



SamToFastq Documentation

Description: Converts a SAM or BAM file to FASTQ format.
Author: Picard team
Contact: Marc-Danie Nazaire, gp-help@broadinstitute.org

Summary

This module converts a file in SAM format (or its binary equivalent, BAM format) to FASTQ format. SAM is a standard short read alignment that allows aligners to attach custom tags to individual alignments. For more information on the SAM format, see the specification: <http://samtools.sourceforge.net/SAM-1.3.pdf>. For more information on the FASTQ format, see the [FASTQ specification](http://maq.sourceforge.net/fastq.shtml) (<http://maq.sourceforge.net/fastq.shtml>).

This module implements the SamToFastq function from Picard. For more information, see the Picard Web site: <http://picard.sourceforge.net/>.

Parameters

Name	Description
input.file (required)	The SAM or BAM file.
per.read.group (required)	Whether to output one FASTQ file per read group. Default: no
re-reverse.bases (required)	Whether to re-reverse bases and qualities of reads with negative strand flag set. Default: no
include.non.pf.reads (required)	Whether to include non-purity-filtered reads. Default: no
clipping.attribute (optional)	The attribute that stores the position at which the SAM record should be clipped. That is, if an attribute-value pair specifying the clipping position for each record is present in the SAM/BAM file, <i>clipping.attribute</i> can be set to the two-letter name of the attribute.

GenePattern

clipping.action (optional)	Action to take with clipped reads when a clipping attribute is specified. Options include: <ul style="list-style-type: none">• trim reads and qualities• change bases to Ns• change base quality values Default: trim reads and qualities
clipping.action.quality.value (optional)	Value to which base quality values will be set when the <i>clipping action</i> is "change base quality values."
fastq.output.file.1 (optional)	The FASTQ output file name for a single-end FASTQ or the first end of a paired-end FASTQ. For more information on the FASTQ format, see the specification: http://maq.sourceforge.net/fastq.shtml . NOTE: This parameter is required if <i>per read group</i> is no and must not be set if <i>per read group</i> is yes.
fastq.output.file.2 (optional)	The FASTQ output file name for the second end of a paired-end FASTQ. NOTE: This parameter must not be set if <i>per read group</i> is yes.

Output Files

1. FASTQ file(s)

The number of FASTQ files depends on the options specified: if the SAM file is for paired-end reads and if the *per read group* parameter is set to yes. For more information on the FASTQ format, see the specification:

<http://maq.sourceforge.net/fastq.shtml>.

Platform Dependencies

Module type:	Data Format Conversion
CPU type:	any
OS:	any
Language:	Java, minimum version 1.6