illumina®

APPLICATION NOTE

Antibody library screening with the MiSeq™ i100 Series

Increased insert sizes enable rapid, accurate screening of display libraries to support antibody discovery applications

Bring antibody discovery applications in-house with a rapid NGS workflow

Generate high-quality, equivalent data faster than long-read sequencing technologies

Achieve full scFv coverage for display library screening with the MiSeq i100 25M 1000 cycles kit

Introduction

Antibody discovery is a significant endeavor in the biotechnology and pharmaceutical industries. Development and engineering of new monoclonal antibodies (mAbs) fuel a diverse array of applications. These include therapeutic antibodies for immuno-oncology and other disease areas, in support of diagnostics and companion diagnostics, and as research tools and reagents for different assays. The field has expanded rapidly, driven by advances in tumor immunobiology, next-generation sequencing (NGS), single-cell profiling, and protein engineering.

Regardless of the application, antibody discovery typically involves the following steps:

- **Antibody screening**—Identification of candidates done *in vivo,* with hybridoma generation and B-cell cloning in mice or single-cell screening of B- or T-cells, *in vitro* with microbe-based or mammalian display platforms, or *in silico* using computational tools to inform design
- **Lead optimization**—Engineering of candidates to improve protein performance, eg, affinity maturation, biophysical properties, cross-reactivity, etc
- **Functional assays**—Characterization and downselection of thousands of candidates to tens of high-value leads based on assay performance

Often, this process is iterative, with multiple rounds of screening, optimization, and functional characterization required before a final candidate is selected.

Technical advances have led to *in vitro* methods where large antibody libraries (10^9–10^{11} variants) are screened for binding of a target antigen. Typically, these diverse libraries are made up of single-chain variable fragments (scFvs), rather than mAbs. An scFv is a recombinant antibody fragment comprised of a single polypeptide that contains the variable light chain (VL) and variable heavy chain (VH) of an antibody, connected by a short linker peptide (Figure 1). Their small size enables scFvs to be expressed on the coats of bacteriophages in a "phage display" (Figure 2).[1]

A phage display library is screened by incubating the library with an immobilized target antigen. After scFvs with low binding affinity for the antigen are washed away, candidate scFvs are eluted and amplified by infection and proliferation in bacteria. After several rounds of "panning," the library is enriched for scFv clones with high binding affinity for the target antigen (Figure 2).[1] These leads are taken forward into optimization and functional characterization. Sequences of the final scFv leads can be cloned into vectors to produce full antibodies, or used as-is or as components in other engineered proteins for research or clinical applications.

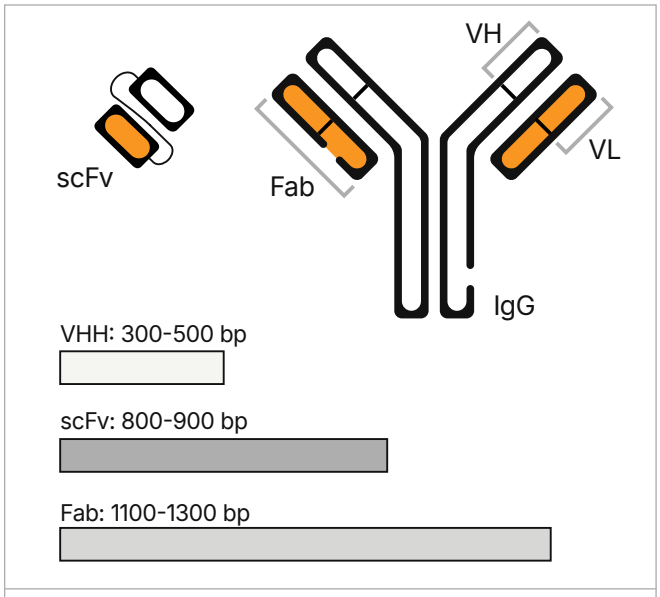

Figure 1: Structure of scFv antibody fragments
An scFV fragment consists of a single polypeptide that includes the variable light (VL) and variable heavy (VH) chains of a monoclonal antibody linked by a short peptide linker; at 800–900 bp in length, the scFv DNA sequence is larger than the variable heavy domain of the heavy chain (VHH) and smaller than the fragment antigen-binding (Fab) region.

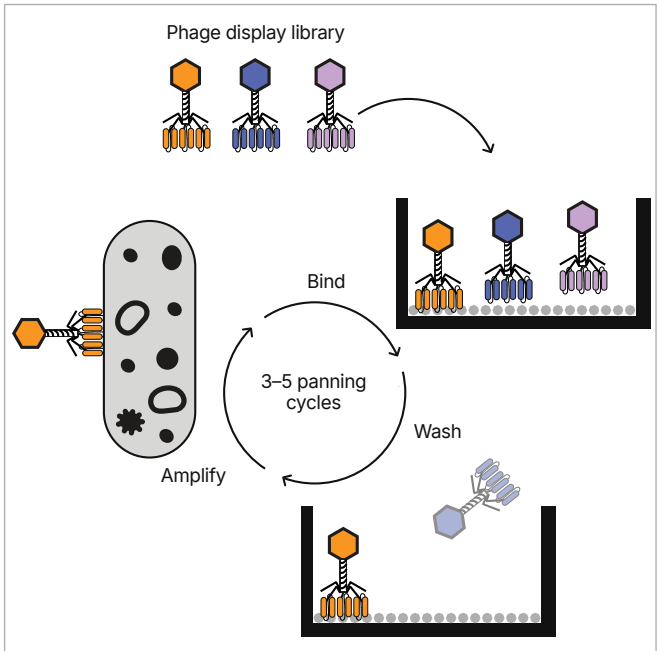

Figure 2: Phage display library panning
Phage display libraries undergo several rounds of panning, consisting of incubation with an immobilized target antigen, washing, and amplification to enrich for clones with high binding affinity for the antigen.

Historically, phage display libraries were analyzed by Sanger sequencing of the phage DNA encoding the scFv. While effective, Sanger sequencing can be time-consuming and limited in scale, offering only a limited view of a diverse phage library.[2] In contrast, the massively parallel nature of NGS can allow for significantly more insights into the actual diversity of a phage library. One limitation is the short read length of NGS platforms, typically ≤ 300 bp for most short-read sequencing platforms, including Illumina sequencing systems; this read length is insufficient to provide full coverage of the scFv sequence (Figure 1).[3] While long-read sequencing technology enables comprehensive scFv sequencing, its low coverage depth may be unable to cover the diversity of a phage library.[4] In contrast, Illumina short-read sequencing delivers higher coverage and data quality. Industry leading short read XLEAP-SBS™ chemistry offers speed and high accuracy to fit the need of antibody screening.

This application note demonstrates that 2 × 500 bp sequencing on the MiSeq i100 Series achieves high-quality, full-length scFv coverage in rapid turnaround time. This scFv sequencing solution on the MiSeq i100 Series enables rapid antibody screening compared to on-market long-read technologies.

Methods

In collaboration with Twist Biosciences, samples were collected from internal phage display library panning and screened with a streamlined NGS workflow (Figure 3).

Samples

The discovery campaign targeted a cell-surface tumor-associated antigen, leveraging its extracellular accessibility for efficient affinity-based enrichment. Antibody selection was performed using the Twist Human Repertoire scFv Library (10^9 diversity), a mixed-germline framework library generated by amplifying full-length VDJ (variable, diversity, joining) regions from ten healthy donors spanning diverse genders and ethnic backgrounds, thereby preserving authentic human antibody architecture and developability. The scFv format, in which VL and VH domains are linked as a single polypeptide, supports compact, high-density display on filamentous bacteriophage and enables efficient antigen engagement during iterative binding, washing, and bacterial amplification. A three-arm, five-round panning workflow was executed to explore distinct selection pressures in parallel and drive convergence toward high-affinity, target-specific clones.

Library preparation

Fifty nanograms of enriched scFv amplicon were used as input for library construction with the Twist Mechanical Fragmentation with Universal Adapter Kit protocol. To preserve full-length inserts, the protocol was modified by omitting the fragmentation step, and the amplicons were carried directly into adapter ligation and subsequent library-prep stages. Adapter-ligated products were purified and amplified with sample-specific indexing primers to generate sequencing-ready libraries. Final libraries were quantified fluorometrically, assessed for size distribution by capillary electrophoresis, and pooled equimolarly prior to sequencing.

Sequencing

Following enrichment, Twist validated leads via NGS, which is essential to overcome the throughput limitations of traditional Sanger sequencing and to provide the deep coverage needed to capture entire VDJ segments and preferential germline pairings. Libraries were sequenced on the MiSeq i100 Plus System (Illumina, Catalog no.

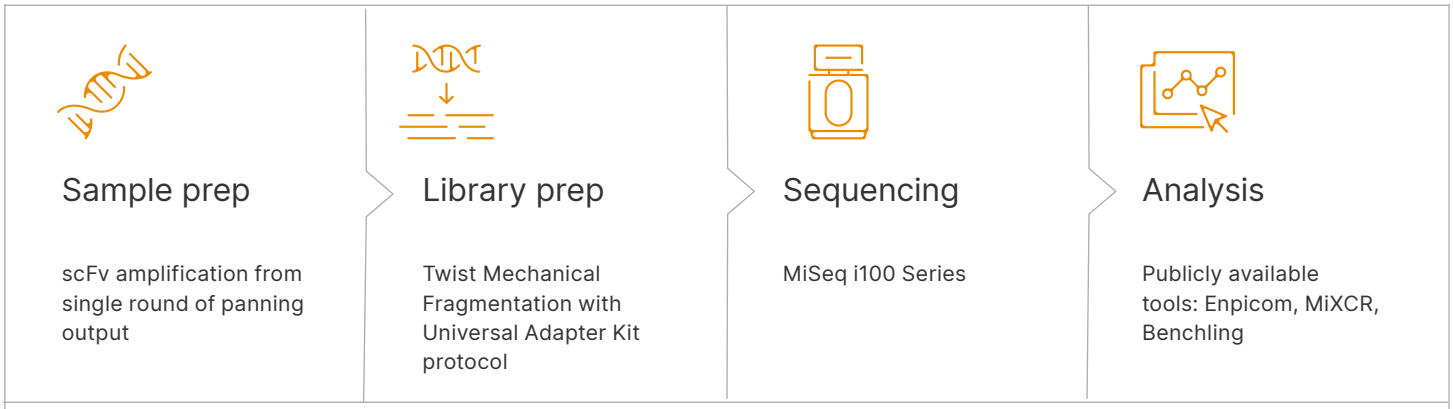

Figure 3: Antibody screening on the MiSeq i100 Series workflow

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

20115695) using the MiSeq i100 Series 25M Reagent kit (1000 cycles) (Illumina, Catalog no. 20148254) with a run configuration of 2 × 500 bp, which delivers high-accuracy, full-length sequence coverage across the ~800–900 bp scFv insert size, enabling precise clonal tracking, cluster network analysis, and the rapid identification of high-affinity leads for downstream therapeutic reformatting.

Data analysis

Analysis of sequencing data was performed by Twist Biosciences using proprietary analysis software (not disclosed). Data analysis can also be performed using publicly available tools such as the ENPICOM Platform, MiXCR, and Benchling analysis solutions.

Phage display library clone identification

Sequencing of scFv phage library panning outputs with the MiSeq i100 Series enabled identification of individual enriched clones and related clones in cluster networks (Figure 7A). The greater depth provided by 2 × 500 bp sequencing on the MiSeq i100 Series enabled study of clonal clustering and enrichment profiles during library panning, which can enhance lead identification and optimization. Also, scFv sequencing with the MiSeq i100 Series enables capture and analysis of entire VDJ segments from heavy and light chains, including preferred germline pairings (Figure 7B). This enables rapid insights into clonotypes that can accelerate antibody discovery and engineering workflows.

Results

Library prep and sequencing QC

An aliquot of the prepared libraries were run on the 4200 TapeStation system (Agilent, Catalog no. G2991BA) to determine average library size and quality control (QC). The resulting library trace showed a single peak at ~1000 bp (Figure 4). It is important to make sure that smaller fragments are removed to optimize 2 × 500 bp sequencing.

Read the Maximizing performance on the MiSeq i100 Series technical note for more information on library loading optimization.

Sequencing on the MiSeq i100 Series resulted in an average %Q30 greater than 90% and a total number of paired-end (PE) reads that exceeded the 25M specification of the flow cell (Table 1). High-quality reads were obtained throughout the sequencing run (Figure 5) with a merged read insert size of ~850 bp (Figure 6A) that represent annotated, unique sequences (Figure 6B). These results support high-confidence downstream analysis.

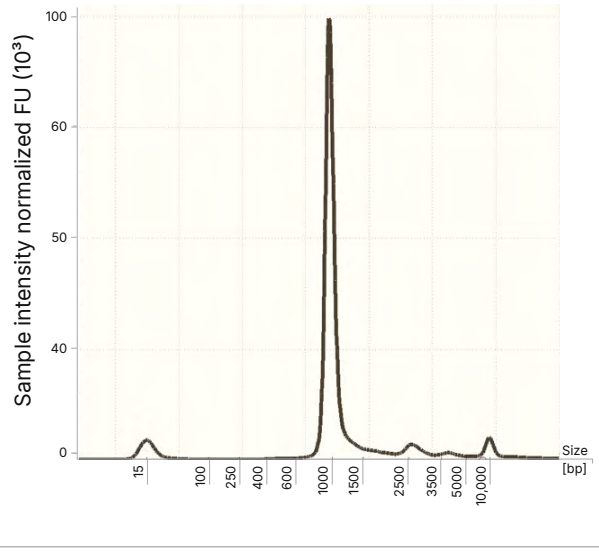

Figure 4: QC trace indicates high-quality libraries
Library QC with the Fragment Analyzer showed a single peak at ~1000 bp.

Table 1: Sequencing metrics pass output and quality specifications

Run	Run configuration	Sequencing run time	% ≥ Q30	PhiX aligned	PhiX error rate
MiSeq i100 2 × 500 bp	10 × 10 × 501 × 501	23 hr 29 min	94.7%	5.6%	0.37%

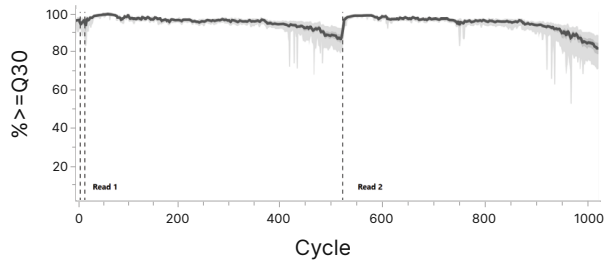

Figure 5: %Q30 by sequencing cycle indicates high-quality reads

High-quality reads were obtained with %Q30 scores > 90% throughout the sequencing run.

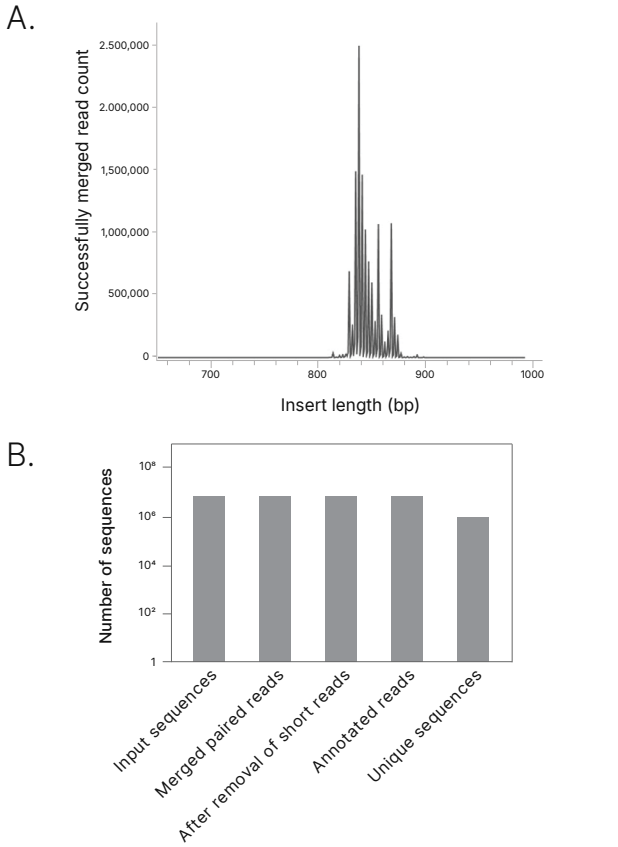

Figure 6: High-quality reads of expected size

Analysis showed (A) a merged read insert size of ~850 bp and (B) quality reads that are merging and not being mistakenly discarded from downstream analyses.

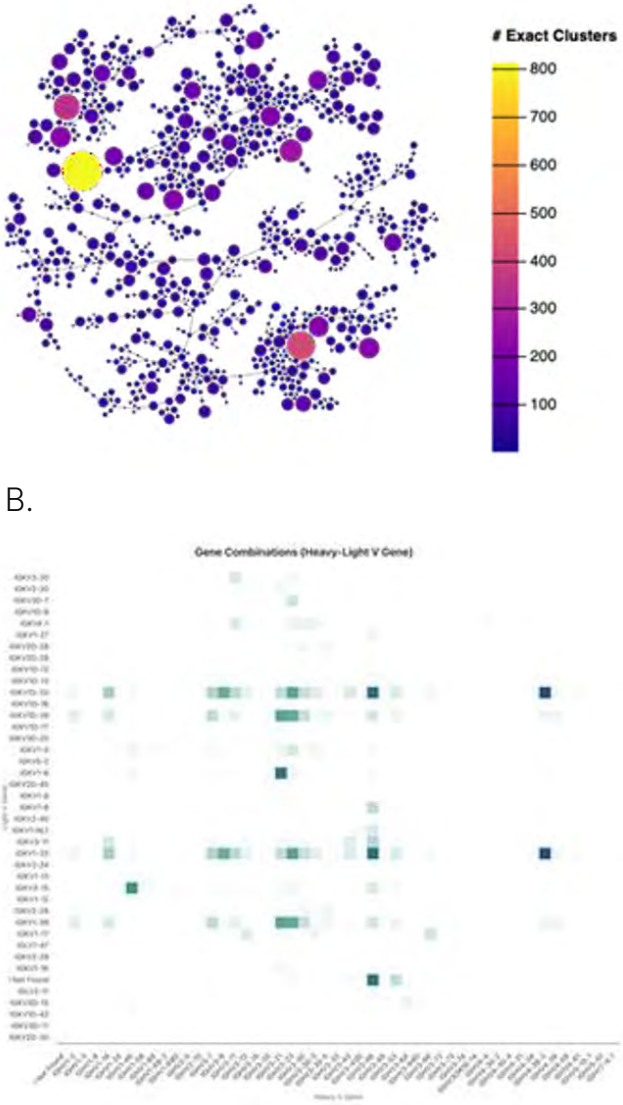

Figure 7: Clonal identification and insights

Sequencing of scFv panning outputs with the MiSeq i100 Series resulted in (A) identification of related clones, visualized as an enriched cluster network using Twist proprietary software and (B) preferences in germline pairings for entire VDJ segments, with the pair count of each heavy chain with each light chain represented in the scatter plot.

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

Data concordance with long-read technology

Comparison of sequencing results showed high concordance between the MiSeq i100Series 2 × 500 bp run and data generated using a native long-read sequencing platform (Figure 8). The equivalent performance of the MiSeq i100 Series combined with its rapid turnaround time (sequencing runs can be completed in < 24 hr, compared to several days for native long-read sequencing) provide a clear advantage.

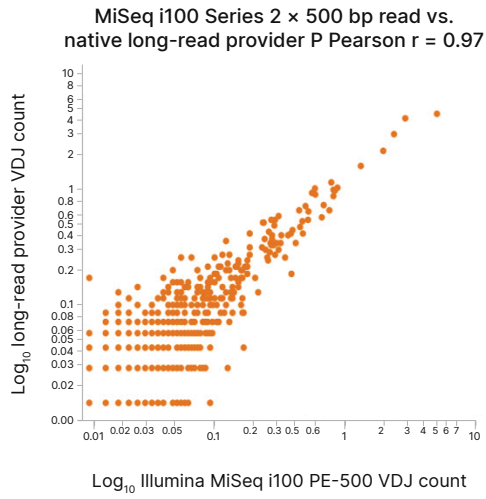

Figure 8: Concordant sequencing results
Comparison of base calls with 2 × 500 bp sequencing on the MiSeq i100 Series showed high concordance with a long-read sequencing provider, with a Pearson correlation coefficient of r = 0.97, indicating a strong linear relationship.

Summary

The MiSeq i100 Series offers simplified operation, flexible configurations, and fast run times with exceptional performance. The MiSeq i100 25M Reagent Kit (1000 cycles) unlocks new applications, such as scFv sequencing for phage display library screening and antibody discovery. With 2 × 500 bp sequencing on the MiSeq i100 Series, customers can conduct scFv discovery, generating high-quality results rapidly in-house, without the need to outsource screening to long-read sequencing providers.

Learn more

MiSeq i100 Series

Twist antibody discovery services

References

1. Kenkel B. Antibodies 101: Single Chain Fragment Variables (scFvs). 2021. Available at: blog.addgene.org/antibodies-101-single-chain-fragment-variables-scfvs. Accessed March 18, 2026.
2. Rouet R, Jackson KJL, Langley DB, Christ D. Next-Generation Sequencing of Antibody Display Repertoires. *Front Immunol.* 2018;9:118. Published 2018 Feb 2. doi:10.3389/fimmu.2018.00118
3. Yang W, Yoon A, Lee S, Kim S, Han J, Chung J. Next-generation sequencing enables the discovery of more diverse positive clones from a phage-displayed antibody library. *Exp Mol Med.* 2017;49(3):e308. Published 2017 Mar 24. doi:10.1038/emm.2017.22
4. Nannini F, Senicar L, Parekh F, et al. Combining phage display with SMRTbell next-generation sequencing for the rapid discovery of functional scFv fragments. *MAbs.* 2021;13(1):1864084. doi:10.1080/19420862.2020.1864084

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

© 2026 Illumina, Inc. All rights reserved. All trademarks are the property of Illumina, Inc. or their respective owners. For specific trademark information, see www.illumina.com/company/legal.html.
M-GL-04206 v1.0