Problem: store alignments in a standard format

Solution:

Sequence Alignment/Map Format Specification

The SAM/BAM Format Specification Working Group

5 Feb 2020

The master version of this document can be found at https://github.com/samtools/hts-specs. This printing is version dfc3e48 from that repository, last modified on the date shown above.

1 The SAM Format Specification

SAM stands for Sequence Alignment/Map format. It is a TAB-delimited text format consisting of a header section, which is optional, and an alignment section. If present, the header must be prior to the alignments. Header lines start with '@', while alignment lines do not. Each alignment line has 11 mandatory fields for essential alignment information such as mapping position, and variable number of optional fields for flexible or aligner specific information.

Example alignment

```
Coor         12345678901234    567890123456789012345678901234
ref          AGCATGTTAGATAA**GATAGCTGTGCTAGTAGGCAGTCAGCGCCAT

+r001/1          TTAGATAAAGGATA*CTG
+r002           aaaAGATAA*GGATA
+r003       gcctaAGCTAA
+r004                          ATAGCT..............TCAGC
-r003                                 ttagctTAGGC
-r001/2                                          CAGCGGCAT
```