

# *SAM* file format

# SAM file contents

```
@HD VN:1.4 SO:coordinate
@SQ SN:chr13 LN:114364328
@SQ SN:chr17 LN:83257441
@SQ SN:chr20 LN:64444167
@RG ID:NOID PG:tmap SM:NOSM
@PG ID:tmap CL:map3 -f ref/chr_merged.fa -r reseq_reads.fastq -i fastq -o 2 -s reseq_reads_mapped.bam VN:3.4.1

JXTYJ:00295:00658 0 chr13 16131323 125 16M1095 * 0 0 CATTCCAAGTAATTTTCTGTTTGGTCAAAACAATTCCTTCAAACAAAATGAACAGA
AAGGCAGTTGGGTAGTGATTACTGCCATTTAGCGTAACCTTGCAATTCCTTGTGTTTCTAGTATTTGGC CBD>CDF>BCC=CEE8CCBAA9A=BCCED5A<CD8A:@: ?
>B49844(39:6BABDD596;;@A9>@2?..---626:=77/5BD9BEGDCC;BA;BBBBB;C=C=A?@BB<ACBA>=>?/919 RG:Z:NOIDPG:Z:tmap MD:Z:1
6 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00949:00263 0 chr13 17166350 0 16M1655 * 0 0 ACGCAGCAAACAGAGTGAAATGCCATTTTTCATATTCTGCTTTCCATTTATATGTGCTT
TATTTATACATGCTCCAGGGCATGTTTGCTTGGTCCCTTTTGGGCAGTAAGTTCCAAACCTTCTCAATTCTGTCCTTCAAATAAAGGTCTAGGCCCTTCTGAGAAGT
CTCAGGATGTGG :76;;@:99+:7995:99909;9454444'4555=6;;;666*65>993<?;<;;655)5;;0;888>>><9B999.:+5;44490443-33333(
3:;(33(366884:<6:2:929-3-48:;;:A;BB@@@5?3889/>@A9A;A>>>B<CE9A;@9@=?=0377BCCBA8?@?@8 RG:Z:NOID PG:Z:tmap MD:Z:
16 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:16 XT:i:1
JXTYJ:00427:00223 16 chr13 18193502 50 60S18M * 0 0 CTGGGTGCTTCAGTGCAGATATTGAAATACAGAGTGCCTCTTCTCTCACTGCAAGC
GGGAGCTTGGCATTGTGCTG AA9AA;CB>8>ACCA???AAAA:A78>>>@@@>>>BBBBBB<??@BCABAA@A7>>5+:.99A<<B>B@B>FFBCAA RG:Z:NOID
PG:Z:tmap MD:Z:16A0A0 NM:i:2 AS:i:10 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00327:00477 0 chr13 18222206 56 18M38S * 0 0 CCCCCTCTGCACCAGCCTCTGGCCTTCCGCTTCTGTGCTGGTTCCTCGGTCTC 99
92:CB<BBBBBA<ABB<BB<=<181:0918<>;A>>>CCDC=B<A?47+*((( RG:Z:NOID PG:Z:tmapMD:Z:18 NM:i:0 AS:i:18 XA:Z:map3-1 X
S:i:10 XT:i:1
```

# SAM header

```
@HD VN:1.4 SO:coordinate
@SQ SN:chr13 LN:114364328
@SQ SN:chr17 LN:83257441
@SQ SN:chr20 LN:64444167
@RG ID:NOID PG:tmap SM:NOSM
@PG ID:tmap CL:map3 -f ref/chr_merged.fa -r reseq_reads.fastq -i fastq -o 2 -s reseq_reads_mapped.bam VN:3.4.1
```

```
JXTYJ:00295:00658 0 chr13 16131323 125 16M1095 * 0 0 CATTCCAAGTAATTTCTGTTTGGTCAAACAATTCCTTCAAACAAATGAACAGA
AAGGCAGTTGGGTAGTGATTACTGCCATTTAGCGTAACCTTGCATTTCTTGTTTCTAGTATTTGGC CBD>CDF>BCC=CEE8CCBAA9A=BCCED5A@<CD8A:@: ?
>B49844(39:6BABDD596;;@A9>@2?..---626:=77/5BD9BEGDCC;BA;BBBCB;C=C=A?@BB<ACBA>=>?/919 RG:Z:NOIDPG:Z:tmap MD:Z:1
6 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00949:00263 0 chr13 17166350 0 16M1655 * 0 0 ACGCAGCAAACAGAGTGAAATGCCATTTTTCATATTCTGCTTTCCATTTATATGTGCTT
TATTTATACATGCTCCAGGGCATGTTTGCTTGGTCCCTTTTGGGCAGTAAGTTCCAAACCTTCTCAATTCTGTCCCTTCAAATAAAGGCTAGGCCCTTCTGAGAAAGT
CTCAGGATGTGG :76;;@:99+:7995:99909;9454444'4555=6;;;666*65>993<?;<;:655)5;;0;888>>><9B999::+5;44490443-33333(
3::(33(366884:<6:2:929-3-48:::A;BB@@@573889/>@A9A;A>>>B<CE9A;@9@=?@377BCCBA8?@??@8 RG:Z:NOID PG:Z:tmap MD:Z:
16 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:16 XT:i:1
JXTYJ:00427:00223 16 chr13 18193502 50 60S18M * 0 0 CTGGGTTGCTTCAGTGCAGATATTGAAATACAGAGTGCGCTCTTCTCTCACTGCAAGC
GGGGAGCTTGGCATTTGCTG AA9AA;CB>8>ACCA??AAAA:A78>>>@@@>>>BBBBBB<??@BCABAA@A7>>5+:.99A<<B>B@B>FFBCAA RG:Z:NOID
PG:Z:tmap MD:Z:16A0A0 NM:i:2 AS:i:10 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00327:00477 0 chr13 18222206 56 18M38S * 0 0 CCCCACCTCTGCACCAGCCTCCTGGCCTTCCGCTTCTGTGCTGGTTCCTCCGCTCTC
992:CB CB BBBA<ABB<BB<=<181:0918<;A>>>CCDC=B<A747+*'( (( ( RG:Z:NOID PG:Z:tmapMD:Z:18 NM:i:0 AS:i:18 XA:Z:map3-1 X
S:i:10 XT:i:1
```

# SAM body

```
@HD VN:1.4 SO:coordinate
@SQ SN:chr13 LN:114364328
@SQ SN:chr17 LN:83257441
@SQ SN:chr20 LN:64444167
@RG ID:NOID PG:tmap SM:NOSM
@PG ID:tmap CL:map3 -f ref/chr_merged.fa -r reseq_reads.fastq -i fastq -o 2 -s reseq_reads_mapped.bam VN:3.4.1
```

```
JXTYJ:00295:00658 0 chr13 16131323 125 16M109S * 0 0 CATTCCAAGTAATTTTCTGTTTGGTCAAACAATTTCTTCAAACAAAATGAACAGA
AAGGCAGTTGGGTAGTGATTACTGCCATTTAGCGTAACTTGCATTTCTTGTTTCTAGTATTTGGC CBD>CDF>BCC=CEE8CCBAA9A=BCCED5A@<CD8A:@:?
>B49844(39:6BABDD596;;@A9>@2?.....626:=77/5BD9BEGDCC;BA;BBBCB;C=C=A7@BB<ACBA>=>7/919 RG:Z:NOIDPG:Z:tmap MD:Z:1
6 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00949:00263 0 chr13 17166350 0 16M165S * 0 0 ACGCAGCAAAACAGAGTGAAATGCCATTTTTCATATTCTGCTTTCCATTTATATGTGCTT
TATTTATACATGCTCCAGGGCATGTTTGCTTGGTCCCTTTTGGGCAGTAAGTTCCAAACCTTCTCAATTCTGTCCTTCAAATAAAGGCTAGGCCCTTCTGAGAAGT
CTCAGGATGTGG :76;;@:99+:7995:99909;9454444'4555=6;;;666*65>993<?;<;:655)5;;0;888>>><9B999::+5;44490443-33333(
3:::(33(366884:<6:2:929-3-48;;:A;BB@@@573889/>@A9A;A>>>B<CE9A;@9@=?=377BCCBA8?@??@8 RG:Z:NOID PG:Z:tmap MD:Z:
16 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:16 XT:i:1
JXTYJ:00427:00223 6 chr13 18193502 50 60S18M * 0 0 CTGGGTTGCTTCAGTGCAGATATTGAAATACAGAGTGCGCTCTTCTCTCACTGCAAGC
GGGGAGCTTGGCATTGCTG AA9AA;CB>8>ACCA??A AAA:A78>>>@@@>>>BBBBBB<??BCBAA@A7>>5+:.99A<B>B@B>FFBCAA RG:Z:NOID
PG:Z:tmap MD:Z:16A0A0 NM:i:2 AS:i:10 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00327:00477 0 chr13 18222206 56 18M38S * 0 0 CCCCCTCTGCACCAGCCTCTGGCCTTCCGCTTCTGTGCTGGTTCCTCGGTCTC 99
92:CBCBBBBBA<ABB<BB<=<181:0918<>;A>>>CCDC=B<A747+*'(((( RG:Z:NOID PG:Z:tmapMD:Z:18 NM:i:0 AS:i:18 XA:Z:map3-1 X
S:i:10 XT:i:1
```

# SAM body

```
JXTYJ:00295:00658 0 chr13 16131323 125 16M1095 * 0 0 CATTCCAAGTAATTTTCTGTTTGGTCAAAACAATTCCTTCAAACAAAATGAACAGA
AAGGCAGTTGGGTAGTGATTACTGCCATTTAGCGTAAC TTGCATTTCTTGTTTCTAGTATTTGGC CBD>CDF>BCC=CEE8CCBAA9A=BCCED5A@<CD8A:@: ?
>B49844(39:6BABDD596;;@A9>@2? . - - - 626:=77/5BD9BEGDCC;BA;BBBCB;C=C=A?@BB<ACBA>=>?/919 RG:Z:NOIDPG:Z:tmap MD:Z:1
6 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00949:00263 0 chr13 17166350 0 16M1655 * 0 0 ACGCAGCAAAACAGAGTGAAATGCCATTTTTCATATTCTGCTTTCCATTTATATGTGCTT
TATTTATACATGCTCCAGGGGCATGTTTGCTTGGTCCCTTTTGGGCAGTAAGTTCCAAACCTTCTCAATTCTGTCCCTTCAAATAAAGGTCTAGGCCCTTCTCGAGAAGT
CTCAGGATGTGG :76;;@:99+:?995:99909;9454444'4555=6;;;666*65>993<?;<;(;655)5;;0;888>>><9B999::+5;44490443-33333(
3:: (33(366884;<6:2:929-3-48::;A;BB@@@573889/>@A9A;A>>>B<CE9A;@9@=?=@377BCCBA8?@??@8 RG:Z:NOID PG:Z:tmap MD:Z:
16 NM:i:0 AS:i:16 XA:Z:map3-1 XS:i:16 XT:i:1
JXTYJ:00427:00223 16 chr13 18193502 50 60S18M * 0 0 CTGGGTGCTTCAGTGCAGATATTGAAATACAGAGTGCCTCTTCTCTCACTGCAAGC
GGGGAGCTTGGCATTGCTG AA9AA;CB>8>ACCA??AAAA:A78>>>@@@>>>BBBBBB<??@BCABAA@A7>>>5+:.99A<<B>B@B>FFBCAA RG:Z:NOID
PG:Z:tmap MD:Z:16A0A0 NM:i:2 AS:i:10 XA:Z:map3-1 XS:i:-2147483647 XT:i:1
JXTYJ:00327:00477 0 chr13 18222206 56 18M38S * 0 0 CCCCACTCTGCACCACTCTCTGGCCCTCCGCTTCTGTGCTGGTTCCCTCGGTCTC 99
92:CBCBBBBBA<ABB<BB<=<181:0918<>;A>>>CCDC=B<A747+*'( ((( RG:Z:NOID PG:Z:tmapMD:Z:18 NM:i:0 AS:i:18 XA:Z:map3-1 X
S:i:10 XT:i:1
```

From SAM format specification:

Col	Field	Brief description
1	QNAME	Query template NAME
2	FLAG	bitwise FLAG
3	RNAME	Reference sequence NAME
4	POS	1-based leftmost mapping POSition
5	MAPQ	MAPping Quality
6	CIGAR	CIGAR string
7	RNEXT	Ref. name of the mate/next read
8	PNEXT	Position of the mate/next read
9	TLEN	observed Template LENgth
10	SEQ	segment SEQUENCE
11	QUAL	ASCII of Phred-scaled base QUALity+33

# FLAG field

Bit	Description
0x1	template having multiple segments in sequencing
0x2	each segment properly aligned according to the aligner
0x4	segment unmapped
0x8	next segment in the template unmapped
0x10	SEQ being reverse complemented
0x20	SEQ of the next segment in the template being reversed
0x40	the first segment in the template
0x80	the last segment in the template
0x100	secondary alignment
0x200	not passing quality controls
0x400	PCR or optical duplicate
0x800	supplementary alignment

# CIGAR string

Op	Description
M	alignment match (can be a sequence match or mismatch)
I	insertion to the reference
D	deletion from the reference
N	skipped region from the reference
S	soft clipping (clipped sequences present in SEQ)
H	hard clipping (clipped sequences NOT present in SEQ)
P	padding (silent deletion from padded reference)
=	sequence match
X	sequence mismatch



Reference: AGTCGATTAGGAA  
Read: AGTCGATTTAGAA

Reference: AGTCGATTAGGAA  
Read: ACTCGATTTAGAA  
**6M**

Reference: AGTCGA\*TTAGGAA  
Read: ACTCGATTTAGAA  
**6M 1I**

Reference: AGTCGA\*TTAGGAA  
Read: ACTCGATTTAGAA  
**6M 1I4M**

Reference:

AGTCGA\*TTAGGAA

Read:

ACTCGATTTAG\*AA

**6M 1I4M1D**

Reference:

AGTCGA\*TTAGGAA

Read:

ACTCGATTTAG\*AA

**6M 1I4M1D2M**