Illumina Technology, Overview for IABS Conference

Workshop December 3r ,2024

Dr. Felix Krombholz

Sr. Technical Specialist, Germany

December 3rd, 2024

Brief facts about Illumina

1

illumina®

Who we are

We are a global genomics and human health leader innovating the future of precision health.

We develop DNA sequencing and array-based life sciences technologies to enable boundless research discovery and personalized health. Our products help pioneer advances in oncology, genetic and infectious diseases, reproductive health, and beyond. Our technology empowers continued innovation toward positive and impactful people- and planet-healing solutions.

$4.43B USD

2023 core revenue

~9250

Number of employees*

Jacob Thaysen

Chief executive officer

San Diego, United States

Headquarters

1998

Year founded

*As of December 30, 2023

illumina®

Key Applications Fit: Flagship Sequencers

MiSeq™ i100 & i100 Plus

NextSeq™ 1000 & 2000

NovaSeq™ X & X Plus

- Small Whole-Genome Sequencing
- Targeted Gene Sequencing
- Targeted Gene Expression Profiling
- 16S Metagenomic Sequencing
- Exome Sequencing
- Transcriptome Sequencing
- Cell-Free Sequencing*
- Single-Cell or Spatial Profiling
- Shotgun Metagenomics
- Methylation & Chromatin Analysis^
- Protein Expression Profiling
- Large Whole-Genome Sequencing

- DNA
- RNA
- Epigenetics
- Other

* Cell-free sequencing includes noninvasive prenatal testing (NIPT) and liquid biopsy.
^ Chromatin analysis includes assay for transposase accessible chromatin (ATAC-Seq), chromatin immunoprecipitation (ChIP-Seq), and chromatin conformation capture (Hi-C).

illumina®

For Research Use Only. Not for use in diagnostic procedures.

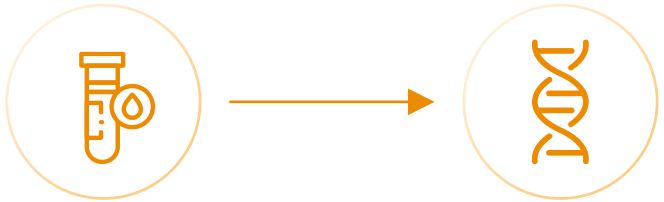
Illumina Sequencing Overview

2

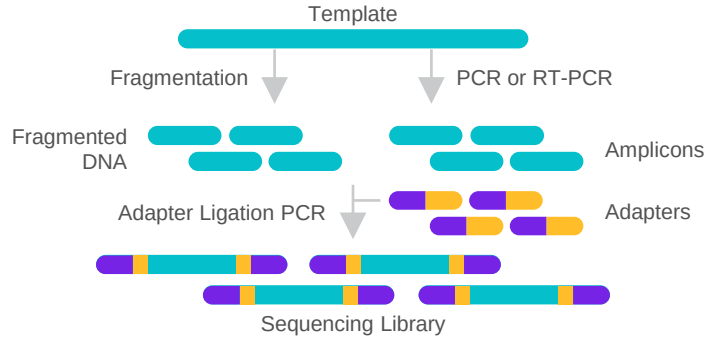
illumina®

What Does a General NGS Workflow Look Like?

Step 1
Extraction

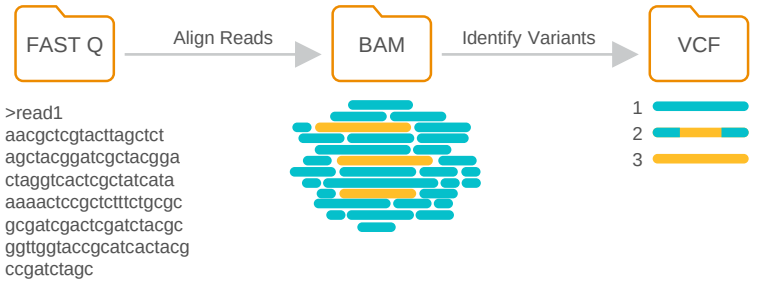


Step 2
Library prep



Step 3
Sequencing

Step 4
Analysis



NGS analysis is now very accessible! Your local core lab likely offers expert help. Or if you prefer to analyze your own data, there are user-friendly apps like Correlation Engine and Base Space Sequence Hub

Preparing Samples for Sequencing (Library Prep)

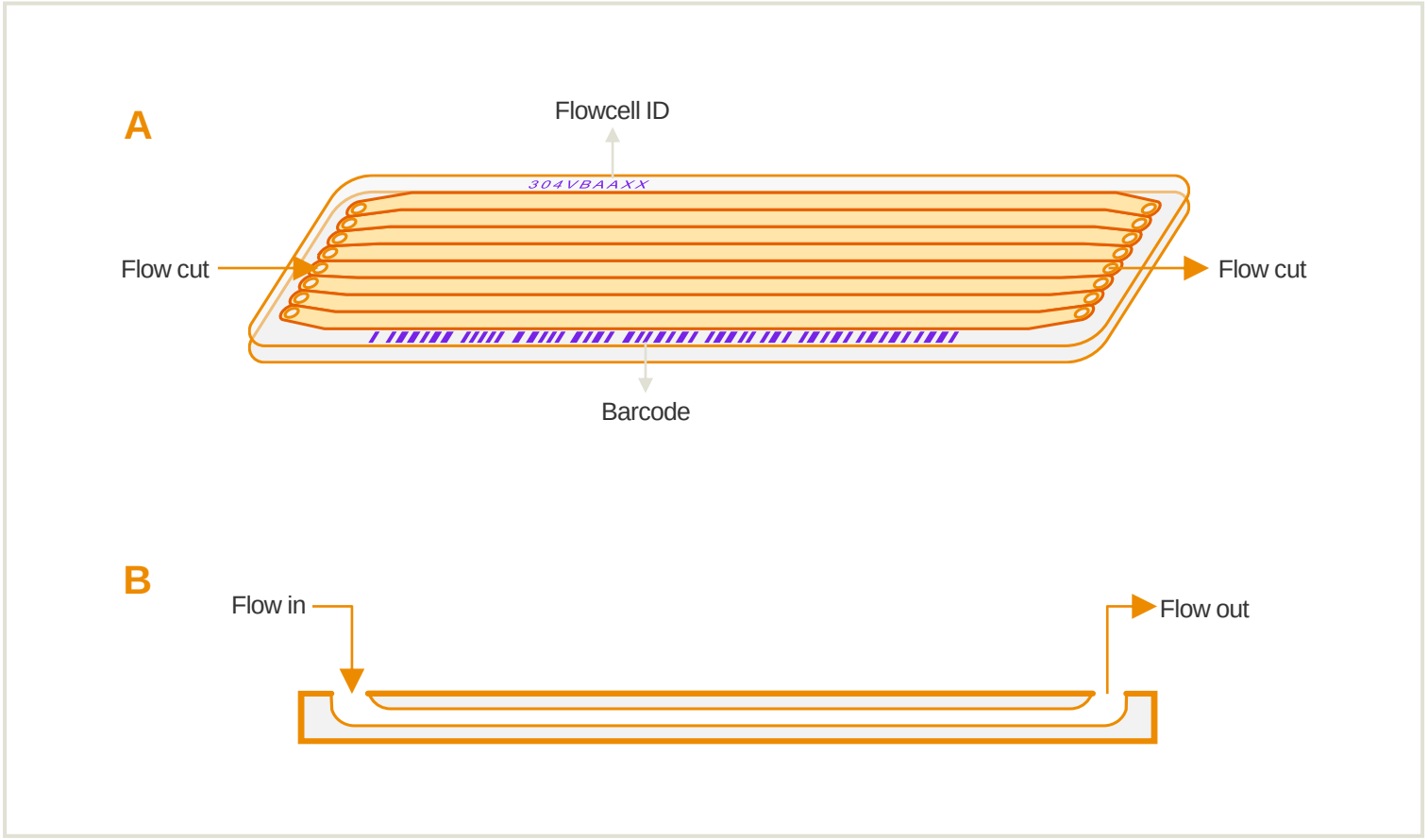
illumina®

Our Goal Is to Make Samples Bind a “Flow Cell”, Which Is Where Nucleotides Are Detected By an Illumina Sequencer

DNA is hybridized onto the surface of the flow cell

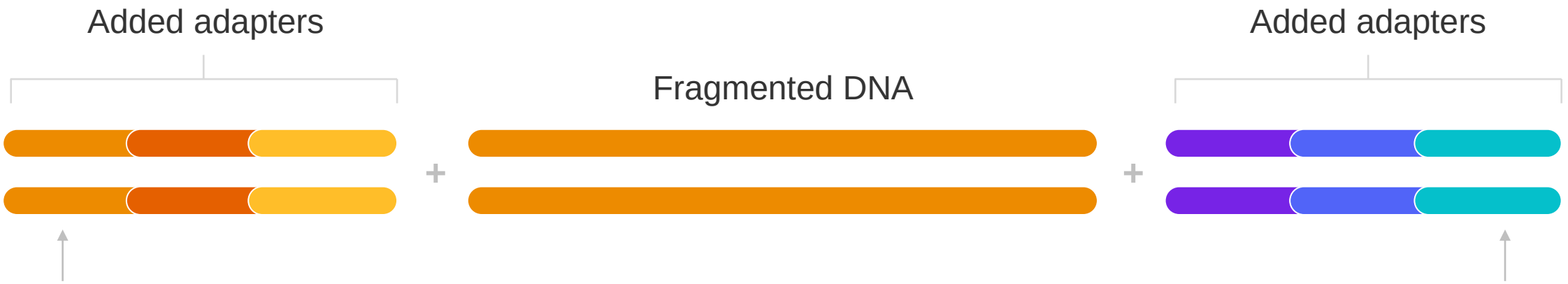
Polymerases extend DNA using fluorescent nucleotides

Cameras detect fluorescent signals across the flow cell



Library Prep Is the Process of Making DNA Compatible With an Illumina Flow Cell to Prepare It for Sequencing

The first step is prepare nucleic acids so that they (1) have barcodes that identify them post-sequencing, and (2) have adapters that make them compatible with the flow cell

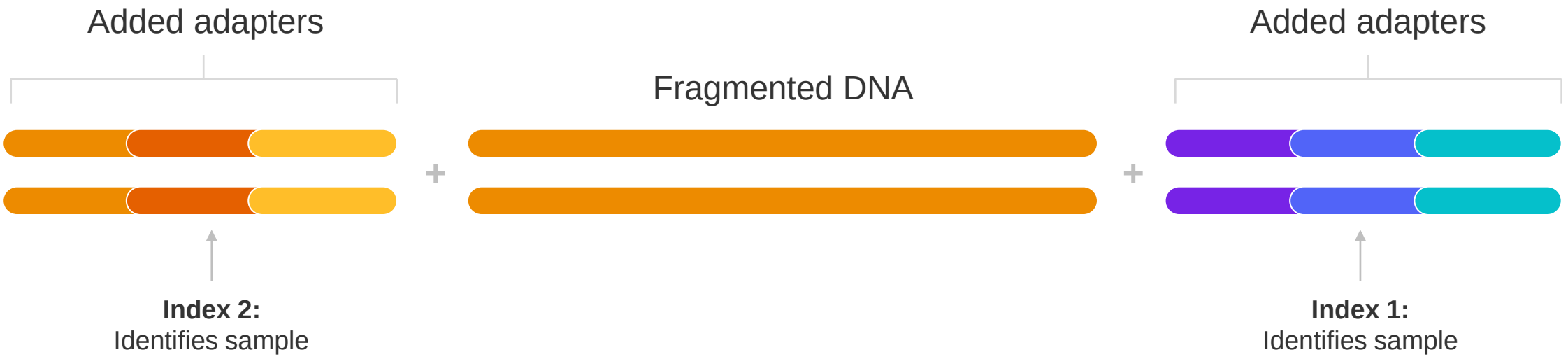


P5 Sequence:
Allows DNA fragment to bind sequencing flow cell so that sequencing by synthesis can occur

P7 Sequence:
Allows DNA fragment to bind sequencing flow cell so that sequencing by synthesis can occur

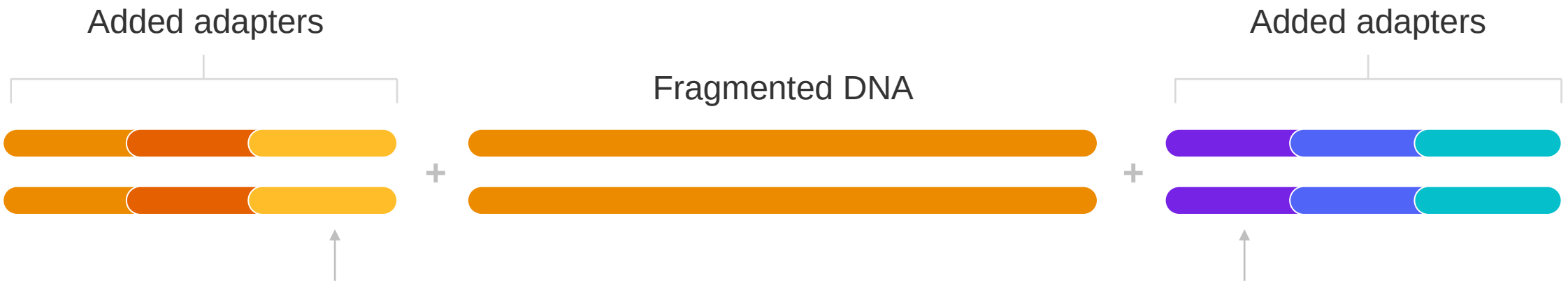
Library Prep Is the Process of Making DNA Compatible With an Illumina Flow Cell to Prepare It for Sequencing

The first step is prepare nucleic acids so that they (1) have barcodes that identify them post-sequencing, and (2) have adapters that make them compatible with the flow cell



Library Prep Is the Process of Making DNA Compatible With an Illumina Flow Cell to Prepare It for Sequencing

The first step is prepare nucleic acids so that they (1) have barcodes that identify them post-sequencing, and (2) have adapters that make them compatible with the flow cell



Read 1 Sequencing Primer:
Primer for polymerase reaction on flow cell for a read in the forward direction

Read 2 Sequencing Primer:
Primer for polymerase reaction on flow cell for a read in the reverse direction

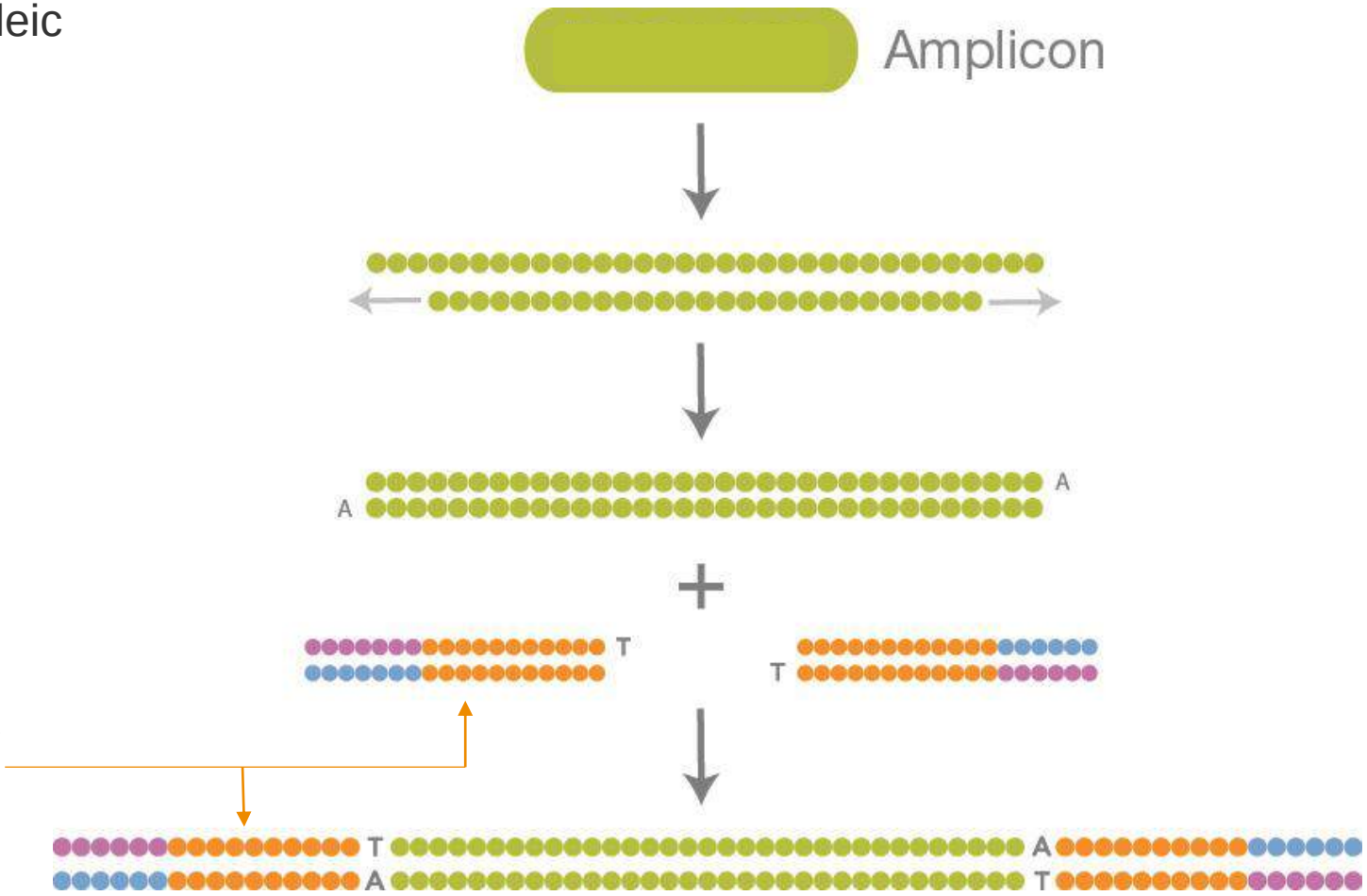
What is Library Prep?

Library Prep – The creation of a pool of nucleic acids that will be sequenced on an illumina sequencer.

What does this process involve? – Library prep is a process that makes nucleic acids compatible with illumina sequencers.

Key Steps:

- PCR amplification of nucleic acids
- Addition of barcodes that identify unique molecules within a pool of nucleic acids
- Addition of adapter sequences to the 5’ and 3’ ends of nucleic acid fragments to make them compatible with an illumina sequencer
  - *Specifically, it allows them to hybridize with a sequencer’s flow cell*



How Does Sequencing by Synthesis Work?

4

illumina®

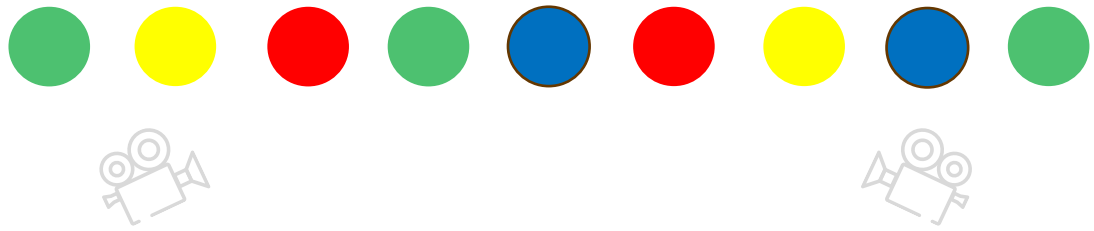
How Does Sequencing by Synthesis Work?

- SBS generally works by adding fluorescently-tagged nucleotides to a growing piece of DNA.
- Since each nucleotide is associated with a certain color (or lack of color), identifying that color also identified what nucleotide was added to a growing piece of DNA
  (i.e. Was that nucleotide an A, T, G, or C?)

Example: Detecting nucleotides by synthesis using four colors, or “four color chemistry”

As each fluorescent nucleotide is added to the growing strand, a camera detects the added color

DNA Sequence: T G C T A C G A T



KEY:

- Green = T
- Red = C
- Blue = A
- Yellow = G

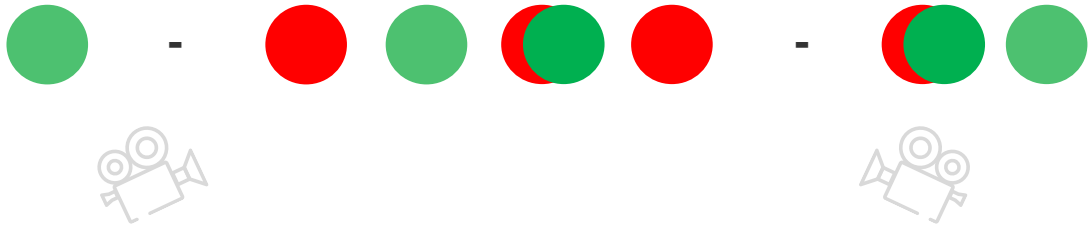
How Does Sequencing by Synthesis Work?

- SBS generally works by adding fluorescently-tagged nucleotides to a growing piece of DNA.
- Since each nucleotide is associated with a certain color (or lack of color), identifying that color also identified what nucleotide was added to a growing piece of DNA
  (i.e. Was that nucleotide an A, T, G, or C?)

Example: Detecting nucleotides by synthesis using two colors, or “two color chemistry”

Two color chemistry may also be used to speed up the sequencing process! Less colors means easier computation and less filter changes on the camera.

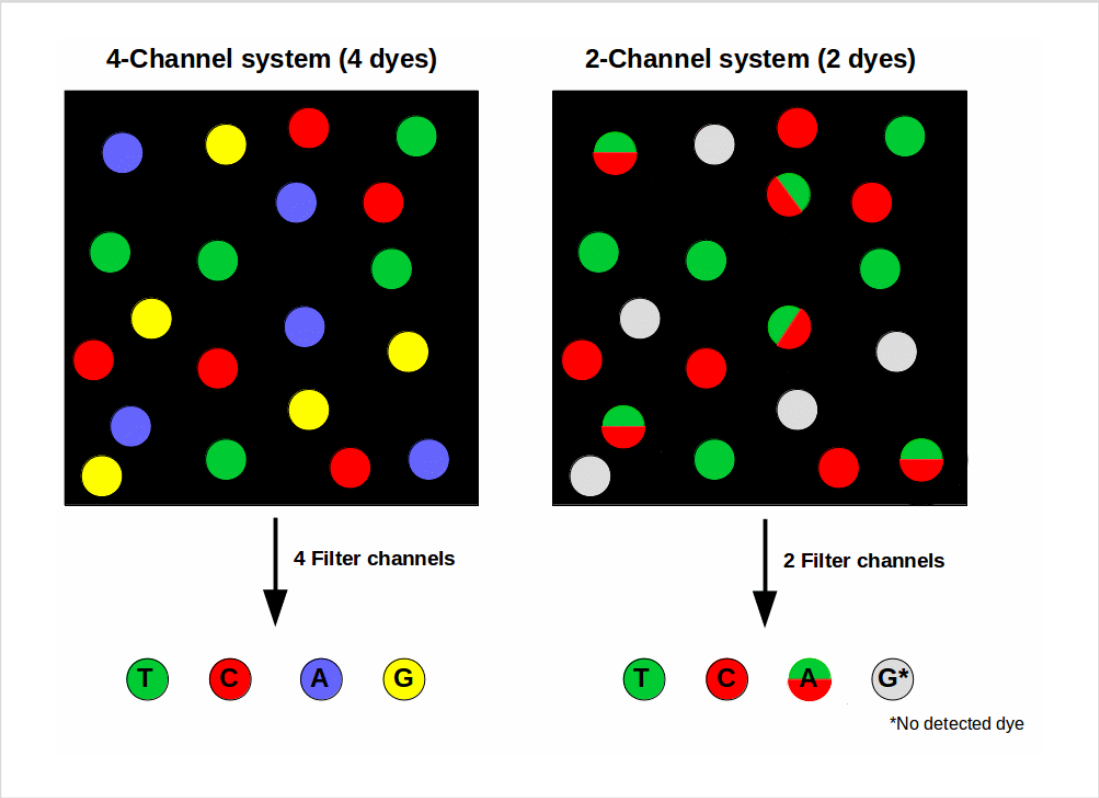
DNA Sequence: T G C T A C G A T



KEY:

- Green = T
- Red = C
- Red/Green = A
- No Color = G

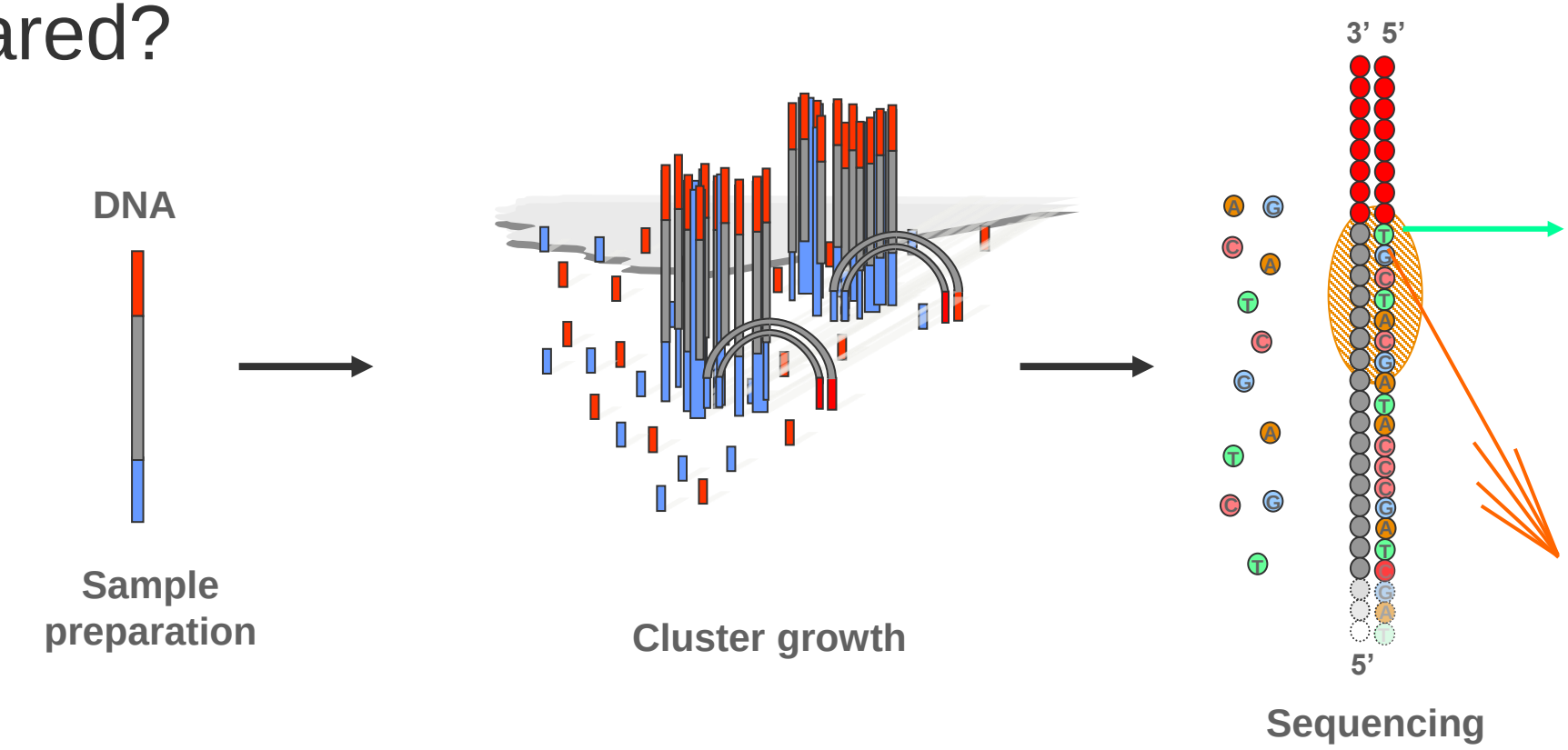
Illumina Sequencing by Synthesis Applies the Color Detection Principle to Many Growing Strands at Once!



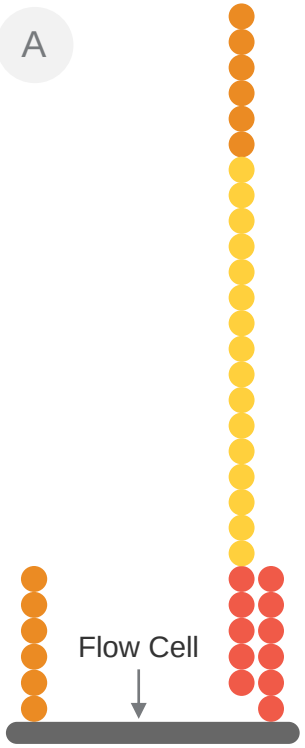
Modern sequencing systems can allow up to 26 Billion of these color-detection SBS processes to occur at once.

That's a lot of sequenced DNA!

What happens on a flow cell now that the sample is prepared?



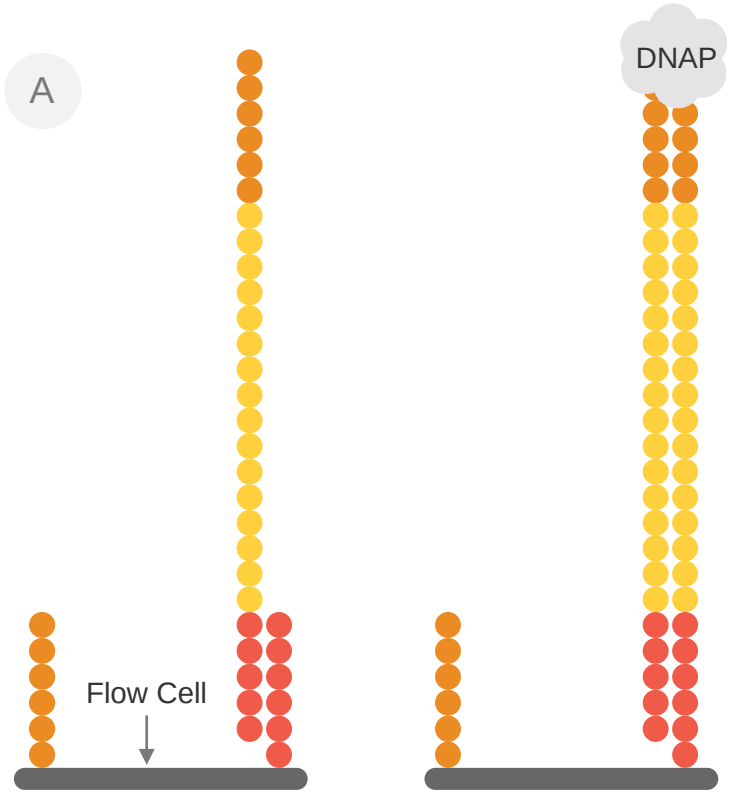
What Happens on a Flow Cell Now That the Sample Is Prepared?



Step A

- Addition of the DNA Library to the flow cell
- Samples bind oligos on the flow cell via adapter complementarity

What Happens on a Flow Cell Now that the Sample Is Prepared?

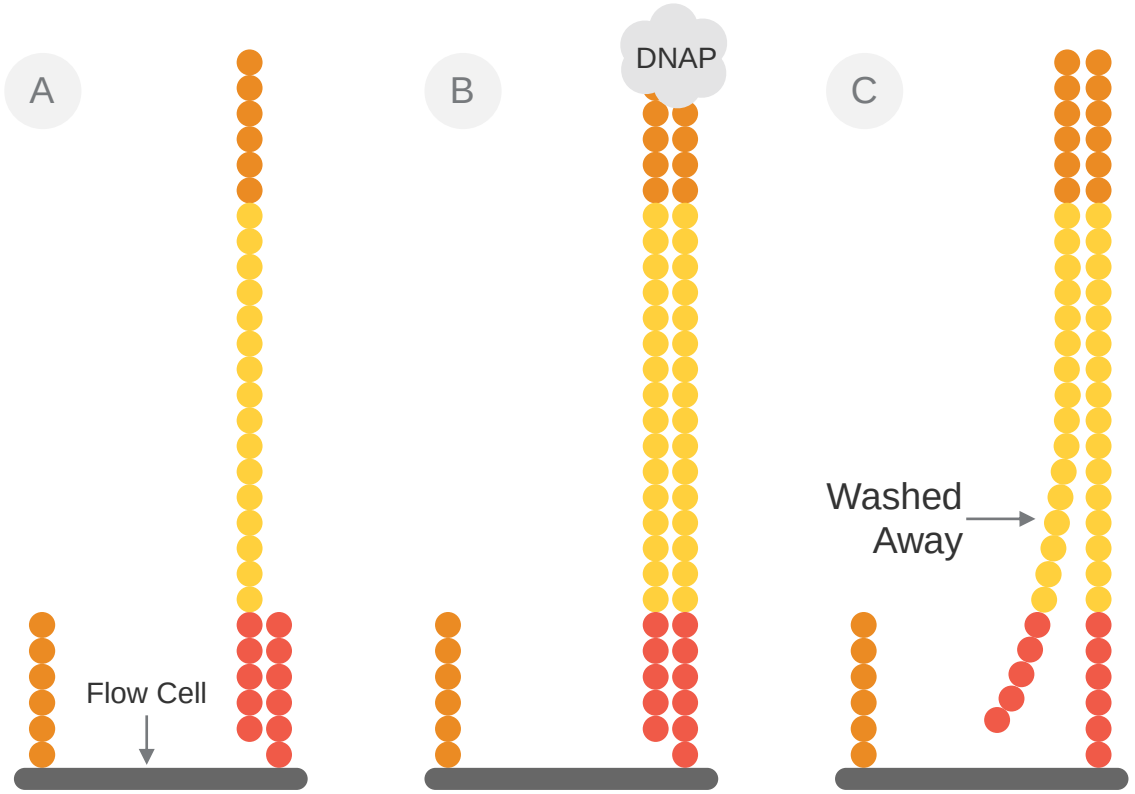


DNAP = DNA Polymerase

Step B

- DNA replication using the hybridized DNA as a template (like a PCR reaction)
- Fluorescent bases are detected as strand grows

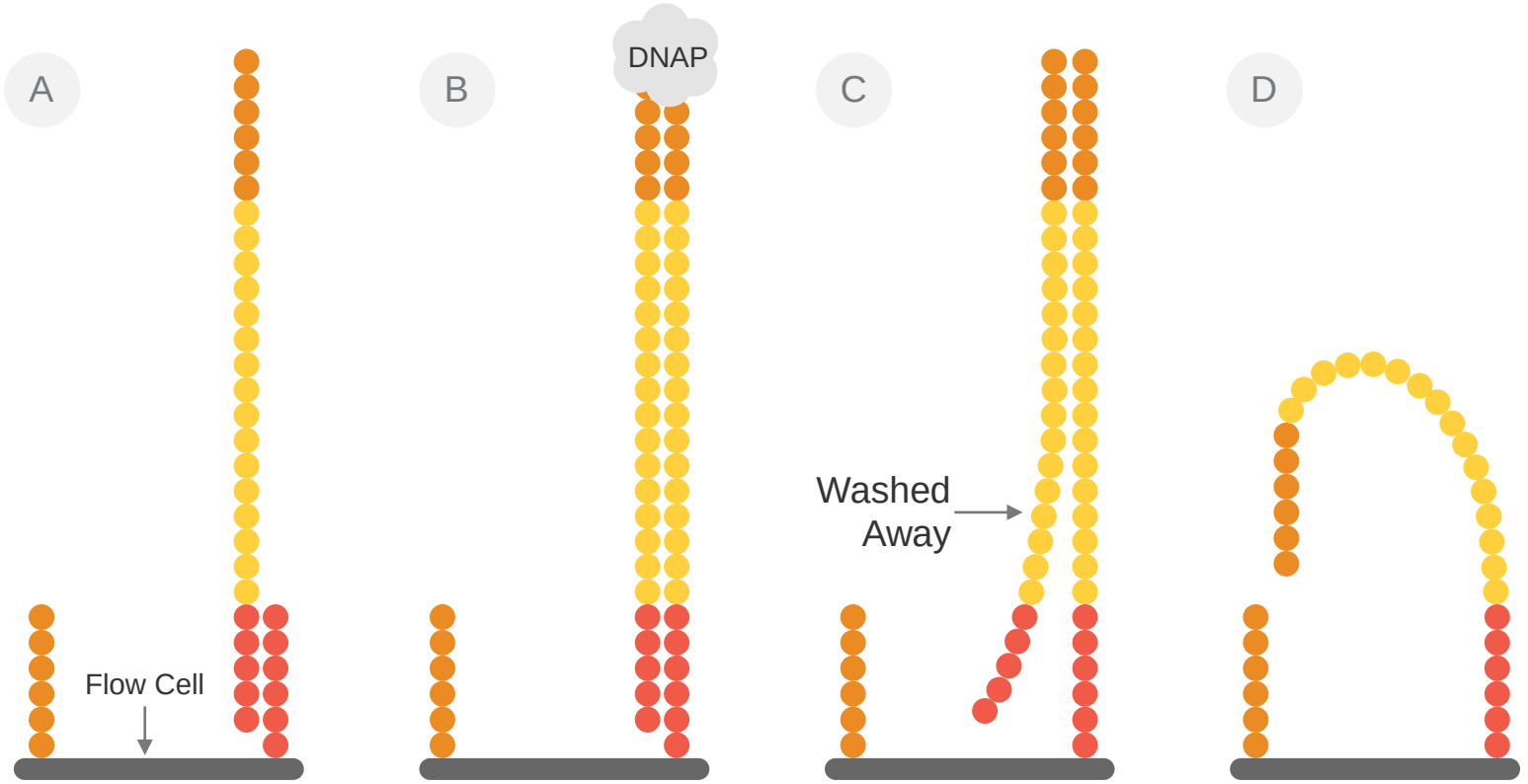
What Happens on a Flow Cell Now that the Sample Is Prepared?



Step C

- The original DNA strand is denatured and then washed away.
- This leaves a newly synthesized strand that is attached to the flow cell.

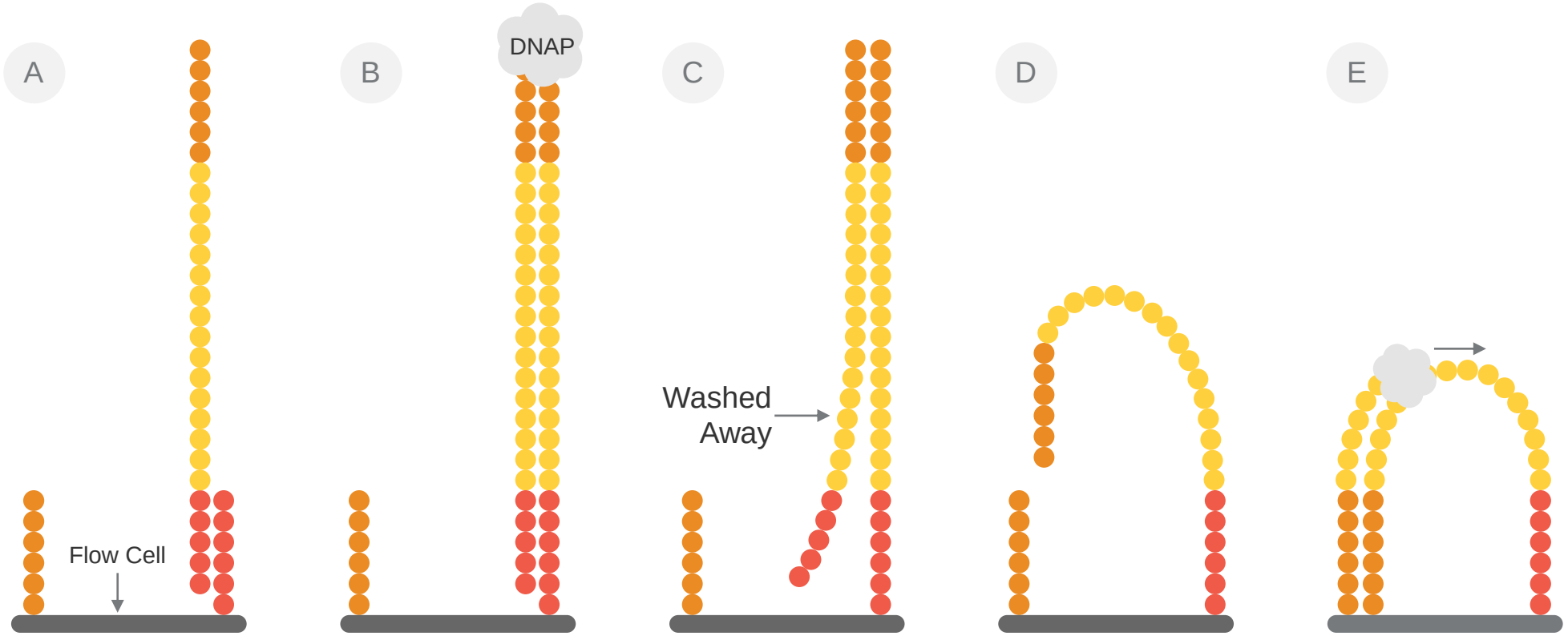
What Happens on a Flow Cell Now that the Sample Is Prepared?



Step D

- Bending of the tethered DNA to bind to an adjacent adapter via complementarity

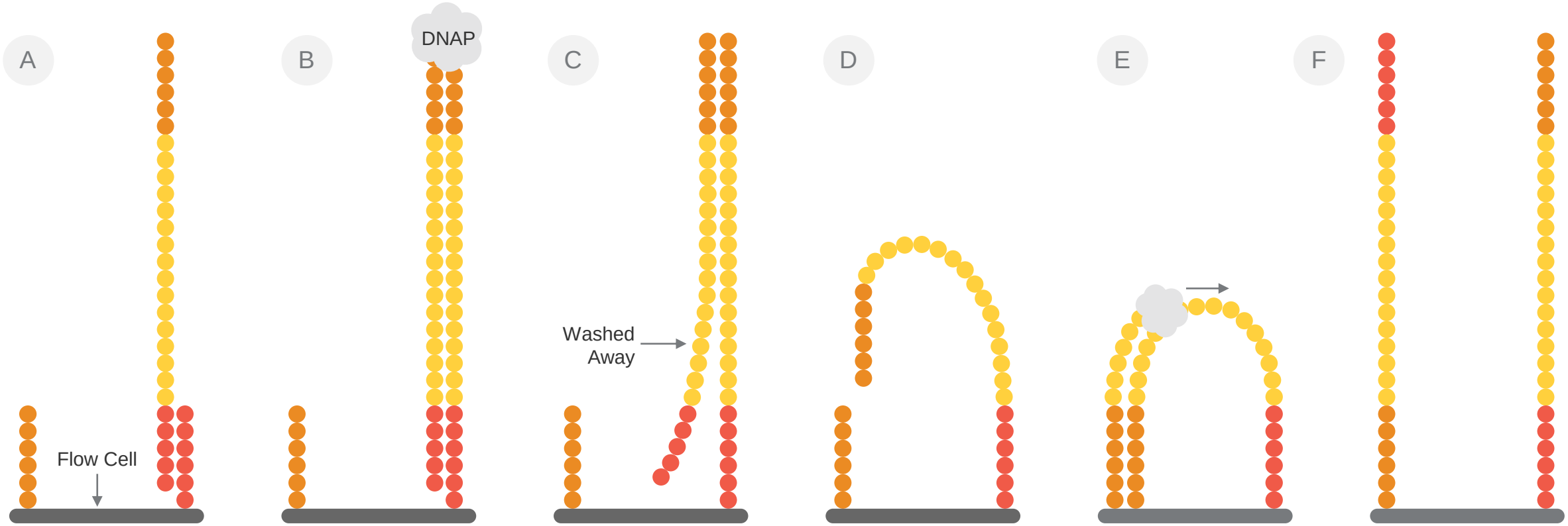
What Happens on a Flow Cell Now that the Sample Is Prepared?



Step E

- Another DNA amplification reaction is completed using the bent DNA as a template
- Fluorescent nucleotides are detected as the strand grows

What Happens on a Flow Cell Now that the Sample Is Prepared?

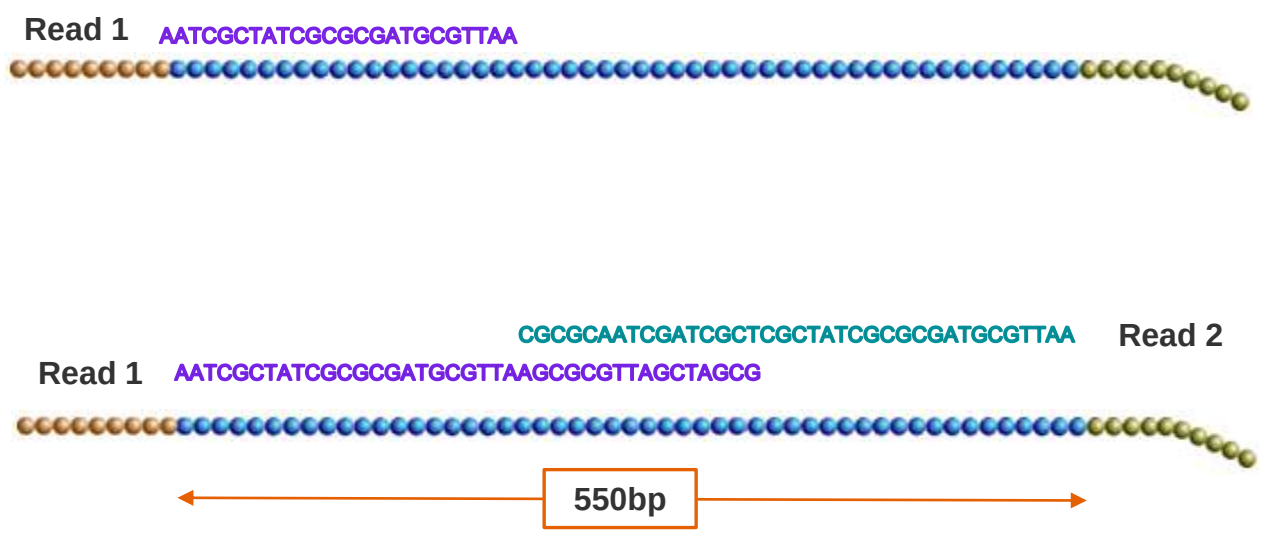


Step F

Denaturation of the dsDNA bridge to result in two ssDNA tethered to the flow cell.

Reads Can Be Single End or Paired End

A unique approach to variant identification

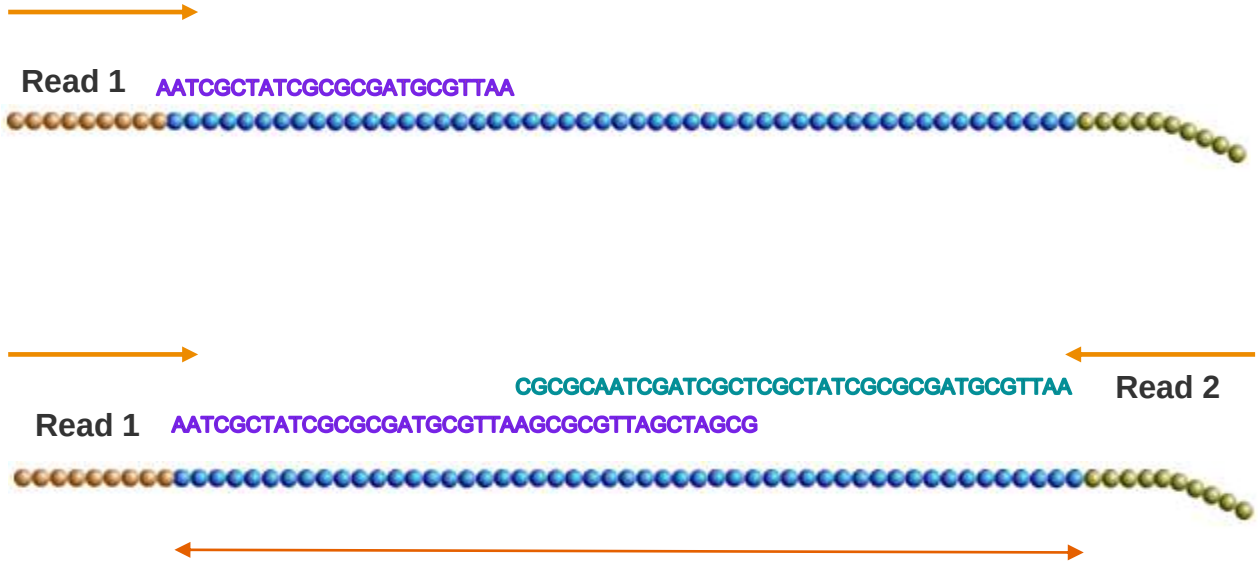


Single end:
The Sequencer reads only in one direction at a given length

Paired end:
The sequencer makes a forward read at a given length, then starts another round of reading from the opposite end of the fragment

Reads Can Be Single End or Paired End

A unique approach to variant identification



Single end:
The Sequencer reads only in one direction at a given length

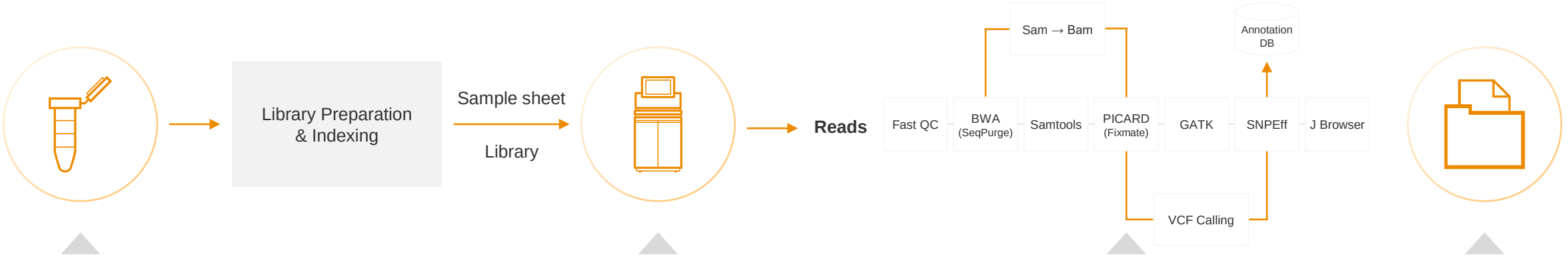
Paired end:
The sequencer makes a forward read at a given length, then starts another round of reading from the opposite end of the fragment

Paired-end reading improves detection of structural rearrangements such as gene insertions, deletions, or inversions. It also improves accuracy in repetitive regions.

Data Analysis

illumina®

It’s Time to Analyze Data !



SAMPLE	SEQUENCER	DATA ANALYSIS	RESULTS

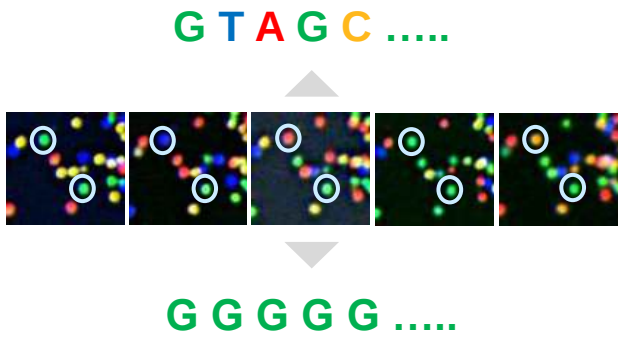
DRAGEN

- Mutation list
- Gene expression levels
- Methylated bases
- Metagenomic profile

Analysis

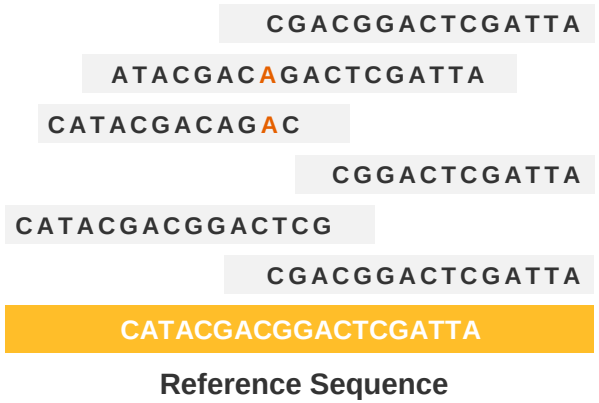
Primary analysis

- Runs in parallel to the sequencing
- Uses the images to provide base calls of the reads



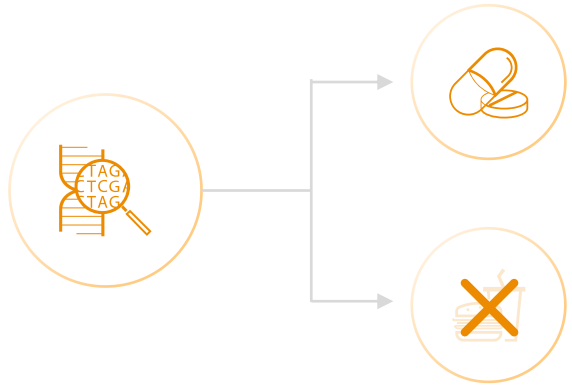
Secondary analysis

- Maps the reads to the reference genome (alignment)
- Analyzes the mapped reads for any differences compared to the reference genome (variant calling)



Interpretation

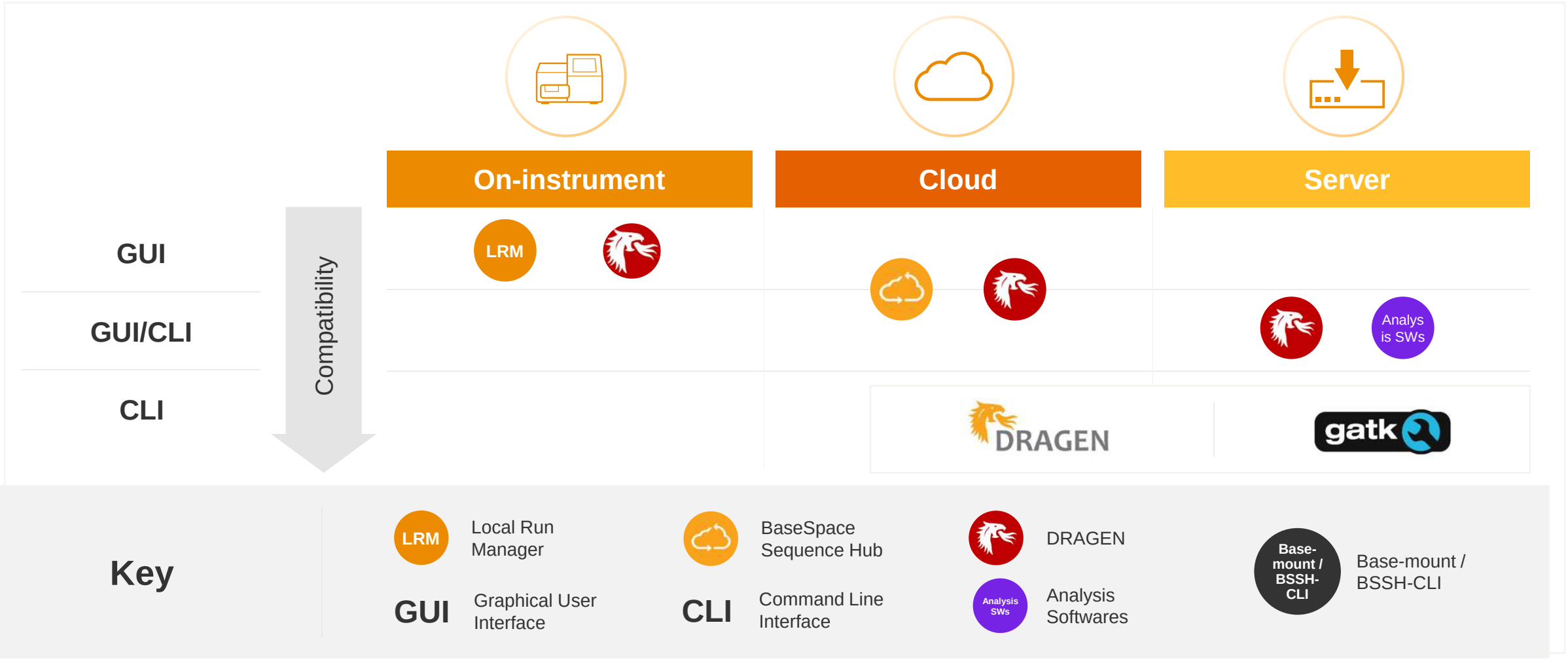
- Leads to knowledge and insights into basic biology
- Potential causes of genetic disease
- Possible treatment or prevention



https://emea.illumina.com/informatics/sequencing-data-analysis.html

OPTION 1

On-instrument, Cloud and On-site Analysis Options



Key

- LRM — Local Run Manager
- BaseSpace Sequence Hub
- DRAGEN
- Base-mount / BSSH-CLI
- **GUI** — Graphical User Interface
- **CLI** — Command Line Interface
- Analysis SWs — Analysis Softwares

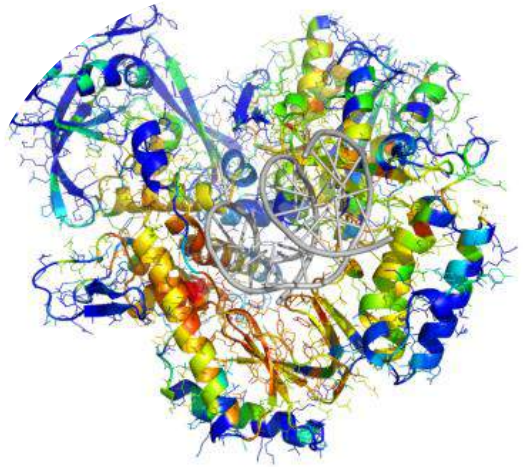
Thank you

illumina®

Appendix

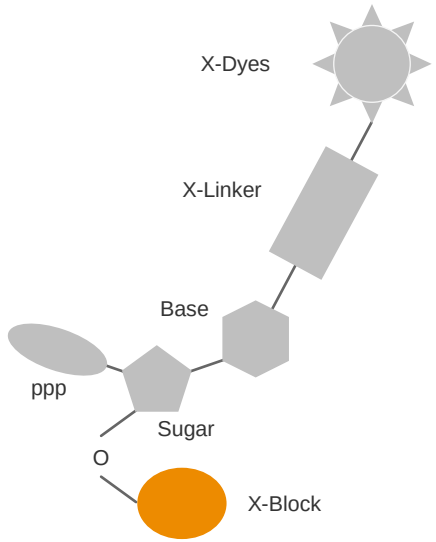
illumina®

XLEAP-SBS chemistry offers the highest level of accuracy and performance on MiSeq i100 Series[1]



Novel Polymerase

Faster incorporation, higher fidelity



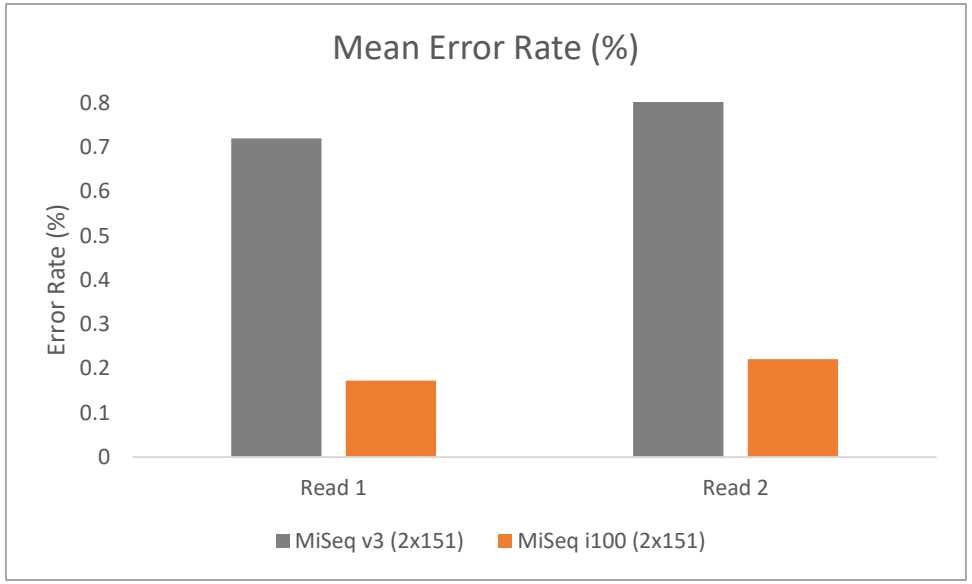
X-Block, X-Linker, X-Dyes

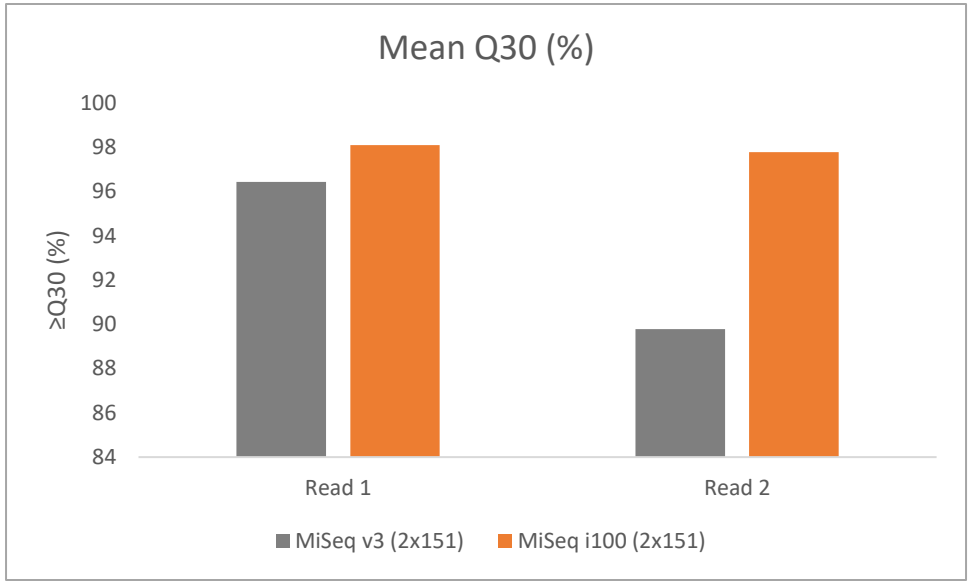
Most resistance to heat,~50x more stable in solution, faster block cleavage

- Faster cycle times
- Greater accuracy
- More Cost-Effective

[1] As compared to standard SBS chemistry on the MiSeq System

XLEAP-SBS Chemistry on the MiSeq i100 Series provides superior data quality



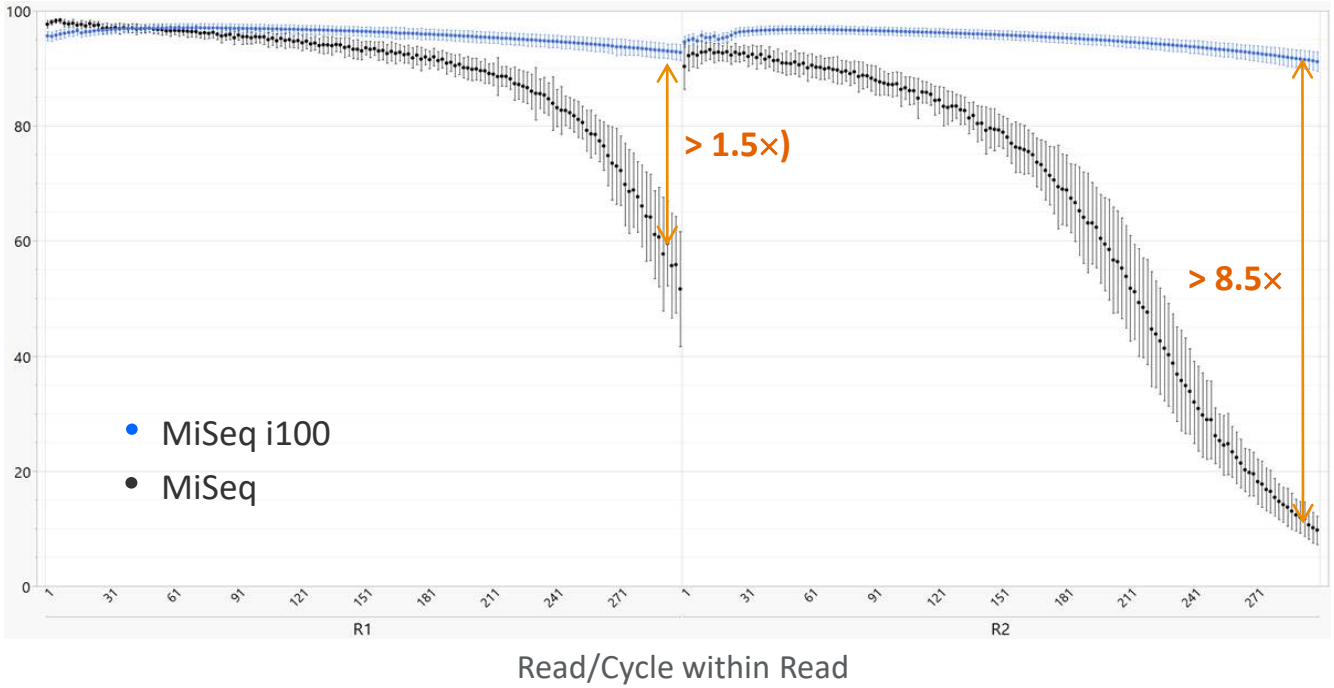


>4x Improved error rates

Increased % bases ≥ Q30 compared to MiSeq Systems

Results were generated using sWGbugpool library with 300 cycle 25M Reagent kits on MiSeq i100 Plus System and 600 cycle 25M Reagent v3 flow cell on MiSeq System
Data generated on a prototype configuration and may change prior to launch

MiSeq i100 engineered for superior data quality throughout the cycle



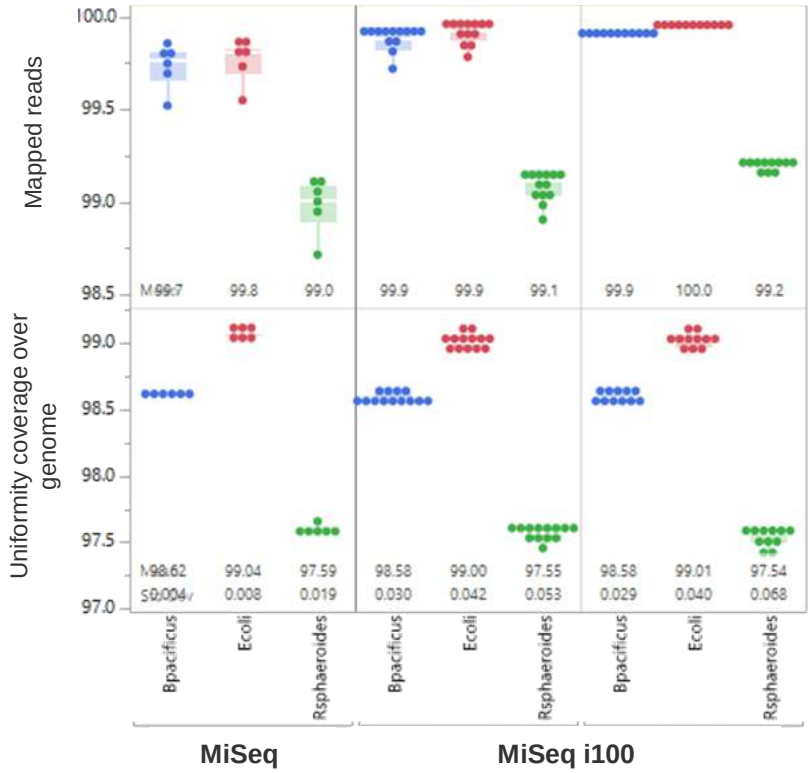
	MiSeq	MiSeq i100
≥Q30 (2×300, average), R1	89.1%	95.8%
≥Q30 (2×300, last 10 cycles), R1	55.5%	92.9%
≥Q30 (2×300, average), R2	65.8%	95.1%
≥Q30 (2×300, last 10 cycles), R2	10.7%	91.4%

Longer read lengths have improved accuracy throughout the whole read.
>1.5× improvement for Read 1, >8.5× improvement for Read 2

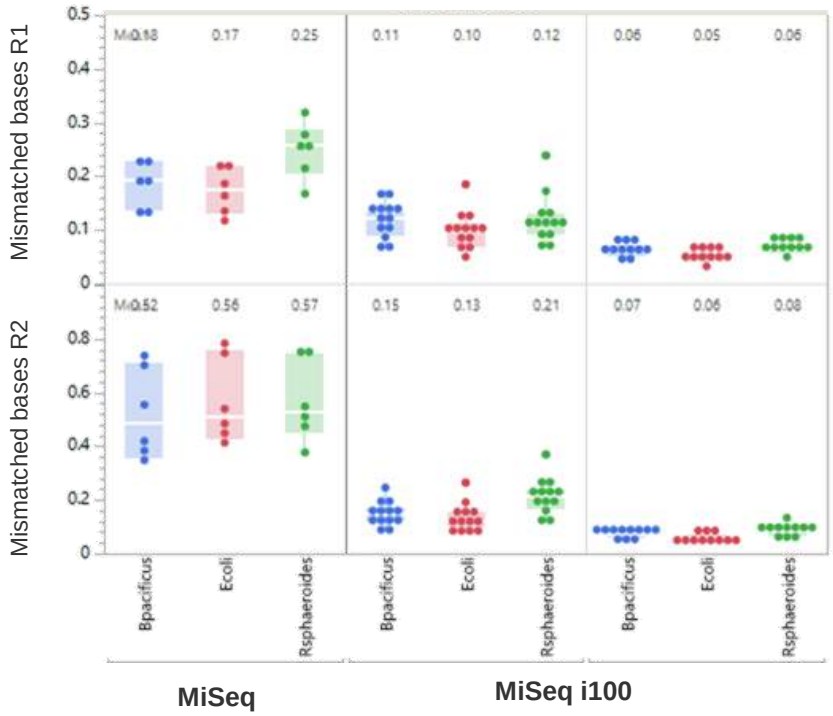
Results were generated using sWGbugpool library with 600 cycle 25M Reagent kits on MiSeq i100 Plus System and 600 cycle 25M Reagent v3 flow cell on MiSeq System
Data generated on a prototype configuration and may change prior to launch

Comparable or improved alignment for microbial small WGS

Mapped reads and coverage uniformity



Mismatched bases



Improved/comparable mapping and lower mismatch rates compared to MiSeq Systems

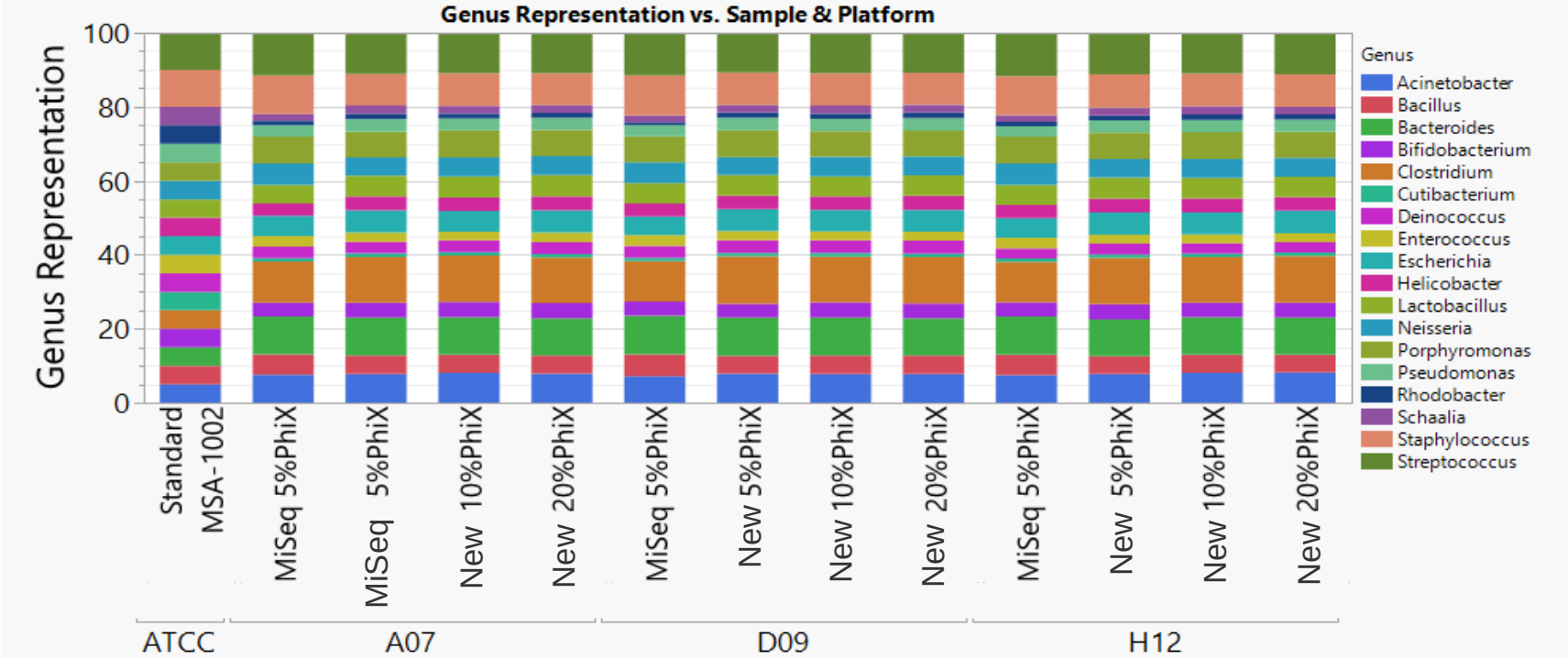
Results were generated using Illumina DNA Prep, 100 ng input
Data generated on a prototype configuration and may change prior to launch

MiSeq i100 Series enables genus-level identification of complex microbial populations

	New series using PhiX 5%
%Q30	93.5
%ER	0.5
%Aligned	4.1

Genus representation of 20 bacteria in reference sample

Genus representation sequenced on MiSeq and new benchtop series



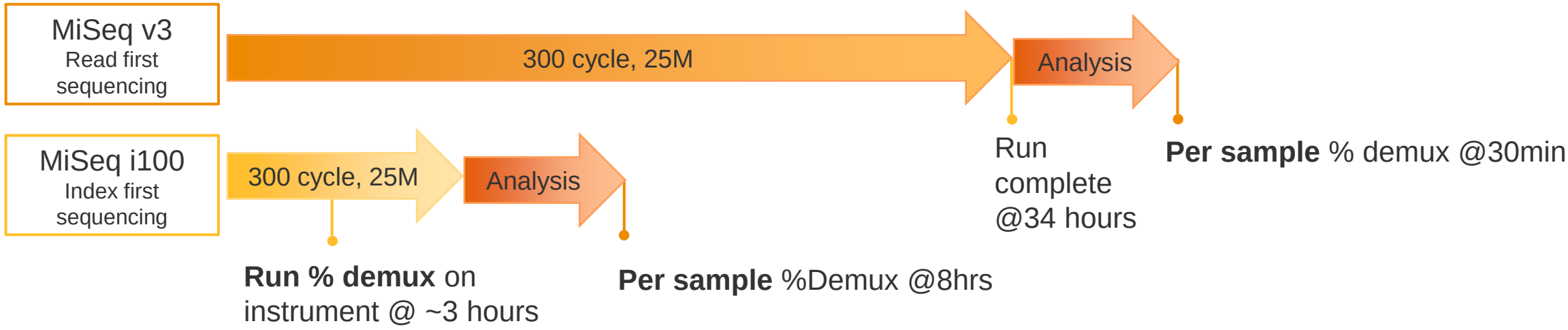
* Data generated on a prototype configuration and may change prior to launch
**PCR amplicon prepped based on 16S Metagenomic protocol

Comparable performance to MiSeq System with 16S metagenomic protocol

Quicker run insights with index first sequencing

Access primary QC results faster informing next steps for rebalancing or planning subsequent runs

Standard sequencing can still be performed, no change in library construction



Onboard and perform NGS workflow with ease

Panels/LP kits

Highly open platform compatible library prep kits from Illumina and third-party vendors and require no additional conversion steps to streamline operations

Automation

Reliable, consistent data with automated library prep on 3rd party automation[1], including Beckman NGenius

Sequencing

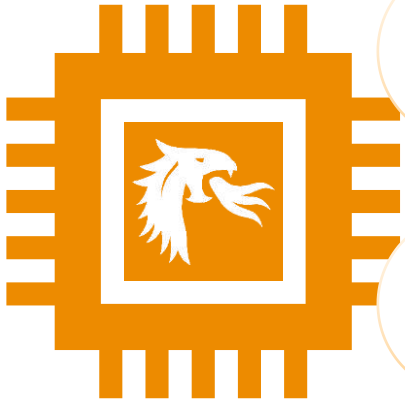
Fastest, simplest benchtop sequencing system

Secondary analysis

Powerful, data analysis pipelines powered by DRAGEN™ onboard with sample-to-analysis workflows

[1] See full lists of compatible 3rd party automation on illumina.com

Streamline data analysis with onboard DRAGEN secondary analysis



DRAGEN BCL Convert

Converts BCL files into FASTQ
Option to auto-launch FASTQ files generated on board to cloud for further analysis
ORA Compression optional

DRAGEN Library QC

Provides quality metrics including Mean Base Quality, Read Length and quality distribution, Mapping, GC % and more. **Rebalancing volumes for pooling libraries to support high-throughput sequencing**

Includes integrated BCL Convert

DRAGEN small WGS

Reference based mapping of small whole genome sequences
Includes pre-set list of 17 organisms or optional custom reference with VCF file generation

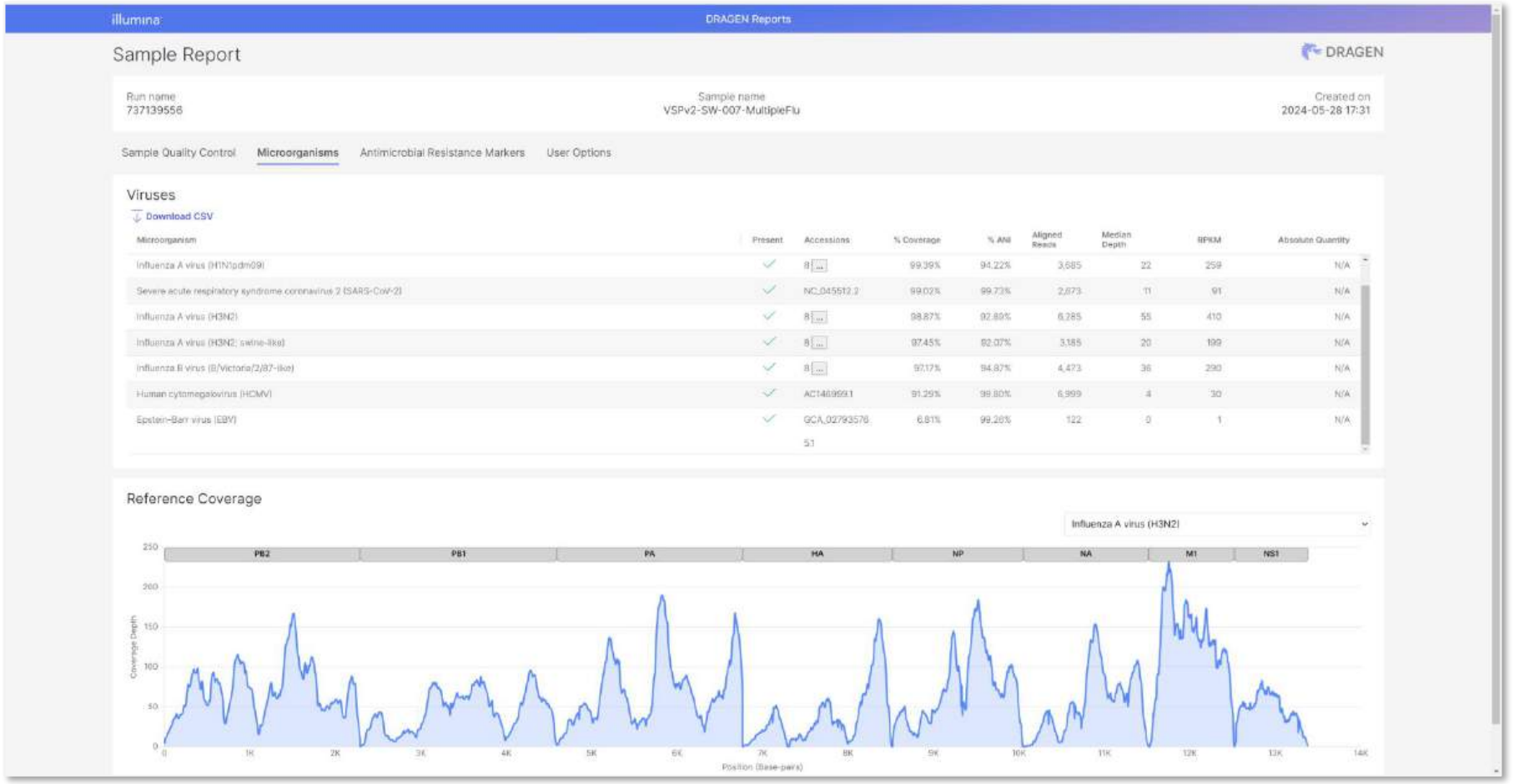
Includes integrated BCL Convert & QC

DRAGEN Microbial Enrichment Plus

Comprehensive secondary analysis of infectious disease and microbiology enrichment panel assays with the same outputs as cloud version (Microbe and/or AMR detection, consensus genome in FASTA format, VCF files, DRAGEN reports and more).

Includes integrated BCL Convert & QC

On board DRAGEN .html report



Example report from **DRAGEN Microbial Enrichment Plus:**

- .html dynamic report with microbial identification, coverage metrics and quantity (if using internal controls)
- Coverage plots for each genome with segments/genes (for segmented viruses)
- Each report includes QC metrics
- Delivers sufficient microbial data for quick, actionable decisions by public health labs and organizations

illumina®

For Research Use Only. Not for use in diagnostic procedures.