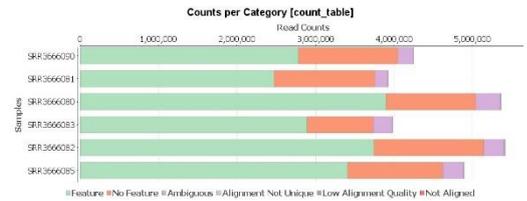


Transcriptomics

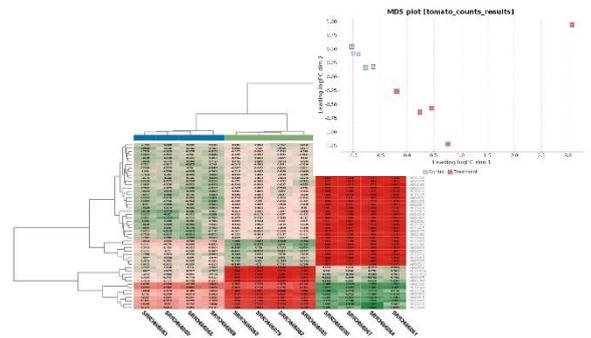
生データからRNA配列を組み立てる

NGSより出力されるデータからクオリティチェックを実行し、リードをフィルタリングし、低品質のベースを削除します。リファレンスゲノムの無い生物種でもDe-Novo Assemblyによって配列を組み立てることができます。



統計

さまざまな統計グラフが、アセンブリおよび定量化プロセスに関する追加情報、および結果の品質評価を提供します。



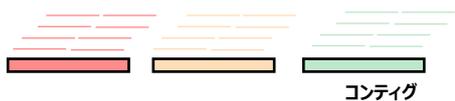
豊富な視覚化

インタラクティブなヒートマップは、異なる遺伝子とサンプルの発現値の違いと類似性を直感的に確認することに役立ちます。その他MDSプロット、ボルケーノプロット、MAプロットなどの多彩なグラフ化が可能です。

スプレッドシート

差次的発現の結果をソートおよびフィルタリングし、統計的基準を調整して重要な遺伝子进行评估し、それらを機能情報と組み合わせて生物学的洞察を理解することが可能です。

№	Tags	Name	FC	logFC	logCPM	P-Value	FDR
23	+	NG00873	9.89874	3.30724	5.70897	2.94112-175	1.71212-173
24	+	NG00702	6.96864	2.80046	6.47333	1.17179-173	6.62029-172
25	+	NG00701	8.92625	3.11907	7.72212	4.95979-172	2.72222-170
26	+	NG00606	-7.40596	-2.42711	6.92197	1.80179-170	9.61430-169
27	+	NG00757	-50.77995	-5.45188	5.25238	6.63465-170	3.18435-169
28	+	NG01463a	-4.27027	-2.44453	9.96006	9.40122-170	4.24792-169
29	+	NG01189	18.48991	3.37970	5.49797	8.46812-168	4.17012-164
30	+	NG00176	39.04366	5.28702	5.51294	1.13162-165	5.43432-164
41	+	NG00945	5.79089	2.53379	8.43118	6.68236-165	3.22462-163
42	+	NG00006	11.24897	3.72778	5.08441	3.59336-164	1.64752-162
43	+	NG01466	-4.98124	-2.31708	8.64767	1.16662-158	5.21182-157
44	+	NG00999	5.96829	2.58455	8.61918	2.49792-154	1.17792-154



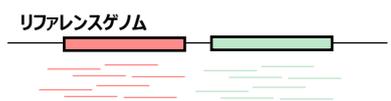
RNA-seq De-Novo Assembly

Trinityで短いリードからリファレンスゲノムなしで配列を再構築します。さらにBUSCOによりアセンブルの完全性を評価できます。



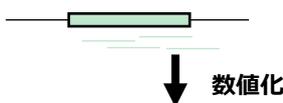
Custer & Predict Coding Regions

作成されたコンティグなどをCD-HITで類似配列をクラスター化しTransDecoderを使用して転写産物配列内のコーディング領域を検出します



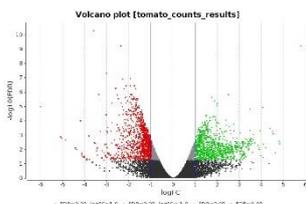
RNA-Seq Alignment

超高速ユニバーサルRNA-seq alignerとして知られるSTARを使用して、RNA-seqデータをリファレンスゲノムに合わせます。



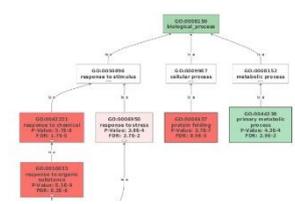
Quantify Expression

リファレンスゲノムの有無にかかわらず、HTSeqまたはRSEMによる遺伝子または転写産物レベルでの発現の定量化が可能です。



Dierencial Expression Analysis

実験条件間または経時的に異なる発現遺伝子を検出します。有名な統計パッケージNOISeq、edgeR、maSigProを使用します。豊富な視覚化により結果を解釈できます。



Enrichment Analysis

Functional Analysisモジュールと組み合わせることによって、過剰および過少発現の生物学的機能を同定することができます。



Transcriptomicsモジュールでなにができる？

【使用したデータセット】

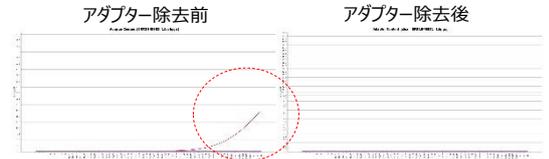
Ascaridia galli (飼育鳥に感染する腸内寄生虫)
コントロール x 3 FLBZ処理 (駆虫薬) サンプル x 3を(fastq)をダウンロードしました。



A.galliがFLBZ耐性メカニズムを発達させる可能性があるので耐性機構について理解することが求められた。

【データの品質チェックと前処理】

OmicsBoxに搭載されたFastQCで品質をチェックします。本データではアダプターがシーケンスに残ったままでしたので、OmicsBoxに搭載されたTrimmomaticでトリミングを行いました(右図)。クオリティチェックでは「PASS」や「FALE」など視覚的に簡単に配列の品質を理解することができ、トリミングもスコアや長さなど様々な処理により配列の品質を上げることができます。

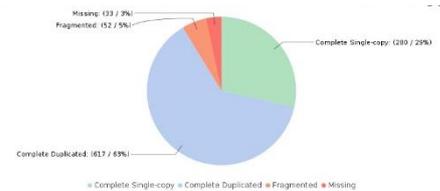


アダプタートリミング後クオリティチェックをしたところ、除去が図からも確認できた。

【新規トランスクリプトームの組み立てと完全性評価】

データ品質を上げた後、OmicsBoxではTrinityを使用したRNA-Seq de novo assemblyを実行できます。この解析によりリファレンスゲノムなしで配列を組み立てることができます。

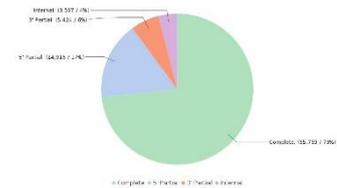
またSuperTranscriptsなど非モデル生物に適したオプションも指定できます。この解析には通常高スペックなシステム要件を必要としますが、OmicsBoxの場合解析はメカのサーバーにて行われるため高価なPCなど初期費用を抑えることができます。作成されたアセンブリの完全性を評価するために、OmicsBoxにはBUSCO (Benchmarking Universal Single-Copy Orthologs)が搭載されています。



290,713のコンティグが作成された。線虫オースログの92%は完全に予想され、5%は部分的に見つかり、3%は欠落した。

【オープンリーディングフレーム予測】

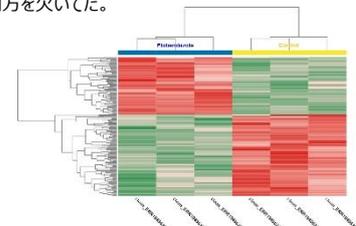
良質なアセンブリデータを取得後、OmicsBoxでトランスクリプト内のORF予測を検索しました。このツールは、TransDecoderに基づいておりPfamドメイン検索データに基づいてデータセットから最適なORFを抽出および分類できます。ORFの予測後、本データでは遺伝子ごとに1つだけのアイソフォームがあったと仮定し独自のフィルタリングにより19,700の遺伝子が下流分析のために保存されました。



90,000個のORFが予測された。転写産物の73.3%が完全で、16.6%が開始コドンの欠落、6.1%が停止コドンの欠落、4.0%が開始コドンと停止コドンの両方を欠いてた。

【機能アノテーションとInterProScan】(Functional Analysis Moduleを使います。)

SwissProtデータベースに対する BlastX→マッピング→アノテーション+CloudIPSにより8,661個の遺伝子の機能アノテーションが得ることができました。



【発現値定量・比較解析】

OmicsBoxではリファレンスゲノムの有無に関わらず発現値の定量が可能です。ここではDe novo assemblyで組み立てフィルター処理されたデータと初めのステップで作成されたクリーンリードを使用して定量化を行いました。さらにOmicsBoxでは実験条件間に異なる発現遺伝子を検出することができます。これにより統計的に高発現遺伝子・低発現遺伝子の検出を行うことができます。結果をヒートマップやボルケーノプロットなどの視覚的な図にまとめることも可能です。

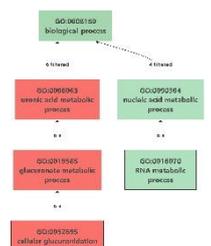
コントロール(未処理)とFLBZ処理サンプルを比較したところ58個の高発現遺伝子と89個の低発現遺伝子を検出した。高発現遺伝子の中ストレス状態、薬物排泄、解糖系、低発現遺伝子の中からATPおよびエネルギー生産に関する遺伝子を検出した。(これらがFLBZへの曝露によって影響を受ける可能性を示唆)

【エンリッチメント解析】(Functional Analysis Moduleを使います。)

OmicsBoxでは発現変動を示した遺伝子についてどのような機能が豊富に含まれるか統計的に解析することができます。

予想に反して、エンリッチメント解析は、主要な薬物排出経路であるグルクロン酸代謝が低発現遺伝子間で過剰に含まれることを示した。

OmicsBoxを使用することで、リファレンスゲノムの無い配列を1から組み立てそれらと比較して薬剤耐性機構について関与する可能性がある遺伝子を推定することができた。



Transcriptomicsモジュールではリファレンスゲノムの有無にかかわらず実績のある様々なRNA-seq解析ができます。

高価なPC無しで
高速・高品質に
解析

173M
使用したリードデータ量

91%
高品質のトランスクリプト

12時間
解析時間