MiSeq™ i100 시리즈로 최적의 성능 확보

성공적인 런을 위한 모범 사례

라이브러리 로딩 최적화

MiSeq i100 시리즈의 플로우 셀 사용 시 적합한 로딩 농도 결정

최적의 성능 확보

라이브러리 삽입 크기 결과를 개선하여 시퀀싱 성능 향상

다양성이 낮은 라이브러리의 사용 지원

PhiX spike-in으로 분석이 어려운 다양성이 낮은 라이브러리의 시퀀싱 지원

연구 전용입니다. 진단 절차에는 사용할 수 없습니다.

소개

MiSeq i100 시리즈는 Illumina 시퀀서 중 가장 간편하고 빠른 벤치탑 시퀀싱을 제공합니다.[1] MiSeq i100 시리즈는 획기적으로 발전된 시스템 디자인, XLEAP-SBS™ chemistry, 통합된 데이터 분석 기능을 기반으로 향상된 사용성, 우수한 데이터 정확도, 매우 빠른 처리 속도를 제공하여 기존 MiSeq 시스템보다 최대 4배 빠른 속도로 결과를 도출합니다.[1] 엔드투엔드 차세대 시퀀싱(next-generation sequencing, NGS) 솔루션의 한 구성 요소인 MiSeq i100 시리즈는 미생물학, 감염병, 종양학 등 핵심 분야에서 전사체학(transcriptomics) 연구, 미생물 유전체학(microbial genomics) 연구, 표적 유전자 시퀀싱(targeted gene sequencing) 연구 등 다양한 연구 애플리케이션에 활용 시 하루 안에 시퀀싱 결과를 제공할 수 있습니다.

라이브러리 로딩을 최적화하면 MiSeq i100 시리즈의 데이터 수율(yield) 및 품질을 크게 향상시킬 수 있습니다. 이 Technical Note는 MiSeq i100 시리즈로 향상된 결과를 얻을 수 있도록 라이브러리 로딩 농도(loading concentration), 라이브러리 품질 및 뉴클레오티드 다양성(nucleotide diversity) 고려 사항을 비롯한 다양한 권장 사항을 안내해 드립니다.

최적의 라이브러리 로딩

로딩 농도란 시퀀싱을 위해 시퀀싱 시스템에 로딩된 라이브러리의 최종 농도를 의미합니다. 라이브러리를 준비한 후에는 사용하는 라이브러리의 종류, 시퀀싱 시스템 및 시약 키트에 적합한 로딩 농도로 라이브러리를 희석해야 합니다.

너무 높거나 낮은 농도로 라이브러리를 로딩하면 시퀀싱 품질 및 수율이 저하될 수 있으며 극한 조건에서는 런 실패를 초래할 수 있습니다. 언더로딩(Underloading)은 나노웰(nanowell) 이용률(% Occupied)을 낮추고 중복 리드(duplicate read) 수를 증가시킬 수 있어, 표적 커버리지(target coverage)를 달성하기 위해서는 더 많은 리드가 필요할 수 있습니다. 반면 오버로딩(overloading)은 낮은 %의 필터 통과 클러스터(PF)를 야기할 수 있습니다. MiSeq i100 시리즈에 최적화된 로딩 농도를 결정하기 위해 Sequencing Analysis Viewer(시스템에서 바로 실행 가능)에서 % Occupied 및 % PF 메트릭스를 플로팅하여 각 런에 대한 언더로딩, 최적의 로딩 또는 오버로딩 결과를 확인합니다. 다음의 실험 방법 예시를 참고해 로딩 농도를 적정(titration)하고 1차 및 2차 메트릭스를 평가하시기 바랍니다.

자세한 정보는 Optimizing library loading for Illumina NGS systems with patterned flow cells를 참조하시기 바랍니다.

최적의 로딩 농도 확인 방법

최적의 로딩 농도를 파악할 때는 다양한 농도를 테스트하는 것이 매우 중요합니다. 이때 % PF 및 % Occupied와 같은 1차 메트릭스와 Duplicates, Insert size, Coverage와 같은 메트릭스를 함께 활용하여 다양한 로딩 농도에서 성능을 측정함으로써 해당 애플리케이션에 "적용 가능한 수율"을 확인합니다.

1단계: 적정 실험 설계

기존 MiSeq 시스템에서의 프로젝트를 MiSeq i100 시리즈로 옮기려면, MiSeq Reagent Kit v2(Illumina, 카탈로그 번호: MS-102-2001) 로딩 농도의 약 10.4배, MiSeq Reagent Kit v3(Illumina, 카탈로그 번호: MS-102-3001) 로딩 농도의 약 6.5배를 중심으로 적정을 진행합니다. 권장되는 중심 농도는 MiSeq i100 시리즈와 어떤 라이브러리 프렙 키트를 사용하는지에 따라 결정됩니다(표 1). 그 외의 모든 경우에는 중심 농도로 100 pM을 사용하는 것이 권장됩니다.

다음 예시에서는 Illumina DNA Prep을 사용하여 준비한 *Bacillus pacificus*, *Cereibacter sphaeroides* 및 *Escherichia coli*의 유전체 샘플로 구성된 라이브러리 풀(pool)을 40 pM, 80 pM 및 120 pM의 농도로 로딩하여 테스트했습니다.

2단계: 나노웰 이용률 및 필터 통과 클러스터 확인

각 로딩 농도에 대해 레인별로 % PF 및 % Occupied 메트릭스를 플로팅하여 어떤 농도값이 언더로딩, 오버로딩, 또는 균형 있는 로딩 결과를 가져왔는지 확인합니다. 이 예시에서는 테스트한 3가지 농도값(40 pM, 80 pM, 120 pM)은 모두 % PF vs % Occupied 플롯에서 최적의 로딩 모양(양의 기울기를 가진 점 구름)을 보여, MiSeq i100 시리즈가 넓은 라이브러리 로딩 농도 범위 내에서 안정적인 결과를 도출할 수 있음이 입증되었습니다(그림 1).

3단계: 중복 리드 확인

중복 리드의 비율(% Duplicates)을 분석해 목표 농도 범위를 좁힙니다. 로딩 농도가 증가하면 중복 리드의 수는 감소하는 경향이 있습니다. 이 예시에서 테스트한 3가지 농도값은 모두 15% 미만의 중복 리드를 보였으며 80 pM 및 120 pM의 농도에서 가장 적게 나타났습니다(그림 2).

4단계: 라이브러리 삽입 크기 분석

라이브러리 삽입 크기를 검토합니다. 실제 라이브러리 및 애플리케이션에 알맞은 최적의 범위는 워크플로우 요구 사항에 따라 다를 수 있습니다. 이 예시에서는 3가지 세균 균주 모두 라이브러리 삽입 크기가 테스트한 농도 범위에 걸쳐 차이를 보였으며, 특히 40 pM과 80 pM 사이의 농도에서 가장 큰 차이가 관찰되었습니다(그림 2).

표 1: MiSeq i100 시리즈를 사용한 적정 계획 시 권장되는 중심 농도

라이브러리 프렙 키트[a,b]	중심 농도
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
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* & *BRCA2* plus CNV Panel	80 pM
PhiX Control v3	120 pM
PhiX Indexed Control (1000 cycles)	120 pM

a. 이중 가닥 DNA(Double-stranded DNA) 라이브러리의 정량에는 형광 측정 assay인 Qubit dsDNA Quantitation High Sensitivity Assay(Thermo Fisher Scientific, 카탈로그 번호: Q32851)를 이용하였으며, 절편의 평균 크기 추정에는 Bioanalyzer High Sensitivity DNA Kit(Agilent, 카탈로그 번호: 5067-4626)를 이용함. 단일 가닥 DNA(Single-stranded DNA) 라이브러리의 정량에는 Qubit ssDNA Assay Kit(Thermo Fisher Scientific, 카탈로그 번호: Q10212)를 이용함

b. 표에 포함되지 않은 종류의 라이브러리는 적정 런을 통해 최적의 로딩 농도를 결정해야 함. 로딩 농도를 최적화할 때는 100 pM을 중심으로 적정 실험을 수행한 후 25~50 pM 단위로 미세 조정해야 함. 삽입 크기가 큰 라이브러리에는 일반적으로 더 높은 로딩 농도가 필요함

5단계: 기타 애플리케이션 의존적 메트릭스(Coverage, Mapping 등) 검토

실제 애플리케이션에 최적화된 성능을 확보하기 위해 추가적으로 2차 분석 메트릭스를 검토합니다. 이 예시에서 % Mapped reads를 보면, 테스트한 3가지 로딩 농도값 모두가 안정적인 결과를 도출한 것을 확인할 수 있습니다(그림 2). 2차 메트릭스는 시퀀싱 시스템 또는 클라우드에서 이용 가능한 DRAGEN™ Small Whole Genome Sequencing 앱으로 생성했습니다.

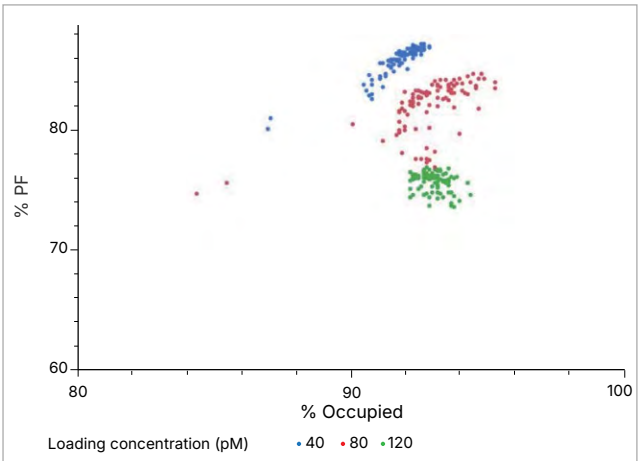

그림 1: 넓은 라이브러리 로딩 농도 범위에 걸쳐 확인된 최적의 나노웰 이용률

40 pM, 80 pM 및 120 pM의 농도로 로딩된 라이브러리를 시퀀싱했을 때 최적의 로딩 모양을 보여, MiSeq i100 시리즈가 넓은 라이브러리 로딩 농도 범위에 걸쳐 안정적인 결과를 도출할 수 있음이 입증됨

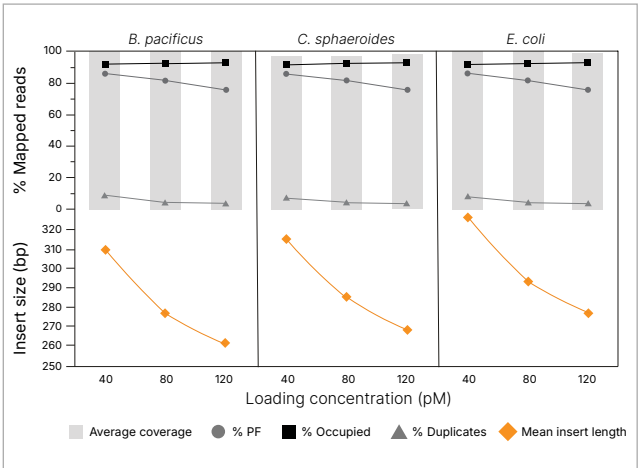

그림 2: MiSeq i100 시리즈를 사용한 시퀀싱 런 최적화

중복 리드 비율, 평균 커버리지, 라이브러리 삽입 크기를 분석한 적정 실험 예시

연구 전용입니다. 진단 절차에는 사용할 수 없습니다.

라이브러리 품질

삽입 길이가 짧은 라이브러리와 라이브러리 준비 과정에서 유입되는 오염물(예: 어댑터 다이머(adapter dimer), 프라이머 다이머(primer dimer), 불완전하게 제작된 라이브러리)은 MiSeq i100 시리즈의 클러스터링 결과에 부정적인 영향을 줄 수 있습니다. 삽입 길이가 짧은 라이브러리는 삽입 길이가 긴 라이브러리보다 효율적으로 클러스터링됩니다. 시퀀싱 리드 길이가 라이브러리 삽입 크기보다 길면, 삽입된 라이브러리뿐만 아니라 어댑터 시퀀스(adapter sequence)를 지나 플로우 셀(flow cell)까지도 시퀀싱될 수 있습니다. 플로우 셀까지 시퀀싱되면, 리드가 염기를 삽입할 템플릿(template, 주형)이 부족해져 강도(intensity)가 감소되며, MiSeq i100 시리즈 사용 시 이는 곧 Q30 점수의 급격한 하락 및/또는 다른 Illumina 2채널 시퀀싱 시스템(NextSeq™ 1000 및 NextSeq 2000 시스템, NovaSeq™ 6000 시스템, NovaSeq X 시리즈)과 마찬가지로 G 베이스 콜(base call)의 증가로 이어질 수 있습니다.

자세한 정보는 How short inserts affect sequencing performance를 참조하시기 바랍니다.

따라서 이렇게 삽입 길이가 짧은 라이브러리나 오염물은 클린업(cleanup) 단계나 크기 선별(size selection) 단계에서 제거하는 것이 매우 중요합니다. MiSeq i100 시리즈를 사용하여 2 × 500 bp의 리드 길이로 최적의 성능을 확보하려면, 평균 라이브러리 삽입 크기는 600 bp에서 1,200 bp 사이여야 하며 삽입 크기가 500 bp 미만인 라이브러리는 전체 라이브러리 양의 1% 미만을 차지해야 합니다. 필요시 기존의 라이브러리 준비 프로토콜에 선택적으로 비드 정제(bead purification) 단계를 추가해 더 효과적으로 삽입 길이가 짧은 라이브러리와 오염물을 제거해 볼 수도 있습니다. 라이브러리가 준비되면, 연구자는 시퀀싱을 시작하기에 앞서 모든 라이브러리의 품질과 순도(purity)를 확인해야 합니다. Agilent사의 Bioanalyzer나 Fragment Analyzer 시스템 또는 TapeStation 시스템을 사용해 라이브러리 무결성(integrity), 평균 라이브러리 삽입 크기 및 오염물을 확인합니다.

다음 섹션에는 MiSeq i100 시리즈의 시퀀싱 성능을 개선하기 위해 적용한 추가적인 비드 정제 단계의 두 가지 예시가 제공되어 있습니다. 첫 번째 예시에서는 삽입 크기 분포가 광범위한 인리치먼트를 거친 라이브러리 중 짧은 라이브러리를 선별적으로 제거함으로써 G 오버콜(overcall)을 없애고 시퀀싱 성능을 높였습니다(그림 3). 두 번째 예시에서는 삽입 길이가 긴 앰플리콘 라이브러리에서 어댑터 다이머 오염물을 제거함으로써 2 × 500 bp 시퀀싱 런 성능 및 Q-Score를 향상시켰습니다(그림 4).

짧은 라이브러리의 제거로 향상된 시퀀싱 성능

이 예시에서는 Illumina Viral Surveillance Panel v2 Kit (Illumina, 카탈로그 번호: 20108081)를 이용하여 준비한 폐수 샘플의 라이브러리를 가지고 0.8×의 비드 대 샘플 비율로 추가적인 비드 정제를 진행했습니다. 추가적인 비드 정제를 통해 길이가 250 bp 미만인 절편(즉, 어댑터 제외 삽입 크기가 < 100 bp인 라이브러리)이 대부분 효과적으로 제거되었으며, 라이브러리 수율은 약 35% 감소했습니다(그림 3).

Viral Surveillance Panel v2 라이브러리는 추가적인 비드 정제를 적용한 조건과 적용하지 않는 조건에서 MiSeq i100 시리즈를 사용하여 2 × 150 bp의 리드 길이로 시퀀싱하였으며, 시퀀싱 데이터는 DRAGEN Microbial Enrichment Plus 앱으로 분석했습니다. 추가적인 비드 정제를 적용하여 준비한 라이브러리를 시퀀싱하였을 때는 변경하지 않은 프로토콜을 그대로 사용했을 때보다 G 오버콜이 감소했고 Mean read length나 % post-quality reads와 같은 2차 메트릭스가 향상되었으며 Microorganism detection count도 증가했습니다(그림 4).

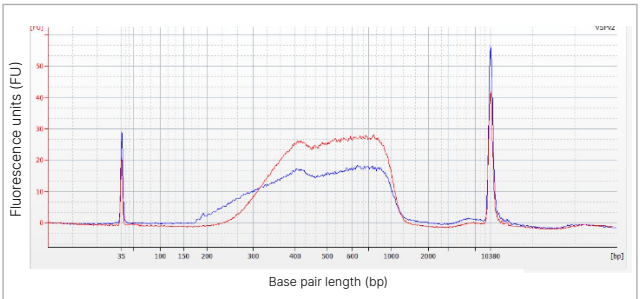

그림 3: 추가적인 비드 정제 시 증가된 라이브러리 삽입 크기
Viral Surveillance Panel v2를 사용해 라이브러리를 준비하는 단계에서 0.8×의 비드 대 샘플 비율 적용 시 추가적인 비드 정제를 적용한 프로토콜(빨간색 선)은 변경하지 않은 프로토콜(파란색 선)과 비교했을 때 길이가 250 bp 미만인 절편 대부분을 효과적으로 제거하였음

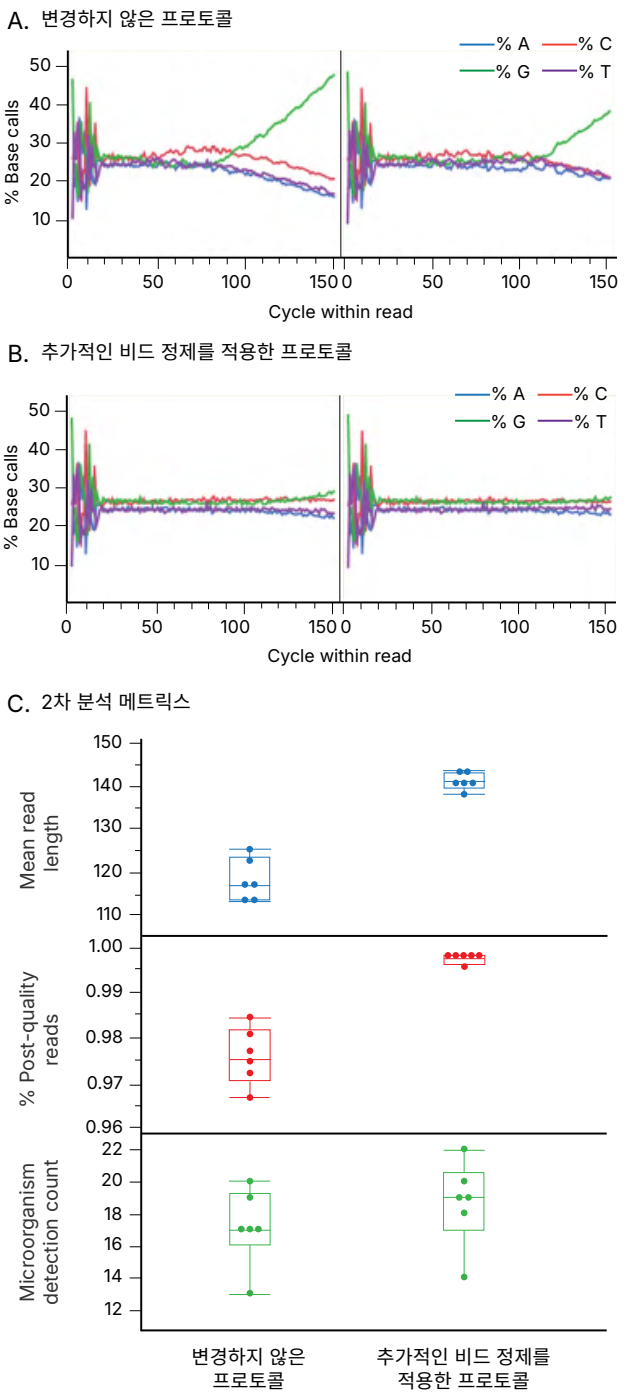

그림 4: 라이브러리 삽입 크기 증가 시 향상된 성능

(A) 변경하지 않은 프로토콜과 (B) 변경한 프로토콜(라이브러리 삽입 크기 증가)에 따라 준비한 라이브러리를 MiSeq i100 시리즈로 시퀀싱했을 때 G 오버콜이 감소하였음. (C) DRAGEN Microbial Enrichment Plus 앱으로 데이터를 분석한 결과, 평균 리드 길이(2 × 300 bp 및 2 × 500 bp 길이의 리드), 고품질 리드 비율(%) 및 미생물 검출 횟수가 모두 증가한 향상된 성능을 보여줌

어댑터 다이머 오염물 제거로 최적화된 시퀀싱 성능

이 예시에서는 표적 삽입 길이가 950 bp인 커스텀 앰플리콘 라이브러리를 가지고 0.6×의 비드 대 샘플 비율로 추가적인 비드 정제를 3번 연속적으로 진행했습니다. TapeStation 시스템으로 라이브러리 QC를 실행한 결과, 추가적인 비드 정제를 통해 절편 길이가 약 250 bp인 어댑터 다이머 오염물이 대부분 제거되었고 표적 라이브러리는 그대로 유지된 것을 확인할 수 있었습니다(그림 5).

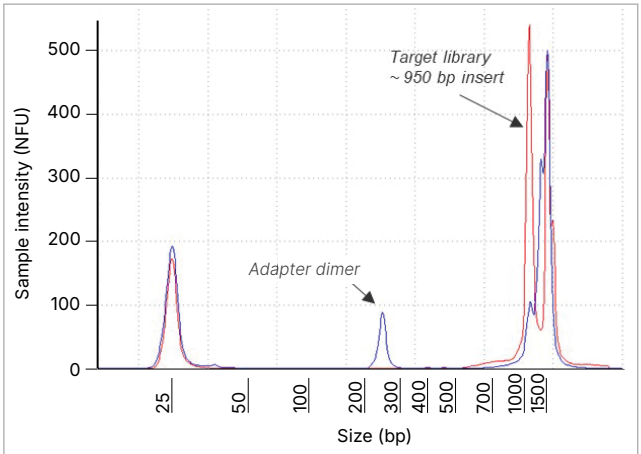

그림 5: 추가적인 비드 정제를 통해 제거된 어댑터 다이머

커스텀 앰플리콘 라이브러리에 추가적인 비드 정제를 적용한 프로토콜(빨간색 선)은 변경하지 않은 프로토콜(파란색 선)보다 효과적으로 어댑터 다이머 오염물을 제거하였음.
NFU = normalized fluorescence unit(정규화 형광 단위)

커스텀 앰플리콘 라이브러리는 추가적인 비드 정제를 적용한 조건과 적용하지 않는 조건에서 MiSeq i100 시리즈를 사용하여 2 × 500 bp의 리드 길이로 시퀀싱하였습니다. 그 결과, 추가적인 비드 정제를 거치지 않은 라이브러리를 시퀀싱한 런에서는 100사이클 무렵부터 시그널 강도의 급격한 감소, Q30 점수의 하락, G 베이스 콜의 증가와 같은 어댑터 다이머로 인한 성능 이슈가 관찰되었습니다. 반면 추가적인 비드 정제를 적용한 라이브러리를 시퀀싱했을 때는 향상된 성능을 확인할 수 있었습니다(그림 6).

연구 전용입니다. 진단 절차에는 사용할 수 없습니다.

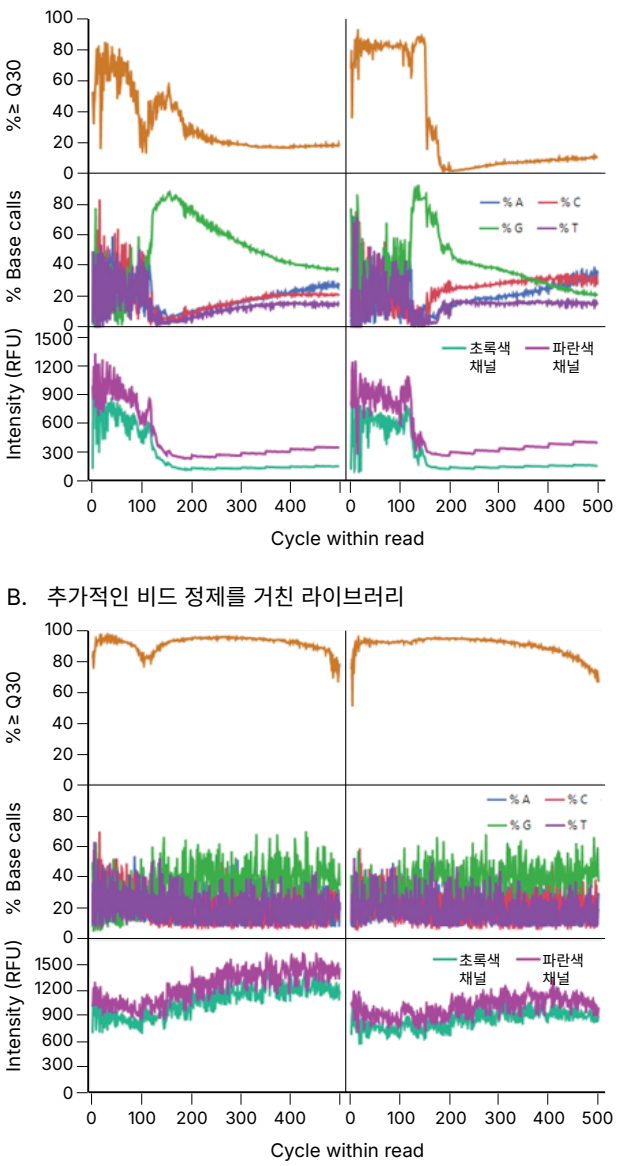

그림 6: 어댑터 다이머 감소에 따라 향상된 시퀀싱 성능

MiSeq i100 시리즈에서 맞춤형 앰플리콘 라이브러리를 2 × 500 bp로 시퀀싱한 결과, (A) 추가적인 비드 정제 전보다 (B) 비드 정제를 통해 어댑터 다이머를 줄인 후 Q-Score가 향상되고 G 오버콜이 감소하는 등 시퀀싱 성능이 크게 향상됨 %≥ Q30 = Q-Score가 30 이상인 염기의 비율, RFU = relative fluorescence unit(상대적 형광 단위)

뉴클레오티드 다양성

뉴클레오티드 다양성은 런의 모든 사이클에 존재하는 각 염기(A, C, G 또는 T)의 상대적 비율을 의미합니다. 뉴클레오티드 균형은 시퀀싱 시스템의 컬러 매트릭스 수정(color matrix correction)과 강도 정규화(intensity normalization)에 중요합니다. Real-Time Analysis는 MiSeq i100 시리즈에 내장되어 있는 적응형 소프트웨어(adaptive software)로, 다양성이 낮은 라이브러리의 정확한 베이스 콜링(base calling)을 위해 세심하게 개발되었습니다. 가장 낮은 %의 PhiX spike-in(≥ 5%)으로 최대한 많은 고품질의 리드를 생성하여 다양성이 낮은 라이브러리를 시퀀싱할 때도 최적의 성능을 확보할 수 있습니다.

이 예시에서 5% 및 20% PhiX spike-in을 사용하여 다양성이 낮은 16S 앰플리콘 라이브러리를 MiSeq i100 시리즈로 시퀀싱했을 때 다양성이 높은 인간 Illumina DNA Prep 라이브러리를 시퀀싱했을 때와 비슷한 수준의 우수한 성능을 확인할 수 있었습니다(그림 7).

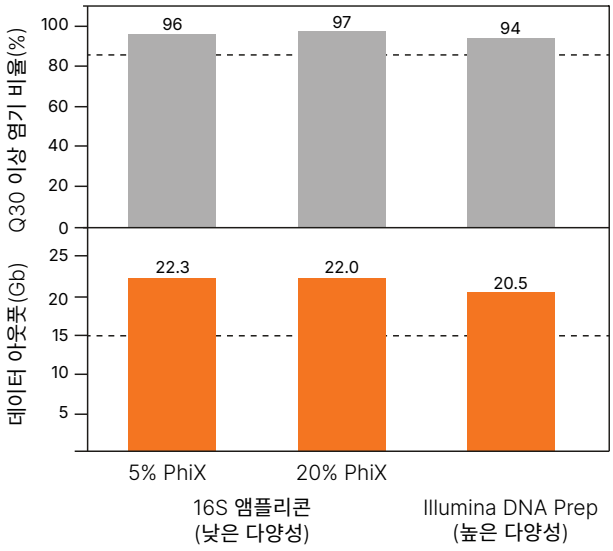

그림 7: 다양성이 낮은 라이브러리의 사용 지원

Q30 이상 염기 비율(%) 및 데이터 아웃풋(Gb)을 통해 확인할 수 있듯이, MiSeq i100 시리즈에 내장되어 있는 소프트웨어가 다양성이 낮은 라이브러리의 시퀀싱 성능을 최적화해 줌. 모든 시퀀싱 런은 MiSeq i100 Series 25M Reagent Kit(600 Cycles)와 2 × 301 bp의 리드 길이를 사용해 수행하였으며, 그림의 파선은 성능 사양을 나타냄. 각 조건의 정확한 메트릭스는 그래프의 각 막대 위에 숫자로 표시되어 있음

시퀀싱 아티팩트

NGS 시스템 간의 고유한 차이로 인해 오류 패턴에 편차가 발생할 수 있습니다. 이러한 시퀀스 특이적 오류(sequence-specific error, SSE)는 유전체의 특정 모티프(motif) 또는 기능에서 발생하는 경향이 있습니다. SSE는 가닥 방향성 오류(stranded error), 낮은 염기 품질, 소프트 클리핑(soft clipping), 또는 이들의 조합으로 나타날 수 있습니다. MiSeq i100 시리즈로 프로젝트를 전환할 때는 이러한 SSE의 영향을 고려해야 합니다. 분석 파이프라인과 설정값에 따라 워크플로우가 SSE에 취약할 수 있으며, 이로 인해 변이 검출 오류가 발생할 수 있기 때문입니다.

SSE 완화를 위한 모범 사례

워크플로우에서 어떤 필터가 사용되는지 파악하고, 변경 사항이 있을 때마다 성능이 워크플로우 사양을 지속적으로 충족하는지 확인하는 것이 중요합니다. 변경 사항에 따라 SSE를 완화하기 위해 필터 설정값을 조정해야 할 수 있습니다. 성능 변화가 관찰되는 경우, 가닥 편향(strand bias) 및 낮은 베이스 품질과 관련된 필터를 사용하면 SSE로 인한 아티팩트(artifact, 인위적 오류)를 줄이는 데 도움이 될 수 있습니다. DRAGEN v2.6 이상 버전의 분석 파이프라인에는 이러한 필터가 기본적으로 적용됩니다. 해당 필터가 기본적으로 적용되지 않는 파이프라인에서는 워크플로우 검증 과정의 일부로 필터를 활성화하고 이를 검증하는 방안을 고려해야 합니다. 사용 가능한 필터링 옵션(표 2), 권장 임곗값 및 특정 필터를 통한 자세한 SSE 완화 방법은 소프트웨어 설명서에서 확인할 수 있습니다. 필요시 개발자에게 추가 지침을 요청할 수 있습니다.

표 2: SSE 필터링 옵션

필터 범주	완화 가능한 문제
가닥/방향성	한쪽 리드 방향(F1R2 또는 F2R1)에 집중되는 오류
소프트 클리핑	클리핑된 리드 말단에 의해 뒷받침되는 낮은 VAF 콜
염기 품질	낮은 염기 품질 또는 리드 말단 클러스터링을 보이는 클러스터 버스트(cluster burst)
체계적 노이즈 파일	특정 위치에서 반복적으로 발생하는 노이즈(체세포(somatic) 분석에 권장됨)

기기, chemistry, 라이브러리 또는 플로우 셀 등의 변경으로 인해 노이즈 프로파일이 달라질 수 있음. 재교육이 필요한 경우에 대비하여 현재 필터와 기존 노이즈 파일의 검증 계획을 수립해야 함. SSE = sequence-specific error(시퀀스 특이적 오류), VAF = variant allele frequency(대립유전자 빈도)

요약

MiSeq i100 시리즈는 획기적으로 진보된 시퀀싱 chemistry와 통합된 데이터 분석 기능을 기반으로 향상된 사용성, 우수한 데이터 정확도, 매우 빠른 처리 속도를 제공합니다. 이 Technical Note에 소개된 모범 사례에 따라 라이브러리 품질 확인, 라이브러리 로딩 농도 최적화, 라이브러리 풀링(pooling)을 진행하면 MiSeq i100 시리즈로 최적의 성능을 확보할 수 있습니다.

상세 정보

MiSeq i100 및 MiSeq i100 Plus 시퀀싱 시스템

참고 문헌

1. Illumina. MiSeq i100 Series specification sheet. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/miseq-i100-specification-sheet-m-gl-02244/miseq-i100-specification-sheet-m-gl-02244.pdf. Published October 2024. Accessed September 10, 2026.

무료 전화(한국) 080-234-5300
techsupport@illumina.com | www.illumina.com

© 2026 Illumina, Inc. All rights reserved.
모든 상표는 Illumina, Inc. 또는 각 소유주의 자산입니다.
특정 상표 정보는 www.illumina.com/company/legal.html을 참조하십시오.
M-GL-03322 v3.0 KOR