



次世代シーケンス解析用ソフトウェアの決定版

CLC Genomics Workbench



フィルジェン株式会社

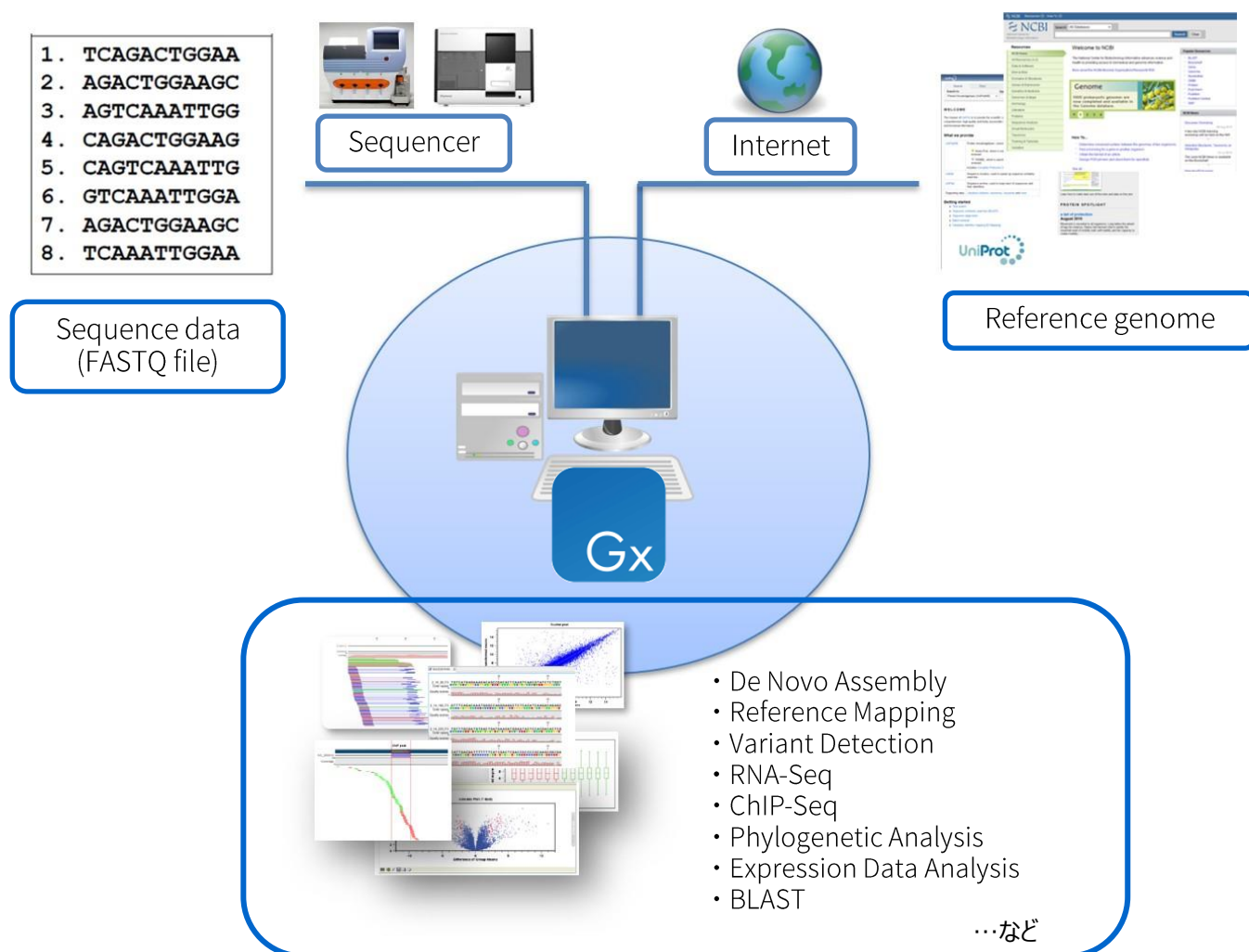
NGSデータ解析の決定版

CLC Genomics Workbenchは国内500ユーザーを突破し、世界中で広く利用されている次世代シーケンズ解析ソフトウェアです。次世代シーケンサーから産出された膨大なデータをシームレスに解析に利用することができ、またユーザーフレンドリーなインターフェースを搭載しているため、簡単な操作で専門的なデータ解析が可能です。医学研究、基礎生物学研究の両分野で使用可能なアプリケーションが搭載されており、幅広い研究用途に利用できます。



統合配列解析ソフトウェア

最先端の解析アルゴリズムと高速なゲノムアセンブル・マッピングツール、豊富なグラフィカル機能や多彩な出力オプションを搭載し、ユーザーフレンドリーかつ直感的なインターフェイスで稼働します。illumina社やMGI社、Thermo Fisher Scientific社の次世代シーケンサーデータに対応し、アセンブルやマッピング、各種アプリケーションに対応した専用解析モジュールを搭載しています。



CLC Genomics Workbenchは次世代シーケンサーの出力データから多彩な解析を実行できます。



- ## ◆リファレンスゲノムへのマッピング

[illegible]

◆見やすく操作性の良い変異リスト

マッピングデータから変異(SNV, InDel, CNV)の検出を行うと、解析結果はテーブル形式で表示されます。テーブルには、各変異のゲノム上の位置やアリル情報、接合性情報、さらにアリルごとのリード配列のカウント情報などを含み、テーブル上でフィルターをかけることで、任意の条件を満たす変異のみをピックアップすることが可能です。また、本ソフトウェアには複数種類の変異検出用ツールが搭載されており、変異をもつリード配列のカバレッジ中の頻度や、解析サンプルの生物種の倍数性などによって、ツールを使い分けることが可能です。いずれのツールも、様々なノイズフィルターが組み込まれており、パラメータを調整することで、検出される変異の偽陽性を減らしたり、逆に検出感度を上げたりすることが可能です。

- Chromosome: 変異の検出された染色体番号
- Region: 変異のポジション
- Type: 変異の種類 (SNV, Insertion, Deletionなど)
- Reference: リファレンスの塩基配列
- Allele: 検出された塩基配列

- Zygosity: 変異の接合性(HeteroかHomoか)
- Count: マップされたリードのうち、変異を有するリードの数
- Coverage: マップされたリード数
- Frequency: 変異の頻度

Rows: 19,104 Table view: Homo sapiens											
Chromosome	Region	Type	Reference	Allele	Reference a...	Length	Zygosity	Count	Coverage	Frequency	Probability
1	887577	SNV	G	A	No		Heterozygous	25	44	56.82	1.00
1	898323	SNV	T	C	No		Homozygous	63	63	100.00	1.00
1	900505	SNV	G	C	No		Heterozygous	14	27	51.85	1.00
1	909309	SNV	T	C	No		Heterozygous	7	15	46.67	1.00
1	911916	SNV	C	T	No		Heterozygous	20	50	40.00	1.00
1	916662	SNV	A	C	No		Heterozygous	31	70	44.29	1.00
1	1007432	SNV	G	A	No		Heterozygous	11	18	61.11	1.00
1	1110987-1110988	Insertion	-	C	No		Heterozygous	22	51	43.14	1.00

# unique en...	Amino acid change in longest tr...	Coding region change in longest tra...	Other varian...	Non-synony...	chr20 (...)	Gene Cards	ENSEMBL	dbSNP
15		ENST00000377873:c.1500G>A	No	No	BFPSP1	BFPSP1	ENSG00000125864	6136118
9		ENST00000377873:c.957-83T>C	No	-	BFPSP1	BFPSP1	ENSG00000125864	6044850
2		ENST00000377873:c.736-45A>G	No	-	BFPSP1	BFPSP1	ENSG00000125864	6111549
8			No	-	BFPSP1	BFPSP1	ENSG00000125864	
24		ENST00000377873:c.439-230G>T	No	-	BFPSP1	BFPSP1	ENSG00000125864	2023224
72			No	No	DSTN	DSTN	ENSG00000125868	
30	ENSP00000246069:p.Ile37fs	ENST00000246069:c.109_110insT	No	Yes	DSTN	DSTN	ENSG00000125868	
77			No	No	DSTN	DSTN	ENSG00000125868	
26	ENSP00000246069:p.Lys44Glu	ENST00000246069:c.130A>G	No	Yes	DSTN	DSTN	ENSG00000125868	

- アミノ酸置換情報や非同義置換情報、オーバーラップする遺伝子名などがアノテーションとして追加
- 青字はデータベースにリンクしており、クリックすることで当該サイトを閲覧することが可能



トランスクリプトミクス解析

- ✓ 遺伝子発現量解析用(RNA-Seq)と、small RNA解析用の2種類のツールが利用可能
- ✓ ヒートマップ作成やクラスタリング、サンプル間の発現変動遺伝子解析など、各種統計解析用のツールも利用可能
- ✓ DNAマイクロアレイデータの解析用ツールも搭載

◆柔軟なデータの表示

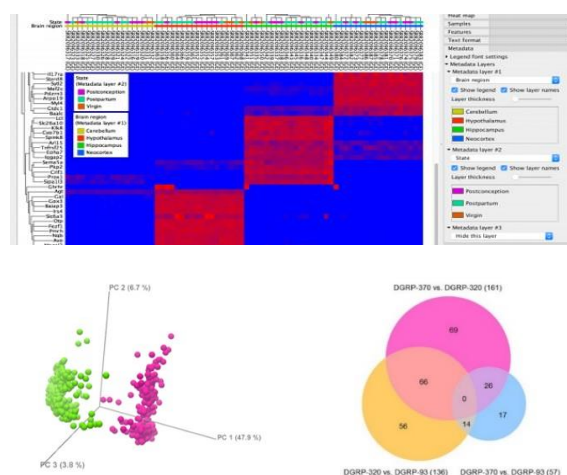
RNA-Seq解析用ツールを使用すると、各遺伝子の発現量の算出を自動で行います。エクソン間をまたいだリード配列のマッピングにも対応し、遺伝子の発現量については、遺伝子ごとのリード配列カウント以外に、RPKMとTPMによる計算が可能です。また1つの遺伝子に複数のスプライスバリエントが存在する生物の場合は、遺伝子ごとの発現量データの他に、転写物(mRNA)ごとの発現量データも得ることができます。さらにペアリード配列データでマッピングを行う場合は、ペアとなっている各リード配列のマッピング状況から、融合遺伝子の候補を予測することが可能です。

また、発現量データとマッピングデータの両方を同時に取得することができます。



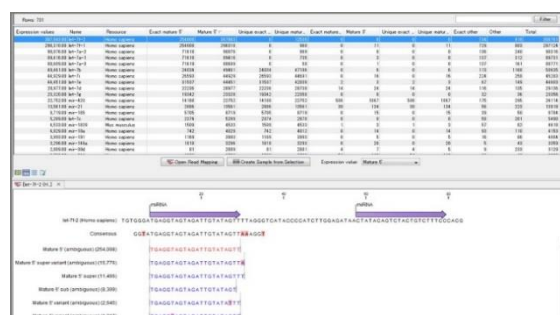
◆比較解析と多彩なグラフ化

RNA-Seq遺伝子発現量データのサンプル間比較解析を行うための様々なツールが使用できます。各サンプルをグループごとに分類し、グループ間で統計的に有意に発現量が変動している遺伝子の検索や、統計データによるボルケーノプロットの表示、主成分分析や二次元クラスタリングとヒートマップ、さらに複数の比較データからベン図の作成、さらにGene Ontologyを用いた遺伝子発現解析が可能です。



◆small RNA Analysis

small RNA Analysis ツールでは、small RNAのデータベースをリファレンス配列として使用し、各small RNAとmicroRNAの発現量を計算します。計算した発現量データは、専用の解析ツールを使用し、サンプル間のデータ補正と、グループ間の発現量比較を行うことができます。



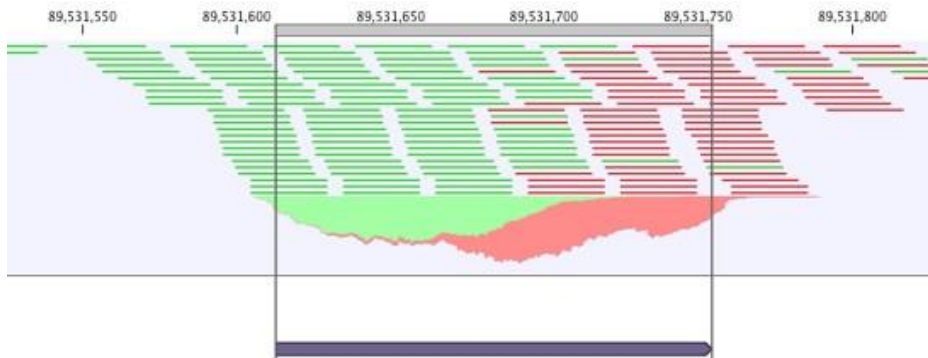


エピゲノミクス解析

- ✓ DNAとタンパク質の結合部位を、ゲノム配列上のChIPピークとして検出
- ✓ 検出したピークは、リスト形式で表示が可能
- ✓ ヒストン修飾やバイサルファイトシーケンスのメチル化解析も可能

◆ChIP-Seq

クロマチン免疫沈降(ChIP)アッセイと次世代シーケンスを組み合わせることで、アッセイに使用した転写因子のゲノム配列上の結合部位を同定します。シーケンスされたリード配列データを、リファレンスゲノム配列にマッピングすることでピークが形成され、専用の解析ツールを使って、ゲノム配列上のすべてのピーク領域をピックアップします。その際、ChIPサンプルとは別にInput DNAなどのコントロールサンプルのシーケンスデータが存在する場合は、コントロールサンプルのデータを考慮して、ChIPサンプルからのピーク検出を行うことが可能です。



◆ピーク検出

検出されたピークの情報は、テーブル形式のデータとして表示されます。テーブルには、各ピークのゲノム配列上の位置とピークの信頼性、さらにピーク近傍の遺伝子名をアノテーションとして付加することが可能です。また、無償プラグイン「Advanced Peak Shape Tools」を使用することで、外部データに基づいたピーク検出なども行うことが可能になります。

- Peak shape score: ChIPピークのスコア
- P-value: ChIPピークの信頼性

Chromosome	Region	Name	Center of peak	Length	Peak shape score	P-value	5' gene	5' distance	3' gene	3' distance
NC_000021	3534587-3534588	Peak	3534587.5	1	1.79	0.04	C2orf196	11885	RPL14P3	420455
NC_000021	3534762-3534763	Peak	3534762.5	1	1.74	0.04	C2orf196	13812	RPL14P3	418533
NC_000021	3542982-3542983	Peak	3542982.5	1	2.18	0.01	C2orf196	99888	RPL14P3	332958
NC_000021	3545442-3545443	Peak	3545442.5	1	1.66	0.05	C2orf196	135552	RPL14P3	296595
NC_000021	3572278-3572279	Peak	3572278.5	1	1.82	0.03	C2orf196	388275	RPL14P3	431893
NC_000021	3598438-3598439	Peak	3598438.5	1	2.84	4.17E-04	LOC268893	89765	RPS20P1	34339
NC_000021	3617048-3617049	Peak	3617048.5	1	1.79	0.04	RPS20P1	151886	PPP1R2P2	105858
NC_000021	3655409-3655410	Peak	3655409.5	1	1.71	0.06	RPL28AP3	43883	SETD4	0
NC_000021	3640758-3640759	Peak	3640758.5	1	2.49	6.66E-03	DBR1	39742	LOC101932886	0
NC_000021	3640800-3640801	Peak	3640800.5	1	2.03	0.01E-03	DBR1	0	DOPEY2	22745
NC_000021	3645588-3645589	Peak	3645588.5	1	1.86	0.03	DBR3	9880	DOPEY2	8894
NC_000021	3645914-3645915	Peak	3645914.5	1	2.08	6.98E-03	DBR3	10141	DOPEY2	7595
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894
NC_000021	3645970-3645971	Peak	3645970.5	1	2.49	6.22E-03	DBR3	10778	DOPEY2	7841
NC_000021	3645955-3645956	Peak	3645955.5	1	2.08	6.98E-03	DBR3	9880	DOPEY2	8894



- ## ◆ De Novo Assembly

Contig sequence		Assembly report																																					
Ecoli FLX (single) contig 1 De Novo Assembly 100% CGCCGCGTGGCGGACGCTCCGCTGCCAGAAATCCACAGCCGGTTCCTCTGGCAGCACTGTAKAGTCTCTTGT 100% CGCCGCGTGGCGGACGCTCCGCTGCCAGAAATCCACAGCCGGTTCCTCTGGCAGCACTGTAKAGTCTCTTGT Ecoli FLX (single) contig 2 De Novo Assembly 100% CCDCAACCCCAACTCTCTCTGTGTAATA 100% CCDCAACCCCAACTCTCTCTGTGTAATA Ecoli FLX (single) contig 3 De Novo Assembly 100% ATAGAAACAATAATAAGATGGAGACCTAACCCAAATTTAGCAAGTGAAGA 100% ATAGAAACAATAATAAGATGGAGACCTAACCCAAATTTAGCAAGTGAAGA Ecoli FLX (single) contig 4 De Novo Assembly 100% CGCGATGGTGTGGAGAACTGGTGACAAAGCCAGCGTTACAAACAAGGAGATATGTTGTTTCGATCGATGGATGG 100% CGCGATGGTGTGGAGAACTGGTGACAAAGCCAGCGTTACAAACAAGGAGATATGTTGTTTCGATCGATGGATGG Ecoli FLX (single) contig 5 De Novo Assembly 100% TGAAGCTATTAAGGCGCAAGTAAGTCTGGT 100% TGAAGCTATTAAGGCGCAAGTAAGTCTGGT Ecoli FLX (single) contig 6 De Novo Assembly 100% CTTTGTCCGGGTAACTCTCTCTCAAGCCAGCGAGGCGGCCAGTCTCTCCAAAGGATGTGCAGCAATCTCGTCCATAT 100% CTTTGTCCGGGTAACTCTCTCTCAAGCCAGCGAGGCGGCCAGTCTCTCCAAAGGATGTGCAGCAATCTCGTCCATAT Ecoli FLX (single) contig 7 De Novo Assembly 100% CCGGTAGAATATTTTGGCTCTGGAATCTATGTCGCGGATACATGACCGCAAGTGA 100% CCGGTAGAATATTTTGGCTCTGGAATCTATGTCGCGGATACATGACCGCAAGTGA Ecoli FLX (single) contig 8 De Novo Assembly 100% ACGAAGGAAGTCAACATCATTTATCGCTAAGAAATGTATGCGATGCTCGC 100% ACGAAGGAAGTCAACATCATTTATCGCTAAGAAATGTATGCGATGCTCGC		1 DH10B-150K_R1 (paired) trimmed (paired) assembly summary report 1.1 Nucleotide distribution <table> <tr> <th>Nucleotide</th><th>Count</th><th>Frequency</th></tr> <tr> <td>Adenine (A)</td><td>1,106,952</td><td>24.7%</td></tr> <tr> <td>Cytosine (C)</td><td>1,148,039</td><td>25.6%</td></tr> <tr> <td>Guanine (G)</td><td>1,128,391</td><td>25.2%</td></tr> <tr> <td>Thymine (T)</td><td>1,102,978</td><td>24.6%</td></tr> <tr> <td>Any nucleotide (N)</td><td>187</td><td>0.0%</td></tr> </table> 1.2 Contig measurements (including scaffolded regions) <table> <tr> <th></th><th>Length</th></tr> <tr> <td>N75</td><td>57.85</td></tr> <tr> <td>N50</td><td>107.63</td></tr> <tr> <td>N25</td><td>176.16</td></tr> <tr> <td>Minimum</td><td>20</td></tr> <tr> <td>Maximum</td><td>326.31</td></tr> <tr> <td>Average</td><td>43.55</td></tr> <tr> <td>Count</td><td>10</td></tr> <tr> <td>Total</td><td>4,488.54</td></tr> </table>		Nucleotide	Count	Frequency	Adenine (A)	1,106,952	24.7%	Cytosine (C)	1,148,039	25.6%	Guanine (G)	1,128,391	25.2%	Thymine (T)	1,102,978	24.6%	Any nucleotide (N)	187	0.0%		Length	N75	57.85	N50	107.63	N25	176.16	Minimum	20	Maximum	326.31	Average	43.55	Count	10	Total	4,488.54
Nucleotide	Count	Frequency																																					
Adenine (A)	1,106,952	24.7%																																					
Cytosine (C)	1,148,039	25.6%																																					
Guanine (G)	1,128,391	25.2%																																					
Thymine (T)	1,102,978	24.6%																																					
Any nucleotide (N)	187	0.0%																																					
	Length																																						
N75	57.85																																						
N50	107.63																																						
N25	176.16																																						
Minimum	20																																						
Maximum	326.31																																						
Average	43.55																																						
Count	10																																						
Total	4,488.54																																						

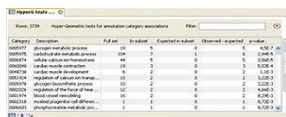
◆BLAST解析

De Novo Assemblyで作成したコンティグ配列データなど、各種配列データを使用して、BLASTによるデータベースの相同性検索を行うことが可能です。解析結果はテーブル形式で表示され、各コンティグ配列と最も相同性の高かったデータを、容易に確認することができます。

Query	Number of hits	Lowest E-value	Accession (E-value)	Description (E-value)
EcoliFLX (single) contig 1 De Novo Assembly	8	0.36	B8DXP5	Fructan 1-xylohydrolase OS=Leymus chinensis GN=1-1-FEH PE=2 SV=1
EcoliFLX (single) contig 2 De Novo Assembly	3	3.49	Q07262	1-aminocyclopropane-1-carboxylate synthase OS=Nicotiana tabacum GN=ACG51 PE=2 SV=1
EcoliFLX (single) contig 3 De Novo Assembly	5	1.32	Q19705	Uncharacterized protein 025R OS=Ivertebrate invertecent virus 3 GN=IV3-025R PE=4 SV=1
EcoliFLX (single) contig 4 De Novo Assembly	0		not available	not available
EcoliFLX (single) contig 5 De Novo Assembly	4	0.05	P93942	14-3-3-like protein A OS=Nicotiana tabacum PE=2 SV=1
EcoliFLX (single) contig 6 De Novo Assembly	3	0.12	Q42766	14-3-3 protein homolog OS=Candida albicans GN=BMH1 PE=1 SV=2
EcoliFLX (single) contig 7 De Novo Assembly	9	0.02	P46731	18K2_MYC4V 18K2_MYC4V 18 kDa antigen 2 OS=Mycobacterium avium PE=3 SV=1
EcoliFLX (single) contig 8 De Novo Assembly	3	2.59	P05080	Putative 194 kDa protein OS=Tobacco rattle virus (strain SYM) PE=4 SV=2
EcoliFLX (single) contig 9 De Novo Assembly	5	1.29	P15711	104K_THEPA 104K_THEPA 104 kDa microneph/hoptry antigen OS=Thelazia parva GN=TP04_043
EcoliFLX (single) contig 10 De Novo Assembly	4	0.75	Q83943	Replication protein 1a OS=Olive latent virus 2 (isolate Italy) GN=ORF1a PE=3 SV=1
EcoliFLX (single) contig 11 De Novo Assembly	1	4.81	Q01526	14-3-3-like protein GF14-12 OS=Zea mays GN=GRF2 PE=2 SV=2
EcoliFLX (single) contig 12 De Novo Assembly	1	3.55	Q19709	Uncharacterized protein 011L OS=Ivertebrate invertecent virus 3 GN=IV3-011L PE=4 SV=1
EcoliFLX (single) contig 13 De Novo Assembly	1	2.61	P70AT7	Fructan 1-xylohydrolase OS=Hordeum vulgare GN=1-FEH PE=2 SV=1
EcoliFLX (single) contig 14 De Novo Assembly	25	0.63	P28726	Replication protein 1a OS=Peanut stunt virus (strain J) GN=ORF1a PE=3 SV=1
EcoliFLX (single) contig 15 De Novo Assembly	6	2.97	Q07262	1-aminocyclopropane-1-carboxylate synthase OS=Nicotiana tabacum GN=ACG51 PE=2 SV=1

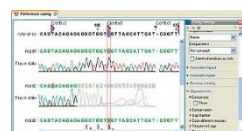
サンガーシークエンスデータ解析

- ✓ CLC Genomics Workbenchは同社製品「CLC Main Workbench」のすべての解析モジュールも搭載
- ✓ DNA、RNA、Proteinなどの配列解析機能に加え、マイクロアレイにも対応



データベース検索

GenBankのすべての配列にアクセスすることができます。
PubMedルックアップ/ウェブベースのシーケンスデータルックアップ



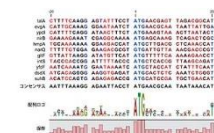
アセンブリ

高度なコンティグアセンブリ/トレースデータのインポート、表示、および編集
シーケンスのトリム、参照シーケンスを使用せずに組み立てる
既存のコンティグへのアセンブル/コンティグの表示、コンティグの編集



プライマーデザイン

さまざまな実験条件設定でプライマーを設計することができます。



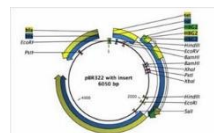
配列アライメント

多重配列アライメント(2つの独自アルゴリズム)
複数の配列アライメントにおいて個々の配列のアノテーションを表示可能
シーケンスロゴにより高度に保存された領域の情報の確認



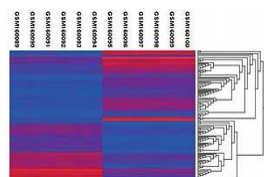
系統樹

UPGMA, NJ, MLの3つのアルゴリズムで系統樹を作成できます。



クローニング

in silicoクローニングのグラフィック表示
Gibson Assemblyにも対応



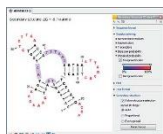
遺伝子発現解析

GEO・Affymetrix Microarray・illuminaデータインポート
変換と正規化/ t検定・ANOVA・FDR補正
ヒートマップ/ヒストグラム/散布図・MAプロット・ボックスプロット・火山プロット
クラスターリング(階層的クラスターリング、k-means、Medoid周辺の分割(PAM))
遺伝子セットエンリッチメント解析(GSEA)/アノテーションテスト



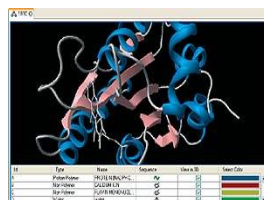
BLAST検索

BLASTを使用したWebベースのシーケンス検索/ local BLASTデータベースの作成
BLAST結果のテーブルビュー/ BLAST検索のグラフィカルビュー



RNA二次構造

二次構造予測/二次構造のグラフィック表示と編集
構造とエネルギー寄与の表形式ビュー
シーケンスビューでの記号表現

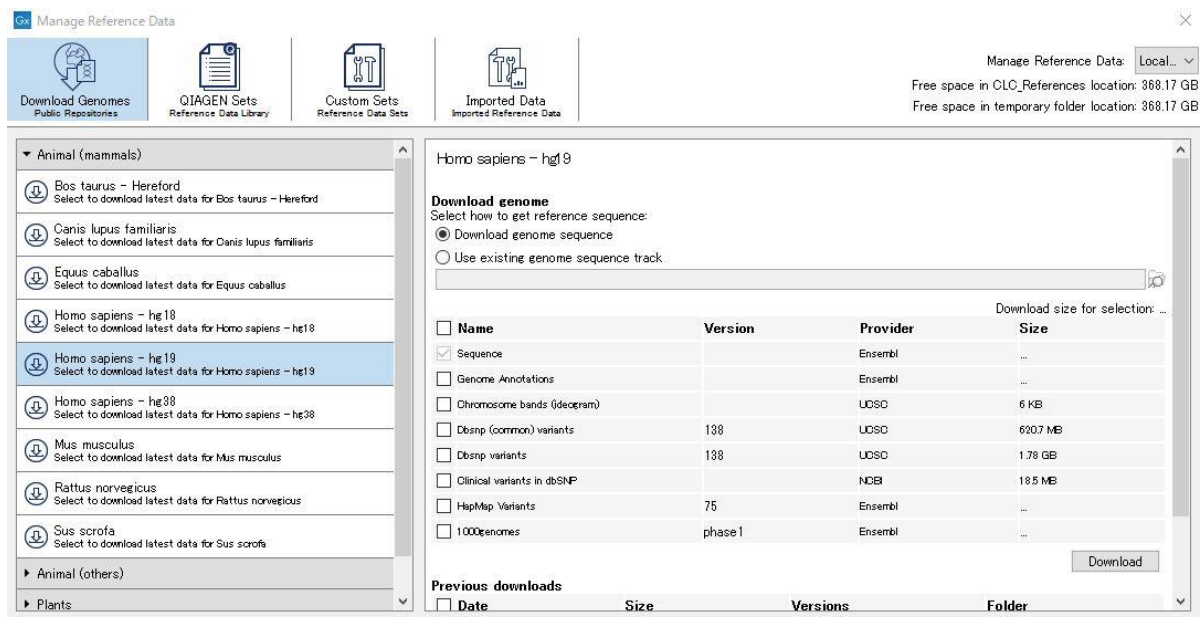


タンパク質分析

統合3次元分子ビューア1/2次元および3次元シーケンスの同時表示
抗原性解析/タンパク質電荷解析/ PEAMドメイン検索/疎水性分析
タンパク質からDNAへの逆翻訳
タンパク質分解的切断検出、シグナルペプチドの予測(SignalP)
膜貫通ヘリックス予測(TMHMM)/二次タンパク質構造予測

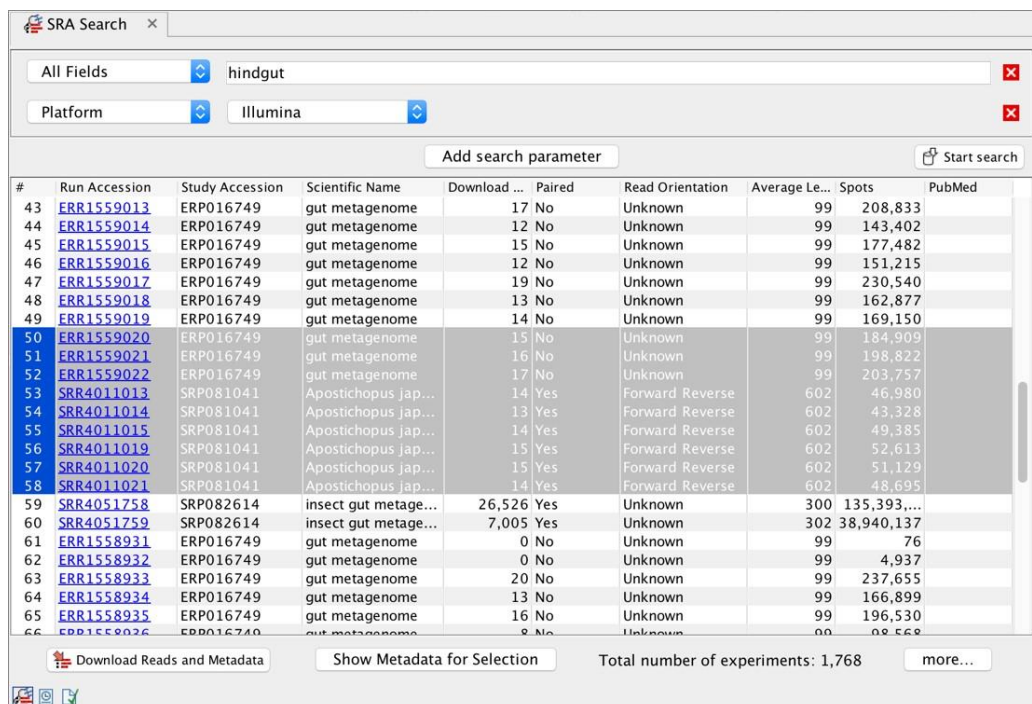
その他の機能

◆ゲノム・アノテーションデータ取得



ゲノム・アノテーションを取得するためのDownload Reference Genome Dataウィザードを搭載し項目を選択していただくだけで、目的とするリファレンスゲノム配列やアノテーション情報を取得できます。

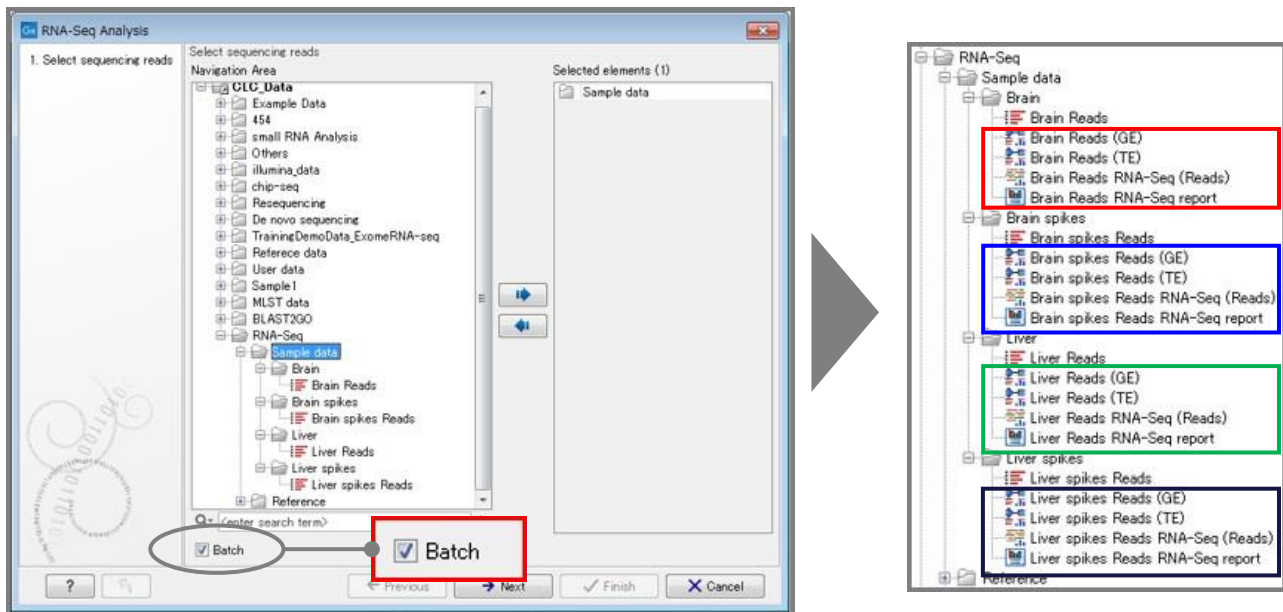
◆SRAダウンロード



NCBIのSRA(Sequence Read Archive)にアクセスし、Accession IDやキーワードなどを指定してデータの検索を行い、登録レコードのリード配列データおよび各サンプルの詳細情報(メタデータ)のダウンロードが可能です。ダウンロードしたデータは、本ソフトウェアでそのまま解析に使用することができます。

その他の機能

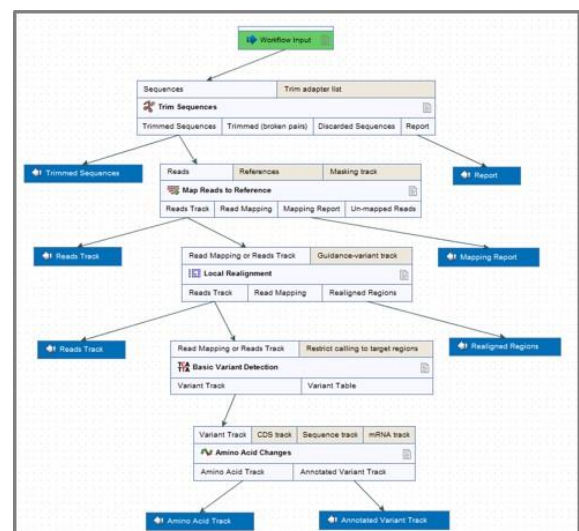
◆バッチ解析



複数のサンプルデータが存在する場合、同一の解析条件で、全サンプルを一括して解析を行うことができます。ワークフロー機能と組み合わせることで、データ解析の作業効率を大幅に向上させることが可能です。

◆ワークフロー作成

複数のツールを組み合わせる解析を行う場合、使用するツールの順番などをワークフローとして登録することが可能です。各ツールのパラメータも任意に設定でき、データ解析をルーチン的に行う場合の解析パイプラインとして使用することができます。また無償プラグイン「Biomedical Genomics Analysis」をインストールすることで、全ゲノム解析、全エクソーム解析、ターゲットアンプリコン解析、全トランスクリプトーム解析などのヒト疾患研究用の変異解析ワークフローを使用できるようになります。



◆分子バーコード解析

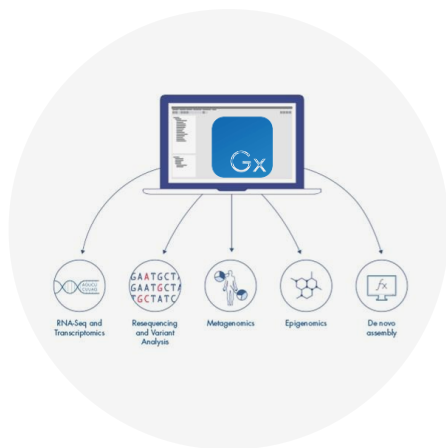
QIAGEN社のQIAseq Panelを使用した、分子バーコード解析に対応しています。パネルの各種デザインファイルを、分子バーコード解析用の専用ツールまたは解析パイプラインにそのまま使用し、ゲノム上の遺伝子変異、RNA-Seqデータによる遺伝子発現解析や融合遺伝子の検出が可能です。



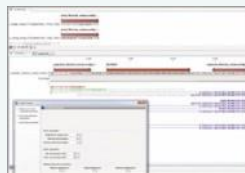
CLC Genomics Workbench Premium

- ✓ CLC Genomics Workbench Premiumでは、CLC Genomics Workbenchの機能に加え、より多彩な機能が使用できます。

CLC Genomics Workbench



新規ゲノム配列決定



CLC Genome Finishing Module

超高速変異解析



CLC LightSpeed Module

微生物メタゲノム解析



CLC Microbial Genomics Module

シングルセル解析



CLC Single Cell Analysis Module



CLC Genome Finishing Module

PacBioロングリードのアセンブルや近縁種ゲノムへのマッピングなど、完全長ゲノム配列決定用のツールを搭載



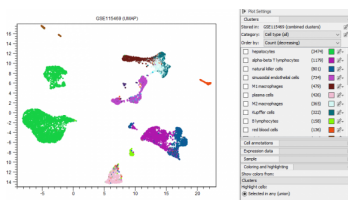
CLC LightSpeed Module

参照ゲノム配列へのマッピングと生殖細胞系列の変異の検出を超高速で実行するツールを搭載。



CLC Microbial Genomics Module

16S rRNAやメタゲノム解析による微生物組成や遺伝子機能解析、および病原菌のタイピングと系統樹解析用のツールを搭載



CLC Single Cell Analysis Module

single-cell (sc) RNA-seqデータ、scTCR-seqデータ
の他、scATAC-seqデータ解析用のツールを搭載



- De Novo Assemblyで作成したコンティグ配列データなどを使用し、コンティグ配列間の重複やギャップ部位の確認、またはデータの結合などによって、より高精度なコンティグ配列を作成することができます。



-
- | Environment | HG001 (34x / 113Gb) | HG002 (30x / 101Gb) |
|-------------|---------------------|---------------------|
| Cloud | 21 min | 19 min |
| Workstation | 25 min | 23 min |
| Laptop | 84 min | 70 min |

10



CLC Microbial Genomics Module

- ✓ 16S rRNAやショットガンメタゲノムデータを用いた菌種組成解析と遺伝子機能組成解析、病原菌のタイピングや疫学解析用のプラグイン
- ✓ 解析パイプラインのワークフローが最初から組み込まれており、簡単な操作で解析が可能
- ✓ バクテリアやウイルスなど、解析に使用するリファレンスゲノムデータなども、専用ツールから簡単にダウンロードが可能

微生物組成解析

- サンプル中の微生物あるいは分類の存在比を解析
- リード配列のQCチェックや、ホストDNAの除去機能も搭載

遺伝子機能組成解析

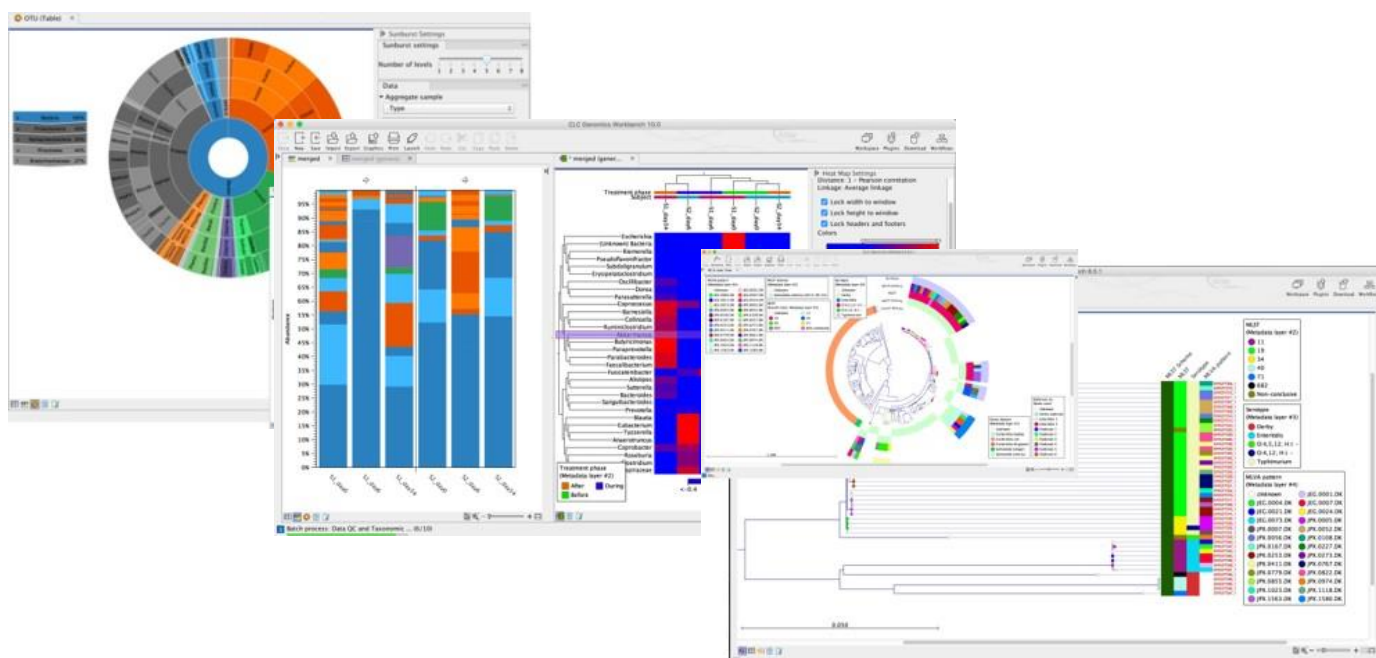
- サンプル中の遺伝子の存在比を解析
- コンティグ配列に遺伝子機能アノテーションを付加

多様性解析

- 微生物組成データを用いて、アルファ多様性やベータ多様性を計算
- OTU配列データを用いたUniFrac解析
- 組成データのサンプル間比較やヒートマップ作成

病原菌タイピングと疫学解析

- MLST法による病原菌のタイピングおよび薬剤耐性の確認
- K-mer Treeによる複数菌種のゲノム配列の類似度の比較
- SNP Treeによる分子系統樹の作成



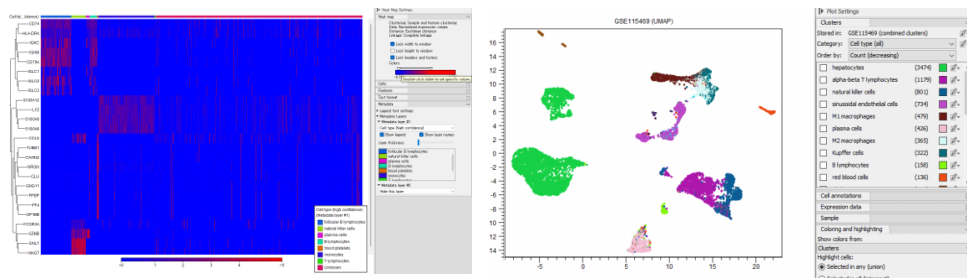


CLC Single Cell Analysis Module

- ✓ Single-cell シークエンスデータ解析用プラグイン
- ✓ 解析パイプラインのワークフローを搭載しており、fastqファイルなどの生データを取り込み、パラメーターを設定するだけでQCから各種グラフの出力まで簡単に解析を行うことができる
- ✓ 解析に使用するリファレンスゲノムデータ(ヒト・マウス)なども、簡単にダウンロードが可能

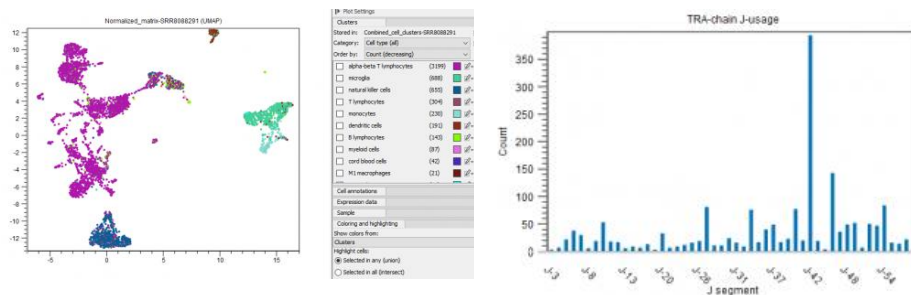
遺伝子発現解析

- 遺伝子発現による細胞のクラスタリングやRNA velocity解析を行うことが可能。
- 解析結果は、ヒートマップ、ドットプロット、バイオリンプロット、次元削減(tSNE, UMAP)、phase portraitなど様々なグラフを出力可能



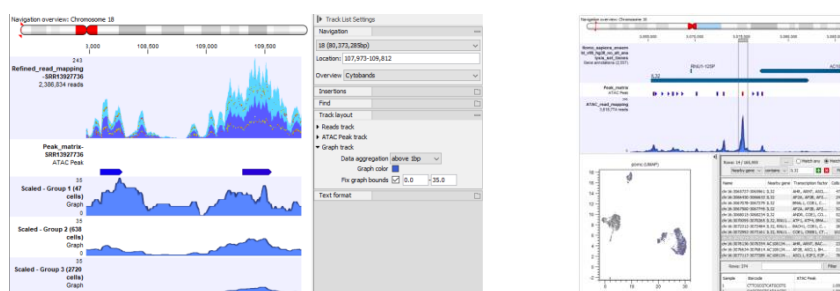
免疫レパトア解析

- TCR-seqデータから単一細胞レベルのクローノタイプ解析を行うことが可能
- レパトア解析専用のワークフローを搭載



クロマチンアクセシビリティ解析

- ATAC-seqデータから単一細胞レベルでクロマチンの状態を解析可能
- 近傍遺伝子や転写因子結合モチーフの探索も可能
- クロマチンアクセシビリティ解析専用のワークフローを搭載



ライセンスタイプ

固定ライセンス

- ✓ Workbenchをインストールした1台のPCでのみ利用可能です。
- ✓ CPUの論理コア数が64コアを超える(Hyper-threadingを含む) PCではご利用いただけません。
- ✓ リモートデスクトップによるアクセスは出来ません。

ネットワークライセンス

- ✓ 研究室内のライセンス管理用サーバーに接続された全てのPCで使用可能です※1。
- ✓ ネットワーク上のPCであれば、複数台にインストールできますが、同時に使用できるのは1ライセンスにつき1ユーザーのみとなります※2。
- ✓ リモートデスクトップによるアクセスが可能です。

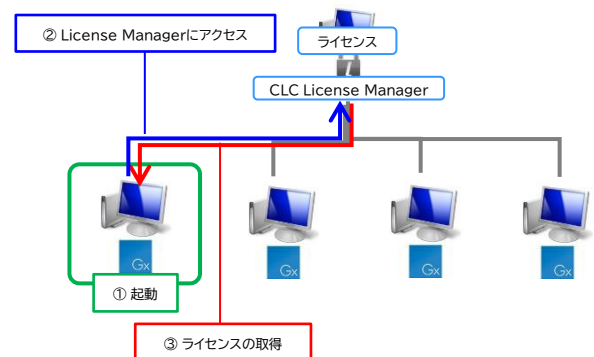
※1: Network License Managerは常時電源が入っている必要があります。

※2: 同時に複数での使用をご希望される場合、その分のライセンスをご購入いただく必要があります。

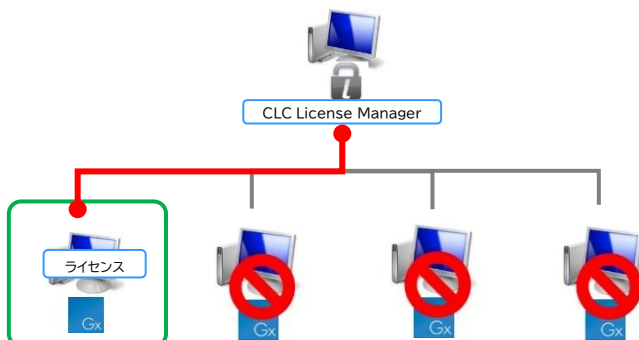
ネットワークライセンスのイメージ

Workbenchを使用するにはライセンスが必要となり、通常、Network License Managerが保有しています。

Workbenchをインストールしているコンピューターの一つを起動すると、ライセンスサーバーにアクセスし、ライセンスを取得します。その後Workbenchが使用可能となります。

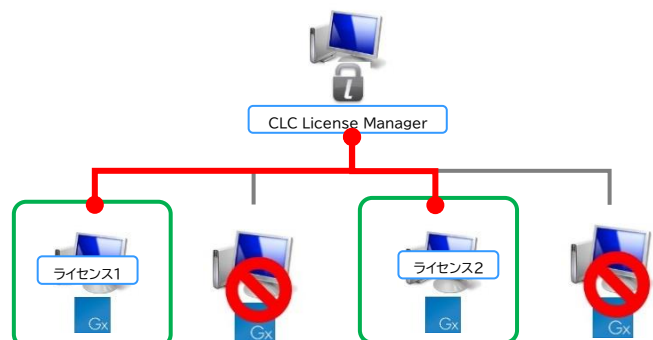


ネットワークライセンスが1ライセンスの場合



ライセンスを取得したコンピューターがWorkbenchを使用している間、他のコンピューターは取得するライセンスが無いいため、それらのWorkbenchは使用することが出来ません。

ネットワークライセンスが2ライセンスの場合



Workbenchを2台同時に使用するには2ライセンス必要となります。同様に、使用台数を増やすには、ライセンスをその台数分だけ追加する必要があります。

インテグレーション&サポートサービス

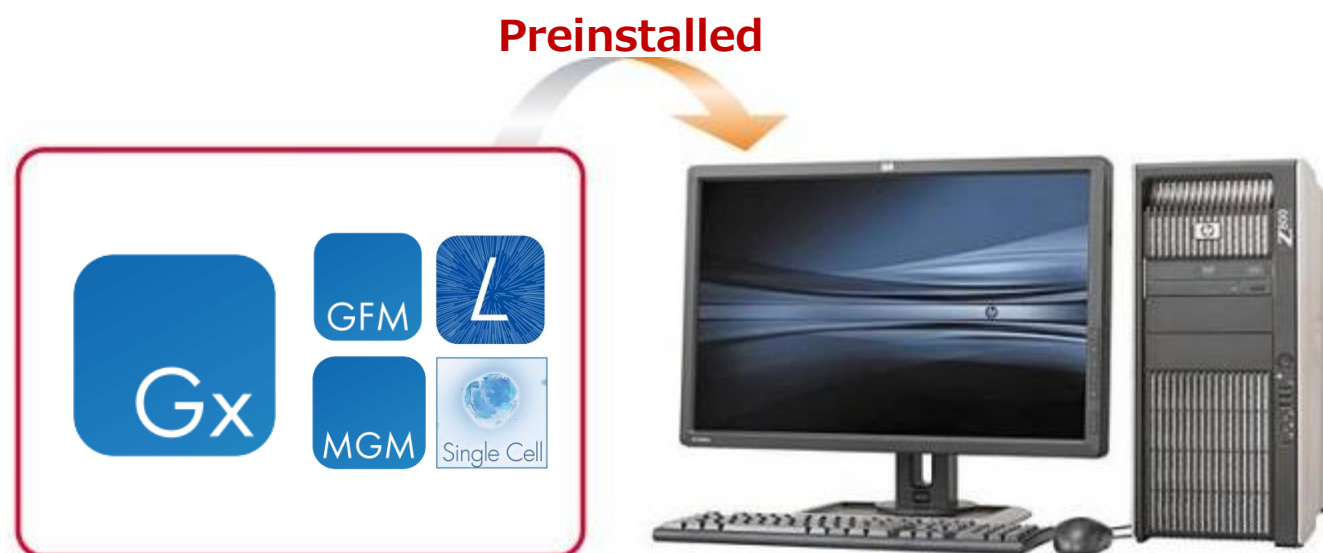
次世代シーケンサーから出力されるデータは膨大であるため、解析を行うコンピューターもそれに対応したスペックのものが必要となります。

「フィルジェンで適切なスペックのコンピューターを選択して、CLC Genomics Workbenchをプレインストールして販売してほしい。」というお客様のご要望にお応えし、

メーカー推奨スペックのコンピューターにソフトウェアをプレインストールしたセット品の販売を行っております。

また、お客様のご希望にあわせたスペックのコンピューターでのセット販売にも対応いたしております。

お気軽に弊社までお問い合わせください。



解析システム仕様例

例1:

細菌・ウイルスのゲノム解析やヒト・マウスのRNA-Seq解析など

OS : Windows11(64bit)

メモリ : 64GB RAM

CPU : インテル®Xeon®x2

例2:

ヒト全ゲノムリシーケンス解析やDe Novoシーケンスなど

OS : Windows11(64bit)

メモリ : 128GB RAM

CPU : インテル®Xeon®x2

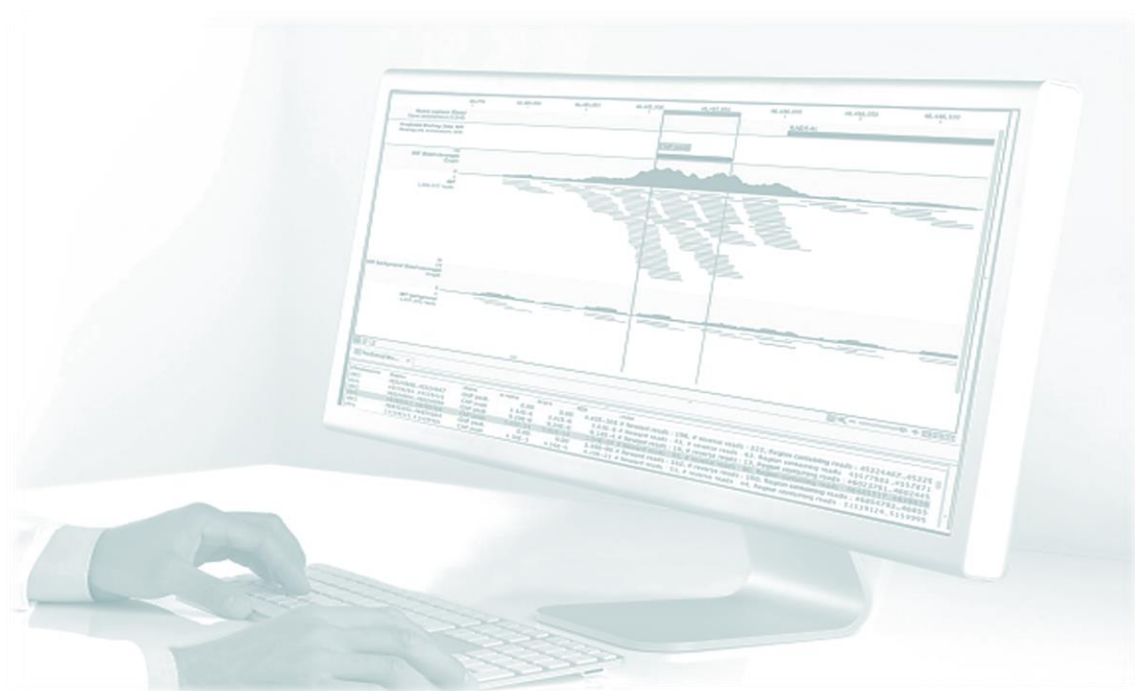
安心・充実のサポート体制



CLC Genomics Workbenchのインテグレーションサービスでは解析ソフトウェアの1回分のオンサイトトレーニングがパッケージされています。弊社スタッフが直接現地に伺いまして、ソフトウェアの操作方法の説明をいたします。

また、弊社ではバイオインフォマティクスソフトウェア製品に関するWebセミナーを開催しております。オンサイトトレーニング実施後でも、こちらのWebセミナーにご参加いただくことで、ソフトウェアに関するサポートを受けることが可能です。

※ お見積り内容、ソフトウェアの種類によっては、対応しない場合もあります。詳細は弊社までお問い合わせください。



国内販売代理店



フィルジェン 株式会社
バイオインフォマティクス部

【お問い合わせ】

〒459-8011 愛知県名古屋市緑区定納山1丁目1409番地

TEL : 052-624-4388 FAX : 052-624-4389

メール : biosupport@filgen.jp URL : <https://filgen.jp/>

代理店

(JUN.,2023)