

GWB CLC Genomics Workbench

다양한 NGS 데이터 분석을 내 손으로 직접!



GENOMICS

- *De novo* assembly of any genome size
- Read mapping for reference genome
- Variant detection (High performance)
- Hybrid assembly (Short + Long)
- Visualization of genome information

TRANSCRIPTOMICS

- RNA-seq analysis
- micro RNA analysis
- Gene expression analysis (PCA, Heatmap, Venn diagram)
- Detection of fusion genes

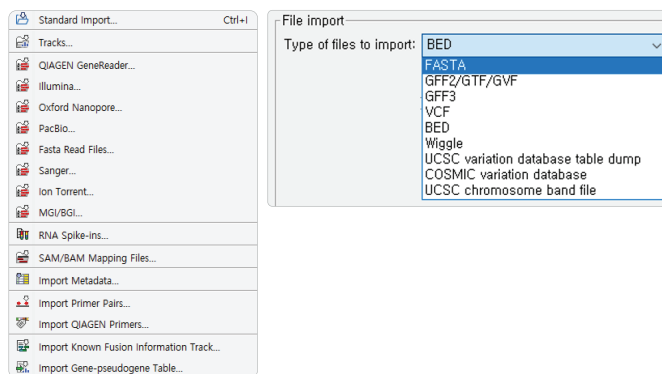
EPIGENOMICS

- ChIP-seq analysis
- Peak finding and peak refinement
- Bisulfite sequencing analysis
- Histone ChIP-seq analysis
- Transcription factor ChIP-seq

Support NGS multi platform

CLC Genomics Workbench는 다양한 차세대염기서열분석(NGS) 플랫폼에서 생성된 데이터를 지원하여 단일 통합 환경에서 유전체 분석을 수행할 수 있도록 설계된 소프트웨어입니다. 주요 NGS 플랫폼에서 생성된 short / long-read 데이터를 모두 지원하여, 효율적인 유전체 분석을 수행할 수 있습니다.

GFF3 / GFF2 / GTF / GVF 포맷의 annotation 파일, BED, WIGGLE, VCF, UCSC chromosome band file 등을 import 할 수 있으며, txt or csv 포맷 등의 발현 정보 파일이나 annotation 파일도 분석에 사용 가능합니다.



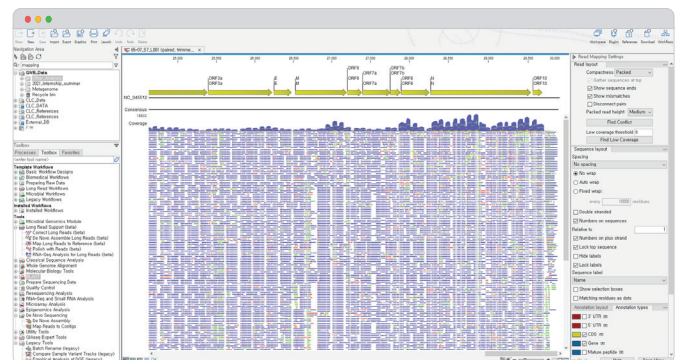
De novo assembly

CLC Genomics Workbench는 다양한 NGS 멀티 플랫폼 데이터를 활용한 *de novo* assembly 분석을 지원합니다. Long-read 데이터를 이용한 assembly가 가능하며, 서로 다른 시퀀싱 데이터를 조합한 hybrid assembly 후 polishing 통합 과정을 통해 신뢰도 높은 결과를 얻을 수 있습니다. 또한 미생물부터 대형 유전체에 이르기까지 유전체 크기에 제한 없이 분석을 수행할 수 있습니다.



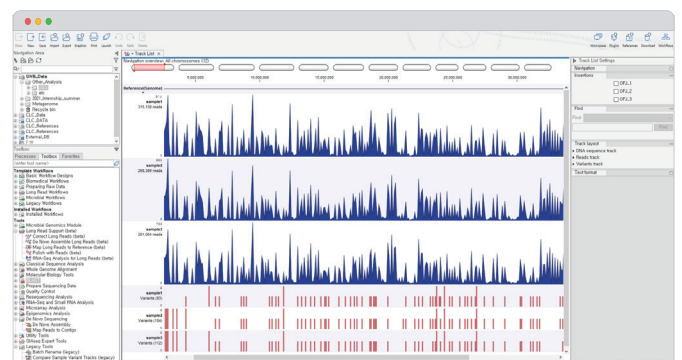
Reference assembly

Reference 유전체에 시퀀싱 데이터를 mapping 함으로써, 이후 변이 분석을 수행하기 위한 기초 데이터를 제공합니다. 유전체 전반의 커버리지와 서열 차이를 다양한 시각화 결과를 통해 직관적으로 확인할 수 있습니다. 또한, 주요 모델 생물의 유전체 데이터는 데이터베이스와 연동하여 쉽게 내려받을 수 있어 분석에 용이합니다.



Variation detection

CLC Genomics Workbench는 Reference Assembly 데이터에 기반한 Variant detection 분석을 제공하며, 시퀀싱 데이터로부터 유전체 서열의 차이를 체계적으로 분석할 수 있도록 지원합니다. Mapping 된 sequencing reads를 이용해 SNP 및 InDel과 같은 변이를 검출할 수 있으며, 변이 빈도 등 다양한 필터링 기준을 적용하여 시퀀싱 과정에서 발생할 수 있는 오류와 실제 변이를 구분할 수 있습니다. 검출된 변이 결과는 시각화 도구와 연계하여 확인할 수 있으며, 유전자 및 단백질 서열 수준에서의 변이 영향 해석에도 활용할 수 있습니다.



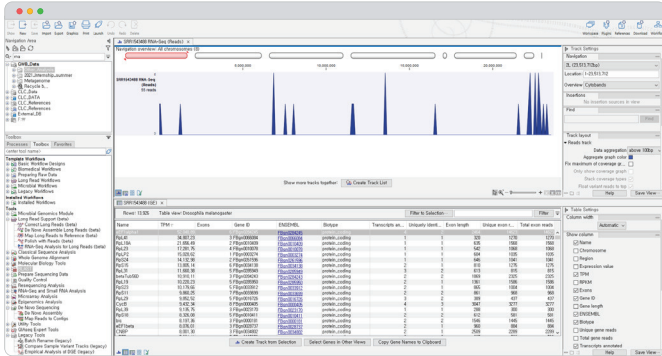
GWB CLC Genomics Workbench

다양한 NGS 데이터 분석을 내 손으로 직접!

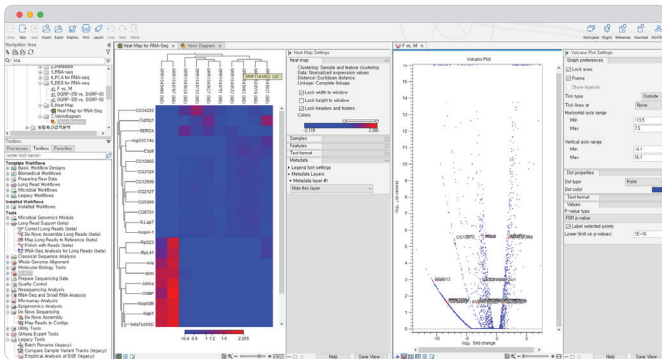


RNA-Seq analysis

CLC Genomics Workbench는 annotation 된 reference 유전체 또는 전사체 서열을 기반으로 RNA-seq 데이터를 분석하여 샘플 간 유전자 발현 수준을 비교할 수 있는 기능을 제공합니다. 시퀀싱 reads를 reference에 mapping 하여 발현량을 계산하며, read count, RPKM, TPM 등 다양한 발현 값 기준을 선택하여 분석 목적에 맞게 활용할 수 있습니다.



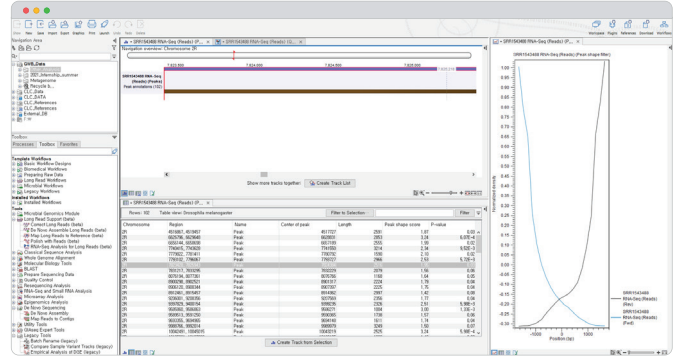
계산된 발현 값과 샘플 추가 정보(Metadata)를 이용해 실험 조건에 따른 그룹 구분이 가능하며, 주성분 분석(PCA)을 통해 샘플 간 전반적인 발현 패턴과 상관관계를 확인할 수 있습니다. 또한 발현 양상이 유사한 유전자 또는 샘플을 대상으로 클러스터링 분석을 수행하고 heatmap으로 시각화할 수 있습니다.



차등 발현 분석을 통해 조건 간 유의미한 발현 변화를 보이는 유전자를 식별할 수 있으며, 분석 결과는 volcano plot 및 venn diagram과 같은 시각화 도구를 통해 직관적으로 확인할 수 있습니다.

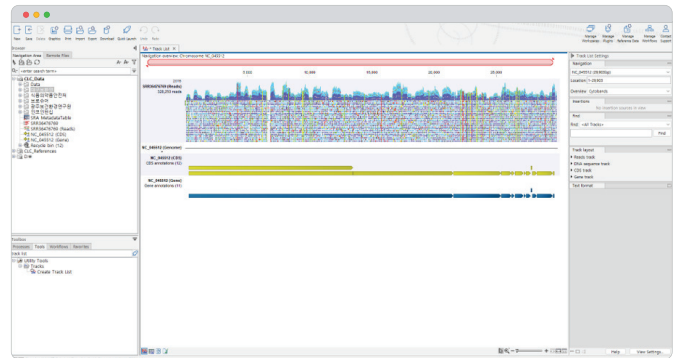
ChIP-Seq analysis

CLC Genomics Workbench는 ChIP-seq 데이터를 기반으로 DNA-단백질 상호작용 영역을 분석할 수 있는 기능을 제공합니다. 시퀀싱 reads를 reference 유전체에 mapping 후, 단백질 결합 신호가 축적된 영역을 식별하기 위한 peak detection 분석을 수행할 수 있습니다. 이를 통해 transcription factor 결합 부위나 histone 변형과 같은 유전체 상의 기능적 조절 영역을 분석할 수 있습니다.



Genome informatics framework

CLC Genomics Workbench에서는 NGS 데이터를 기반으로 다양한 분석을 수행하고, 분석 결과를 서로 연계하여 확인할 수 있습니다. 이때 사용하는 Track tool은 genome 수준의 데이터를 한눈에 확인할 수 있도록, 브라우저 시스템과 유사한 형태로 제공됩니다.



Workflow

CLC Genomics Workbench는 NGS 데이터의 불러오기와 품질 확인부터 분석 및 시각화까지, 전 과정을 하나의 환경에서 수행할 수 있는 End-to-End 워크플로우를 제공합니다.

이를 통해 반복적이고 표준화된 분석을 지원하며, 효율적이고 일관된 NGS 데이터 분석이 가능합니다.