

Illumina CMOS Chip and One-Channel SBS Chemistry

The iSeq™ 100 System combines CMOS technology with innovative one-channel SBS chemistry to deliver high-accuracy data in a compact system.

Introduction

Illumina sequencing platforms leverage proven, highly accurate sequencing by synthesis (SBS) chemistry—the most widely adopted next-generation sequencing chemistry in the world.¹ Illumina SBS allows massively parallel sequencing using a proprietary reversible terminator–based method that enables detection of single bases as they are incorporated into growing DNA strands. Illumina SBS chemistry employs natural competition among all four nucleotides, which reduces incorporation bias and allows more robust sequencing of repetitive regions and homopolymers.² Compared to capillary electrophoresis–based Sanger sequencing, NGS can detect a broader range of DNA variants, including low-frequency variants and adjacent phased variants, with a faster time to results and fewer hands-on steps.^{3,4}

While SBS chemistry forms the foundation of all Illumina sequencing platforms, the optical systems have evolved to include four-, two-, and one-channel detection methods. The iSeq 100 System employs the latest evolution in Illumina SBS technology: one-channel SBS on a complementary metal-oxide semiconductor (CMOS) chip. The one-channel SBS method supports lower sequencing costs and the convenience of an instrument footprint measuring just over one cubic foot.

Highly Accurate SBS Chemistry

Illumina SBS chemistry uses reversibly labeled nucleotides to sequence hundreds of millions of clusters on a flow cell surface in parallel (Figure 1). During each sequencing cycle, a single labeled deoxynucleotide triphosphate (dNTP) is added to the nucleic acid chain. The nucleotide label serves as a terminator for polymerization, so that after each dNTP incorporation, the fluorescent dye is imaged and then chemically cleaved to allow incorporation of the next nucleotide. Base calls are made directly from signal intensity measurements during each cycle. Because all four reversible terminator-bound dNTPs (A, C, T, G) are present as single, separate molecules, natural competition minimizes incorporation bias. Furthermore, reversible termination ensures single base incorporation per template strand per sequencing cycle, which greatly improves sequencing accuracy in homopolymeric regions and greatly reduces raw error rates compared to other technologies.²



To see an animated video of Illumina SBS technology, visit www.illumina.com/SBSvideo.

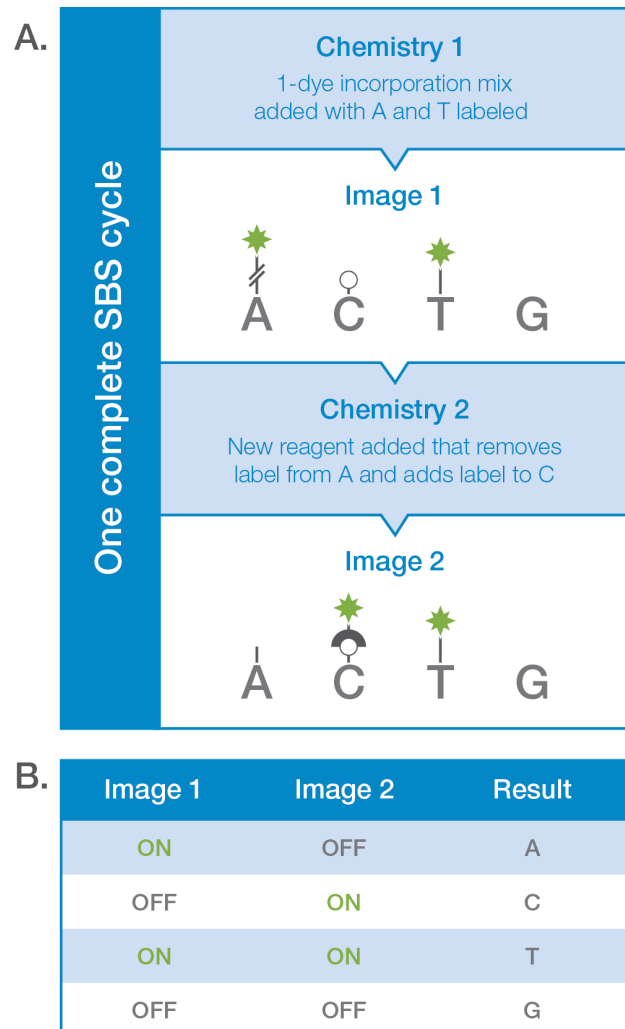


Figure 1: Illumina SBS Chemistry with One-Channel Detection—(A) The iSeq 100 System uses an innovative one-channel chemistry method. Each sequencing cycle includes two chemistry steps and two imaging steps. The first chemistry step exposes the flow cell to a mixture of nucleotides that have fluorescently labeled adenines and thymines. During the first imaging step, the light emission from each cluster is recorded by the CMOS sensor. The second chemistry step removes the fluorescent label from adenine and adds a fluorescent label to cytosine. In both chemistry steps, guanine is dark (unlabeled). The second image is recorded. (B) The combination of Image 1 and Image 2 are processed by image analysis software to identify which bases are incorporated at each cluster position. This sequencing cycle is repeated "n" times to create a read length of "n" bases.

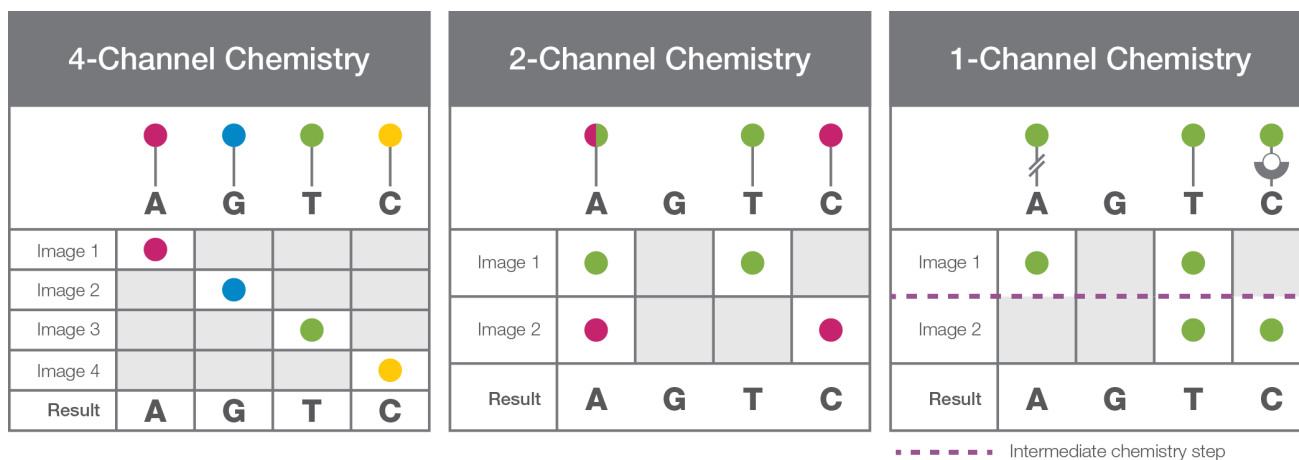


Figure 2: Four-, Two-, and One-Channel Chemistry—Four-channel chemistry uses a mixture of nucleotides labeled with four different fluorescent dyes. Two-channel chemistry uses two different fluorescent dyes, and one-channel chemistry uses only one dye. The images are processed by image analysis software to determine nucleotide identity.

Four-Channel SBS Chemistry

In four-channel SBS, bases are identified using four different fluorescent dyes for each base and four images per sequencing cycle (Figure 2A). The sequencing cycle begins with a chemistry step where all four differentially labeled bases are added to the flow cell. Following nucleotide incorporation, the imaging cycle begins and includes the capture of four distinct images using four different wavelength bands. The images are processed with image analysis software to determine which nucleotides were incorporated at each cluster position across the flow cell. Therefore, with four-channel sequencing, every sequencing cycle requires four dyes and four images to determine the DNA sequence. The MiSeq™ and HiSeq™ Series Systems currently use four-channel SBS.

Two-Channel SBS Chemistry

Rather than using a separate dye for each base, two-channel SBS simplifies nucleotide detection by using two fluorescent dyes and two images to determine all four base calls (Figure 2B). Images are taken using red and green filter bands. Thymines are labeled with a green fluorophore, cytosines are labeled with a red fluorophore, and adenines are labeled with both red and green fluorophores. Guanine is permanently dark. The MiniSeq™, NextSeq™, and NovaSeq™ Systems use two-channel chemistry.

One-Channel SBS Chemistry

The iSeq 100 System combines proven Illumina SBS chemistry with CMOS technology to deliver one-channel sequencing chemistry. The system uses a patterned flow cell with nanowells fabricated over a CMOS chip (Figure 3). Clustering and sequencing occur in the nanowells with direct alignment of single clusters over each photodiode (pixel). Using a CMOS sensor embedded in the consumable is a simple and fast detection method.

Unlike four-channel SBS chemistry, where sequencers use a different dye for each nucleotide, the iSeq 100 System uses one dye, two chemistry steps, and two imaging steps per sequencing cycle. In one-channel chemistry, adenine has a removable label and is labeled in the first image only. Cytosine has a linker group that can bind a label and is labeled in the second image only. Thymine has a permanent fluorescent label and is therefore labeled in both images, and guanine is permanently dark. Nucleotides are identified by analysis of the different emission patterns for each base across the two images (Figure 2C).

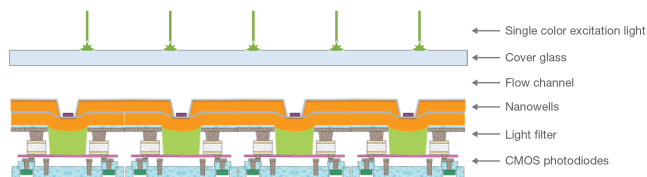


Figure 3: Illumina CMOS Flow Cell—A sequencing library is loaded into the iSeq 100 reagent cartridge, which contains a patterned flow cell fabricated over a CMOS chip. Each well in the flow cell is aligned over a CMOS photodiode. During cluster generation, proprietary ExAmp chemistry ensures that each well in the flow cell generates a single, clonal cluster. During each imaging step, light emissions are detected by the CMOS photodiodes.

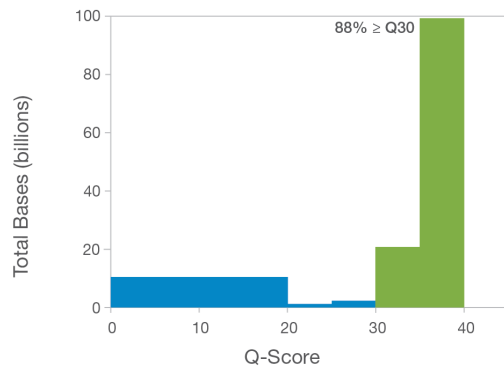
Table 1: Instrument Comparison of Variant Analysis

Systems	Percent On-Target Aligned Reads	Percent Mismatches	Percent Uniformity Of Coverage ^a	Percent Precision (SNP)	Percent Recall (SNP)
Panel 1 (BRCA)^b					
iSeq 100	96.0	0.23	100	90.5	100
MiniSeq	95.8	0.21	100	89.1	100
MiSeq	95.8	0.20	100	89.1	100
Panel 2 (CHP2)					
iSeq 100	97.3	0.20	100	100	100
MiniSeq	97.1	0.17	99.0	100	100
MiSeq	96.8	0.15	100	98.6	100
Panel 3 (Focus)					
iSeq 100	96.5	0.23	99.5	95.2	100
MiniSeq	96.6	0.20	98.9	92.9	100
MiSeq	96.1	0.19	99.0	94.0	100

a. Percent > 0.2x mean

b. The **BRCA** panel includes 265 amplicons spanning across all exon regions of the *BRCA1* and *BRCA2* genes; the **CHP2** panel includes 207 amplicons covering 50 oncogenes and tumor suppressor genes; the **Focus** panel includes 269 amplicons that span 53 genes for multivariant analysis.

A. NextSeq 550 System Run



B. iSeq 100 Run

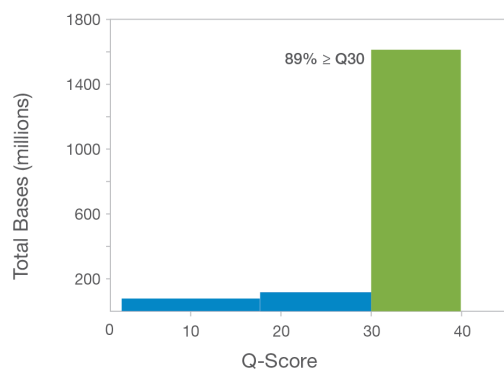


Figure 4: Quality Scores from the NextSeq 500 and iSeq 100 Systems—(A) Two-channel SBS technology on the NextSeq 550 System delivers high-quality data, with a sequencing accuracy of 88% of sequenced bases \geq Q30. Human cell line mixture samples were pooled and run with a 2×101 bp configuration. (B) One-channel SBS chemistry on the iSeq 100 System delivers comparable high-quality data. A metagenomics microbial pool sequenced on a 2×151 bp run provides 89% of bases \geq Q30.

SBS Accuracy Across All Systems

Illumina SBS chemistry delivers highly accurate sequencing data at virtually any coverage level, the highest yield of error-free reads, and the highest percentage of base calls above Q30* in the industry.^{2,6,7} As with other Illumina platforms, the iSeq 100 System generates a high percentage of bases above Q30 (Figure 4). To assess how iSeq 100 System data compares to data generated on the MiniSeq and MiSeq Systems for variant identification, three amplicon panels were run and analyzed on all three systems (Table 1). Data analysis was performed using Burrows-Wheeler aligner (BWA),⁸ Pices variant caller,⁹ and Hap.py Benchmarking software.¹⁰ Precision and recall were calculated with the Platinum Genomes reference data set.¹¹ The results show comparable, high-quality alignment, coverage, and variant analysis metrics across all three systems.

Because all Illumina sequencing systems use the same reversible terminator-based SBS technology, researchers can confidently transition their research from one Illumina sequencing system to another. Regardless of whether data are generated by four-, two-, or one-channel SBS, results can be easily compared and analyzed in BaseSpace™ Sequence Hub, the Illumina genomics computing environment, or across a wide array of third-party analysis tools.

Summary

Illumina SBS chemistry offers highly accurate base-by-base sequencing and robust coverage across the genome. The latest evolution in SBS technology can be found in the iSeq 100 System, which combines CMOS sensor detection with one-channel SBS technology. With simplified chemistry and optics, an entry-level pricepoint, and the convenience of a small footprint, the iSeq 100 System enables the power of NGS to become an everyday research tool for virtually any laboratory.

*A quality score (Q-score) is a prediction of the probability of an error in base calling. A Q-score of 30 (Q30) is widely considered a benchmark for high-quality data.⁵

Learn More

To learn more about the iSeq 100 System, visit www.illumina.com/iseq

To learn more about microbial or mitochondrial sequencing on the iSeq 100 System, read the [Microbial WGS with the iSeq 100 System](#) or the [Mitochondrial DNA Sequencing on the iSeq 100 System](#) Application Notes

To learn more about patterned flow cells, visit the [Patterned Flow Cell](#) page

References

1. Data calculations on file. Illumina, Inc., 2015.
2. Bentley DR, Balasubramanian S, Swerdlow HP, et al. [Accurate Whole Human Genome Sequencing using Reversible Terminator Chemistry](#). *Nature*. 2008;456(7218):53-59.
3. Precone V, Monaco VD, Esposito MV, et al. [Cracking the Code of Human Diseases Using Next-Generation Sequencing: Applications, Challenges, and Perspectives](#). *Biomed Res Int*. 2015;161648.
4. Shokralla S, Porter TM, Gibson JF, et al. [Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform](#). *Sci Rep*. 2015;5:9687.
5. Illumina (2011). [Quality Scores for Next-Generation Sequencing](#). Accessed November 30, 2017.
6. Ross MG, Russ C, Costello M, et al. [Characterizing and measuring bias in sequence data](#). *Gen Biol*. 2013;14:R51.
7. Nakazato T, Ohta T, Bono H. [Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive](#). *PLoS One*. 2013;22;8:e77910.
8. Li H, Durbin R. [Fast and accurate short read alignment with Burrows-Wheeler transform](#). *Bioinformatics*. 2009;25:1754-60.
9. Pisces Variant Caller. github.com/Illumina/Pisces/wiki. Accessed December 4, 2017.
10. Hap.py Benchmarking. www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/hap-py-benchmarking.html. Accessed December 4, 2017.
11. Platinum Genomes. Eberle, MA et al. [A reference data set of 5.4 million phased human variants validated by genetic inheritance from sequencing a three-generation 17-member pedigree](#). *Genome Res* 2017;27:157-164.