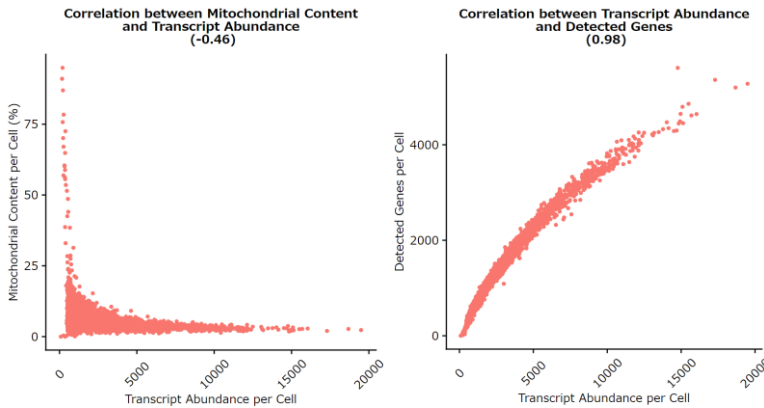


a) QC 情報

シングルセルRNA-seqにおける遺伝子発現データ（1次解析結果）から、RNA分子数・遺伝子数・ミトコンドリアRNA量などを確認し、ノイズの除去を行います。これらの指標は、解析されたサンプル及び得られたデータの品質を評価するための有用な情報となります。ここでは、上記3つの指標における、散布図およびバイオリンプロットをご提供いたします。

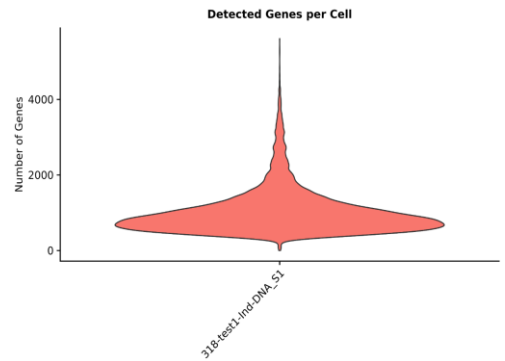
● 散布図

- ✓ Correlation between Mitochondrial Content and Transcript Abundance (左下)
- ✓ Correlation between Transcript Abundance and Detected Genes (中央)



● バイオリンプロット

- ✓ Detected Genes per Cell (右)
- ✓ Mitochondrial Content per Cell
- ✓ Transcript Abundance per Cell



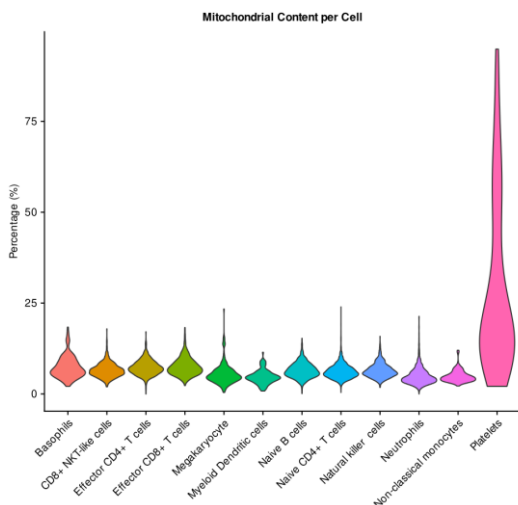
b) 機能解析 & 次元削減

詳細な遺伝子発現を評価するため、データセットを正規化（SCTransform*1）し、主成分分析やクラスタリング解析を行います。その後、各クラスターに対するアノテーション付加により、代表的な細胞種や各細胞が由来する臓器を推測します。

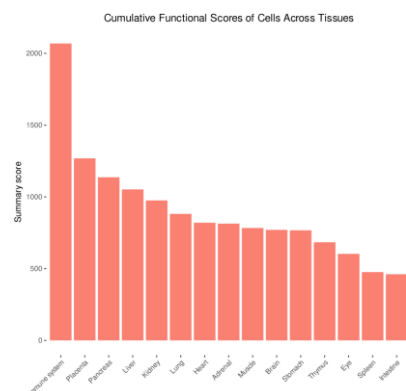
*1: Hafemeister, Christoph, and Rahul Satija. "Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression." *Genome biology* 20.1 (2019): 1-15.

● 機能解析

各クラスターを横軸に取り、いくつかの尺度に注目したバイオリンプロット



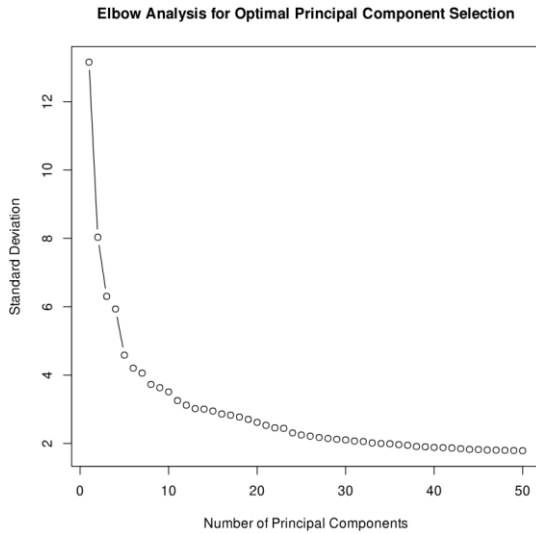
各細胞が由来する臓器の推定結果を示す棒グラフ



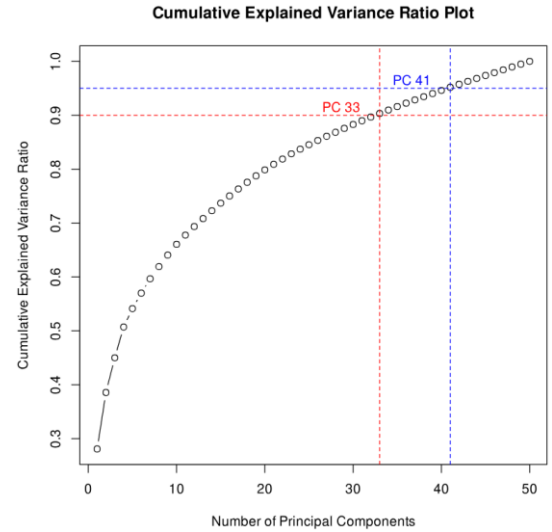
- * その他に、以下のタイトルのバイオリンプロットをご提供いたします。
- ✓ Transcript Abundance per Cell
 - ✓ Detected Genes per Cell

●次元削減

次元削減において重要となる、主成分の数を決定するためのエルボープロット。



各主成分が寄与する分散の累積グラフ。データの大部分を説明するに十分な主成分の数を定め、後段の次元削減に活用します。

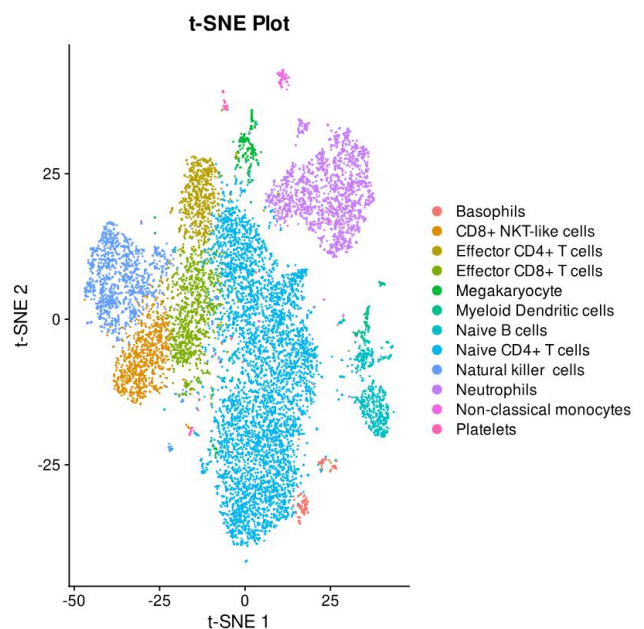
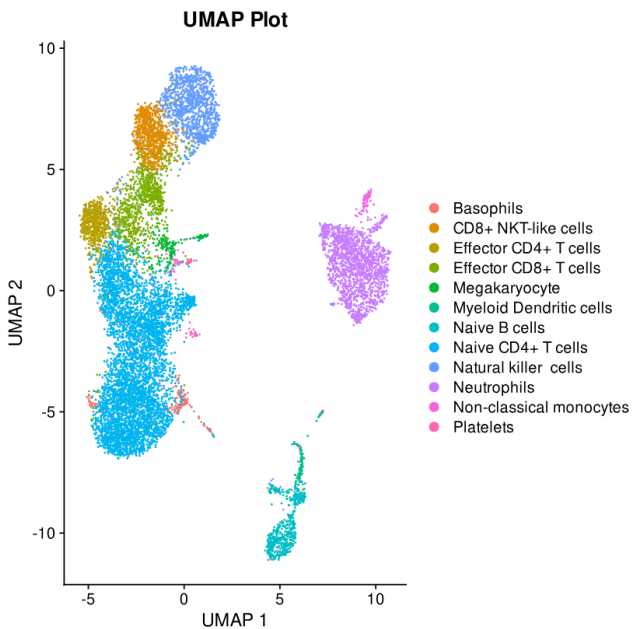


UMAPおよびt-SNEは、多次元データを2Dまたは3Dのグラフに変換します。

これらのグラフでは、類似の遺伝子発現を有する細胞が互いに近くに表示されます。クラスタリングや機能アノテーションの結果と統合することで、解析データが有する生物学的な特徴を視覚的に理解することが可能となります。

UMAPは、データが有する遺伝子発現パターンにおいて、グローバルな構造の理解を促進します（左図）。

t-SNEは局所的な構造の理解に適していると考えられています（右図）。各クラスターにおいて、細胞タイプの推定に基づくアノテーション結果を提供します。



c) DEG 解析 & ヒートマップ

●DEG解析

DEG (Differential Expression Genes) 解析では、特定の категорияにおいて、遺伝子発現が特異的な遺伝子を統計的に抽出します。当解析においては、クラスタリングによって得られた各クラスターをカテゴリーとし、DEGをウィルコックス検定によって同定します。

クラスター16における p値によって昇順にソートしたDEG
(解析結果はエクセルファイルによる表形式によって提供)

| p_val | avg_log2FC | pct.1 | pct.2 | p_val_adj | cluster | gene |
|-------|------------|-------|-------|-----------|---------|--------|
| 0 | 2.197 | 0.845 | 0.063 | 0 | 16 | GNLY |
| 0 | 1.381 | 0.867 | 0.347 | 0 | 16 | CCL5 |
| 0 | 1.099 | 0.687 | 0.131 | 0 | 16 | PRF1 |
| 0 | 1.038 | 0.557 | 0.110 | 0 | 16 | GZMA |
| 0 | 0.911 | 0.526 | 0.048 | 0 | 16 | KLRD1 |
| 0 | 0.814 | 0.480 | 0.050 | 0 | 16 | GZMB |
| 0 | 0.649 | 0.473 | 0.067 | 0 | 16 | NKG7 |
| 0 | 0.390 | 0.253 | 0.006 | 0 | 16 | SH2D1B |
| 0 | 1.686 | 0.973 | 0.318 | 0 | 15 | CCL5 |

●ヒートマップ

DEG解析後、クラスターごとに順位が高い複数の遺伝子に注目して遺伝子発現量を視覚化し、ヒートマップを作成します。ヒートマップでは、クラスターごとに上位5個の遺伝子を選択します。

(縦軸：遺伝子、横軸：クラスターラベル)

