illumina®

TECHNICAL NOTE

Maximizing performance on the MiSeq™ i100 Series

Library loading optimization steps to ensure run success

Optimize library loading

Determine optimal loading concentration for MiSeq i100 Series flow cells

Maximize performance

Improve insert size representation to maximize sequencing performance

Support for low-diversity

Sequence low-diversity libraries by adjusting library complexity with PhiX

Introduction

The MiSeq i100 Series offers the simplest, fastest benchtop sequencing. Breakthrough advancements in system design, XLEAP-SBS™ chemistry, and integrated data analysis deliver enhanced usability, high data accuracy, and exceptional speed, generating results up to 4× faster than the original MiSeq System. As part of an end-to-end NGS solution, the MiSeq i100 Series provides same-day results for various applications, including transcriptomics, microbial genomics, and targeted gene sequencing studies in key areas, such as microbiology, infectious disease, oncology, and more.

When transitioning projects to the MiSeq i100 Series from another sequencing system, optimizing library loading can help maximize data yield and quality. This technical note provides recommendations to optimize results on the MiSeq i100 Series, including guidance on library loading concentration, library quality, and nucleotide diversity considerations.

Optimal library loading

Loading concentration refers to the final concentration of a library loaded onto an instrument for sequencing. After libraries are prepared, they are diluted to the loading concentration appropriate for the library type, sequencing system, and reagent kit.

Loading libraries at a concentration that is too high or too low may lead to lower sequencing quality and yield, and, possibly, run failures in extreme conditions. Underloading may result in a low percentage of nanowell occupancy (% Occupied) and higher duplicate reads, which then requires more reads to achieve target coverage. In contrast, overloading may result in a low percentage of clusters passing filter (PF). To determine the optimal loading concentrations on the MiSeq i100 Series, the % Occupied and % PF metrics can be plotted in Sequencing Analysis Viewer to determine if a run was underloaded, optimally loaded, or overloaded. The approach in the following example experiment can be used to titrate the loading concentration to assess primary and secondary metrics.

To learn more, read Optimizing library loading for Illumina NGSsystems with patterned flow cells

Determining optimal loading concentration

When finding the optimal loading concentration, it is critical to test a wide range of concentrations. Use primary metrics like % PF and % Occupied with secondary metrics like duplicates, insert size, and coverage to measure performance at various loading concentrations to determine the "usable yield" for a given application.

Step 1: Design titration experiment

For transitioning projects from the original MiSeq System to the MiSeq i100 Series, center titrations at ~10.4× the MiSeq Reagent Kit v2 loading concentration and ~6.5× the MiSeq Reagent Kit v3 loading concentration. Recommended centerpoint concentrations vary for different library preparation kits for use with the MiSeq i100 Series (Table 1). For all other cases, it is recommended to use 100 pM for the centerpoint concentration.

In this example, a library pool consisting of bacterial genome samples from *Bacillus pacificus*, *Cereibacter sphaeroides*, and *Escherichia coli* prepared using Illumina DNA Prep was tested at loading concentrations of 40 pM, 80 pM, and 120 pM.

Step 2: Assess nanowell occupancy and clusters PF

Plot the % PF vs % Occupied metrics from the sequencing run for each loading concentration to determine which concentrations resulted in underloading, overloading, or balanced loading. In this example: all three concentrations tested (40 pM, 80 pM, 120 pM) display optimal loading shape (a cloud of points with a positive slope) in the % PF vs % Occupied plot, demonstrating that the MiSeq i100 Series can achieve robust results within a broad library loading concentration range (Figure 1).

Step 3: Assess duplicates

Narrow the target concentration range by analyzing the percent of duplicates. Duplicates tend to decrease with increasing loading concentration. In this example, while all three concentrations tested have duplicates less than 15%, 80 pM and 120 pM had the lowest amount (Figure 2).

Step 4: Analyze insert size

Review the insert sizes. The optimal range for your library and application may vary depending on your workflow requirements. In this example, insert sizes for all three bacterial strains vary across the concentration range tested, with the greatest difference observed between 40 pM and 80 pM (Figure 2).

Table 1: Recommended centerpoint concentrations for titration design with the MiSeq i100 Series

Library preparation kit	Centerpoint concentration
Illumina DNA Prep	80 pM
Illumina DNA Prep with Enrichment	60 pM
Illumina RNA Prep with Enrichment	80 pM
Illumina DNA PCR-Free	120 pM
TruSeq DNA PCR-Free	120 pM
TruSeq DNA Nano	120 pM
Illumina Viral Surveillance Panel v2	80 pM
Illumina Microbial Amplicon Prep—Influenza A/B	80 pM
Respiratory Pathogen ID/AMR Enrichment Panel	80 pM
Urinary Pathogen ID/AMR Panel	80 pM
TruSight RNA Pan Cancer	80 pM
16S rRNA Amplicon	80 pM
Pillar oncoReveal Myeloid Panel	80 pM
Pillar oncoReveal Essential MPN Panel	80 pM
Pillar oncoReveal Multi-Cancer v4 with CNV Panel	80 pM
Pillar oncoReveal *BRCA1* & *BRCA 2* plus CNV Panel	80 pM
PhiX Control v3	120 pM
PhiX Indexed Control (1000 cycles)	120 pM

a. Double-stranded DNA libraries were quantified using the fluorometric Qubit dsDNA Quantitation High Sensitivity assay (Thermo Fisher, Catalog no. Q32851), and Bioanalyzer High Sensitivity DNA Kit (agilent, Catalog no. 5067-4626) for average fragment size estimation. Single-stranded DNA libraries were quantified using the Qubit ssDNA Assay Kit (Thermo Fisher, Catalog no. Q10212).

b. 16S rRNA Amplicon libraries prepared using workflow described in document 16S Metagenomic Sequencing Library Preparation (Part # 15044223 Rev.B).

Step 5: Review other application-dependent metrics (coverage, mapping, etc)

Review additional secondary analysis metrics for optimal performance of your application. In this example, the percent-mapped metric shows robust results for all three loading concentrations tested (Figure 2). Secondary metrics were generated with the DRAGEN™ Small Whole Genome Sequencing app that is available as an on-instrument and on-cloud solution.

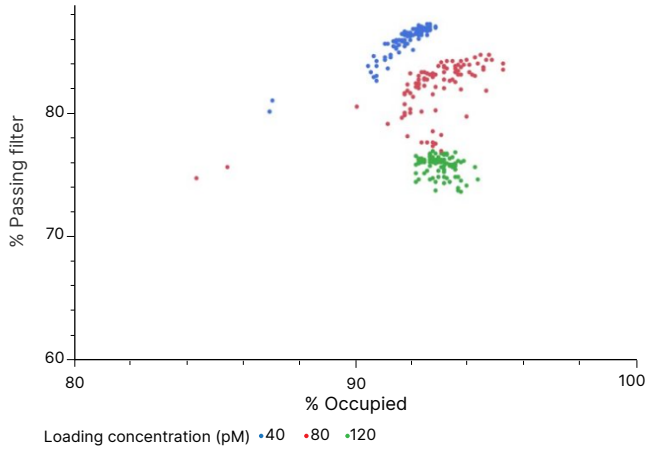

Figure 1: Optimal nanowell occupancy across a broad library loading concentration range

Sequencing of libraries loaded at 40 pM, 80 pM, and 120 pM displayed an optimal loading shape, demonstrating that the MiSeq i100 Series can achieve robust results across a broad range of loading concentrations.

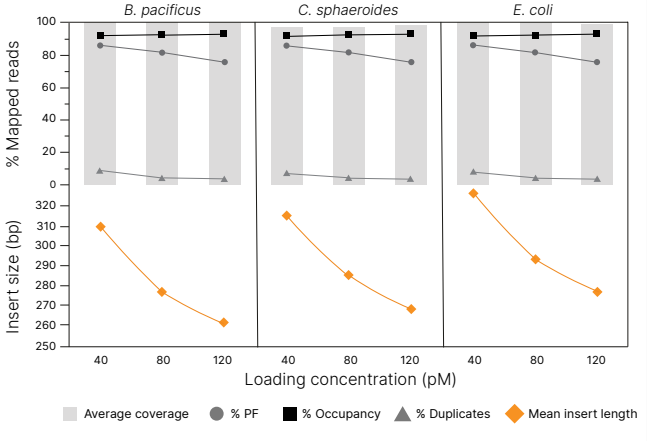

Figure 2: Optimizing sequencing performance on the MiSeq i100 Series

Example titration experiment looking at duplicates, average coverage, and insert size.

Library quality

Short inserts and contaminants introduced during library preparation, including adapter dimers, primer dimers, and partial library constructs, can negatively impact clustering on the MiSeq i100 Series. Short inserts cluster more efficiently than longer inserts. If the sequencing read length is longer than the library insert size, then sequencing will proceed through the insert into the adapter sequence and potentially into the flow cell. When sequencing continues into the flow cell, the read runs out of template for incorporation of bases, causing an intensity drop that on the MiSeq i100 Series can result in one or both of the following effects: a sharp decline in Q30 scores; an increase in G base calls (similar to other Illumina 2-channel instruments such as the NextSeq™ 1000 and NextSeq 2000 Systems, NovaSeq™ 6000 System, and NovaSeq X Series).

To learn more, read How short inserts affect sequencing performance

It is critical to remove these short inserts and contaminants during cleanup or size-selection steps. For optimal performance on the MiSeq i100 Series at 2 × 500 bp read length library mean insert size should range between 600–1200 bp while inserts less than 500 bp should constitute less than 1% of total library mass. If necessary, short inserts and contaminants can be more effectively removed by adding an optional bead purification step to the library preparation protocol. After library preparation is complete and before sequencing, users should verify the quality and purity of all libraries. Use an Agilent Bioanalyzer, Fragment Analyzer system, or TapeStation to check for library integrity, average insert size, and contaminants.

Provided are two examples of additional bead purification procedures used to improve sequencing performance on the MiSeq i100 Series. In the first example, the short inserts of enriched libraries with a broad insert size distribution were selectively removed to eliminate G overcall and improve sequencing performance. In the second example, the removal of adapter dimer contaminants in long insert amplicon library is shown to improve run performance and quality scores for 2 × 500 bp sequencing.

Removing short inserts improves sequencing performance

In this example, libraries of wastewater samples prepared with the Illumina Viral Surveillance Panel v2 Kit were treated with an additional round of bead purification using a bead to sample ratio of 0.8×. The additional round of bead purification effectively removed most of fragments < 250 bp (corresponding to library inserts < 100 bp without adapters), with a total reduction in library yield of ~ 35% (Figure 3).

Viral Surveillance Panel v2 libraries, with and without the additional round of bead purification, were sequenced on the MiSeq i100 Series at a read length of 2 × 150 bp and analyzed with the DRAGEN Microbial Enrichment Plus app. Sequencing of libraries prepared with the additional round of bead purification, as compared to the unmodified protocol, resulted in reduction of G overcall and improved secondary metrics, including mean read length and % post-quality reads, and increased microorganism detection (Figure 4).

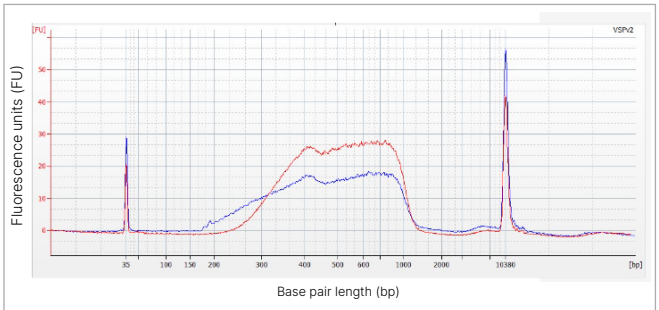

Figure 3: Increased insert size with additional bead purifications
Additional rounds of bead purification (red line) effectively removed most fragments less than 250 bp, compared to the unmodified protocol (blue line).

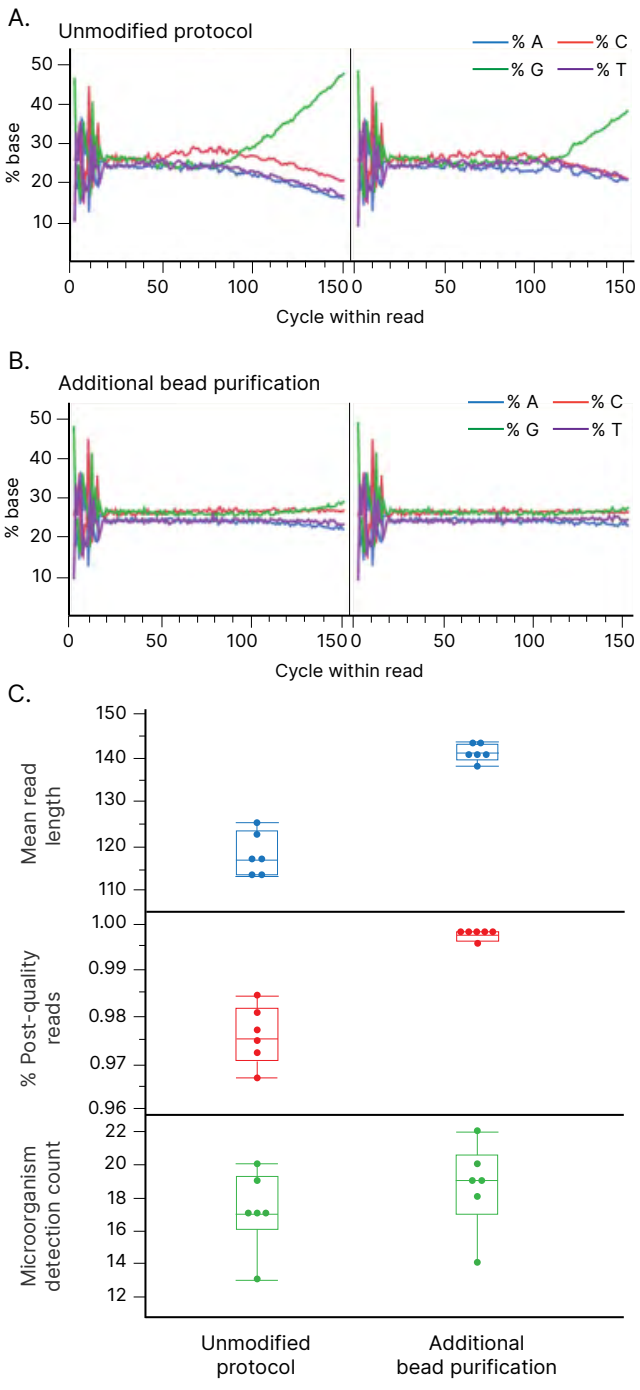

Figure 4: Improved performance with increased insert sizes

Sequencing of libraries generated following a (A) unmodified and (B) modified protocol (with increased insert sizes) on the MiSeq i100 Series reduced G overcall; (C) analysis with the DRAGEN Microbial Enrichment Plus App showed improved performance, including increased mean read length (for 2 × 300 and 2 × 500 bp reads), percent post-quality reads, and microorganism detection.

Removing adapter dimer contaminants optimizes sequencing performance

In this example, custom amplicon libraries with a target insert length of 950 bp were treated with an additional three successive rounds of bead purification using a bead to sample ratio of 0.6x. Library QC on the Tapestation shows that the additional rounds of bead purification selectively removed most of the adapter dimer contaminants at ~ 250 bp fragment length while retaining the target library (Figure 5).

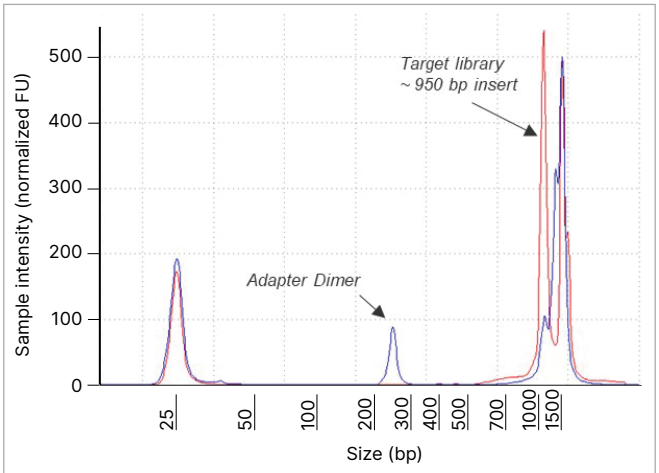

Figure 5: Adapter dimer removal with additional bead purifications

Additional rounds of bead purification of custom amplicon libraries (red line) effectively removed adapter dimer contaminants, compared to the unmodified protocol (blue line).

The custom amplicon libraries, with and without the additional bead purification, were sequenced on the MiSeq i100 Series at read length of 2 × 500 bp. The sequencing run with the original libraries shows performance issues starting at around cycle 100 that are attributable to the presence of the adapter dimers, characterized by sharp drops in signal intensity, declining Q30 scores and increase in G basecalling. Sequencing of the libraries with the additional bead purifications shows improved performance (Figure 6).

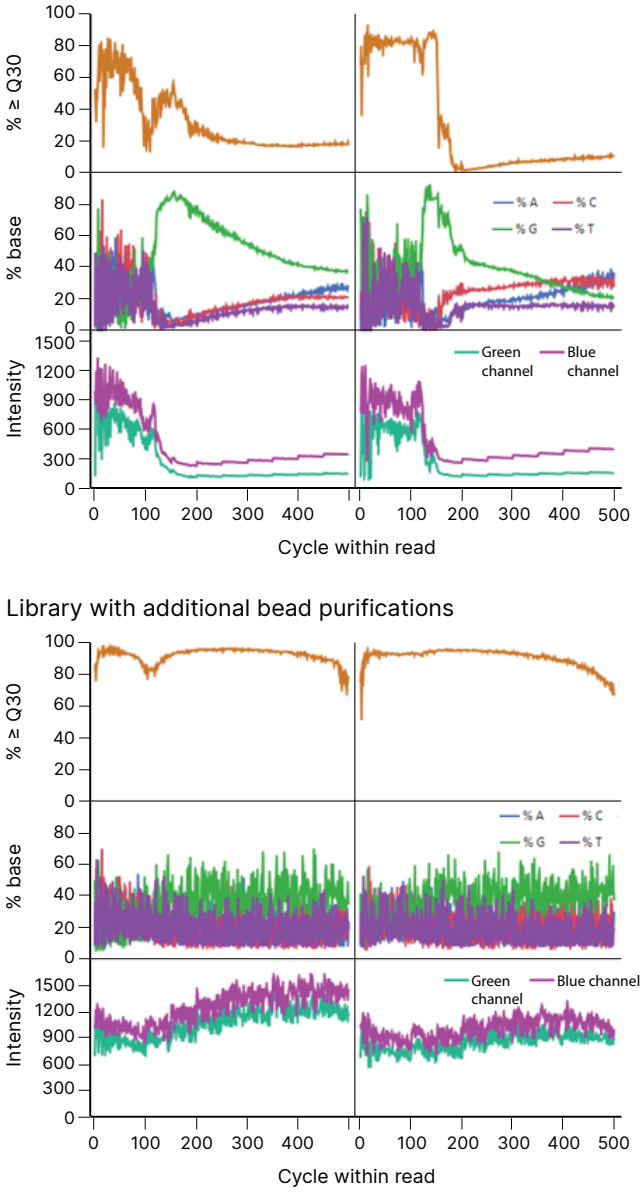

Figure 6: Rescued sequencing performance with adapter dimer reduction

MiSeq i100 Series 2 × 500 bp sequencing of custom amplicon libraries treated with additional bead purifications to reduce adapter dimers showed significant performance improvements, with increased quality scores and a reduction in G overcall.

Nucleotide diversity

Nucleotide diversity indicates the relative proportion of each base (A, C, G, or T) present in every cycle of the run. Nucleotide balance is important for color matrix correction and intensity normalization by the sequencing system. The adaptive Real-Time Analysis software onboard the MiSeq i100 Series has been carefully developed for accurate basecalling of low-diversity libraries. Optimal performance of low-diversity library sequencing can be achieved with a minimum % PhiX spike-in (≥ 5%) to maximize the number of high-quality reads.

In this example, low-diversity 16S amplicon libraries with 5% and 20% PhiX spike-in sequenced on the MiSeq i100 Series show robust performance comparable to the performance of high-diversity human Illumina DNA Prep libraries (Figure 7).

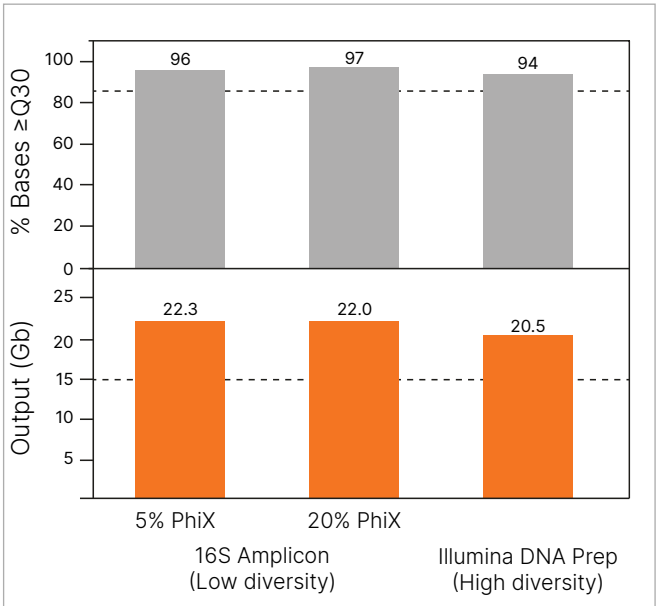

Figure 7: Support for low-diversity libraries

Software onboard the MiSeq i100 Series optimizes sequencing performance for low-diversity libraries, as seen by % bases ≥ Q30 and Gb output. All runs were sequenced at 2 × 301 bp read length using the MiSeq i100 Series 25M Reagent Kit (600 cycles), with dashed lines representing performance specifications.

Summary

Breakthrough advancements in sequencing chemistry and integrated data analysis on the MiSeq i100 Series deliver enhanced usability, high data accuracy, and exceptional speed. Following the best practices outlined in this technical note to assess library quality, optimize loading concentration, and pool libraries can maximize performance on the MiSeq i100 Series.

Learn more

MiSeq i100 and MiSeq i100 Plus Sequencing Systems

illumina®

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

M-GI-03322 v2.0