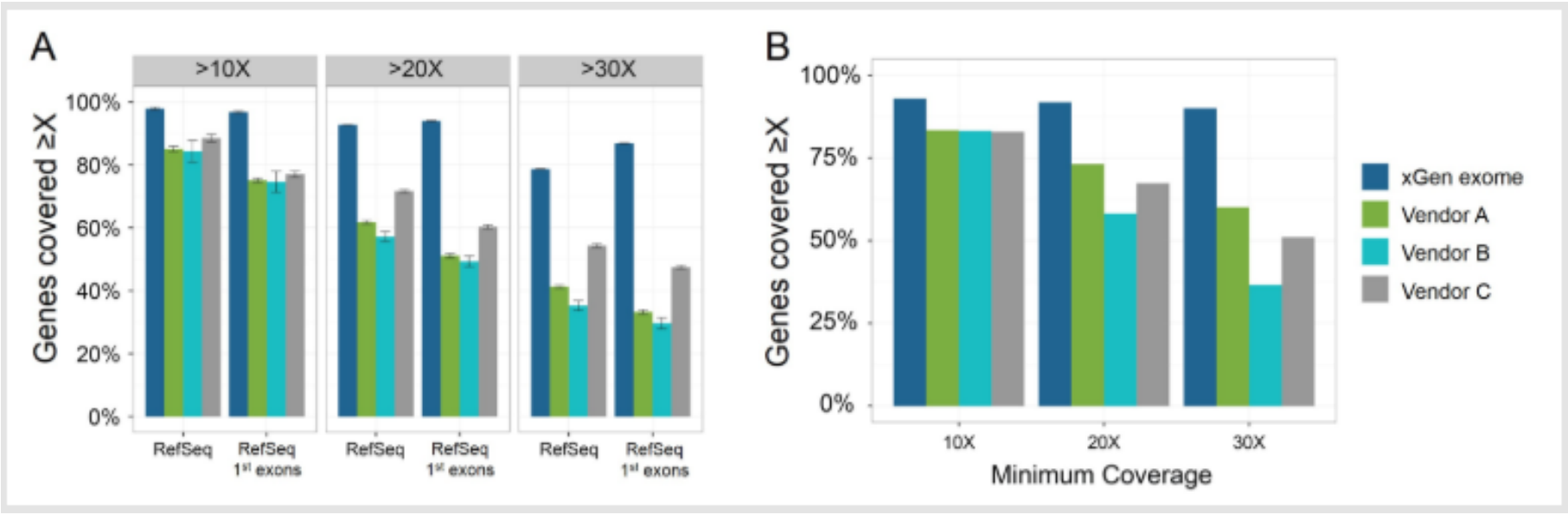
Экзомы



xGen® Exome Research Panel

- 429,826 xGen Lockdown® Probes
- 39 Mb target region (19,396 genes)
- covers 51 Mb of end-to-end tiled probe space
- GMP standards
- QC of each probe



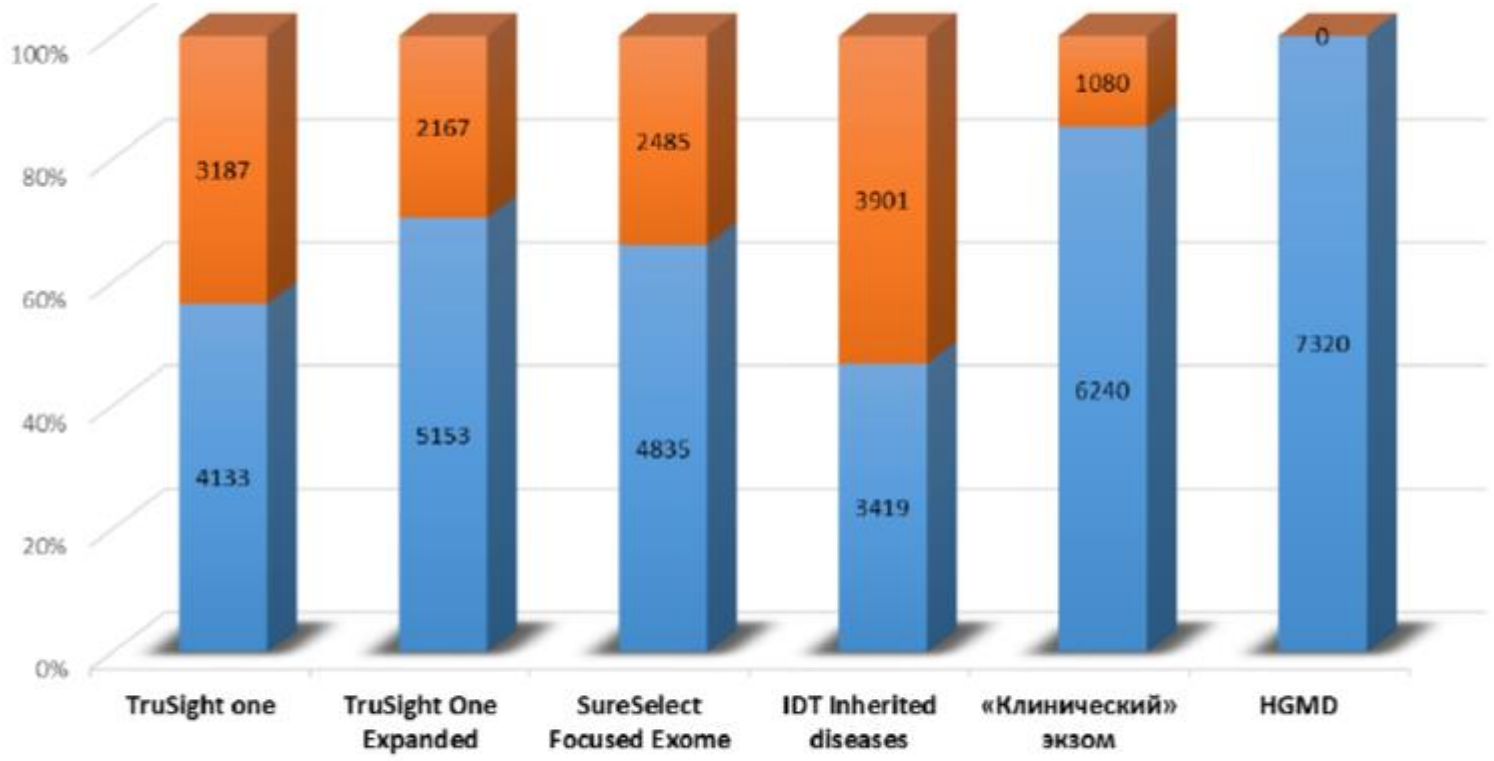


SureSelect Human All Exon V7 is the most comprehensive exome

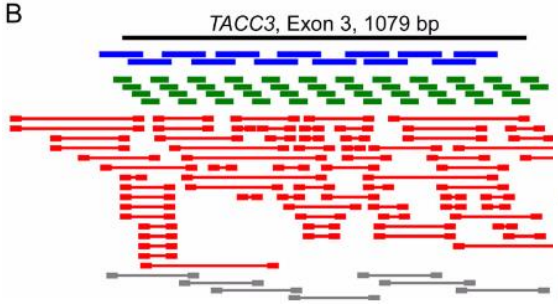
	RefSeq	GENCODE	CCDS	UCSC Known Genes
SureSelect Human All Exon V7	**99.3%**	**99.6%**	**99.6%**	**99.6%**
Vendor ID	97.3%	97.1%	98.3%	94.5%
Vendor R	96.9%	97.2%	97.5%	96%
Vendor I	98.8%	99.1%	99.5%	98.3%
Vendor T	96.9%	97.1%	99.9%	93.7%

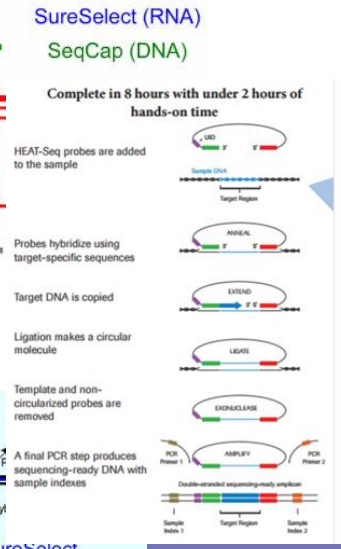
«Клинический» экзом

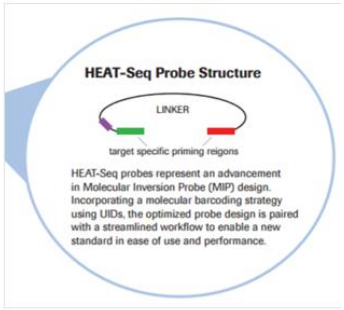
– совокупность генов, патогенные варианты в которых приводят к развитию наследственных (моногенных) заболеваний

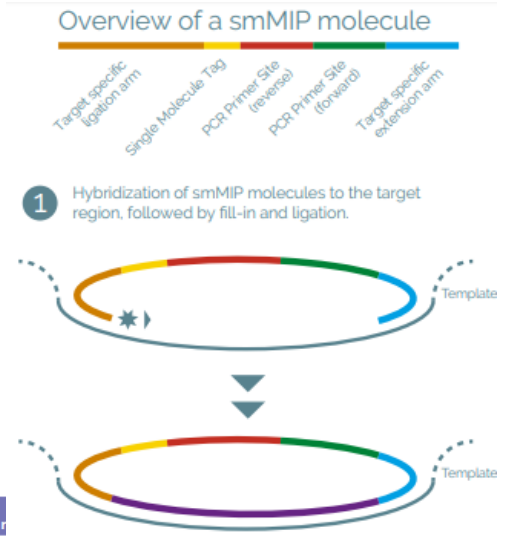


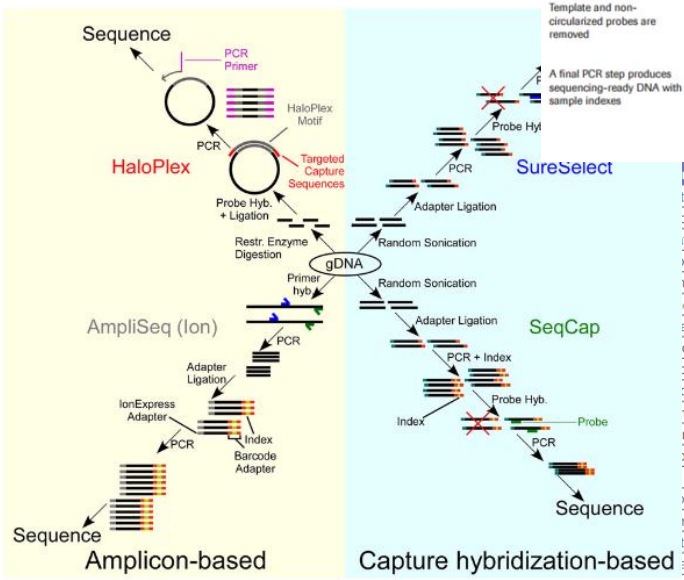
751 ген из базы HGMD (2018) не представлен ни в одном экзоме



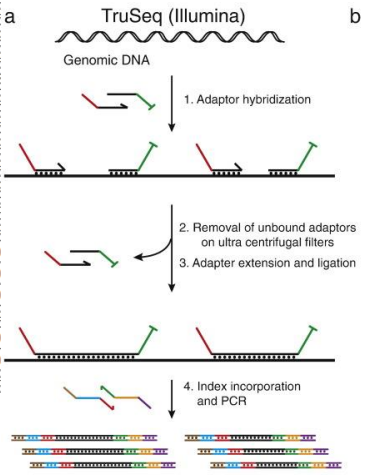


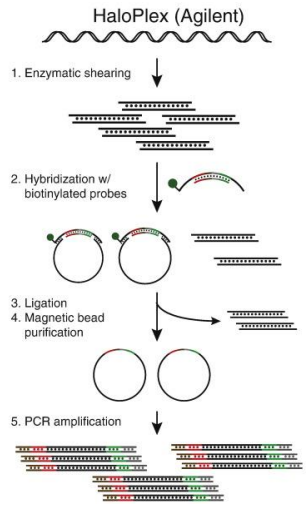




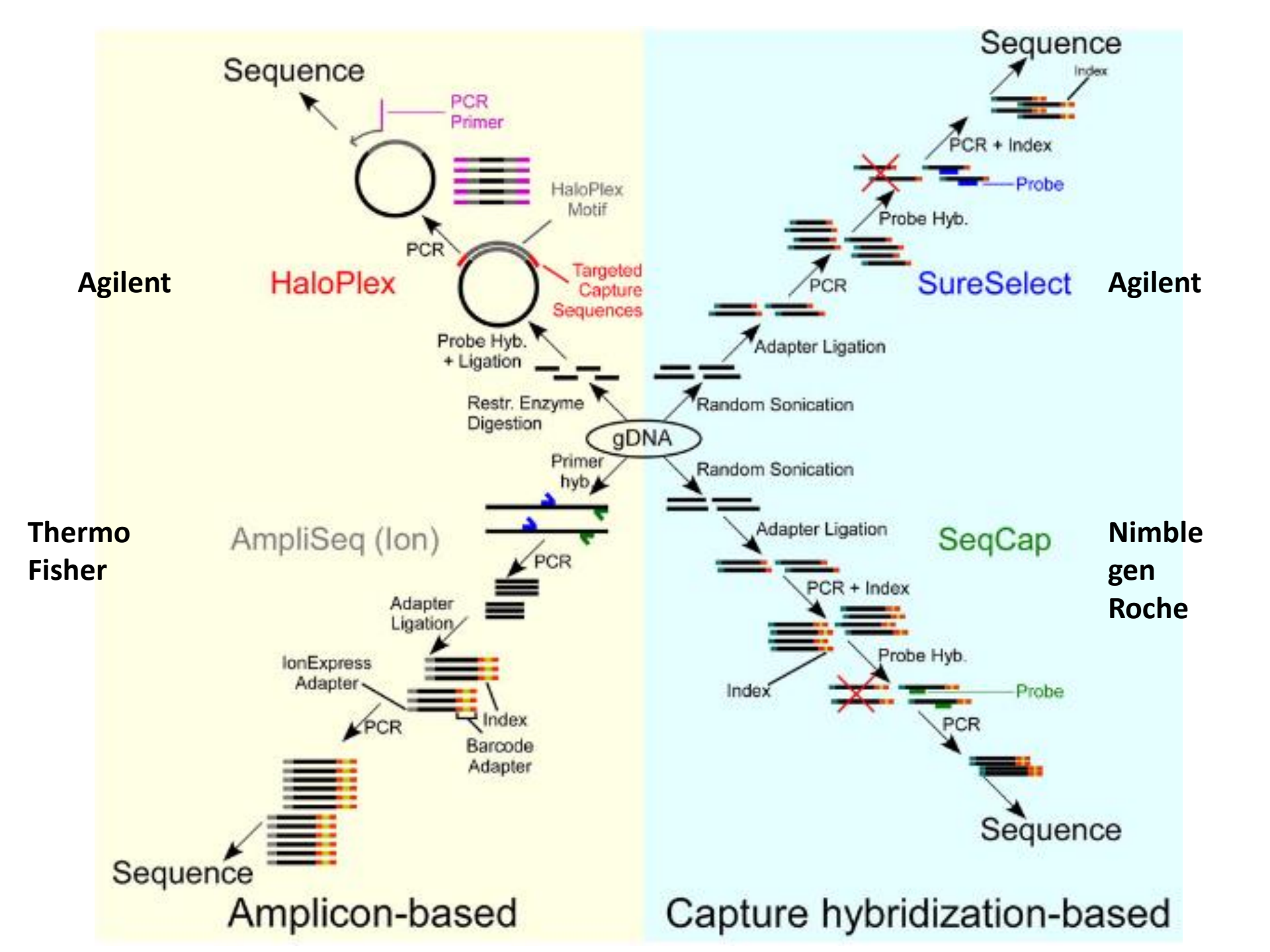


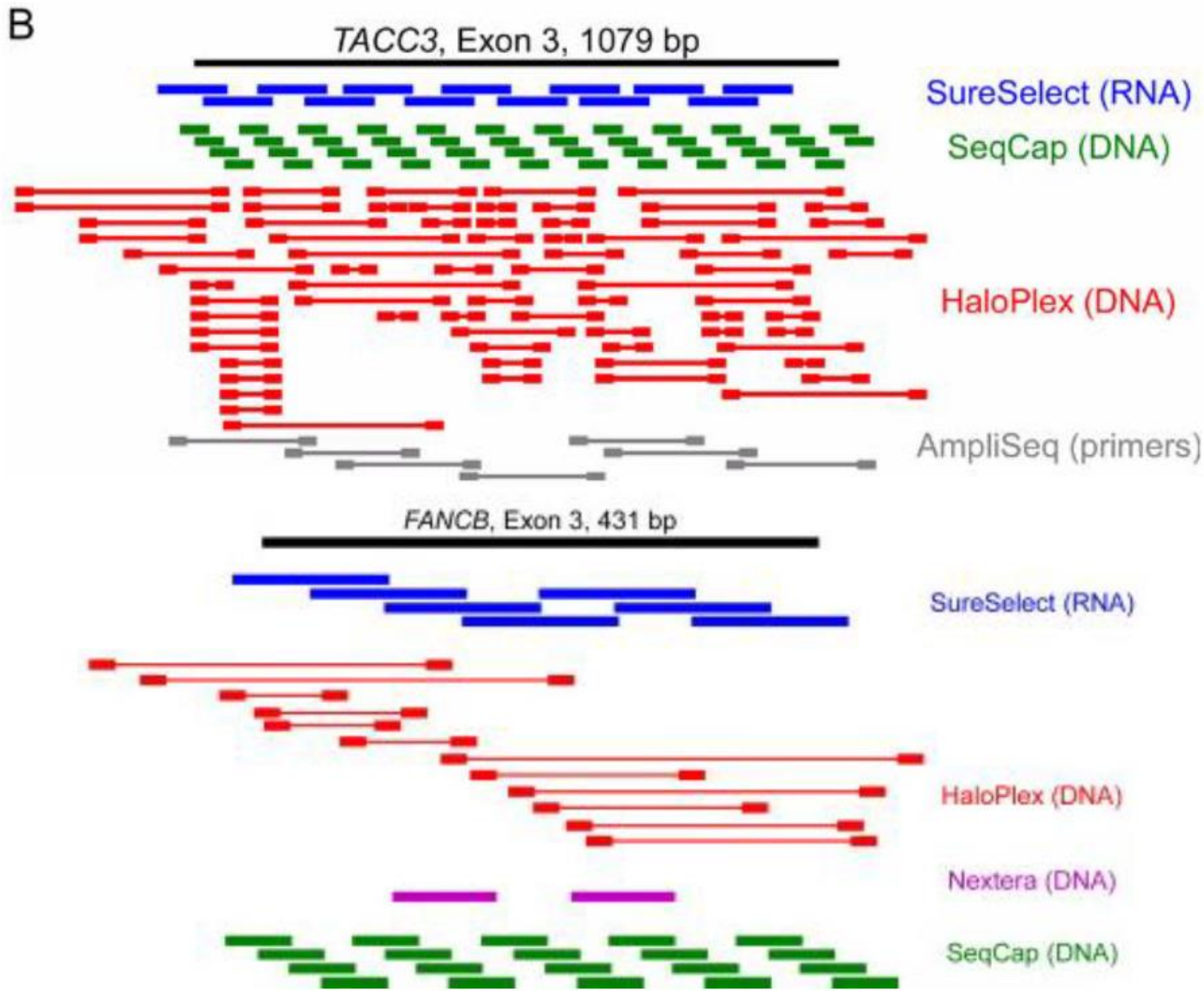
mpliSeq panel	covered	coverage	Panel uniformity
ted Disease Panel	325	97%	91%
ovascular Research Panel	404	99%	90%
tology Research Panel	394	99%	94%
logical Research Panel	751	99%	92%
almic Research Panel	316	99%	89%
ess Research Panel v2	128	99%	94%
atology Research Panel v2	222	99%	96%	96%
orphia-Dysplasia Research Panel v2	519	99%	96%	95%
rine Research Panel v2	340	99%	93%	96%
ointestinal Research Panel v2	194	99
Errors of Metabolism Research Panel v2	594	99
ry Immune Deficiency Research Panel v2	264	99
nary Research Panel v2	131	98
Research Panel	155	99
sy Research Panel	386	99
n Research Panel v2	236	99
ted Cancer Research Panel	134	99
ac Arrythmias and Cardiomyopathy Research Panel	92	99
Hearing Loss Research Panel v1	63	96
Dementia Research Gene Panel	17	99
Noonan Research Panel	14	10
TP53 Research Panel	1	10
BRCA1 & *BRCA2* Research Panel	2	10
Ovarian Cancer Research Panel	41	99
CFTR Research Panel	1	10
Pharmacogenomics Research Panel	40	10
Exome RDY Panel	19,072	96



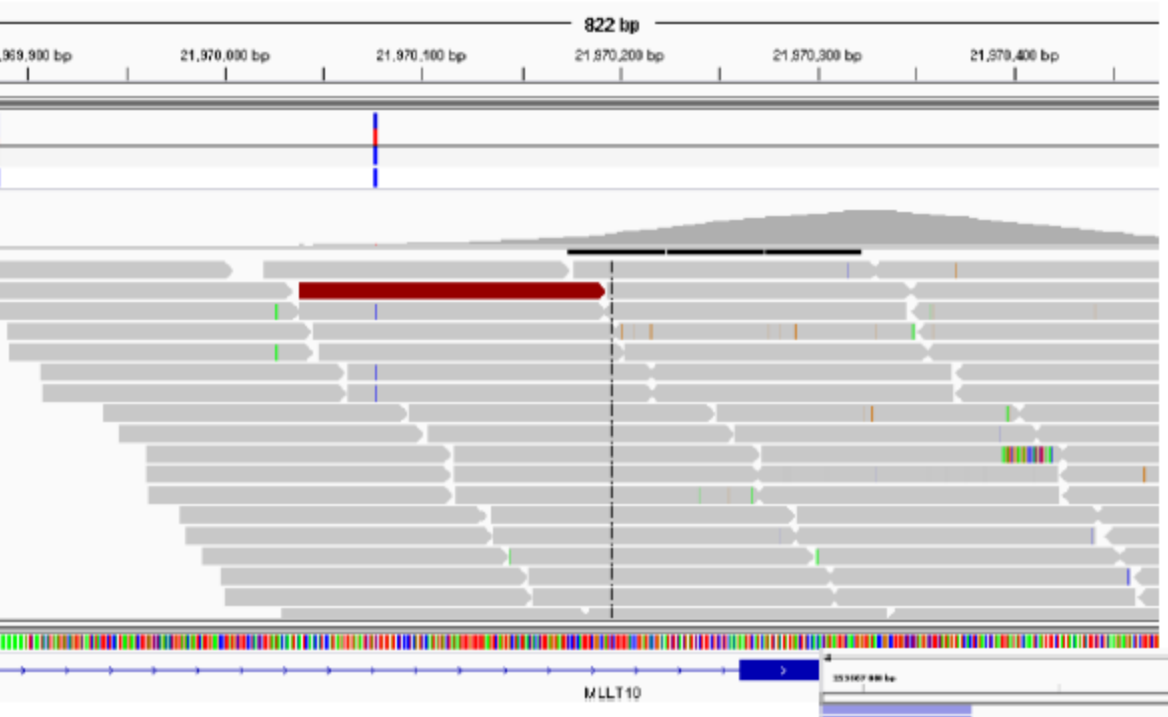


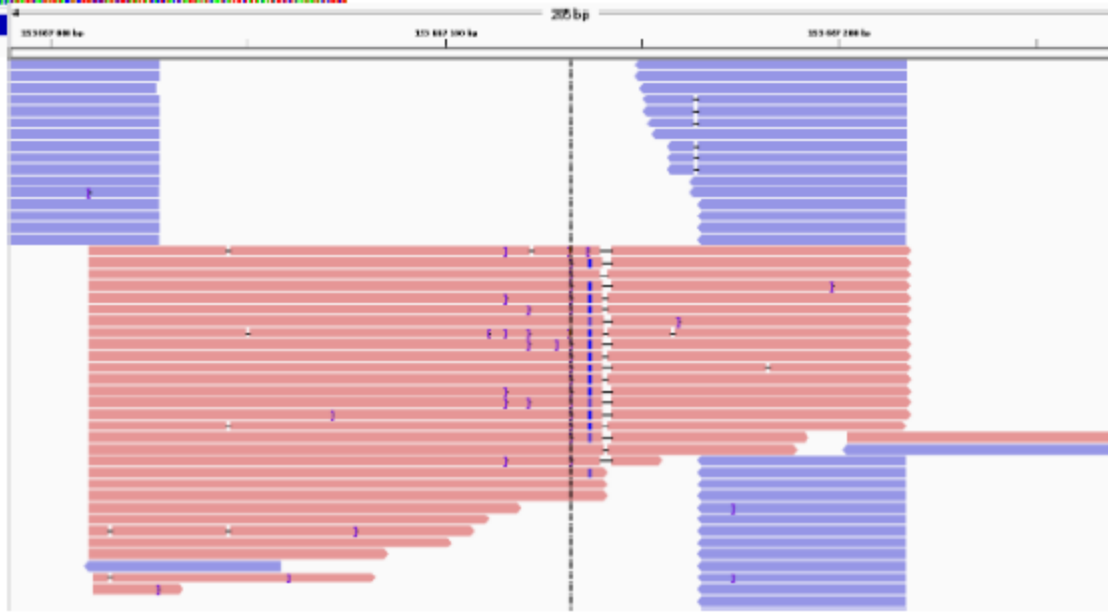
Библиотеки панелей генов

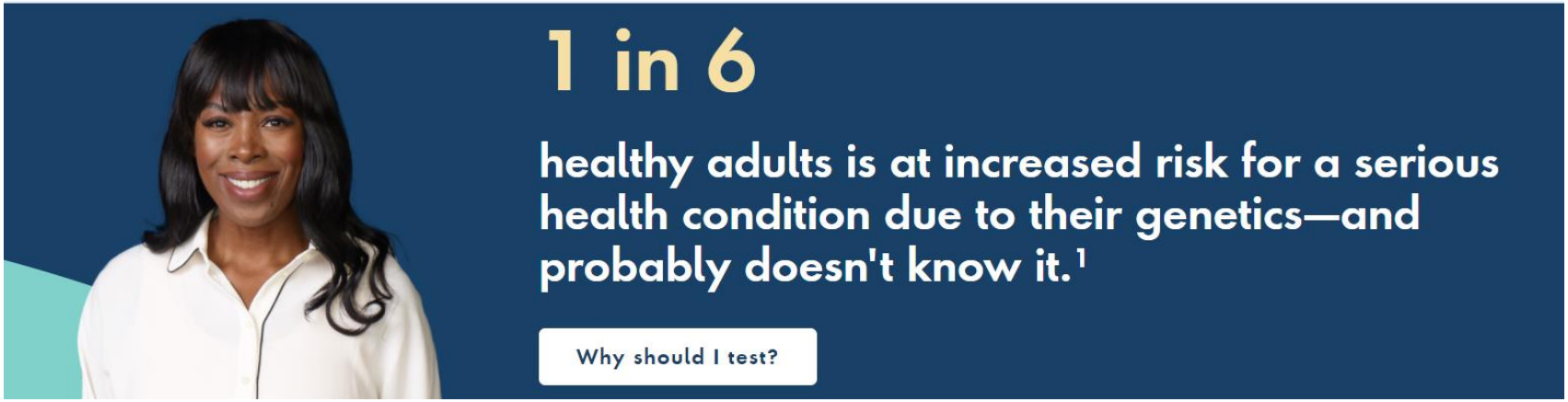

Agilent
HaloPlex
Sequence
PCR Primer
HaloPlex Motif
PCR
Targeted Capture Sequences
Probe Hyb. + Ligation
Restr. Enzyme Digestion
gDNA
Primer hyb
AmpliSeq (Ion)
PCR
Adapter Ligation
IonExpress Adapter
PCR
Index
Barcode Adapter
Sequence
Amplicon-based
Thermo Fisher
Random Sonication
Adapter Ligation
PCR
Probe Hyb.
PCR + Index
Probe
Index
Sequence
SureSelect
SeqCap
Capture hybridization-based
Agilent
Nimble gen Roche


B
TACC3, Exon 3, 1079 bp
SureSelect (RNA)
SeqCap (DNA)
HaloPlex (DNA)
AmpliSeq (primers)
FANCB, Exon 3, 431 bp
SureSelect (RNA)
HaloPlex (DNA)
Nextera (DNA)
SeqCap (DNA)

Амплификация или гибридизация?







Find the right test for you

Click on a test below to start your order.

Invitae Cancer Screen

$250

Looks at 61 genes associated with common cancers, including breast, ovarian, and prostate cancer.

Invitae Cardio Screen

$250

Looks at more than 75 genes to assess your risk of developing an inherited form of heart disease, including hereditary high cholesterol and blood pressure.

Invitae Genetic Health Screen

$350

Looks at up to 147 genes, including all the genes in the Cancer and Cardio Screens—as well as a few actionable inherited conditions.

7

Загрузим данные в кабинет

Мы загрузим результаты теста
оповещение на электронную

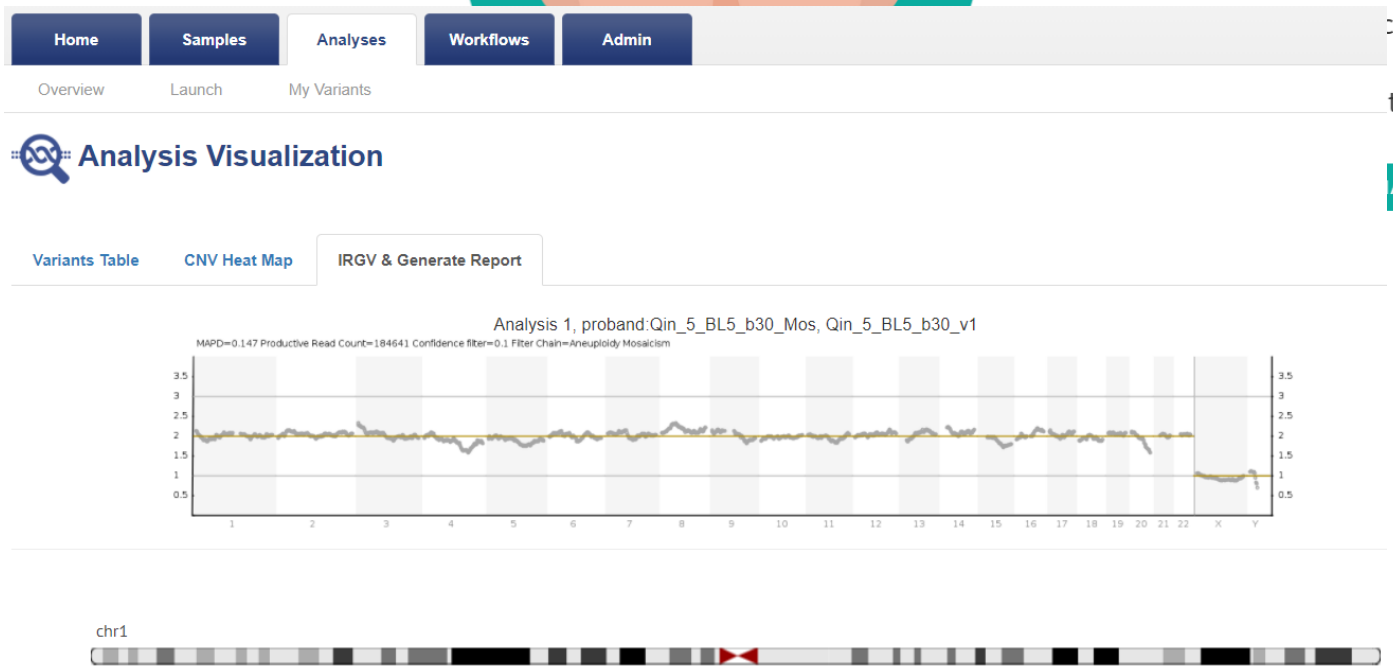
Здоровье ⟶

Питание

Meet Gia, our virtual genetic information assistant

clinical chatbot, Gia facilitates
with patients, including intake of
tomatic delivery of results.

A



Софт для работы с данными NGS

1. Платформо-зависимый
    1. Оценить качество секвенирования
    2. Анализировать результат, и, прежде всего, аннотировать
2. Предметно(задаче)-ориентированный
    1. Аннотация
    2. Приоритетизация и фильтрация
    3. Интепретация

От панелей к геному

- Полный экзом
- «Клинический» экзом
- Небольшие панели
- Онкопанели

- Геном
  - 600 $ -> 100 $
  - CNV
  - Любые перестройки
  - Идеальное покрытие
  - Любые мутации

Проблемы и ограничения при анализе омных данных:

1. Allele dropout – при обогащении методом ПЦР, гетерозиготная мутация, например, рядом с праймером, праймер не работает => ложный гомозиготный немутированный вариант, а мутацию потеряем (или наоборот).

2. Пробелы в покрытии - все наборы для экзомного обогащения не поднимают полностью те участки, которые должны поднять. Мутации в таком участке теряем.

3. Неисчерпывающие списки генов в наборах обогащения. «Клинический экзом», в несколько тысяч генов, но интересные гены, потенциально связанные с заболеванием могут быть в "клинический экзом" не включены.

4. Ошибки пробоподготовки, заметные на QC: сбитый GC-content, плохое качество прочтений

5. Повторы и псевдогены

6. Strand bias: разные варианты на прямых и обратных прочтениях и вообще представленность цепей

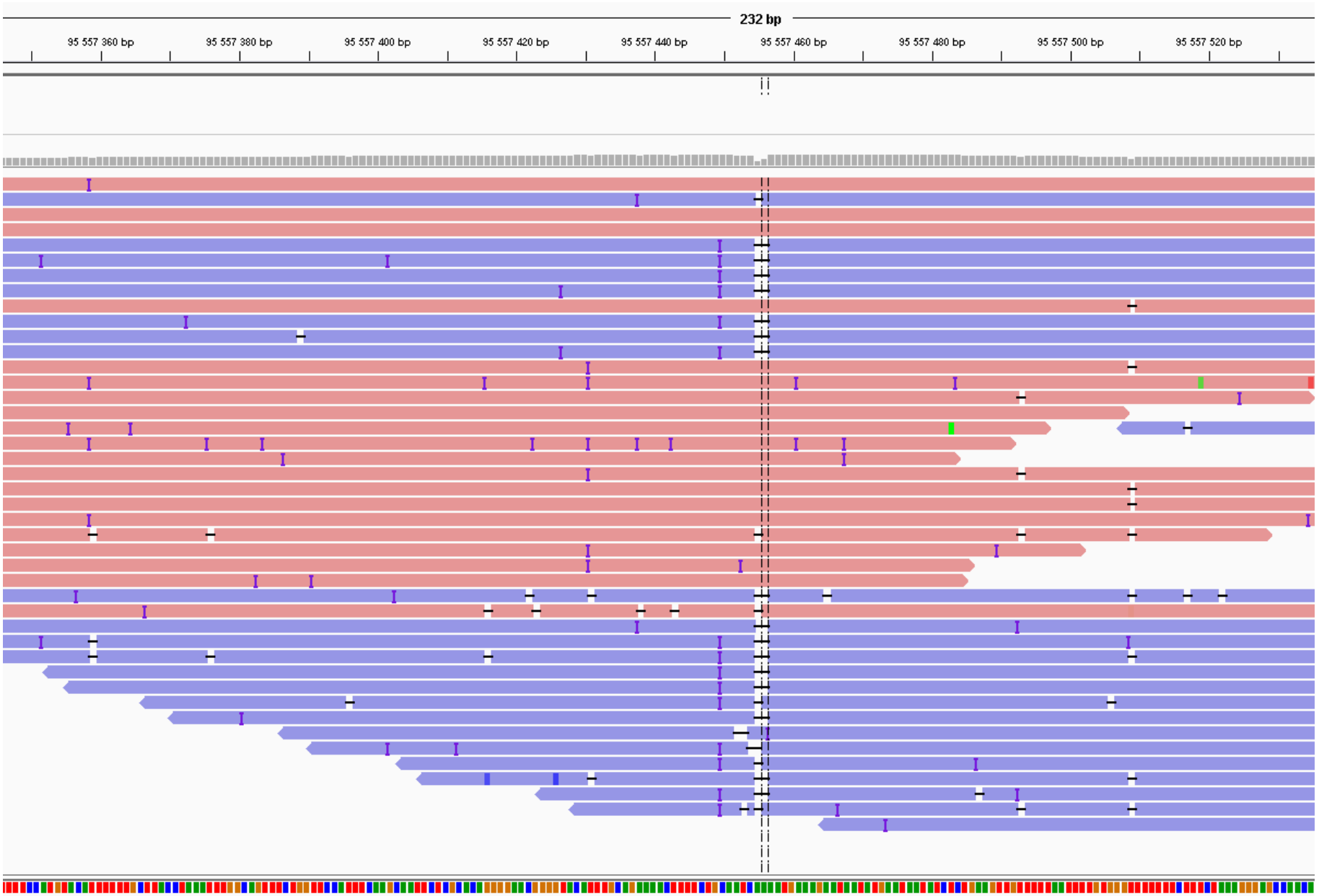
7. Ошибки на гомополимерах (Thermosifher) и падение качества прочтения к концу прочтения и сложность чтения GC (Illumina, BGI)

8. CNV и тандемные повторы

Strand bias


100 bp
106 857 310 bp
106 857 320 bp
106 857 330 bp
106 857 340 bp
106 857 350 bp
106 857 360 bp
106 857 370 bp
106 857 380 bp
106 857 390 bp

Strand bias


232 bp
95 557 360 bp
95 557 380 bp
95 557 400 bp
95 557 420 bp
95 557 440 bp
95 557 460 bp
95 557 480 bp
95 557 500 bp
95 557 520 bp