

FASTA statistics of the 00010 transcript set

Jens Allmer [ORCID](#)

unreviewed

Jens Allmer (2026). FASTA statistics of the 00010 transcript set. Version 1. Internet of Science. local:article/fasta-stats-demo. <https://ioscience.org/a/fasta-stats-demo/v1/>
Produced in the Internet of Science; reproducible: FASTA statistics: summary table, per-record table and length histogram.

The 00010 transcript set (`experimental/fasta/00010.fasta`) holds ten de novo-assembled transcripts. This article shows what one run of the workflow `fasta-stats-report` produces: `FASTAStatistics` computes the statistics once and three endpoints turn them into a summary table, a per-record table and a histogram. Every caption links to the record of that run, and re-running the workflow on the same data reproduces the same content hashes.

1 Summary

statistic	value
number_of_elements	10
total_sequence_length	9,618
avg_seq_len	961.8
min_seq_len	610
max_seq_len	1,570
residue:A	2,486
residue:C	2,259
residue:G	2,205
residue:T	2,668

Table 1. FASTA statistics Record count, length summary and residue counts of the input file. [provenance](#) [open workflow](#)

2 Every record

header	length	A	C	G	T
>pdj2001denovo	1,124	343	220	164	397
>pdj2002denovo	610	172	118	86	234
>pdj2003denovo	912	218	262	205	227
>pdj2004denovo	1,570	432	333	358	447
>pdj2005denovo	836	148	225	288	175
>pdj2006denovo	750	178	217	197	158
>pdj2007denovo	1,198	335	252	280	331
>pdj2008denovo	1,090	233	325	290	242
>pdj2009denovo	615	169	125	120	201
>pdj2010denovo	913	258	182	217	256

Table 2. Per-record statistics Length and residue counts of every record.
provenance open workflow

3 Length distribution

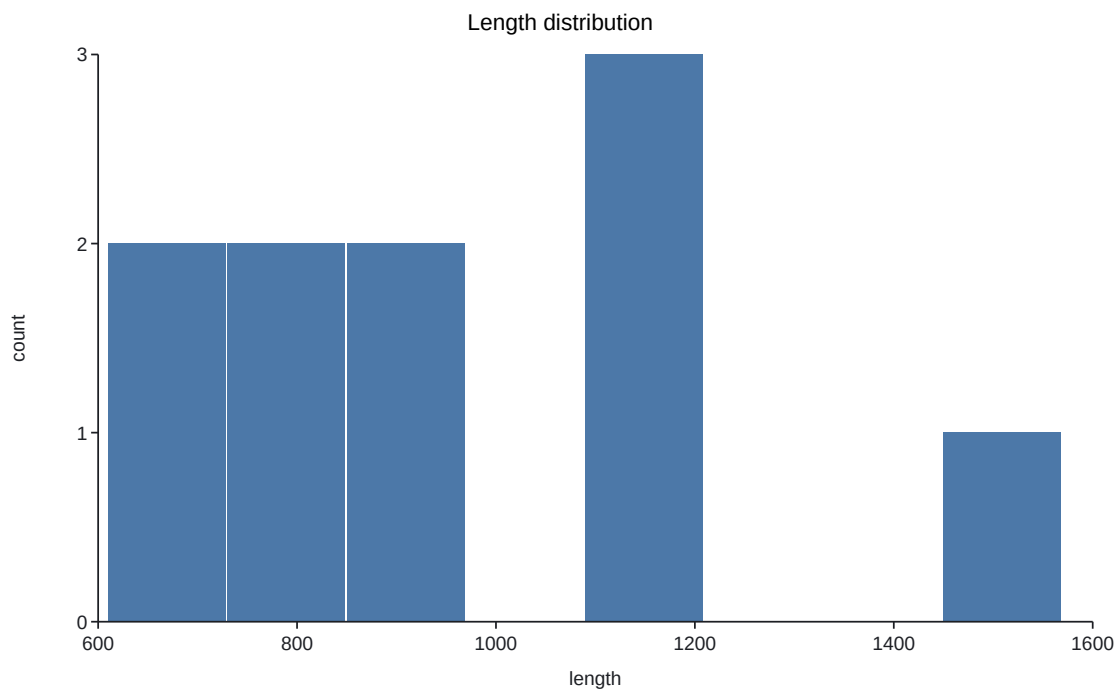


Figure 1. Length distribution Histogram of the record lengths. provenance open workflow

Contributions

Authors are the contributors whose shares, largest first, reach 80 % (ADR 0029): Jens Allmer 100 % (writing 71 of 71 words).

Data curation	Jens Allmer
Software	Jens Allmer
Writing – original draft	Jens Allmer

This article rests on

Components	FASTAReader, FASTAStatistics, FigureArtefact, FileSource, TableArtefact
Datasets	experimental/fasta/00010.fasta