

農学生命情報科学特論 II (NGS ハンズオン講習会第 3 部) 課題

以下に示す計 5 課題のうち 2 つを行え (3 つ以上行ってもよいが加点はしない)。

1. マッピング系 (2016 年 08 月 02 日分)

スライド 5-39 では、乳酸菌 (*Lactobacillus casei* 12A 株) RNA-seq データを、同じ乳酸菌ゲノム配列にマップした。同一種 (同じ株) のリファレンスゲノムがない場合を想定し、別の乳酸菌 (例えばスライド 98 の *Lactobacillus hokkaidonensis* L00C260^T) など近縁種のゲノム配列をリファレンスとしてマッピングを実行せよ。用いたリファレンス配列、手順、結果と考察を示せ。

・ *Lactobacillus hokkaidonensis* L00C260^T のゲノム配列は、例えば以下の URL から取得可能 :

http://bacteria.ensembl.org/Lactobacillus_hokkaidonensis_jcm_18461/Info/Index

・ マッピングに用いるプログラムの種類は問わない

2. ゲノムサイズ推定系 (2016 年 08 月 03 日分)

KmerGenie を用いて (スライド 7-36)、長さ 60 bp、20,000 リードからなる仮想 NGS データ ([kadai_20160720.fasta](#)) のゲノムサイズを推定し、得られた結果について簡単に考察せよ。

入力ファイルの場所 : http://www.iu.a.u-tokyo.ac.jp/~kadota/kadai_20160720.fasta

3. *de novo* アセンブリ系 1 (2016 年 08 月 04 日分)

スライド 5-83 までは、主に forward 側リードのトリミングとアセンブリ結果の評価を行った。reverse 側リードについても複数の条件でトリミングを行ったデータを作成し、Rockhopper2 による *de novo* transcriptome assembly を行い、得られた結果について考察せよ。

用いたトリミング条件、手順、結果と考察を示せ。

4. *de novo* アセンブリ系 2 (2016 年 08 月 04 日分)

スライド 108-130 は、トリミング前のデータを用いた Trinity による *de novo* transcriptome assembly であった。(主に forward 側) リードのトリミングを行ったデータを用いて Trinity を実行し、得られた結果について考察せよ。用いたトリミング条件、手順、結果と考察を示せ。

5. 自由課題

NGS ハンズオン講習会第 3 部で学んだ内容を生かした課題を自由に設定し、レポートとしてまとめよ。例 1 : Bridger or BinPacker のインストールに取り組んだ。講義資料と異なり、ここをこう変えたらインストールに成功した。手順に従い実行したらこのような結果が得られたなど。例 2 : TIGAR2 の Z 値 (カウントデータ) と、Bowtie2 で 1 か所にのみマップされるリードペアを独自にカウントした結果の比較を行ったなど。

課題提出先 : report@iu.a.u-tokyo.ac.jp および kadota@iu.a.u-tokyo.ac.jp (全角の@を半角の@に変えてください)

タイトル : 特論 II 課題

提出期限 : 2016 年 8 月 30 日

- ・ メール本文に、所属・氏名を明記してください。また、課題の添付を忘れないようにしてください。
- ・ 「NGS ハンズオン講習会の受付番号」、「アグリバイオの受講 ID (5 桁)」、東大生は学生証番号などもできるだけ示してください。
- ・ 24 working hours 以内に受領メールをお送りします。返事が届かない場合は、'NGS ハンズオン講習会事務

局' <NGS @ biosciencedbc.jp> 宛てに問い合わせてください。