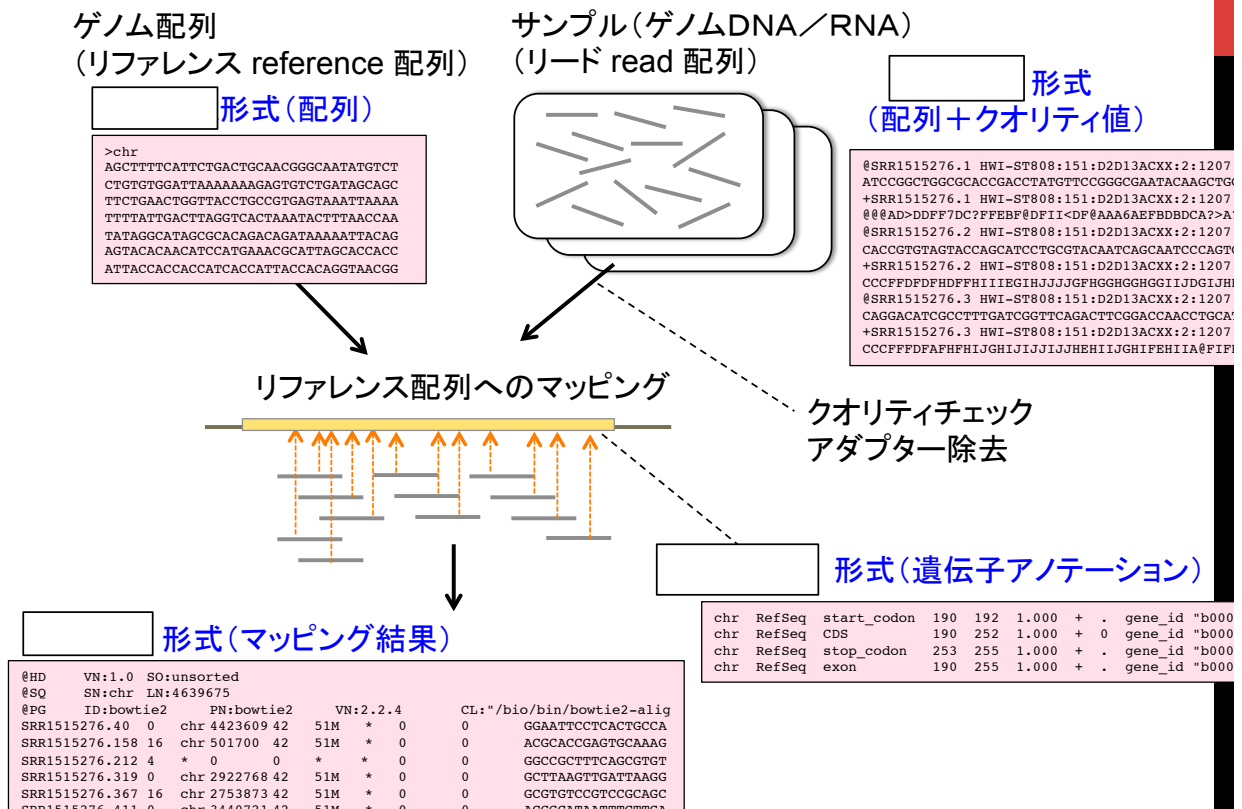


ショートリードのマッピング



復習: cutadaptによるアダプターの除去

入力

- リード配列 (FASTQ 形式; paired-end)
etec_1.fq
etec_2.fq
- アダプター配列 (それぞれを3'端から除去)

Adapter1: GATCGGAAGAGCGGTT

Adapter2: GATCGGAAGAGCGTCG

◆ アダプター配列除去の実行

除去後のデータ (FASTQ形式) は etec_1.cut.fq, etec_2.cut.fqとする)

\$ cutadapt

復習: **bowtie2** 用インデックスの作成

入力

- ゲノムデータ (FASTA形式)
eco_o139.fa 腸管毒素原性大腸菌(ETEC) O139:H28のゲノム配列

◆ bowtie2用インデックスの作成(インデックス名は etec)

\$ bowtie2-build

復習: **bowtie2**の実行 (paired-end)

入力

- リード配列 (FASTQ 形式; paired-end; アダプターを除去したもの)
etec_1.cut.fq
etec_2.cut.fq
- リファレンス配列のインデックス名
etec (先ほど作ったもの)

◆ bowtie2によるマッピングの実行 (出力: etec_bowtie2.sam)

\$ bowtie2

復習: **SAM**から**BAM**への変換

入力

- SAMファイル（さきほどbowtie2によって作成されたもの）
etec_bowtie2.sam

- ◆ SAMからBAMへ変換する（出力ファイル名: etec_bowtie2.bam）

\$ samtools

- ◆ 作成したBAMファイルをヘッダ付きでSAMに変換してlessで表示する

\$ samtools

復習: **BAM**ファイルのインデックスづけ

入力

- BAMファイル（さきほどSAMからの変換によって作成されたもの）
etec_bowtie2.bam

- ◆ リファレンス配列上の位置の順にソートする
（出力ファイル: etec_bowtie2_sorted.bam）

\$ samtools

- ◆ ソートされたBAMファイルに対してインデックスを作成する

\$ samtools

- ◆ インデックスを使って、リファレンスの染色体配列（染色体名: ETEC_chr）の
10000-12000 の範囲にマッピングされた結果のみを表示する

% samtools