

機能ゲノム学 (2014 年 5 月 14 日) 課題の回答とコメント

課題 1: GSE7623 の「24 サンプル全てを用いて RMA 前処理法を実行した結果」と「最初の 2 サンプルのみを用いて RMA 前処理法を実行した結果」において、赤枠内の同一サンプル間で得られる数値が異なる理由を述べよ。

「RMA 法はサンプル間の平均値を使うため、サンプル数が変わると数値が異なるため」とか「quantile 正規化法を採用しているから」とか「サンプル数の違いによる平均値の違い」みたいなことが書かれていれば正解です。

課題 2: RMX (RobLoxBioC) についても同様の比較を行い、正規化の特徴について簡潔に述べよ (per-array basis or multi-array basis)。ヒントは教科書 p39 の表 2-1。

「per-array basis」とか「グローバル正規化法を採用している」とか「アレイごとに独立して正規化している」とか「MAS5 に近い方法」とか「サンプル数が変化しても結果が変わらない」みたいなことが書かれていれば正解です。「Summarization score 算出時に Tukey's biweight のかわりに rmx estimator を利用している」という記述も OK です。

自由記載欄