

ILLUMINA DNA PCR-FREE Prep을 활용한 집단유전체학 연구를 지원하는 ILLUMINA Genomics Architecture

WGS 워크플로우
자동화를 통해
라이브러리 준비 시
일관적인 성능 제공



소개

ILLUMINA의 차세대 시퀀싱(next-generation sequencing, NGS) 기술은 고품질의 정확한 데이터를 제공하고 유전체학(genomics), 전사체학(transcriptomics), 후성유전학(epigenomics) 등 다양한 연구 분야에 활용이 가능합니다. NGS 워크플로우는 라이브러리 준비, 시퀀싱 그리고 데이터 분석 및 해석 단계로 구성되어 있습니다. ILLUMINA의 제품 포트폴리오는 NGS 워크플로우의 각 구성 요소에 적용할 수 있는 다양한 제품을 제공하고 있어 광범위한 애플리케이션을 지원합니다. 단, ILLUMINA 제품을 처음 사용하시는 일부 고객분들은 ILLUMINA에서 제공하는 제품이 매우 다양하기 때문에 각 구성 요소를 하나의 워크플로우에 통합하는 작업(유전 물질의 분리부터 번이 보고까지)에 많은 시간과 노력이 할애된다는 의견을 주셨습니다. 게다가 대부분의 ILLUMINA 워크플로우가 자동화를 지원하지 않지만, 기존 절차에 자동화된 방법을 통합하고 최적화하려면 기술적인 전문 지식이 필요하다는 점에 부담감을 느끼시는 고객분들도 간혹 계십니다.

ILLUMINA Genomics Architecture(IGA)는 표준화된 모듈식의 **유연한 프레임워크**를 통해 이러한 문제를 해결함으로써 자동화를 지원하며 **DNA 라이브러리 준비부터 데이터 분석까지 지원하는 NGS 워크플로우**를 연구 및 임상 연구에 신속하게 도입하고 적용할 수 있도록 해 줍니다. IGA는 전장 엑솜 시퀀싱(whole-exome sequencing, WES), 전장 유전체 시퀀싱(whole-genome sequencing, WGS), 집단유전체학(population genomics, PopGen) 프로젝트에 DNA-라이브러리 준비부터 데이터 분석까지 전 단계를 아우르는 워크플로우를 도입한 연구자들을 지원했던 경험을 토대로 개발되었습니다(그림 1).

IGA의 장점을 활용한 대표적인 프로젝트로는 **SG100K 프로젝트**가 있습니다. SG100K는 아시아인의 유전체 다양성에 대한 이해의 폭을 넓히는 것을 목표로 ILLUMINA와 Precision Health Research, Singapore(PRECISE)가 함께 100,000명의 건강한 싱가포르인의 유전체를 시퀀싱하는 프로젝트입니다. 이 Application Note는 SG100K 프로젝트의 일환으로 NGS 워크플로우를 자동화하기 위해 다수의 영국 및 싱가포르 기관에서 IGA 내 스크립트를 평가하는 독립적인 작업자들이 내부적으로 생성한 WGS 데이터를 소개합니다.

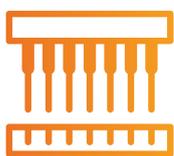
방법

IGA는 시퀀싱 및 데이터 분석의 자동화와 간소화를 위해 라이브러리 준비 및 소프트웨어 통합 전용 로봇 스크립트를 제공합니다.

라이브러리 준비

SG100K 프로젝트에 등록된 건강한 싱가포르인에게서 수집한 혈액 샘플로부터 고품질의 유전체 DNA(genomic DNA, gDNA)를 400 ng 추출하여 ILLUMINA DNA PCR-Free Prep(ILLUMINA, 카탈로그 번호: 20041794)으로 시퀀싱 라이브러리를 준비했습니다. 라이브러리 준비 단계는 Hamilton사의 STAR 리퀴드 핸들링(liquid handling) 플랫폼을 이용해 자동화했습니다.

라이브러리 준비



ILLUMINA DNA PCR-Free Prep
& 리퀴드 핸들링 로봇

시퀀싱



NovaSeq™ 6000 시스템

데이터 분석



BaseSpace™ Sequence Hub
DRAGEN™ Germline 파이프라인

그림 1: ILLUMINA Genomics Architecture의 WGS 워크플로우 — IGA는 ILLUMINA DNA PCR-Free Prep을 이용한 자동화된 라이브러리 준비 단계, NovaSeq 6000 시스템을 이용한 시퀀싱 단계 그리고 DRAGEN Germline 파이프라인을 이용한 데이터 분석 단계를 통합하여 DNA 라이브러리 준비부터 데이터 분석까지 모든 단계를 아우르는 WGS 워크플로우를 제공합니다.

시퀀싱

준비한 라이브러리는 NovaSeq 6000 시스템(Illumina, 카탈로그 번호: 20012850)에서 2 × 151 bp 런(run)을 구성하여 시퀀싱했습니다. 30× 커버리지(coverage)로 S4 플로우 셀(Illumina, 카탈로그 번호: 20028312)당 24개의 샘플을 시퀀싱했습니다. IGA를 이용한 자동화된 워크플로우의 한 구성 요소인 Clarity™ LIMS 소프트웨어가 벌크 풀링(bulk pooling), 변성(denaturation) 및 NovaSeq 6000 시스템으로의 라이브러리 로딩을 실행하도록 리퀴드 핸들링 플랫폼에 명령어를 전달하고, 자동으로 시퀀싱 런을 시작하는 데 필요한 정보를 전송했습니다.

데이터 분석

시퀀싱 런이 완료된 후 데이터는 DRAGEN Germline 파이프라인 v3.7.8을 통한 데이터 분석을 위해 자동적으로 BaseSpace Sequence Hub로 스트리밍되었습니다. 통계 분석과 그래프 플로팅에는 JMP 소프트웨어가 사용되었습니다.

결과

Illumina DNA PCR-Free Prep을 이용한 WGS 워크플로우를 자동화 및 간소화하는 IGA의 스크립트를 평가하기 위해 다수의 싱가포르 및 영국 기관에서 독립적인 작업자들이 생성한 시퀀싱 데이터를 분산 분석(analysis of variance, ANOVA)을 통해 비교했습니다. 비교 결과를 보면, 다수의 기관에서 각기 다른 기기를 사용해 IGA 워크플로우의 스크립트를 적용한 모든 독립적인 작업자들의 데이터에서 우수한 성능과 낮은 가변성(variability)을 확인할 수 있습니다(그림 2 및 표 1).

표 1: IGA 사용 시 확인된 데이터 일치성

파라미터	평균값 ± 표준 편차
주요 성능 매트릭스(샘플 191개^a)	
평균 상염색체 커버리지	36.39 ± 5.9%
커버리지가 ≥ 15×인 유전체 %	94.72 ± 7.8%
Q30 염기(base)	102.1 ± 16.7 Gbp
상염색체 검출력(callability) %	97.05 ± 7.1%
인서트 길이(Insert length) 중앙값	459.4 ± 20.2 bp
샘플 오염 추정치	0.001 ± 0.0002
SNV 분석^b(샘플 99개)	
SNV 재현율(recall)	99.84 ± 0.11%
SNV 정밀도(precision)	99.83 ± 0.03%
Indel 재현율	99.58 ± 0.31%
Indel 정밀도	99.66 ± 0.15%

a. Coriell에서 추출한 DNA
b. NA12878을 하나의 진리(truth) 샘플로 사용해 얻은 데이터에 근거함.

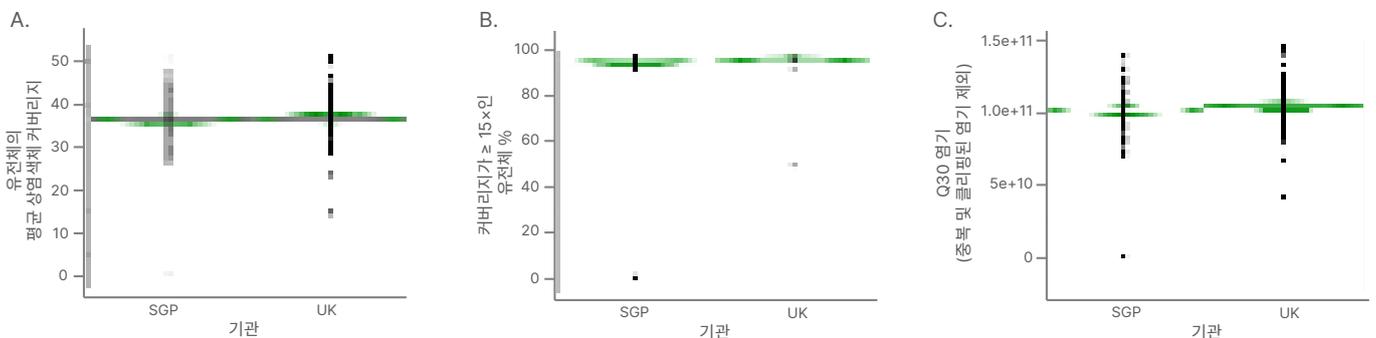


그림 2: 다기관 독립적인 작업자를 통해 확인한 데이터 성능 및 일치성 — 분산 분석을 통해 싱가포르(SGP) 및 영국(UK) 기관에서 생성한 시퀀싱 데이터의 (A) 평균 상염색체 커버리지, (B) 커버리지가 ≥ 15×인 유전체 % 그리고 (C) Q30 염기(중복(duplicate) 및 클리핑된 염기(clipped base) 제외)를 측정된 결과, 매우 일관적인 성능이 확인됨.

요약

IGA는 자동화를 지원하며 DNA 라이브러리 준비부터 데이터 분석까지 전 단계를 아우르는 NGS 워크플로우를 연구자가 신속하게 도입할 수 있도록 표준화된 모듈식의 유연한 프레임워크를 제공합니다. WGS 연구에 Illumina DNA PCR-Free Prep과 IGA를 활용하면 하루 안에 라이브러리를 준비하고 S4 플로우 셀당 24개의 샘플을 풀링한 후 30x 유전체 커버리지의 시퀀싱을 완료할 수 있습니다. SG100K 집단유전체학 연구 프로젝트에도 활용되고 있는 IGA는 다기관 독립적인 작업자들이 매우 일관적으로 강력한 성능을 경험할 수 있도록 해 줍니다. 이러한 결과는 워크플로우의 자동화와 NGS 방법 사용 시 IGA가 신뢰할 수 있는 성능을 제공함을 보여줍니다.

상세 정보

[Illumina Genomics Architecture](#)

[Illumina DNA PCR-Free Prep](#)

[NovaSeq 6000 시스템](#)

[DRAGEN Secondary Analysis](#)



무료 전화(한국) 080-234-5300

techsupport@illumina.com | www.illumina.com

© 2023 Illumina, Inc. All rights reserved. 모든 상표는 Illumina, Inc. 또는 각 소유주의 자산입니다.

특정 상표 정보는 www.illumina.com/company/legal.html을 참조하십시오.

M-KR-00211 KOR