ショートリードのマッピングとデータのフォーマット

ゲノム(リファレンス)配列

フォーマット(配列)

>chr

AGCTTTTCATTCTGACTGCAACGGGCAATATGTCT CTGTGTGGATTAAAAAAAAGAGTGTCTGATAGCAGC TTCTGAACTGGTTACCTGCCGTGAGTAAATTAAAA TTTTATTGACTTAGGTCACTAAATACTTTAACCAA TATAGGCATAGCGCACAGACAGATAAAAATTACAG AGTACACAACATCCATGAAACGCATTAGCACCACC

サンプル(リード) 配列

フォーマット(配列+クオリティ値)

@SRR1515276.1 HWI-ST808:151:D2D13ACXX:2:1207:3625:88631 length=51
ATCCGGCTGGCGCACCGACCTATGTTCCGGGCGAATACAAGCTGGGTGAAG
+SRR1515276.1 HWI-ST808:151:D2D13ACXX:2:1207:3625:88631 length=51
@@AD>DDFF7DC?FFEBF@DFII<DF@AAA6AEFBDBDCA?>A?B=>B::
@SRR1515276.2 HWI-ST808:151:D2D13ACXX:2:1207:3871:88513 length=51
CACCGTGTAGTACCAGCATCCTGCGTACAATCAGCAATCCCAGTCCTCCCC
+SRR1515276.2 HWI-ST808:151:D2D13ACXX:2:1207:3871:88513 length=51
CCCFFDFDFHDFFHIIIEGIHJJJJGFHGGHGGHIJJDGIJHHGGGHIH
@SRR1515276.3 HWI-ST808:151:D2D13ACXX:2:1207:3950:88530 length=51
CAGGACATCGCCTTTGATCGGTTCAGACTTCGGACCAACCTGCATTTTCAG
+SRR1515276.3 HWI-ST808:151:D2D13ACXX:2:1207:3950:88530 length=51
CCCFFFDFAFHFHIJGHIJIJJIJJHEHIIJGHIFEHIIA@FIFHGGIIGI

リファレンス配列へのマッピング

chr. Po

クオリティチェック アダプター除去

chr RefSeq start_codon 190 192 1.000 + . gene_id "b0001"; transcript_id "b0001";
chr RefSeq CDS 190 252 1.000 + 0 gene_id "b0001"; transcript_id "b0001";
chr RefSeq stop_codon 253 255 1.000 + . gene_id "b0001"; transcript_id "b0001";
chr RefSeq exon 190 255 1.000 + . gene_id "b0001"; transcript_id "b0001";

フォーマット(マッピング結果)

| VN:1.0 | so: | SO:unsorted | | | | | | | |
|------------|---|--|--|--|---|---|--|---|--|
| SN:chr | LN:4639675 | | | | | | | | |
| ID:bowtie2 | PN:bowtie2 VN:2.2.4 CL:"/bio/bin/bowtie2-alig | | | | | | lig | | |
| 10 | 0 | chr | 4423609 | 42 | 51M | * | 0 | 0 | GGAATTCCTCACTGCCA |
| .58 | 16 | chr | 501700 | 42 | 51M | * | 0 | 0 | ACGCACCGAGTGCAAAG |
| 212 | 4 | * | 0 | 0 | * | * | 0 | 0 | GGCCGCTTTCAGCGTGT |
| 319 | 0 | chr | 2922768 | 42 | 51M | * | 0 | 0 | GCTTAAGTTGATTAAGG |
| 367 | 16 | chr | 2753873 | 42 | 51M | * | 0 | 0 | GCGTGTCCGTCCGCAGC |
| 11 | 0 | chr | 3440721 | 42 | 51M | * | 0 | 0 | ACGGCATAATTTCTTGA |
| 134 | 0 | chr | 4198737 | 42 | 51M | * | 0 | 0 | GCGCGGTACGCATCTGG |
| | SN:chr | SN:chr LN: ID:bowtie2 PN: 0 58 16 212 4 319 0 667 16 411 0 | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 0 chr .58 16 chr .12 4 * .19 0 chr .667 16 chr .11 0 chr | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 VN:2.2.4 10 0 chr 4423609 158 16 chr 501700 122 4 * 0 139 0 chr 2922768 16 chr 2753873 111 0 chr 3440721 | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 VN:2.2.4 CL:' 10 0 chr 4423609 42 158 16 chr 501700 42 122 4 * 0 0 139 0 chr 2922768 42 1667 16 chr 2753873 42 111 0 chr 3440721 42 | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 VN:2.2.4 CL:"/bio/b 10 0 chr 4423609 42 51M 158 16 chr 501700 42 51M 12 4 * 0 0 * 1319 0 chr 2922768 42 51M 16 chr 2753873 42 51M 11 0 chr 3440721 42 51M | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 VN:2.2.4 CL:"/bio/bin/ 10 0 chr 4423609 42 51M * 158 16 chr 501700 42 51M * 122 4 * 0 0 * * 139 0 chr 2922768 42 51M * 1667 16 chr 2753873 42 51M * 111 0 chr 3440721 42 51M * | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 VN:2.2.4 CL:"/bio/bin/bow 10 0 chr 4423609 42 51M * 0 158 16 chr 501700 42 51M * 0 12 4 * 0 0 * * 0 139 0 chr 2922768 42 51M * 0 1607 16 chr 2753873 42 51M * 0 1611 0 chr 3440721 42 51M * 0 | SN:chr LN:4639675 ID:bowtie2 PN:bowtie2 VN:2.2.4 CL:"/bio/bin/bowtie2-a 10 0 chr 4423609 42 51M * 0 0 158 16 chr 501700 42 51M * 0 0 1212 4 * 0 0 * * * 0 0 1319 0 chr 2922768 42 51M * 0 0 1367 16 chr 2753873 42 51M * 0 0 1411 0 chr 3440721 42 51M * 0 0 |

フォーマット

(遺伝子アノテーション)

バイナリ化 フォーマット (マッピング結果)