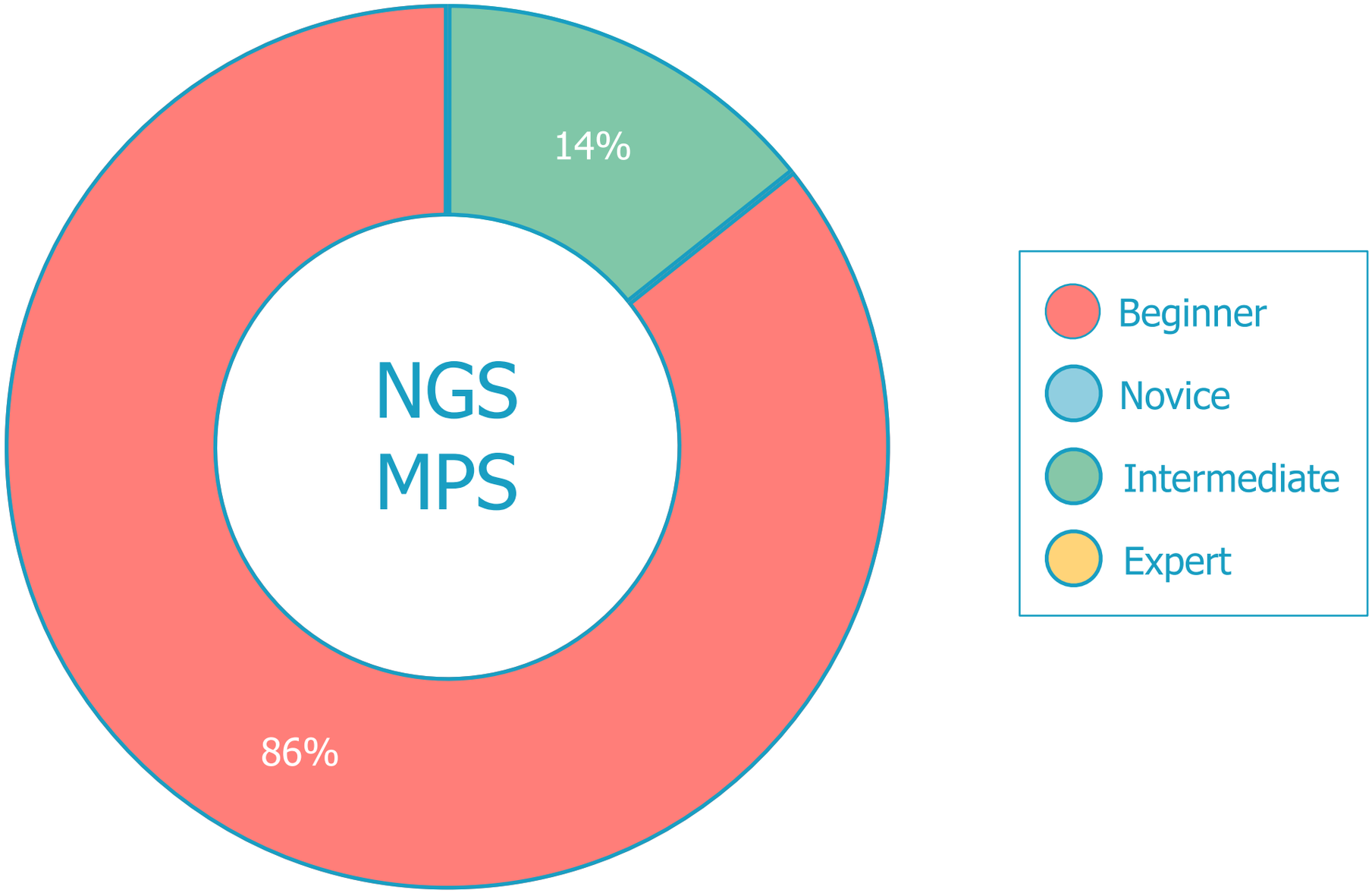
GDC
Genetic Diversity Centre Zurich

Genetic Diversity: Analysis

Next Generation Sequencing

Monday, June 16, 2025

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich


14%
NGS
MPS
86%
Beginner
Novice
Intermediate
Expert

Genetic Diversity Centre Zurich

First Generation Sequencing

Sanger Sequencing

Second Generation Sequencing

Next Generation Sequencing (NGS)

Third Generation Sequencing

Single Molecule Sequencing

Fourth Generation Sequencing

??? Sequencing

GDC Genetic Diversity Centre Zurich

First Generation Sequencing
Sanger Sequencing

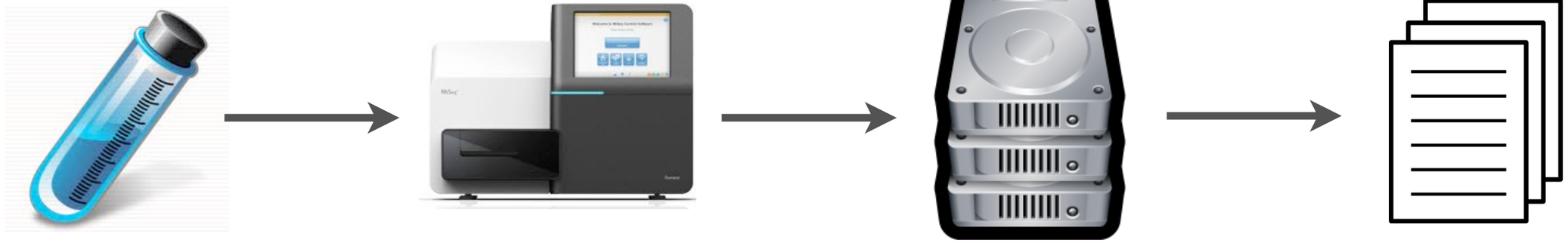
Second Generation Sequencing
Next Generation Sequencing (NGS)

Third Generation Sequencing
Single Molecule Sequencing

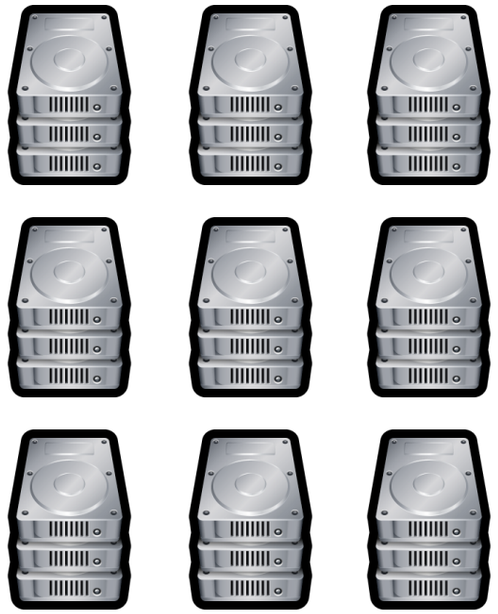
Fourth Generation Sequencing

Massive Parallel Sequencing

Next Generation Sequencing **Hype**



Next Generation Sequencing **Reality**




BREAKTHROUGH
AT THE HUMAN GENOME PROJECT
EUREKA! THE NEXT ONE'S A "G"!
SNAK-O
HIMELBLAU


nature
the human genome
Science
THE HUMAN GENOME

Actually, that's the coffee machine...this is the next-gen sequencer.

The **First Law of Technology** says we invariably **overestimate** the **short-term impact** of a truly transformational discovery, while **underestimating** its **longer-term effects**.

https://www.scientificamerican.com/

"The Human Genome Project has had a considerable effect on research and society more generally, but questions about what a human genome reference is today and how it can benefit human health remain to be resolved."

Source: Rood and Regev (2021) The legacy of the Human Genome Project. Science Vol 373, Issue 6562, pp. 1442-1443.

Short-Term Impact (Overestimated)

- **Excitement about Personalised Medicine:** There was immense hype about how sequencing the human genome would lead to rapid, personalised treatments for diseases based on a person's genetic profile.
- **Gene Editing Breakthroughs:** Technologies like CRISPR were expected to immediately revolutionise medicine by fixing genetic disorders in real-time.
- **Direct Applications in Healthcare:** We imagined widespread and immediate use of genetic information to prevent and treat complex diseases (like cancer or Alzheimer's).

Long-Term Effects (Underestimated)

- **Ethical Implications of Gene Editing:** The long-term effects of technologies like CRISPR are still being worked out. Potential risks include **germline editing** (changing genes in embryos), which could have unintended consequences for future generations. The ethical dilemma about "designer babies" and the possibility of exacerbating social inequality are long-term issues that were not fully anticipated in the excitement of gene editing breakthroughs.
- **Genetic Data Privacy and Security:** With the advent of genomic sequencing, there are increasing concerns about how genetic data is stored and who owns it. This data could be misused for **discrimination** (e.g., by employers or insurance companies), leading to privacy issues we didn’t foresee in the early days.
- **Socioeconomic Divide:** Access to genomic technologies may widen the gap between those who can afford personalised medicine and those who cannot. This could lead to **genetic inequality** where only the wealthy benefit from cutting-edge medical advancements.
- **Unintended Consequences in Evolution:** If gene-editing technologies are applied to populations on a large scale, it could have unforeseen effects on human evolution. We could unintentionally alter aspects of our species' genetic makeup in ways that are difficult to predict, and that may not be beneficial in the long run.
- **Ecological Impact of Genetic Modifications:** In the broader context of genetic engineering, modifying organisms (including humans) could have unpredictable consequences for the environment and ecosystems. For instance, releasing genetically modified organisms into the wild could disrupt natural biodiversity or lead to ecological imbalances.

Long-Term Risks:

- **Genetic Discrimination:** As genetic data becomes more available, the risk of individuals being discriminated against based on their genetic makeup increases. Insurance companies, employers, or even governments could use genetic information to make decisions about people's lives.
- **Ethical Dilemmas in Germline Editing:** Editing genes in embryos raises issues about what constitutes "acceptable" genetic changes and the possible creation of genetic "classes" based on desirable traits.
- **Loss of Genetic Diversity:** If genetic interventions become widespread, there's a potential risk of reducing the genetic diversity of the human population, which could make us more vulnerable to diseases or environmental changes in the future.
- **Unpredictable Medical Outcomes:** While gene therapy and editing offer great promise, they could have **long-term unintended effects**—such as unknown side effects or long-term health problems that manifest in the future generations who undergo genetic modifications.

Sequencing technologies have undergone a revolution - costs have plummeted, throughput has soared and it's now routine to generate terabytes of genomic data in a matter of days. **But despite this explosion in data, the quality and biological insight of genome studies hasn't kept pace.** Many assemblies remain fragmented or poorly annotated, especially in non-model organisms. Challenges such as high error rates, uneven coverage and bioinformatic bottlenecks persist.

Perhaps more critically, the scientific approach itself has changed. Where biology was once driven by clear questions and hypothesis testing, it's increasingly driven by data-first exploration. While this has opened up new avenues, it has also led to unfocused 'fishing expeditions', reproducibility problems and the temptation to value patterns over mechanisms. In other words, **we're drowning in data but still thirsty for understanding**.

GDC Genetic Diversity Centre Zurich

(A-)Typical Workflow

Genetic Diversity Centre Zurich

??? Question

Design

DNA or RNA (or Protein)

Library Prep

Sequencing

Genetic Diversity Centre Zurich

MPS data (fastq, fastq.gz, bam, pod5)

Data Set(s)

Data Conversion

Data Storage (Archive)

Quality Check (QC)

Filtering / Trimming

GDC Genetic Diversity Centre Zurich

Error Correction

Merge Read Pairs

(High(er)-Quality) Reads

Assembly

Contig

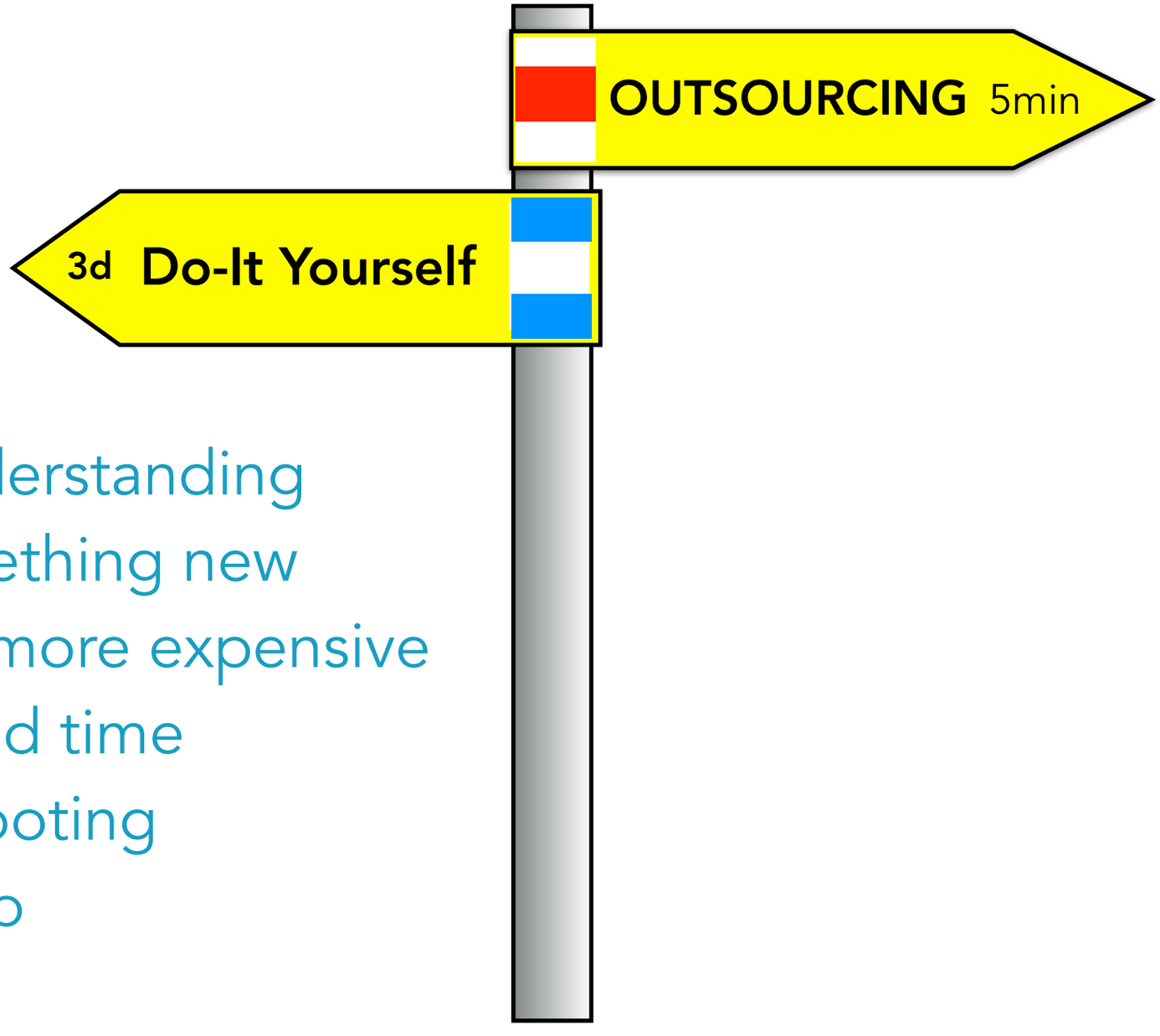
Mapping

Reference

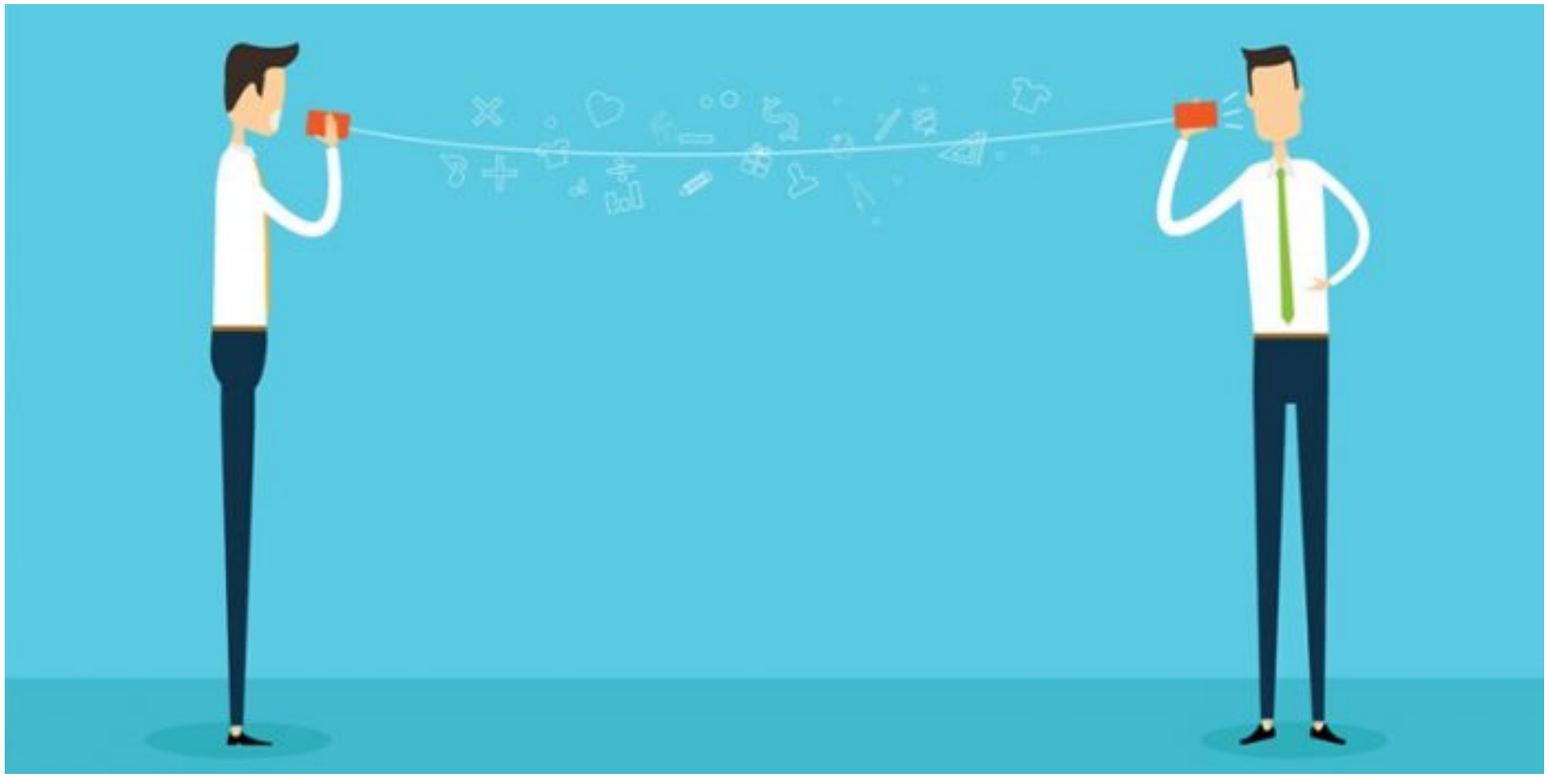
Clustering

Scaffolding / Gap-closing

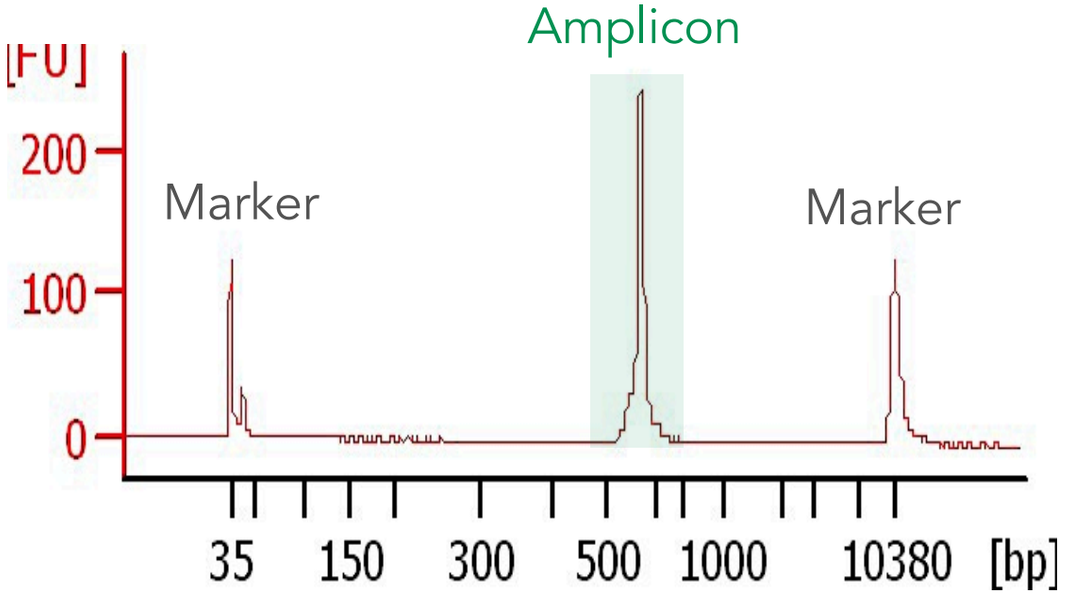
Data Analysis

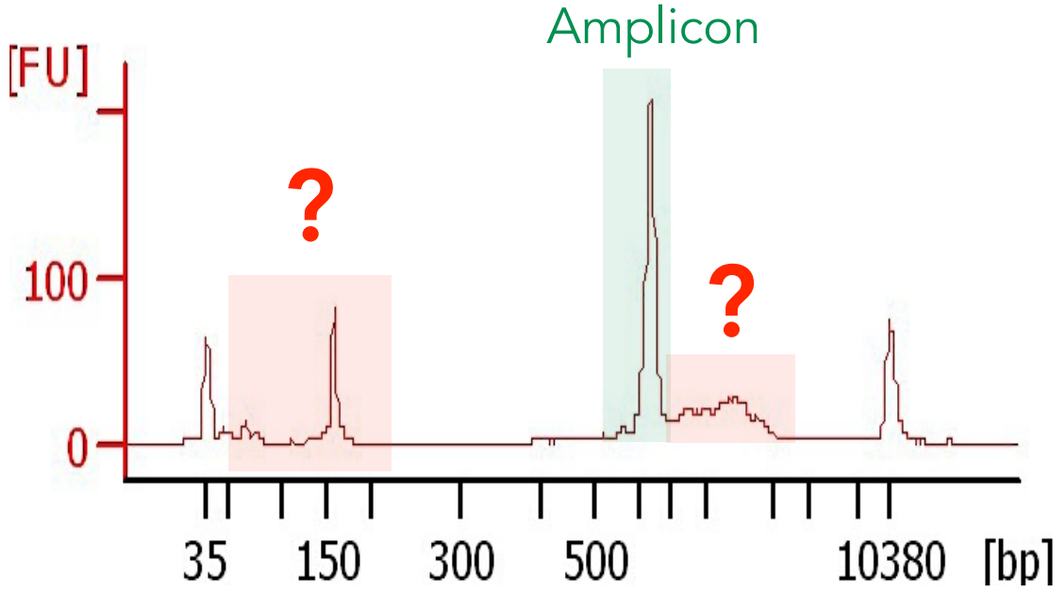


- better understanding
- learn something new
- often not more expensive
- turn-around time
- troubleshooting
- data mixup



Example: Fragment Length Analysis

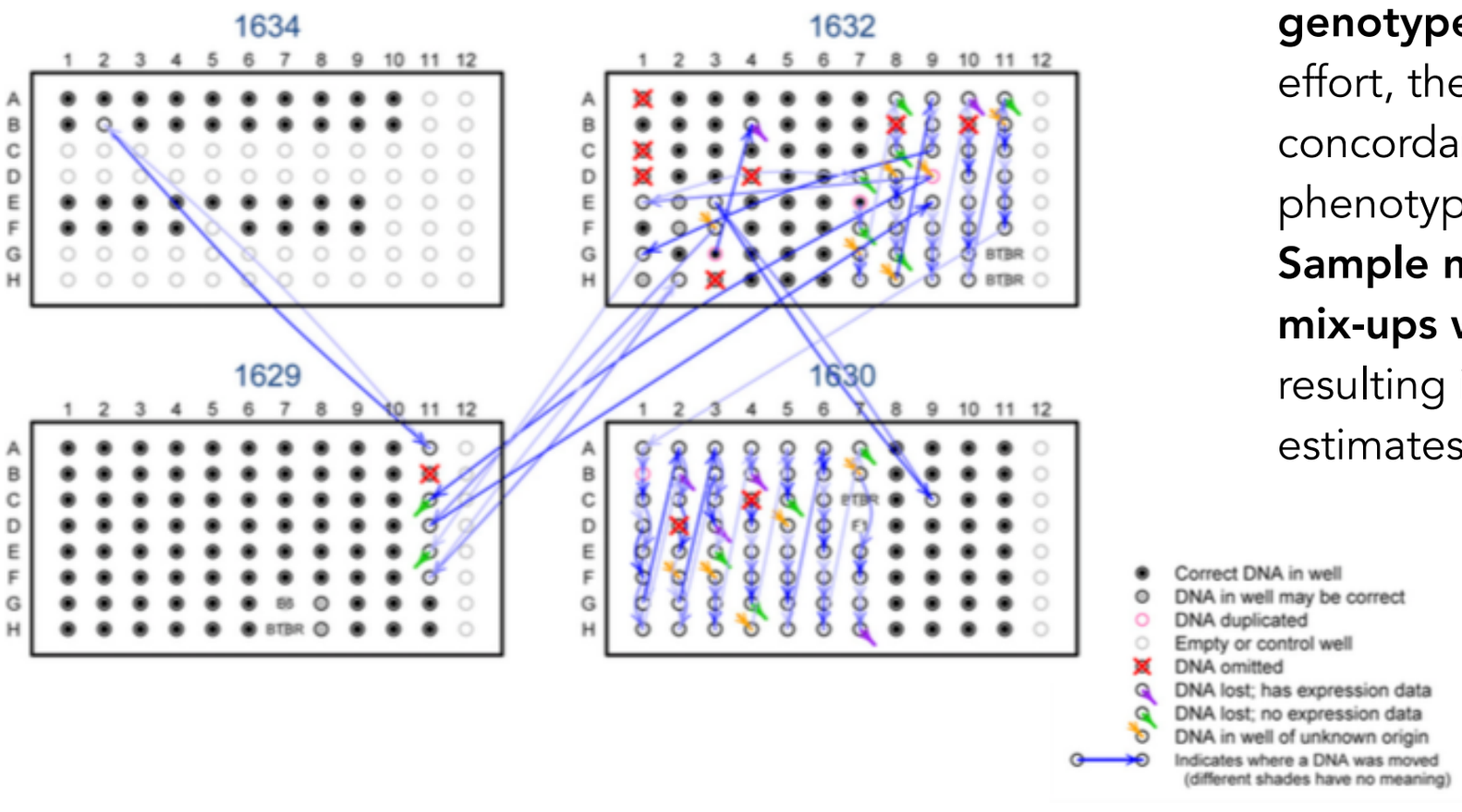




GDC Genetic Diversity Centre Zurich

Identification and Correction of Sample Mix-Ups in Expression Genetic Data: A Case Study

Karl W. Broman,*,[2] Mark P. Keller,† Aimee Teo Broman,* Christina Kendziorski,* Brian S. Yandell,‡,§ Śaunak Sen,**,[1] and Alan D. Attie†
*Department of Biostatistics and Medical Informatics, †Department of Biochemistry, ‡Department of Statistics, and §Department of Horticulture, University of Wisconsin, Madison, Wisconsin 53706, and **Department of Epidemiology and Biostatistics, University of California, San Francisco, California 94107



"To map the genetic loci influencing a complex phenotype, one seeks to establish an **association between genotype and phenotype**. In such an effort, the maintenance of the concordance between genotyped and phenotyped samples and data is critical. **Sample mislabeling and other sample mix-ups will weaken associations**, resulting in reduced power and biased estimates of locus effects."

Sequencing Technologies

Genetic Diversity Centre Zurich

1 Sanger (chain termination)

2 Roche 454 Pyrosequencing (pyrophosphate)

Ion Torrent (semiconductor technology)

Illumina (fluorescent - sequencing by synthesis (SBS))

Singular Genomics (fluorescent - rapid SBS)

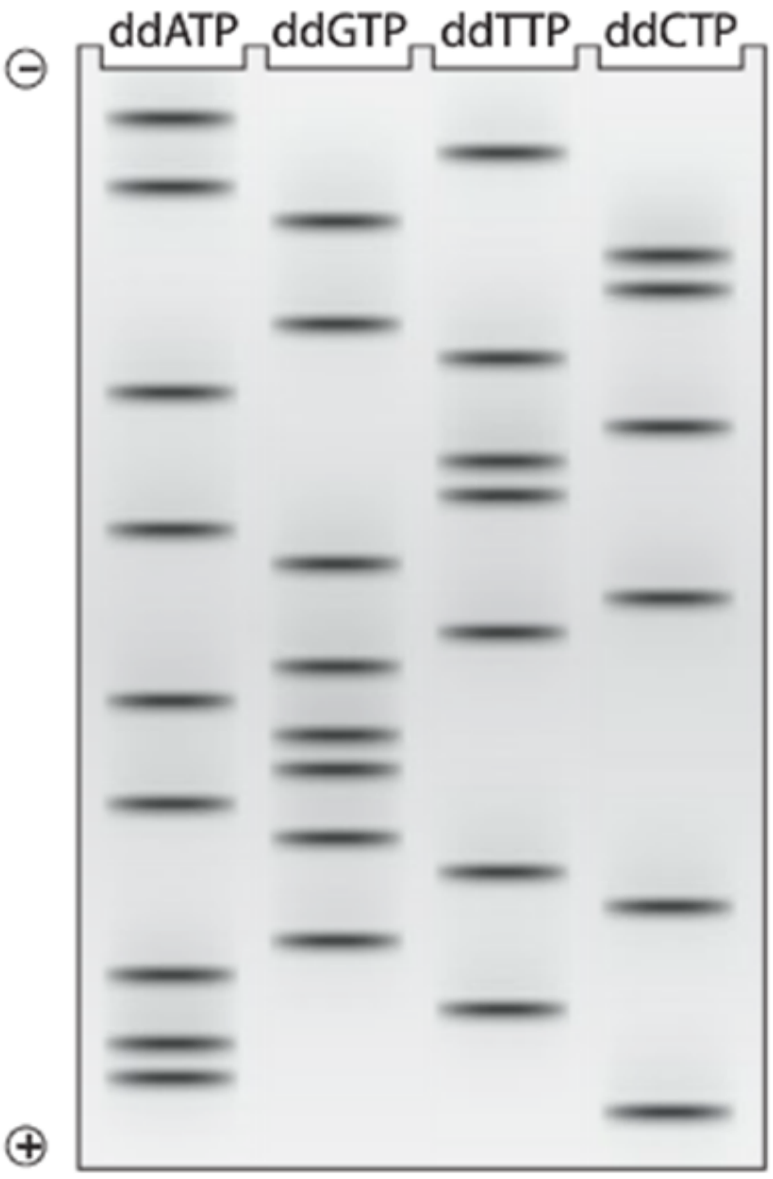
Aviti (fluorescent - sequencing by avidity)

3 **PacBio** (fluorophore)

Nanopore (ionic current)

Helicos - SeqLL (fluorescent)

4 Bionano - Saphyr (third-generation optical mapping)



The Nobel Prize in Chemistry 1980

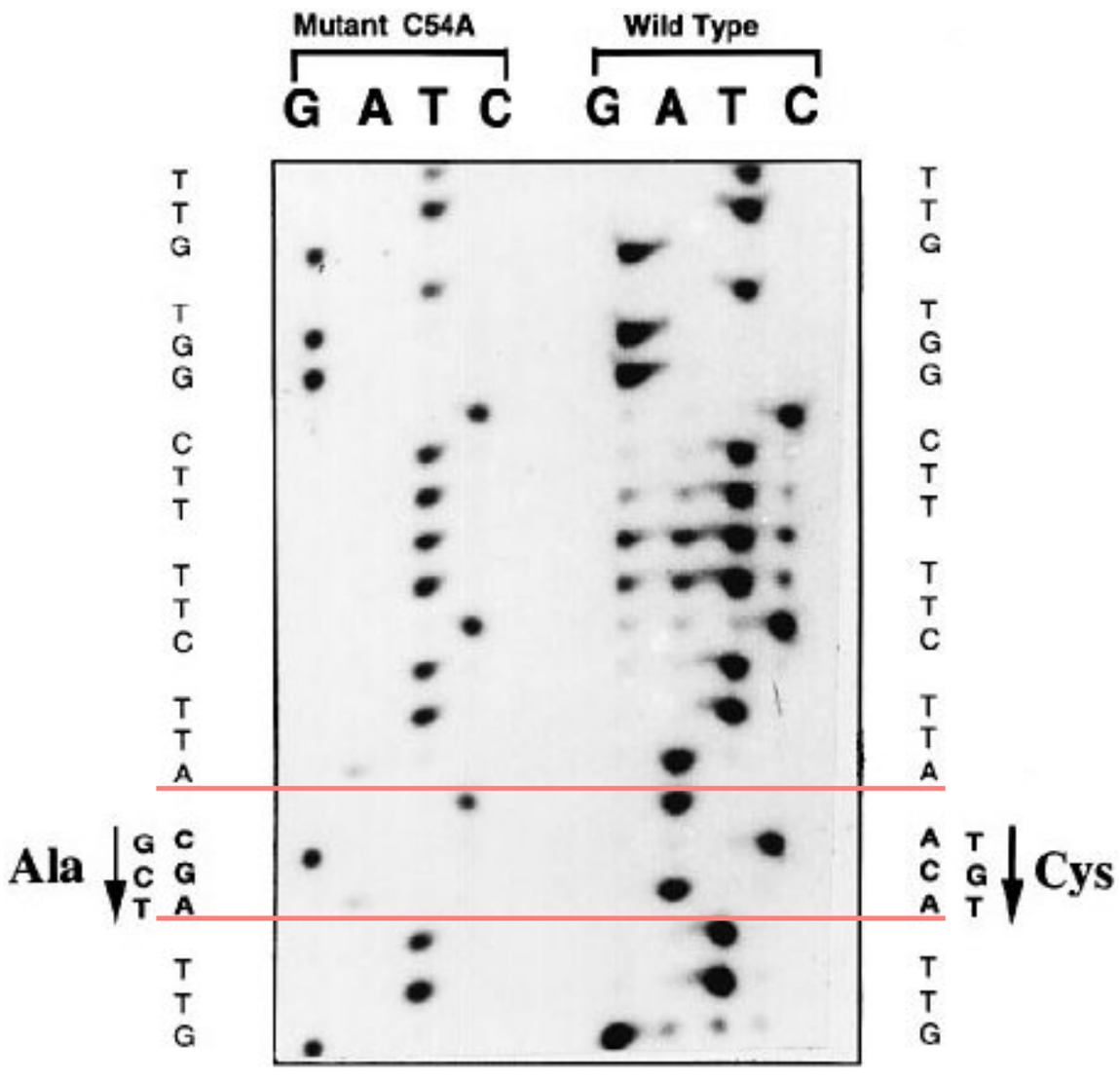
Paul Berg, Walter Gilbert, Frederick Sanger

Paul Berg

Walter Gilbert

Frederick Sanger

The Nobel Prize in Chemistry 1980 was divided, one half awarded to Paul Berg *"for his fundamental studies of the biochemistry of nucleic acids, with particular regard to recombinant-DNA"*, the other half jointly to Walter Gilbert and Frederick Sanger *"for their contributions concerning the determination of base sequences in nucleic acids"*.



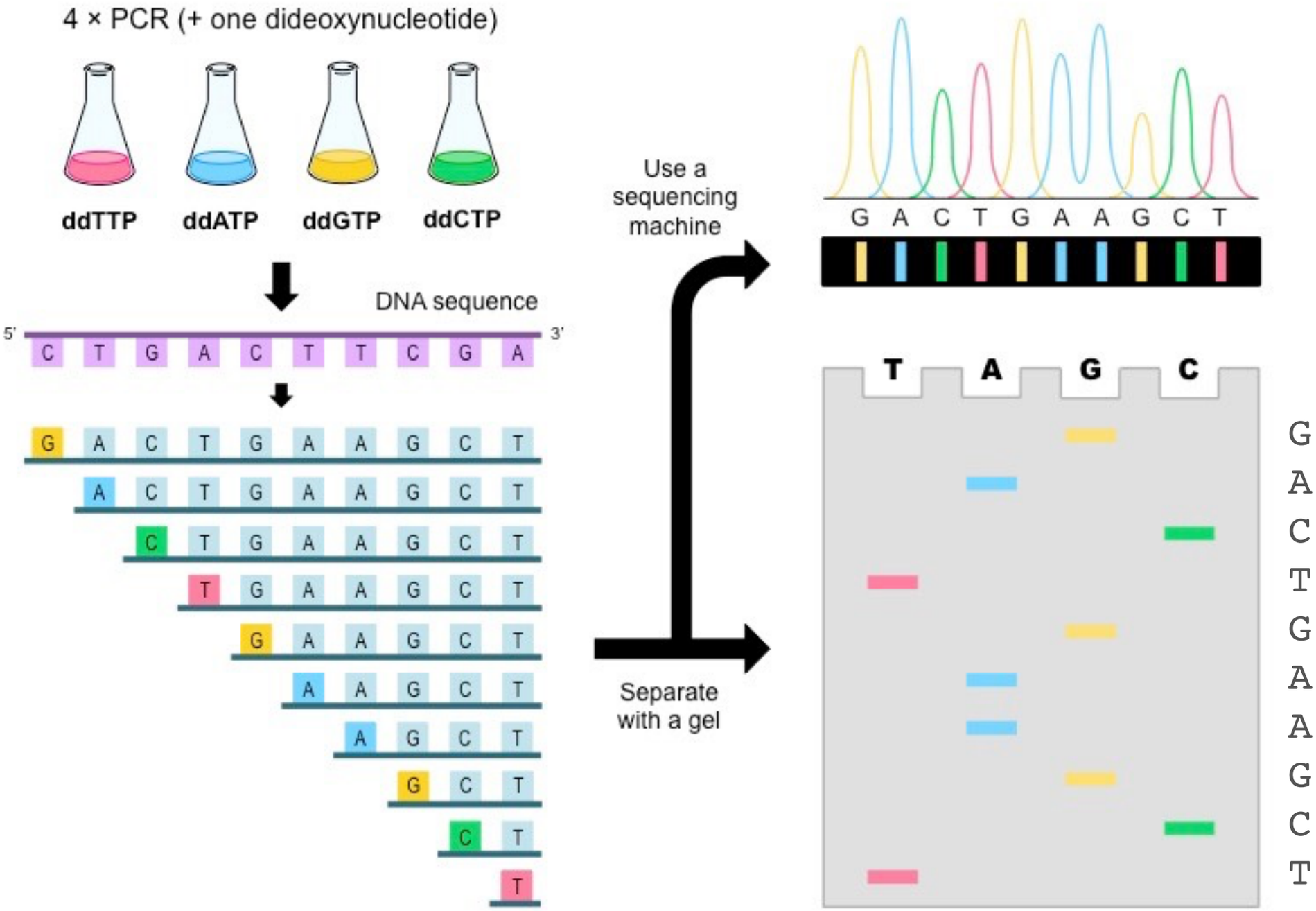
Protein engineering of BamHI restriction endonuclease: replacement of Cys54 by Ala enhances catalytic activity

November 1998 · Protein Engineering 11(10):931-5

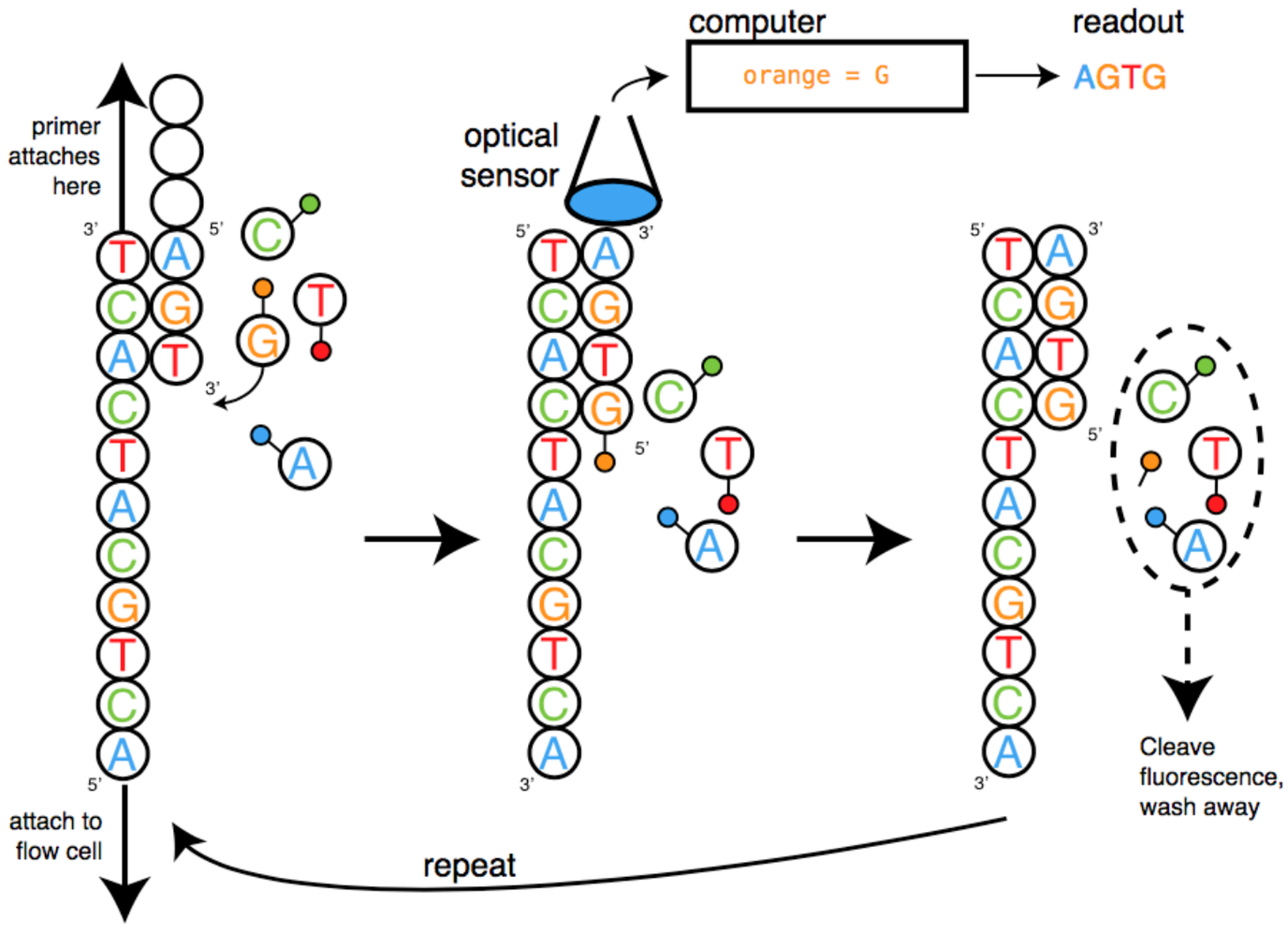
DOI: 10.1093/protein/11.10.931

Source · PubMed

Autoradiogram of DNA sequencing gel covering the mutation region. The mutation was confirmed by full sequencing of the gene.


4 × PCR (+ one dideoxynucleotide)
ddTTP
ddATP
ddGTP
ddCTP
DNA sequence
5'
3'
C T G A C T T C G A
G A C T G A A G C T
A C T G A A G C T
C T G A A G C T
T G A A G C T
G A A G C T
A A G C T
A G C T
G C T
C T
T
Use a sequencing machine
Separate with a gel
G A C T G A A G C T
T A G C
G A C T G A A G C T

Sequencing by Synthesis (fluorescent)



Sequencing by Synthesis. dNTP fluorescence is translated to a base call.

Illumina Systems

From genome-wide discovery to targeted validation and screening									
	Sequencing Systems							Sequencing and Arrays	Arrays
Instrument	NovaSeq™ X and NovaSeq X Plus Systems	NovaSeq 6000 and NovaSeq 6000Dx Systems	NextSeq™ 1000 and NextSeq 2000 Systems	MiSeq™ i100 and MiSeq i100 Plus* Systems	MiSeq and MiSeqDx Systems	MiniSeq™ System	iSeq™ 100 System	NextSeq 550 and NextSeq 550Dx Systems	iScan™ System
Technology	XLEAP-SBS™ chemistry DRAGEN onboard	Sequencing by synthesis (SBS) chemistry	XLEAP-SBS chemistry DRAGEN onboard	XLEAP-SBS chemistry DRAGEN onboard	SBS chemistry	SBS chemistry	SBS chemistry CMOS technology	SBS chemistry Infinium™ BeadChip	Infinium BeadChip
Features and Applications	Vast application breadth at the highest throughput, enabling the most data-intensive methods at production scale	**NovaSeq 6000 System** – A broad range of applications, enabling data-intensive methods **NovaSeq 6000Dx System** – High-throughput, FDA-regulated instrument for IVD testing and clinical research	Small whole-genome, exome, transcriptome, targeted panels, multiomic single-cell, spatial, and metagenomic sequencing	Fastest, simplest sequencing Small whole-genome, targeted gene panel sequencing, metagenomics, transcriptome profiling	**MiSeq System** – Small genome, amplicon, targeted gene panel sequencing **MiSeqDx System** – FDA-regulated instrument for IVD testing and clinical research	Targeted DNA and RNA sequencing	Targeted, bacterial, and viral sequencing	**NextSeq 550 System** – **Small whole-genome**, exome, transcriptome sequencing, and CNV analysis **NextSeq 550Dx System** – FDA-regulated instrument for IVD testing and clinical research	SNP and whole-genome genotyping, CNV analysis, gene regulation, epigenetic analysis, gene expression analysis, and cytogenetic analysis

*MiSeq i100 Plus System available 2H 2025

GDC Genetic Diversity Centre Zurich

MGI	 Sequencers +	 Sequencers +	 Sequencers +	 Sequencers +
Product Model	DNBSEQ-T7	DNBSEQ-G400	DNBSEQ-G400* For HotMPS Only	DNBSEQ-G50
Features	Ultra-high Throughput	Adaptive	Adaptive	Effective
Applications	Whole Genome Sequencing, Deep Exome Sequencing, Transcriptome Sequencing, and Targeted Panel Projects.	WGS, WES, Transcriptome sequencing, etc.	WGS, WES, Transcriptome sequencing, etc.	Small whole genome sequencing, targeted DNA/RNA panels, low-pass whole genome sequencing
Flow Cell Type	FC	FCL & FCS	FCL	FCL & FCS
Lane/Flow Cell++	1 lane	2 or 4 lanes	4 lanes	1 lane
Operation Mode	Ultra-high Throughput	High Throughput	High Throughput	Medium Throughput
Max. Throughput / RUN	6TB	1440GB	720GB	150GB
Effective Reads / Flow Cell	5000M	1500-1800M	1500-1800M	500M / 100M
Average run time	24~30 hours for PE150 sequencing	FCS: 13~37 hours FCL: 14~109 hours	15.5-50.5 hours	9~40 hours
Min. Read Length	PE100	SE50	SE50	SE50
Max. Read Length	PE150	SE400/PE200	PE100	PE150

MGI Tech is the manufacturing sister of China's largest genome sequencing company, BGI Genomics.

GDC Genetic Diversity Centre Zurich

Singular Genomics (fluorescent - rapid SBS)

G4

A highly versatile benchtop sequencer targeting applications where accuracy, speed, flexibility, and scale matter most

Key milestones

- Launched 2021
- First units expected ship in Q2 2022

PX

Integrated in situ platform for multiomic analysis in single cells and tissues

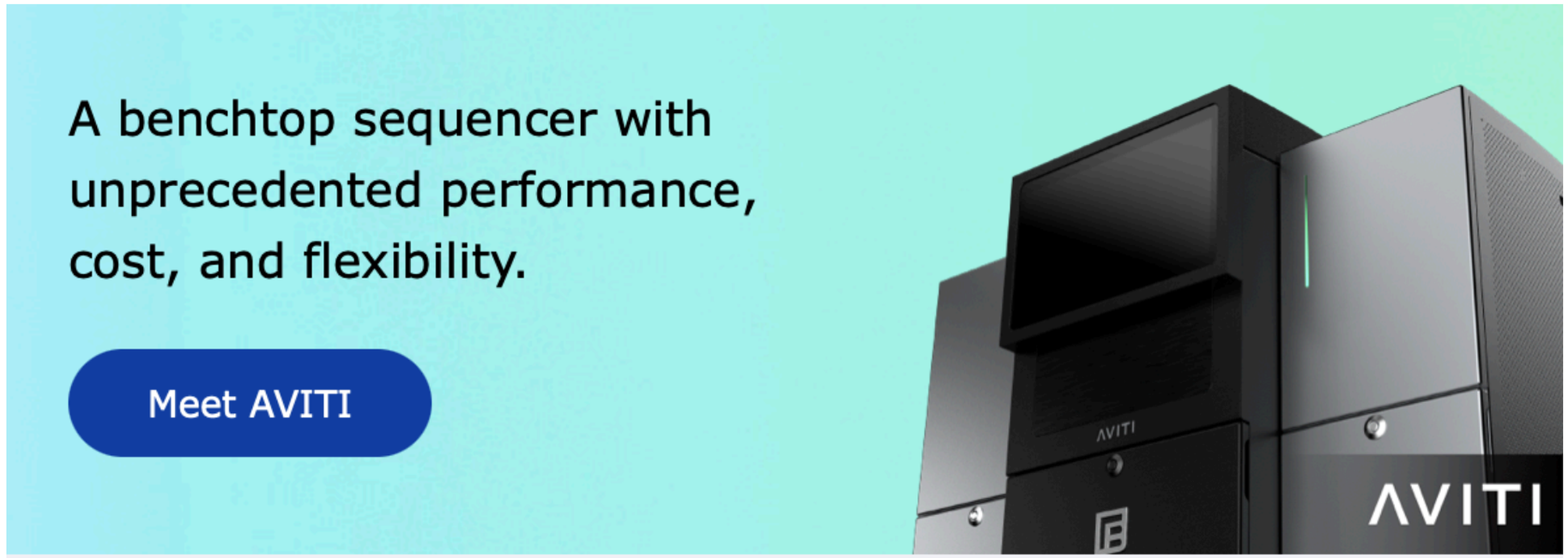
PX image is for illustrative purposes only

Key milestones

- Early access program to begin 2022
- Commercial launch expected in 2023

The **G4X™ Spatial Sequencer** by Singular Genomics is an advanced platform that combines high-throughput sequencing with spatial transcriptomics. This enables researchers to map gene expression directly within the context of tissue. Based on the G4™ Sequencing platform, it provides high-resolution spatial analysis, flexible throughput and rapid processing times. Designed to support research in cancer, neuroscience, and developmental biology, where spatial information is essential for understanding complex biological systems, the G4X is compatible with various tissue types. By combining sequencing power with spatial insight, the G4X is opening up new frontiers in functional genomics.

Element Biosciences (fluorescent - sequencing by avidity)



Performance

%Q30 > 90 at 2x150

1B reads/flow cell*

600 Gb+ output/run

Flexibility

Dual flow cells

Flexible start

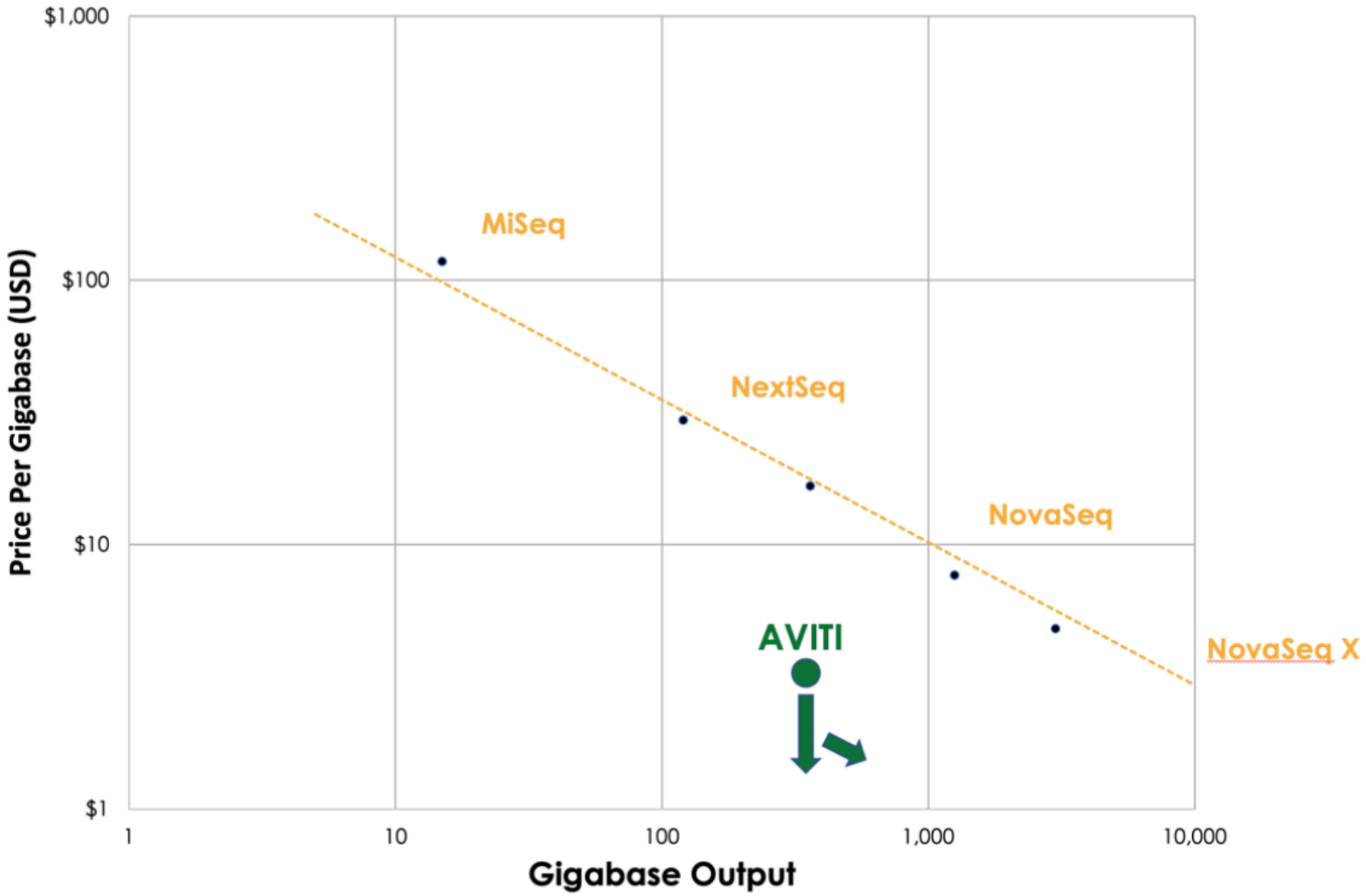
Tunable read throughput

Cost

$289K/instrument

$1680/300 cycles
$1080/150 cycles

Leasing/financing options available


$1,000
$100
$10
$1
Price Per Gigabase (USD)
1
10
100
1,000
10,000
Gigabase Output
MiSeq
NextSeq
NovaSeq
NovaSeq X
AVITI

GDC Genetic Diversity Centre Zurich

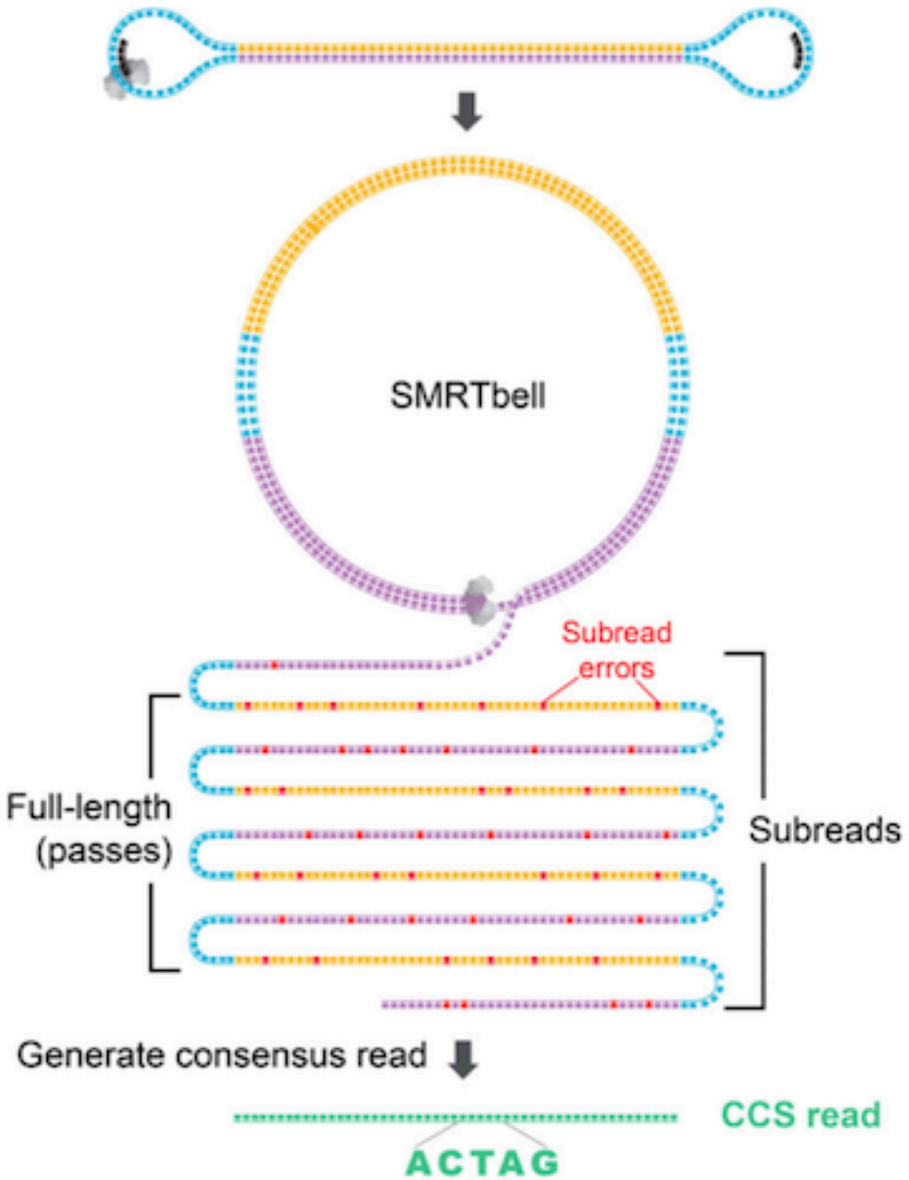
PacBio

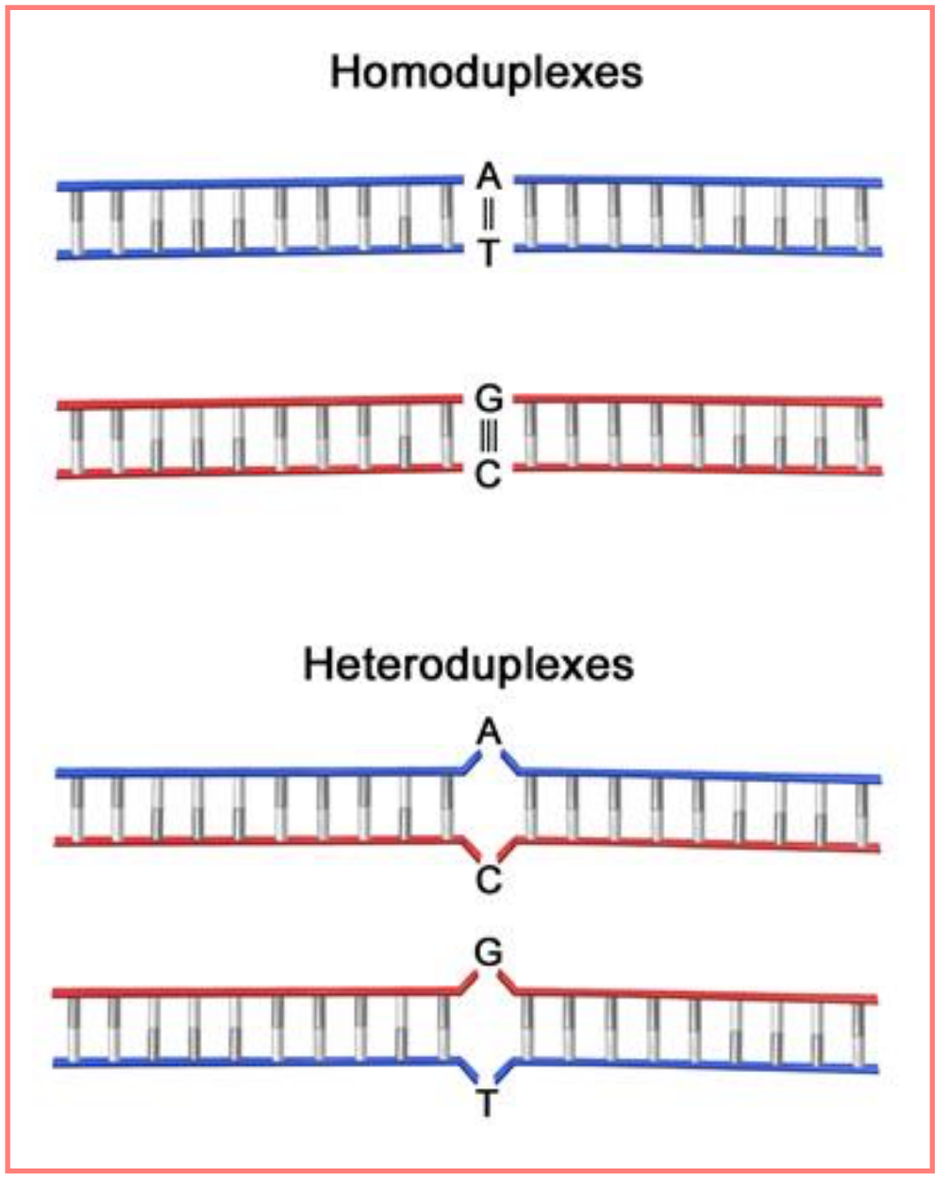
ENTERING THE ERA OF WORLD-CLASS NGS ACCURACY WITH PACBIO HIFI AND SBB TECHNOLOGY

VEGA
Low-Throughput
Long-Read

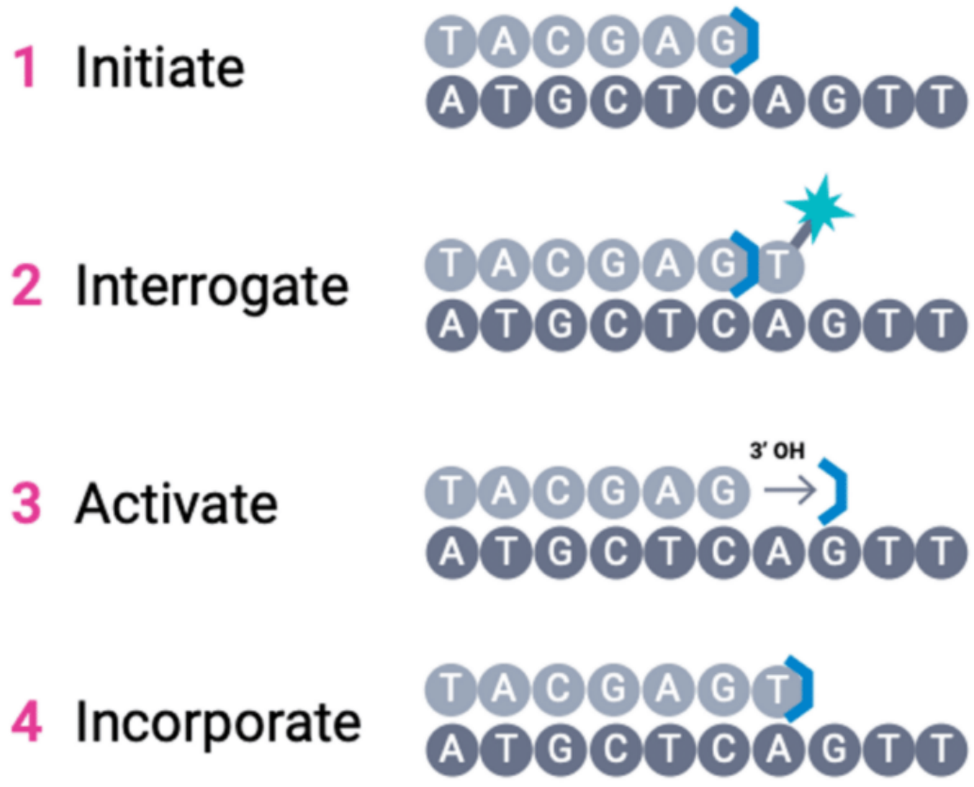
ONSO
Bench-Top
Short-Read

REVIO
High-Throughput
Long-Read


SMRTbell
Subread errors
Full-length (passes)
Subreads
Generate consensus read
CCS read
ACTAG

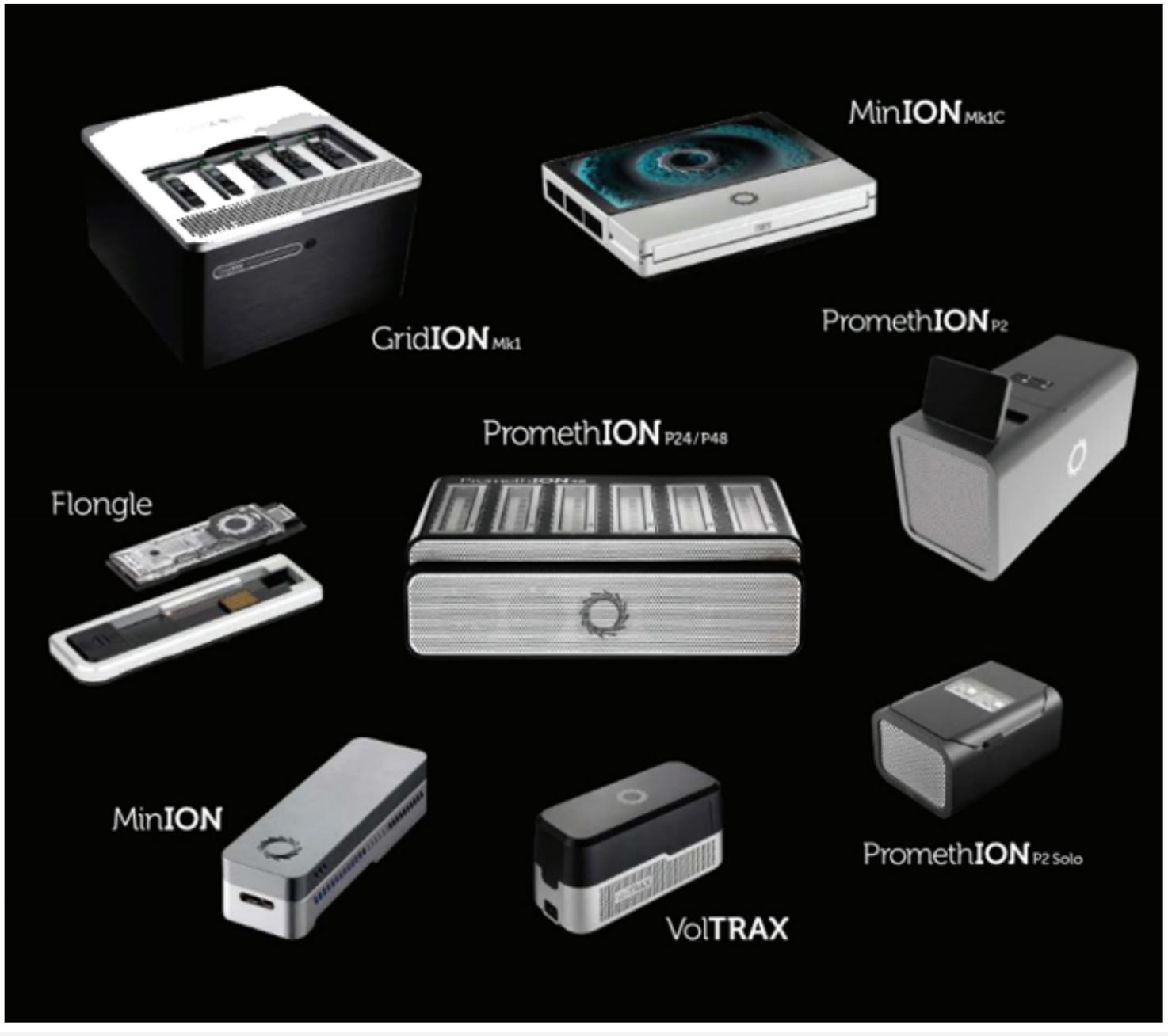

Homoduplexes
A
T
G
C
Heteroduplexes
A
C
G
T

Sequencing by binding (SBB)

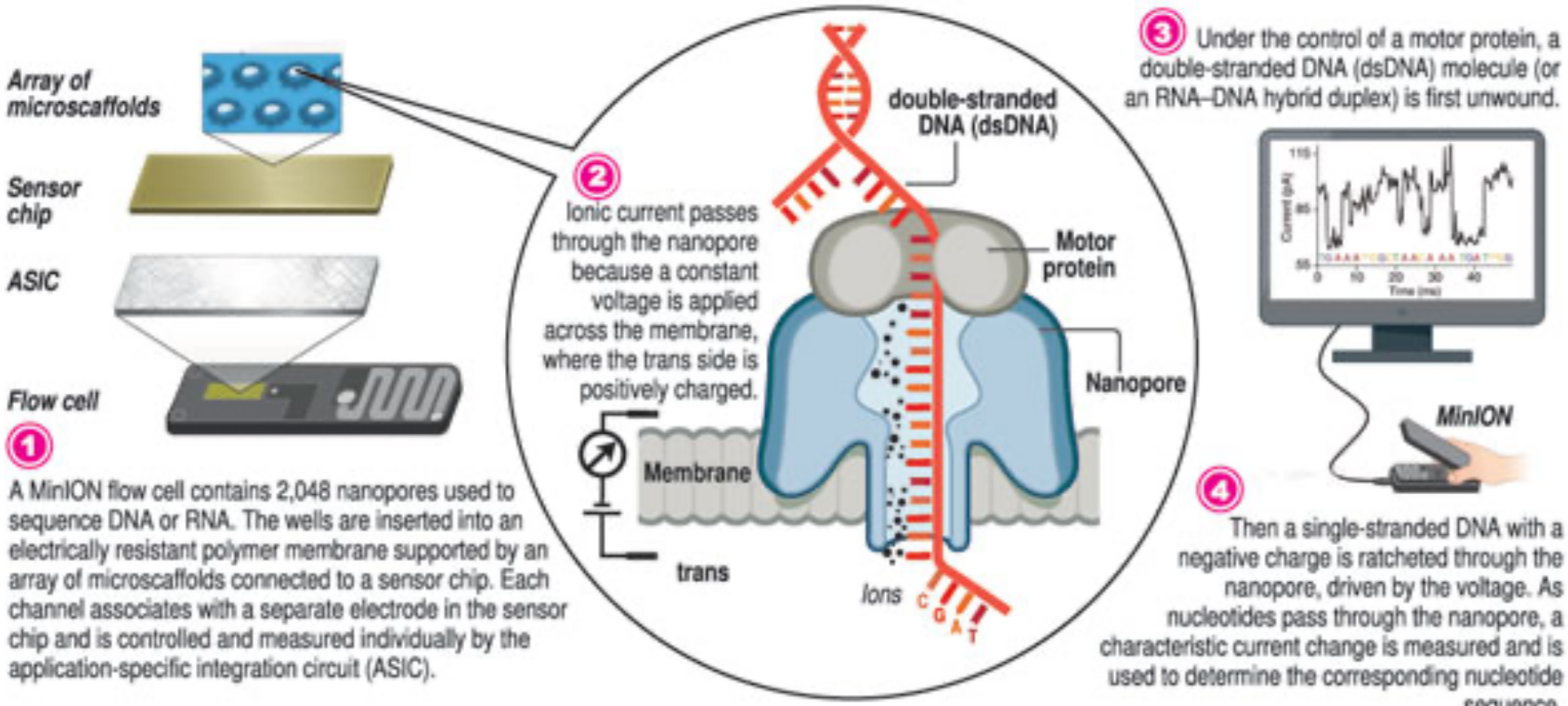


Genetic Diversity Centre Zurich

Oxford NANOPORE Technologies



https://www.nanoporetech.com



Source: The Sunday Times - News: Nanopore sequencing – rapid insights in real time (2022)


The Economist
Open-source medical devices
Dealing with nuclear waste
Burt Rutan, a maverick in flight
Technology Quarterly
December 1st 2012
The dream of the medical tricorder
the hand held devices inspiring medical add-ons for smartphones

Theranos was a Silicon Valley biotech startup founded by Elizabeth Holmes in 2003. It promised to revolutionise medicine by performing hundreds of blood tests from just a few drops of blood, all using a small, sleek machine - pictured here. It was billed as a real-world medical tricorder, seemingly fulfilling the dreams of science fiction.

But the technology never really worked. For years, the company hid the truth behind a wall of secrecy, aggressive legal tactics and charismatic marketing. Investors, the media and even powerful political and business figures bought into the vision - believing more in the narrative of disruption than the scientific evidence.

GDC Genetic Diversity Centre Zurich

Bionano

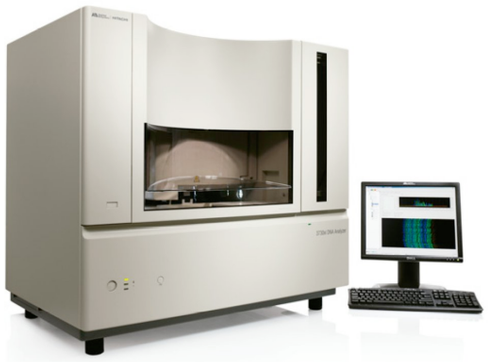
(optical mapping)

Optical mapping is a technique for constructing ordered, genome-wide, high-resolution restriction maps from single, stained molecules of DNA, called "optical maps".

Optical Sequencing

Optical sequencing is a single molecule DNA sequencing technique that follows sequence-by-synthesis and uses optical mapping technology. During synthesis, fluorochrome-labeled nucleotides are incorporated through the use of DNA polymerases and tracked by fluorescence microscopy.


PromethION




AVITI


NovaSeq 6000


DNBSEQ-T7

What is the best NGS platform?

The best platform for a particular application depends on several factors, such as project goals, sample type, sequencing depth, budget and bioinformatics support. Here are some common NGS platforms:

1. Illumina (e.g., NovaSeq, NextSeq, MiSeq): Illumina is currently the most widely used NGS platform. HiSeq instruments offer high-throughput sequencing, making them suitable for large-scale projects. MiSeq is a smaller benchtop sequencer that is more cost-effective for smaller projects or labs with lower sequencing needs.

2. Element Biosciences (Aviti System). Aviti is an advanced DNA sequencing platform. It is designed to provide high-quality, accurate, and cost-effective sequencing for various genomic applications.

3. BGI Genomics (MGISEQ/T7, DNBSEQ): BGI Genomics is a cost-effective and therefore attractive option for large-scale projects.

4. Pacific Biosciences (PacBio Sequel II): PacBio uses Single Molecule Real-Time (SMRT) sequencing and is advantageous for long read sequencing. It enables the sequencing of longer DNA fragments, facilitating the assembly of complex genomes and the detection of structural variation.

5. Oxford Nanopore Technologies (MinION, GridION, PromethION): ONT sequencing uses nanopore-based technology to provide long reads and real-time data analysis. It is portable and has been used for field applications such as rapid pathogen identification and monitoring.

It is important to evaluate the specific requirements of your project and consider factors such as read length, sequencing depth, accuracy, cost and data analysis needs when determining the best NGS platform to use. As the technology is advancing rapidly, it is advisable to consult the latest information and experts in the field to make an informed decision based on the most up-to-date information available.

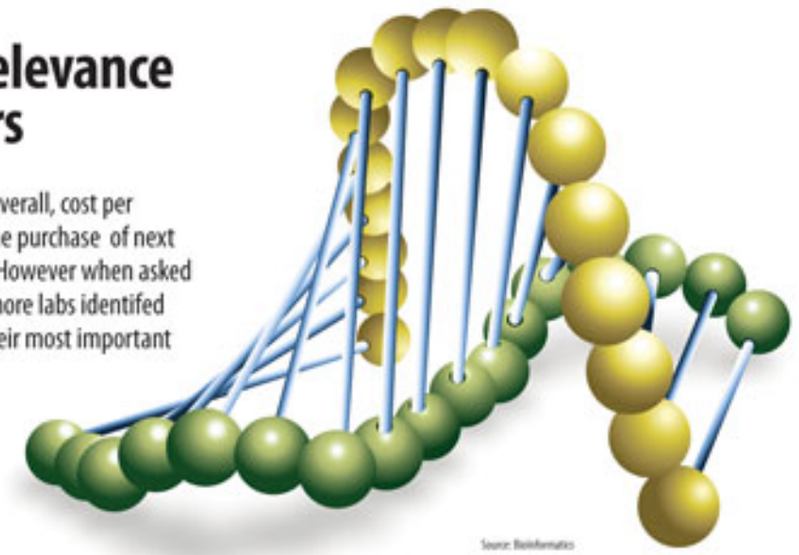
The all-in-one MPS platform does not exist (yet)!

- Research question
- Budget (including storage and analysis)
- Read / sequence length
- Number of reads / coverage
- Possible contaminants
- Quality and quantity of template
- Number of samples
- Availability

GEN Genetic Engineering & Biotechnology News

NGS: Cost and Relevance Are Key to Buyers

Recent market research shows that, overall, cost per base was the most cited concern in the purchase of next generation sequencing instruments. However when asked to identify their top three concerns, more labs identifed "Appropriate to My Application" as their most important criteria.



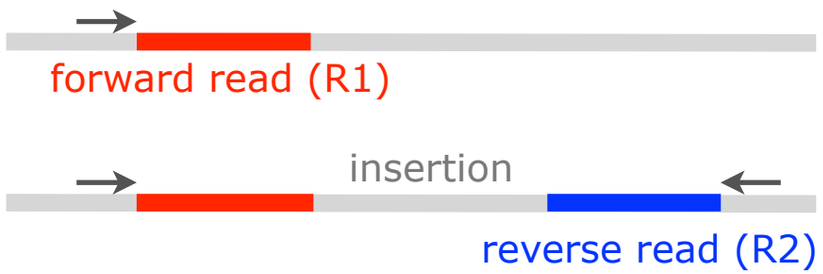
The 10 Most Critical Platform Attributes as Defined by Purchasers

1. Cost per base	43%	6. Read length	24%
2. Sequencing data quality	34%	7. Instrument cost	18%
3. Appropriate for my application	32%	8. Number of reads	17%
4. Reproducibility/accuracy	31%	9. Available software analysis tools	16%
5. Amount of DNA/RNA needed per experiment	25%	10. Instrument reliability	16%

Sequencing Data

GDC Genetic Diversity Centre Zurich

Short Read Sequence Types



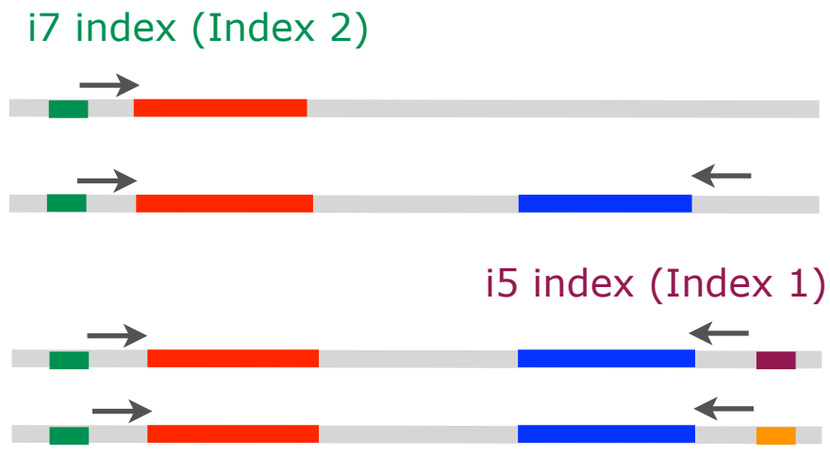
Single Reads (SR) – only R1

Paired-End Reads (PE) – R1 and R2 with insert in between

Overlapping PE – helpful in amplicon sequencing

Extended SR – less common, but valid

Short Read Index Types



Each sample gets a unique i5 + i7 combination.

Single Index (typically on i7 adapter)

PE with Single Index – also a valid older method

Paired-End with UDI – this is the modern best practice

GDC Genetic Diversity Centre Zurich

Index Hopping in Illumina Sequencing

What is Index Hopping?

Index hopping (or index mis-assignment) occurs when the index sequence from one DNA fragment is incorrectly associated with the read data of another fragment. This leads to the mis-assignment of sequencing reads to the wrong sample.

How Does It Happen?

During library preparation or clustering, free-floating adapters or index strands can attach to the wrong DNA fragments.

This is especially problematic on patterned flow cells (e.g., HiSeq 4000, NovaSeq) and with single-indexed libraries.

Why It Matters?

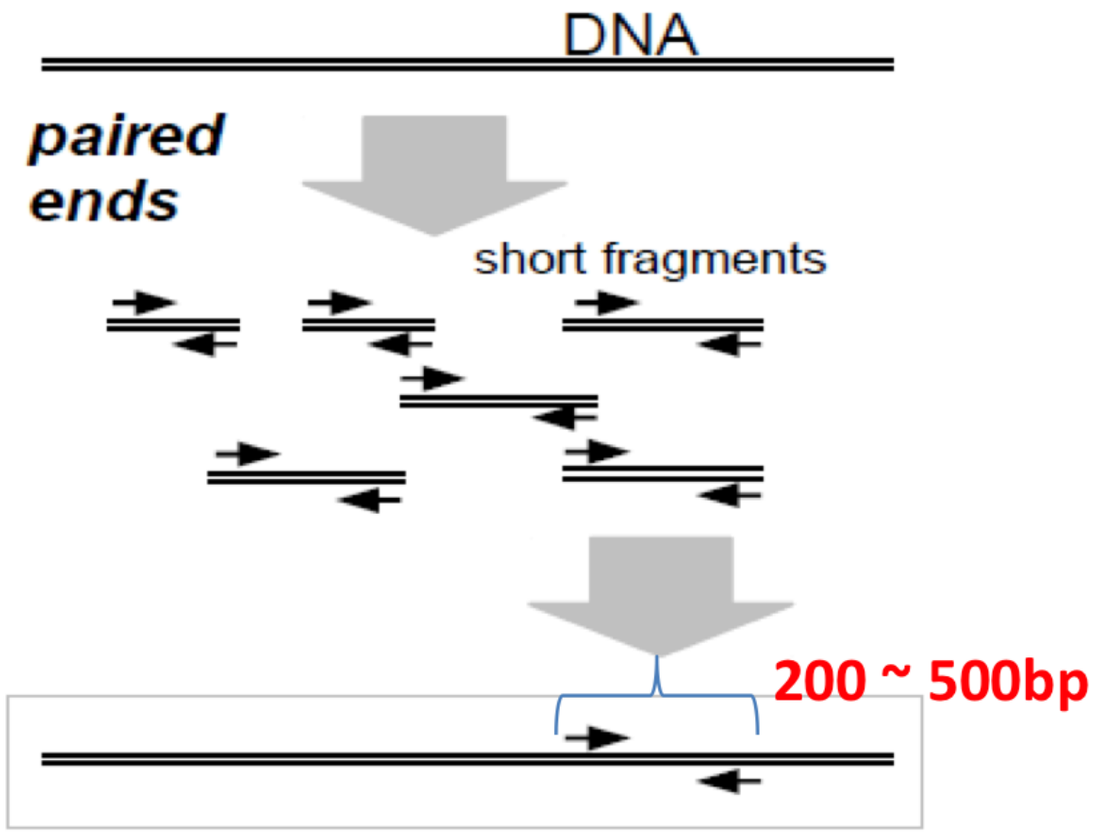
Mis-assigned reads can contaminate samples, especially in multiplexed runs.

Low-frequency variants or rare taxa in microbiome studies can be incorrectly reported.

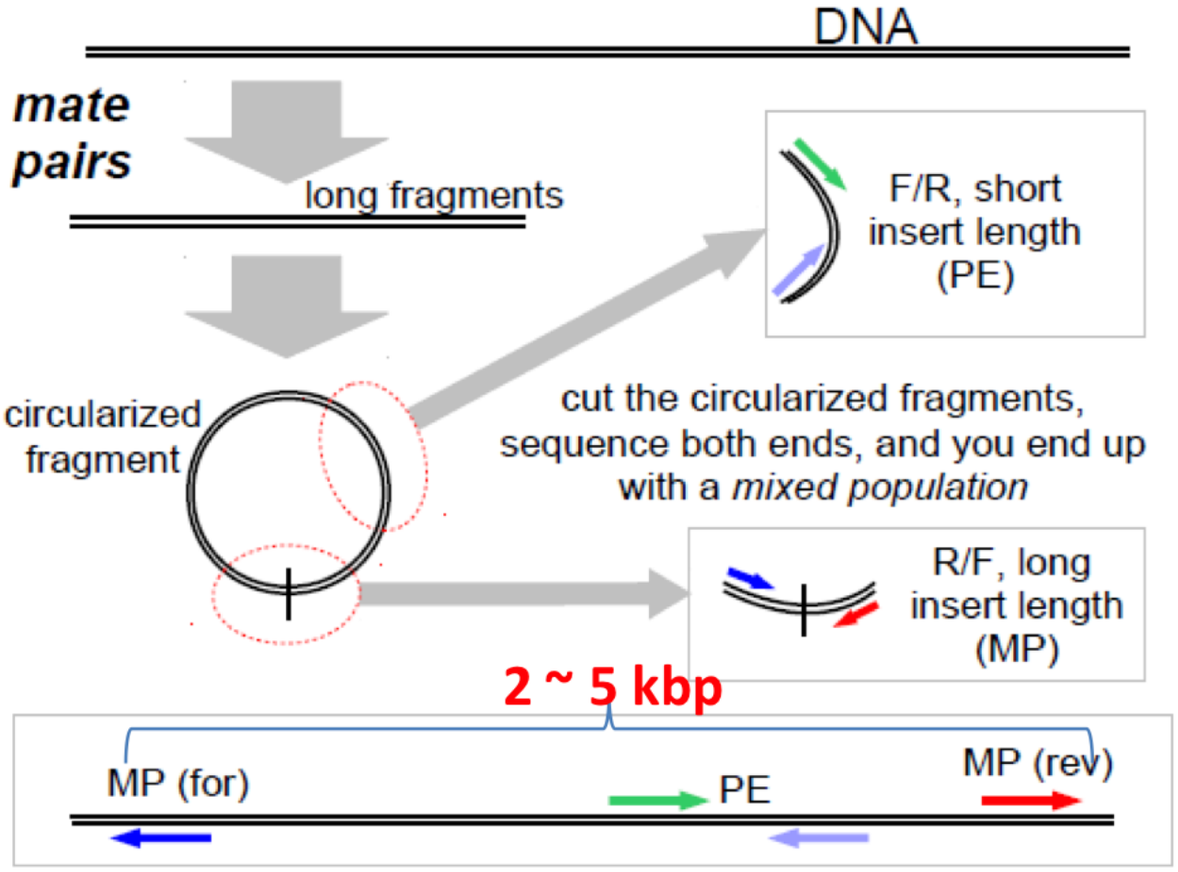
GDC Genetic Diversity Centre Zurich

Sequence Read Data

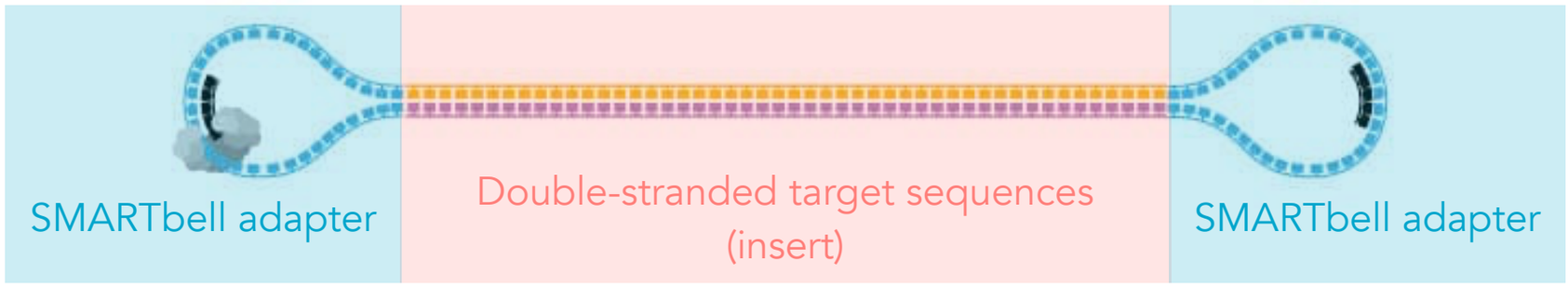
paired-end (PE)

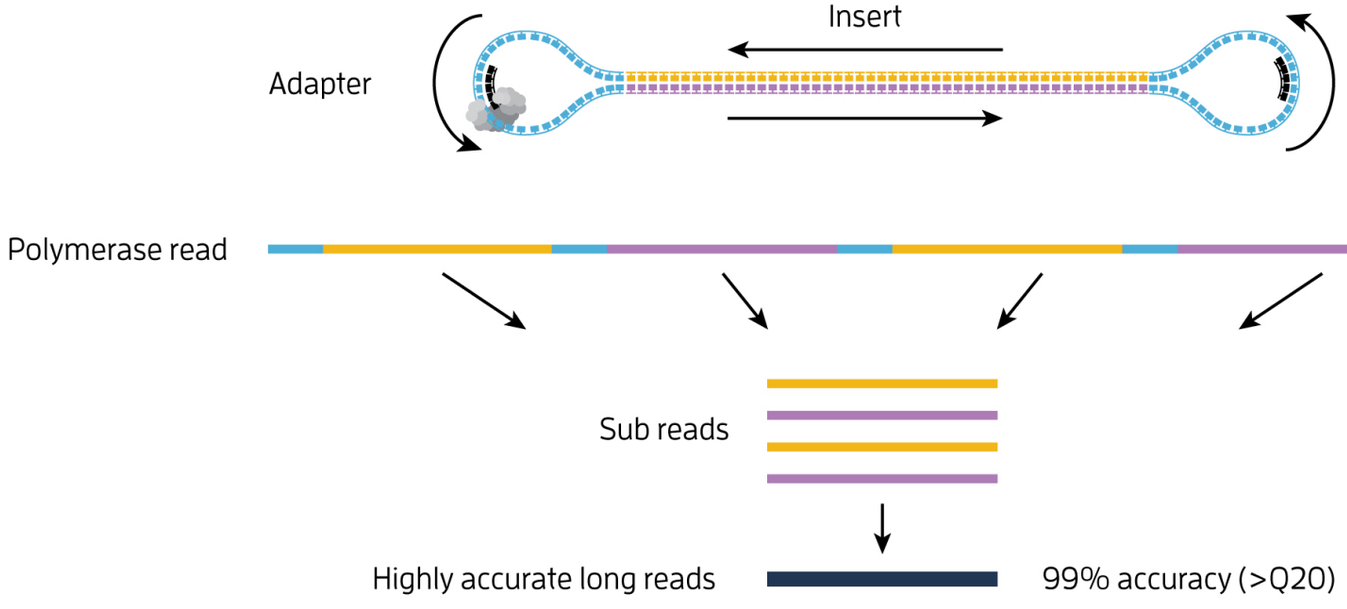


mate-pair (MP)



PacBio SMRTbell Library





PacBio's **HiFi sequencing** technology generates **high-accuracy reads** by sequencing the **same DNA molecule multiple times**. This is done using **circular consensus sequencing (CCS)**: the polymerase loops around a circularized DNA insert, producing **multiple subreads** from a single strand. These subreads are then **combined into a HiFi read** with high per-base accuracy (>Q20–Q30).

GDC Genetic Diversity Centre Zurich

HiFi Read Generation: Homoduplex vs. Heteroduplex

PacBio's HiFi sequencing technology generates high-accuracy reads by repeatedly sequencing a circularized DNA molecule. This process works best with **homoduplexes**, where both strands are identical, allowing the polymerase to produce consistent subreads and enabling a reliable consensus (HiFi read).

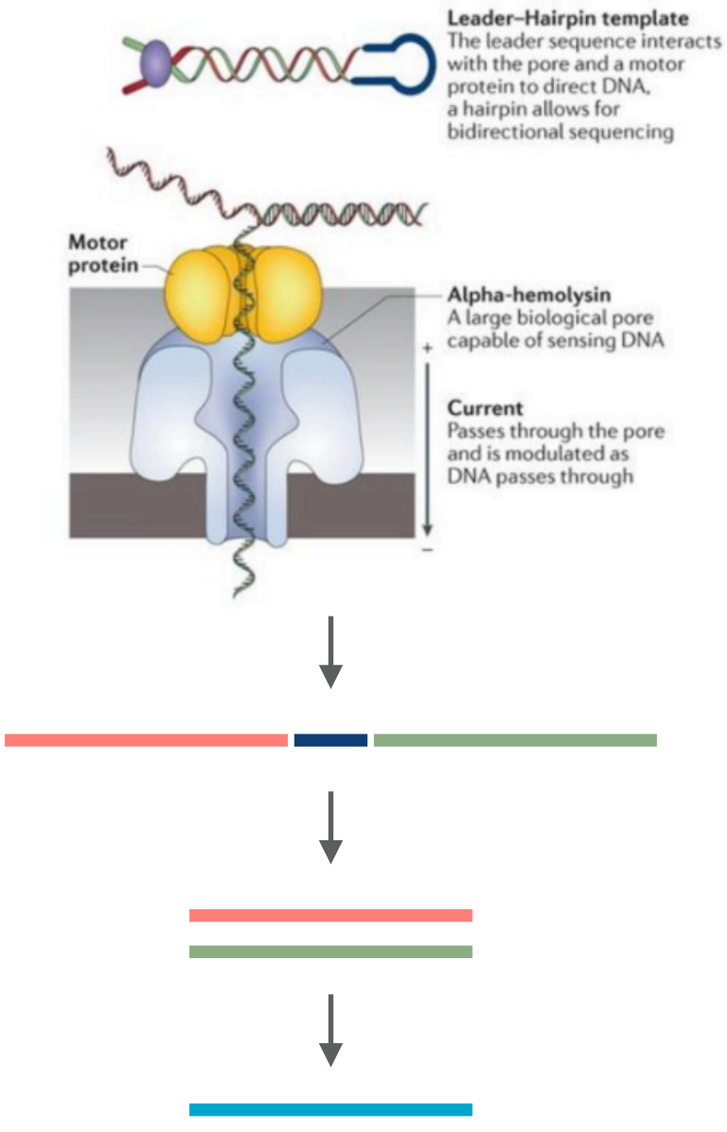
In amplicon sequencing, PCR products from similar but non-identical templates can anneal into **heteroduplexes**. These mismatches between strands cause subread inconsistencies if the polymerase switches strands, leading to reduced accuracy, ambiguous consensus, or failure to produce a HiFi read.

Handling of Heteroduplexes in PacBio Systems

Older PacBio Systems (e.g., Sequel II/IIe): Previously, PacBio's software could detect heteroduplexes and generate separate HiFi reads for each strand. This approach allowed users to analyze forward and reverse reads independently when strand differences were significant.

Revio System: The Revio system employs advanced algorithms, including DeepConsensus+, for real-time data processing. However, current documentation does not specify whether it supports separate HiFi read generation for each strand in heteroduplex cases.

ONT Sequencing



1. Leader-Hairpin Template
A special leader and hairpin adapter directs the double-stranded DNA through the pore. This setup enables sequencing of both strands (2D sequencing).

2. Motor Protein + Nanopore
A motor protein controls the movement of DNA through a biological nanopore (e.g., alpha-hemolysin). As DNA passes, each nucleotide modulates the ionic current in a characteristic way.

3. Basecalling
The signal trace from both strands is split and interpreted individually.

4. Consensus Generation
The forward (template) and reverse (complement) strand signals are aligned.

5. High-Accuracy Read
A final consensus read (2D read or duplex read) is produced, offering improved accuracy compared to single-pass reads.

Data Submission

The European Nucleotide Archive (ENA) captures and presents information relating to experimental workflows that are based around nucleotide sequencing. A typical workflow includes the isolation and preparation of material for sequencing, a run of a sequencing machine in which sequencing data are produced and a subsequent bioinformatic analysis pipeline. ENA records this information in a data model that covers input information (sample, experimental setup, machine configuration), output machine data (sequence traces, reads and quality scores) and interpreted information (assembly, mapping, functional annotation).

Sequence Read Archive (SRA) makes biological sequence data available to the research community to enhance reproducibility and allow for new discoveries by comparing data sets. The SRA stores raw sequencing data and alignment information from high-throughput sequencing platforms, including Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD System®, Helicos Heliscope®, Complete Genomics®, and Pacific Biosciences SMRT®.

Data Deposition Example from the Literature:

Mushegian *et al*. (2018) **Environmental sources of bacteria and genetic variation in behavior influence host-associated microbiota**. AEM doi:10.1128/AEM.01547-18.

Sequence data are deposited in the **European Nucleotide Archive of the EBI under accession number PRJEB30308** (http://www.ebi.ac.uk/ena/data/view/PRJEB30308). Data tables, OTUs sequences and code used for analysis can be found on **Github** at https://github.com/amusheg/Daphnia-microbiota-behavior and will be deposited in **Dryad** upon publication.

EMBL-EBI

Services | Research | Training | About us

ENA European Nucleotide Archive

Search

Examples: BN000065, histone

Advanced Sequence

Home | Search & Browse | Submit & Update | Software | About ENA | Support

Contact Helpdesk

Study: PRJEB30308

Microbiota associated with Daphnia exhibiting genetic variation in behavior

View: Project XML Study XML

Download: Project XML Study XML

Name
Microbiota of browsing Daphnia

Submitting Centre
Universitaet Basel

Secondary accession(s)
ERP112744

Description

In many organisms, host-associated microbial communities are acquired horizontally after birth. This process is believed to be shaped by a combination of environmental and host genetic factors. We examined whether genetic variation in animal behavior could affect the composition of the animal's microbiota in different environments. The freshwater crustacean Daphnia magna is primarily planktonic, but exhibits variation in the degree to which it browses in benthic sediments. We performed an experiment with clonal lines of D. magna showing different levels of sediment-browsing intensity exposed to either bacteria-rich or bacteria-poor sediment or whose access to sediments was prevented. We find that the bacterial composition of the environment and genotype-specific browsing intensity together influence the composition of the Daphnia-associated bacterial community. Exposure to more diverse bacteria did not lead to a more diverse microbiome, but greater abundances of environment-specific bacteria were found associated with host genotypes that exhibited greater browsing behavior. Our results indicate that, although there is a great deal of variation between individuals, behavior can mediate genotype-by-environment interaction effects on microbiome composition.

Navigation | Read Files | Portal | Attributes

Bulk Download Files (If the downloader app doesn't open, please try using Firefox to launch it.)

Download: 1 - 512 of 512 results in TEXT

Select columns

Showing results 1 - 10 of 512 results

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)	Submitted files (FTP)	Submitted files (Galaxy)	NCBI SRA file (FTP)	NCBI SRA file (Galaxy)	CRAM Index files (FTP)	CRAM Index files (Galaxy)
PRJEB30308	SAMEA5166093	ERS2973813	ERX2993334	ERR2990925	1869227	bacterium	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB30308	SAMEA5166094	ERS2973814	ERX2993335	ERR2990926	1869227	bacterium	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB30308	SAMEA5166095	ERS2973815	ERX2993336	ERR2990927	1869227	bacterium	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB30308	SAMEA5166096	ERS2973816	ERX2993337	ERR2990928	1869227	bacterium	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		

GDC Genetic Diversity Centre Zurich

NCBI Resources How To Sign in to NCBI

BioProject

BioProject PRJEB30308 Search

Create alert Advanced Browse by Project attributes Help

Display Settings: Send to:

Microbiota of browsing Daphnia

Accession: PRJEB30308 ID: 516850

Microbiota associated with Daphnia exhibiting genetic variation in behavior

In many organisms, host-associated microbial communities are acquired horizontally after birth. More...

Accession	PRJEB30308
Scope	Monoisolate
Submission	Registration date: 24-Jan-2019 **Universitaet Basel**

Project Data:

Resource Name	Number of Links
SEQUENCE DATA	
SRA Experiments	512
OTHER DATASETS	
BioSample	512

▾ SRA Data Details

Parameter	Value
Data volume, Gbases	22
Data volume, Mbytes	14805

Related information

BioSample

SRA

Recent activity Turn Off Clear

PRJEB30308 (1)

BioProject
Microbiota of browsing Daphnia

BioProject
The European Nucleotide Archive in 2017

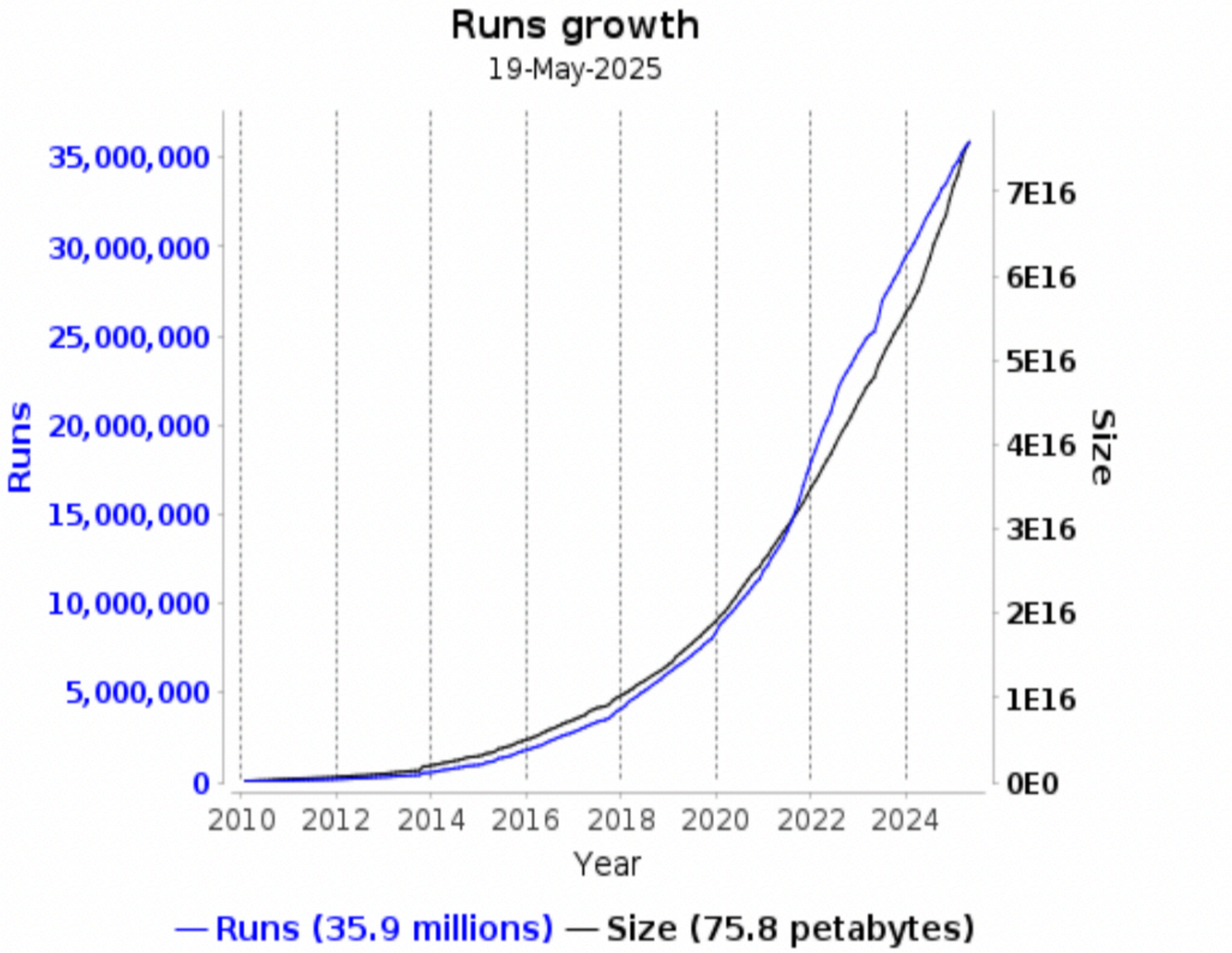
A Benchmark Study on Error Assessment and Quality Control of CCS Reads Derived

Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects

See more...

ENA European Nucleotide Archive

https://www.ebi.ac.uk/ena/browser/about/statistics



GDC Genetic Diversity Centre Zurich

1

Choose the MPS technology according to your needs.

2

Keep your raw data safe and submit it as early as possible.

3

Keep your sequence files zipped.