

BovineSNP50 v3 BeadChip

소의 유전체 전체에 걸쳐 전략적인 위치에 균등한 간격으로 분포하는 53,218개의 SNP 포함

장점

- **우수한 콜 비율 및 정확도**
평균 콜 비율 99% 이상, 재현성 99.9%
- **포괄적이고 균일한 커버리지**
37.4 kb(중앙값)의 간격으로 고르게 분포하는 SNP
- **간단한 워크플로우**
PCR과 라이게이션(ligation)이 필요 없는 프로토콜
- **높은 처리량**
최대 24개의 샘플 동시에 연구 가능

소개

Illumina는 United States Department of Agriculture(USDA, 미국농동청), Agricultural Research Service(ARS, 농업연구청), University of Missouri 그리고 University of Alberta와의 협업을 통해 3세대 BovineSNP50 v3 BeadChip을 개발했습니다(그림 1). 이 고밀도의 전장 유전체 지노타이핑 어레이(genome-wide genotyping array, 표 1 참고)는 소의 유전자 변이 연구 시 소의 유전체 전체를 고르게 포괄하는 53,218개의 단일 염기 다형성(single nucleotide polymorphism, SNP) 프로브를 포함하고 있습니다. BovineSNP50 v3 BeadChip은 전장 유전체 데이터 기반의 선별, 양적 형질 유전자좌(quantitative trait locus, QTL)의 확인, 개체별 유전적 장점의 평가, 그리고 결측치 대체(imputation) 방법을 이용한 육우 및 젖소의 비교 유전학 연구 등에 활용해 볼 수 있습니다.

24개의 샘플을 처리하는 이 BeadChip은 iScan™ 시스템, 통합 분석 소프트웨어와 Infinium HTS assay 사용 시 매우 우수한 콜 비율(call rate)을 보이며 유연한 콘텐츠 배포와 유전자 복제수 변이(copy number variation, CNV)의 검출 및 측정을 지원합니다. 또한 PCR 없이 하나의 튜브로 샘플을 준비하므로 노동력을 절약할 수 있을 뿐만 아니라 샘플 취급 중 발생할 수 있는 오류도 현저히 줄일 수 있습니다.

BovineSNP50 v3 BeadChip의 콘텐츠

BovineSNP50 v3 BeadChip은 *Bos taurus* 소 참조 유전체(bovine reference genome)와 Bovine Genome Consortia와 같은 공공 소스로부터 얻은 가치가 높은 콘텐츠를 제공합니다. Illumina의 과학자들은 협업자들과 함께 모든 품종의 소에 대한 유전자 변이 연구 시 유익한 정보를 제공해 주는 고밀도의 SNP 지노타이핑 마이크로어레이를 개발했습니다. Bovine HapMap Consortium에서 기술한 검증된 흔한 SNP(즉, 소수 대립유전자 빈도(minor allele frequency, MAF)가 0.05 이상인 SNP)를 표적화하기 위해 12,000개가 넘는 프로브를 설계했습니다. 또한 BeadChip 개발자들은 Btau, Baylor College of Medicine의 연구진이 조립한 전장 유전체 샷건 리드(whole-genome shotgun read)를 비롯한 공공 리소스에서 흔한 SNP에 대한 데이터를 마이닝했습니다. 추가로 미국 Clay Center에 위치한 U.S. Meat Animal Research Center의 연구자들이 발견한 혈통 표지자(parentage marker)와 같은 소스에서 얻은 콘텐츠⁶와, USDA ARS의 연구진이 Holstein 세균 인공 염색체(bacterial artificial chromosome, BAC) 시퀀스 데이터와 소 유전체 어셈블리(bovine genome assembly)를 비교하여 발견한 SNP도 콘텐츠에 포함했습니다(표 2).



그림 1: BovineSNP50 v3 BeadChip — 소의 유전체 전체에 걸쳐 균등한 간격으로 분포하는 53,218개의 SNP로 구성된 BovineSNP50 v3 BeadChip

표 1: 제품 정보^a

| 항목 | 설명 |
|-----------------------------------|------------------------------|
| 종 | <i>Bos taurus</i> (소) |
| 총 표지자 수 | 53,218개 |
| 커스텀 비드(bead) 수용력 | 600,000개 |
| BeadChip당 처리 샘플 수 | 24개 |
| DNA 사용량 | 200 ng의 유전체 DNA(genomic DNA) |
| Assay chemistry | Infinium HTS |
| 지원 기기 | iScan 시스템 |
| iScan 시스템의 최대 샘플 처리량 ^a | 주당 약 5,760개의 샘플 |
| 샘플당 스캔 소요 시간 | 30분 |

a. 근사치, 스캔 소요 시간 및 최대 처리량은 랩과 시스템 구성에 따라 상이할 수 있음.

BovineSNP50 BeadChip에 포함되어 있는 24,000개가 넘는 SNP 프로브는 세 모집단으로 모아진 경제적으로 중요한 육우 및 젖소의 샘플을 시퀀싱함으로써 새롭게 발견된 SNP 유전자좌를 표적화합니다. Illumina의 과학자들은 협업자들과 함께 *Bos taurus taurus* 유전체의 약 2%를 딥 시퀀싱하여 62,000개가 넘는 추정(putative) SNP를 발견했습니다. 이 새로운 데이터 소스에서 SNP의 간격(spacing), 예상 MAF, Infinium HTS assay 성능을 기준으로 23,800개가 넘는 SNP를 선정했습니다. 또 BeadChip 개발자들은 포괄적인 커버리지(coverage)를 위해 가장 큰 매핑되지 않은 콘티그(contig)를 표적화하는 유전자좌를 선택했습니다. BovineSNP50 v3 BeadChip을 구성하는 53,218개의 SNP 프로브는 19종의 일반 육우 및 젖소에서 그 성능이 입증되었습니다(표 3).

이 BeadChip은 여러 검사 품종에 걸쳐 고르게 분포하는 다형적(polymorphic)인 SNP를 표적화합니다. 평균적인 프로브 간격은 50.6 kb, 간격 중앙값은 37.4 kb이며, 모든 유전자좌에 걸친 평균적인 MAF는 0.25입니다. 현재 여러 소 품종을 대상으로 한 연관불균형(linkage disequilibrium, LD) 연구에서는 평균적으로

약 70 kb 크기의 하플로타입 블록(haplotype block)을 권장하고 있습니다. 이는 BovineSNP50 v3 BeadChip이 제공하는 해상도가 소의 LD 해상도 범위 내에 있음을 의미합니다. BovineSNP50 v3 BeadChip이 제공하는 53,000개가 넘는 SNP 프로브는 최신 소 참조 유전체 어셈블리인 UMD 3.0에 매핑됩니다.

표 2: BovineSNP50 BeadChip의 콘텐츠 소스

| 소스 | BovineSNP50 v1 프로브 | BovineSNP50 v2 프로브 | BovineSNP50 v3 프로브 |
|---|--------------------|--------------------|--------------------|
| ILLUMINA Genome Analyzer로 일반적인 품종의 소 샘플을 시퀀싱하여 얻은 새로운 SNP | 23,840개 | 24,181개 | 22,299개 |
| Bovine HapMap 데이터 세트 | 12,298개 | 12,342개 | 11,607개 |
| Btau 유전체 어셈블리의 SNP | 9,361개 | 9,404개 | 9,086개 |
| 전장 유전체 샷건 리드 ^a | 5,808개 | 6,038개 | 5,485개 |
| Holstein BAC 시퀀스 데이터 | 1,409개 | 1,411개 | 1,238개 |
| 혈통 표지자 ^b | 116개 | 120개 | 200개 |
| 기타 ^c | 1,169개 | 1,113개 | 3,384개 |
| 합계 | 54,001개 | 54,609개 | 53,218개 |

- a. 여섯 품종의 소(Norwegian Red, Holstein, Brahman, Angus, Jersey, Limousin)로부터 얻은 리드를 Btau2.0과 비교함.
- b. 30개의 중복(duplicate) 포함 엄격한 큐레이션(curation)을 거친 혈통 표지자
- c. Institute for Food and Agricultural Sciences Alberta, INRA 및 French International Institute of Agriculture가 검증한 흔한 SNP를 포함함.

고품질 데이터

BovineSNP50 v3 BeadChip에 포함된 53,218개의 SNP 프로브의 우수한 성능을 확인하기 위해 Infinium HTS assay로 엄격한 기능 검사를 진행했습니다. 전장 유전체 연관성 연구(Whole-genome association study, WGAS)의 성공은 부분적으로는 높은 콜 비율과 정확한 지노타입 콜(genotype call) 결과에 기인합니다. 복합 형질(Complex trait)은 보통 유전적 영향이 상대적으로 적기 때문에 관심 대상인 SNP와 LD에 있는 분석된 SNP의 콜 비율이 낮거나 지노타입 콜이 부정확할 경우 잠재적인 연관성을 놓칠 수 있습니다. 모든 BovineSNP50 v3 BeadChip은 콜 비율이 99%를 넘는 뛰어난 성능을 보입니다(표 3).

Bovine HapMap Consortium에서 제공한 샘플을 이용해 BovineSNP50 v3 BeadChip의 콘텐츠를 내부적으로 검증했을 때 매우 우수한 결과를 확인할 수 있었습니다(표 4). 제품 개발 팀은 샘플을 수집한 품종 중 인접 또는 기저 결실(deletion)이 있는 것으로 보이는 4,290개의 유전자좌를 확인한 후 보존했습니다. 이들 유전자좌가 패널의 다른 유전자좌에 비해 콜 비율이 더 낮음에도 보존된 이유는 관심 대상인 형질 및 향후 유전체 빌드(genome build)의 향상과 생물학적으로 관련이 있는 정보를 제공할 수 있기 때문입니다.

성능 및 콘텐츠 검증 결과는 BovineSNP50 v3 BeadChip이 제공하는 신뢰성과 우수한 데이터 품질을 명확히 보여줍니다. 연구자는 이 BeadChip을 이용하여 관심 있는 표현형(phenotype)의 유전적 장점을 예측하고 다양한 소 품종 간 발생하는 변이의 유전적 근거를 알아볼 수 있습니다. Illumina는 축산업계를 지원하는 연구와 생산에 필요한 혁신적인 도구를 제공하고자 BovineSNP50 v3 BeadChip을 개발했습니다.

표 3: BovineSNP50 v3 BeadChip의 성능

| 파라미터 | 결과 | 제품 사양 |
|-----------------------------------|-------|---------|
| 평균 콜 비율 ^a | 99.7% | > 99% |
| 재현성(Reproducibility) ^b | 100% | > 99.9% |
| 멘델리안 불일치(Mendelian inconsistency) | 0.06% | < 0.1% |

- a. 주요 소 품종 중 17종 274개체, 31개의 트리오(trio), 1회의 반복 실험(replicate) 적용.
- b. 4개의 Holstein 트리오 사용.

표 4: BovineSNP50 v3 BeadChip의 콘텐츠 검증

| 품종 | 샘플 수 | 다형적 유전자좌 수 ^a | 평균 MAF | MAF 중앙값 ^b |
|---------------------------|------|-------------------------|--------|----------------------|
| Angus | 22개 | 41,209개 | 0.22 | 0.23 |
| Beefmaster | 24개 | 43,741개 | 0.23 | 0.23 |
| <i>Bos indicus</i> Gir | 17개 | 25,320개 | 0.11 | 0.03 |
| <i>Bos indicus</i> Nelore | 5개 | 22,422개 | 0.1 | 0 |
| Brahman | 20개 | 33,038개 | 0.13 | 0.08 |
| Charolais | 14개 | 42,685개 | 0.23 | 0.21 |
| Guernsey | 21개 | 37,313개 | 0.2 | 0.19 |
| Hereford | 20개 | 43,902개 | 0.23 | 0.23 |
| Holstein | 22개 | 41,913개 | 0.23 | 0.23 |
| Jersey | 9개 | 36,683개 | 0.18 | 0.17 |
| Limousin | 10개 | 42,732개 | 0.21 | 0.2 |
| N'Dama | 6개 | 29,032개 | 0.14 | 0.08 |
| Piedmontese | 21개 | 42,652개 | 0.23 | 0.24 |
| Red Angus | 10개 | 43,028개 | 0.22 | 0.2 |
| Romagnola | 6개 | 38,521개 | 0.2 | 0.17 |
| Santa Gertrudis | 7개 | 42,675개 | 0.21 | 0.21 |
| Sheko | 9개 | 36,582개 | 0.17 | 0.17 |
| 합계 | 274개 | 47,919개 | 0.26 | 0.26 |

- a. MAF > 0.05
- b. 모든 53,218개의 유전자좌에 대한 값

지노타이핑 솔루션

선택적으로 실험실 정보 관리 시스템(laboratory information management system, LIMS)과 로봇 프로세스 자동화를 사용하면 전체 워크플로우에 걸쳐 샘플을 정확하고 효율적으로 추적할 수 있습니다. 또한 연구자는 Illumina의 GenomeStudio DNA 분석 소프트웨어나 FastTrack Genotyping Services를 통해 이미 보유 중인 샘플로 지노타이핑(genotyping, 유전형 분석)을 수행해 볼 수도 있습니다.

요약

학계를 대표하는 소 연구자들과의 협업을 통해 개발된 BovineSNP50 v3 BeadChip은 소의 유전체 전체에 걸쳐 균등한 간격으로 분포하는 53,000개가 넘는 SNP로 구성되어 있습니다. 24개의 샘플을 처리하는 이 BeadChip은 육우 및 젖소의 전장 유전체 연구 시 비용 대비 효율적인 처리량이 높은 솔루션을 제공합니다.

제품 목록

| BovineSNP50-24 v3 BeadChip ^a | | 설명 | 카탈로그 번호 |
|--|--|---|----------|
| 48 samples | | 각 패키지에는 2개의 BeadChip과 DNA 샘플 48개의 증폭(amplification), 절편화(fragmentation), 하이브리드화(hybridization), 표지(labeling) 및 검출(detection) 시약이 들어 있음. | 20000766 |
| 288 samples | | 각 패키지에는 12개의 BeadChip과 DNA 샘플 288개의 증폭, 절편화, 하이브리드화, 표지 및 검출 시약이 들어 있음. | 20000767 |
| 1152 samples | | 각 패키지에는 48개의 BeadChip과 DNA 샘플 1,152개의 증폭, 절편화, 하이브리드화, 표지 및 검출 시약이 들어 있음. | 20000768 |
| BovineSNP50-24+ v3 BeadChip ^b | | 설명 | 카탈로그 번호 |
| 48 samples | | 각 패키지에는 2개의 BeadChip이 들어 있으며, 표적 연구 시 샘플별로 최대 600,000개의 커스텀 프로브를 추가로 포함할 수 있음. | 20000769 |
| 288 samples | | 각 패키지에는 12개의 BeadChip이 들어 있으며, 표적 연구 시 샘플별로 최대 600,000개의 커스텀 프로브를 추가로 포함할 수 있음. | 20000830 |
| 1152 samples | | 각 패키지에는 48개의 BeadChip이 들어 있으며, 표적 연구 시 샘플별로 최대 600,000개의 커스텀 프로브를 추가로 포함할 수 있음. | 20000831 |

a. 1개의 BeadChip으로 24개의 샘플을 처리하고 약 53,000개의 유전자좌를 분석할 수 있음.
b. 맞춤 콘텐츠를 지원함.

상세 정보

BovineSNP50-24 v3.0 BeadChip에 대한 자세한 정보는 www.illumina.com/products/by-type/microarray-kits/bovine-snp50.html에서 확인하실 수 있습니다.

참고 문헌

1. Ensembl Genome Browser. uswest.ensembl.org/Bos_taurus/Info/Index. Accessed June 2, 2020.
2. The Bovine Genome Consortia. bovinegenome.elsiklab.missouri.edu/bovine_genome_consortium. Accessed June 2, 2020.
3. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS (2005) A genome-wide scalable SNP genotyping assay using microarray technology. *Genet* 37(5): 549-554.
4. Steemers FJ, Weihua Chang W, Lee G, Barker DL, Shen R, et al. (2006) Whole-genome genotyping with the single-base extension assay. *Nat Methods* 3(1): 31-33.
5. <ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/snp/Btau20070913/README>
6. Heaton MP, Keen JE, Clawson ML, Harhay GP, Bauer N, et al. (2005) Use of bovine single nucleotide polymorphism markers to verify sample tracking in beef processing. *J Am Vet Med Assoc* 226(8): 1311-1314.
7. Matukumalli LK, Taylor JF, and Van Tassell CP. Personal communication.
8. Van Tassell CP, Smith TPL, Matukumalli LK, Taylor JF, Schnabel, RD, et al. (2008) Simultaneous SNP discovery and allele frequency estimation by high-throughput sequencing of reduced representation libraries. *Nat Meth* (accepted).
9. Khatkar MS, Zenger KR, Hobbs M, Hawken RJ, Cavanagh JAL, et al. (2007) A Primary Assembly of a Bovine Haplotype Block Map Based on a 15,036-Single-Nucleotide Polymorphism Panel Genotyped in Holstein-Friesian Cattle. *Genetics* 176(2): 763-772.