.PRO Transcriptome Profile

The Profile (.PRO) format is designed to describe the simulated characteristics of each transcript from the reference annotation, one per line. After each step of a simulation run, tab-separated are added to the file.

Column Nr	Name	Value	Description
1	Locus	*chrom*:*start-end*[W|C]	identifier of the transcriptional locus, given by the chromosome (*chrom*), *start* respectively *end* position, and the strand (Watson or Crick).
2	Transcript_ID	String	transcript identifier from the reference annotation.
3	Coding	[CDS|NC]	specifies whether the transcript has an annotated coding sequence (CDS) or not (NC)
4	Length	Integer	the mature length of the transcript after splicing out introns, disregarding the poly-A tail, as annotated in the reference annotation
5	Expressed Fraction	Float	fraction of RNA molecules that represent transcripts that are qualitatively equal to this RNA form
6	Expressed Number	Integer	absolute number of expressed RNA molecules
7	Library Fraction	Float	fraction of cDNA molecules in the final library that have been produced from this transcript
	Library Number	Integer	absolute number of cDNA fragments generated from this transcript
9	Sequenced Fraction	Float	fraction of total reads that have been sequenced from this transcript
10	Sequenced Number	Integer	absolute number of reads sequenced from this transcript
11	Covered Fraction	Float	fraction of the transcript that is covered by reads
12	Chi Square	Integer	chi-square goodness of fit measurement of coverage uniformity
13	Coefficient of Variation	Float	coefficient of variation for transcript coverage