

Sequence lengths of the 00010 transcript set

Jens Allmer [ORCID](#)

unreviewed

Jens Allmer, Example Editor (fixture) (2026). Sequence lengths of the 00010 transcript set. Version 2. Internet of Science. local:article/example-fasta-lengths. <https://ioscience.org/a/example-fasta-lengths/v2/>

Produced in the Internet of Science; reproducible: Sequence lengths of a FASTA file: table and histogram.

The 00010 transcript set (`experimental/fasta/00010.fasta`) holds ten de novo-assembled transcripts from a plant transcriptome assembly. This article reports their lengths; the table and the figure below come from one run of the workflow `fasta-lengths-report`, and each caption links to the record of that run.

1 Lengths

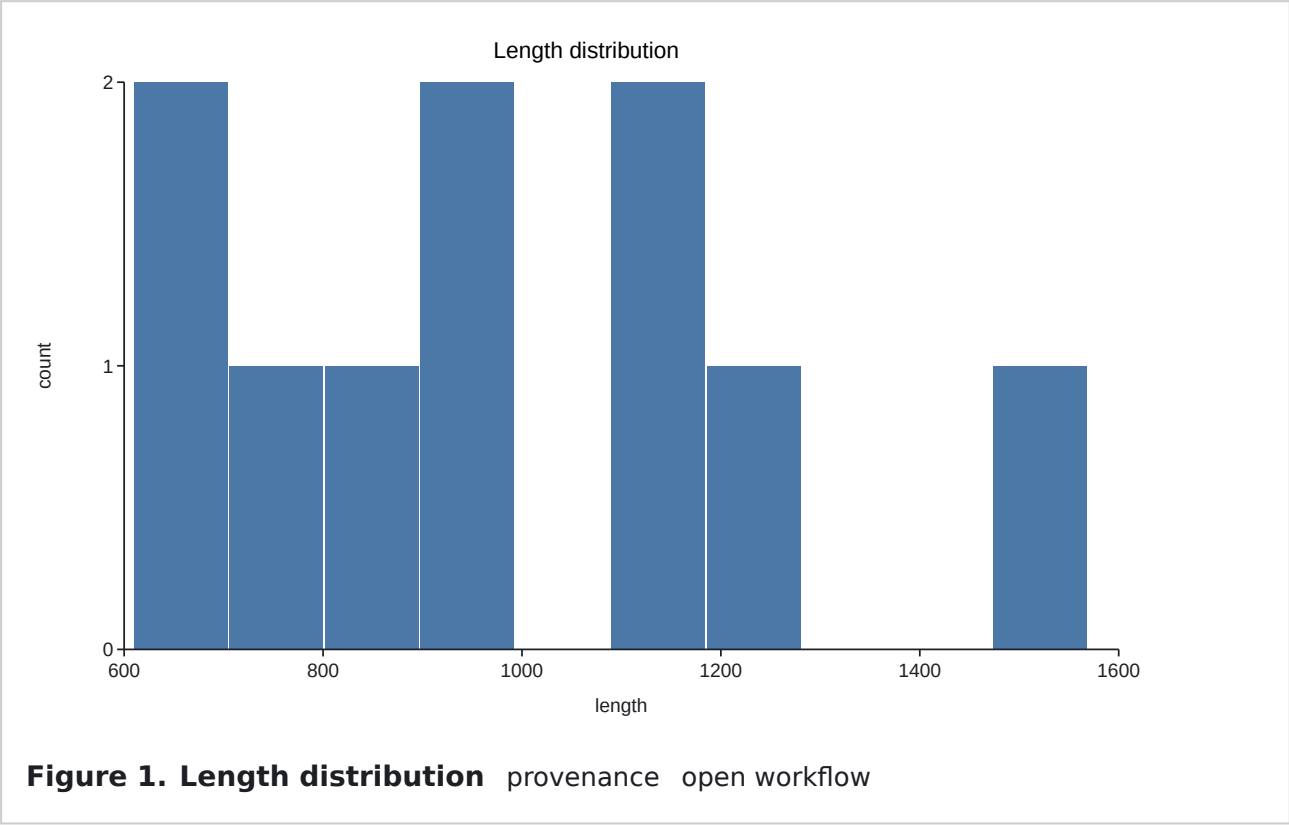
The transcripts are between 610 and 1,570 nucleotides long, with a median of 912.5 and a mean of 961.8 nucleotides.

header	length
pdi2001denovo	1,124
pdi2002denovo	610
pdi2003denovo	912
pdi2004denovo	1,570
pdi2005denovo	836
pdi2006denovo	750
pdi2007denovo	1,198
pdi2008denovo	1,090
pdi2009denovo	615
pdi2010denovo	913

Table 1. Sequence lengths Length of every sequence in the input. [provenance](#) [open workflow](#)

2 Distribution

Most transcripts fall between 600 and 1,200 nucleotides; a single transcript (pdi2004denovo) is longer than 1,400.



3 Reproducing

Every embed names the workflow document, the dataset and the endpoint. Re-running the workflow on the same data gives the same content hash, which is how a reader verifies a figure.

Contributions

Authors are the contributors whose shares, largest first, reach 80 % (ADR 0029): Jens Allmer 81 % (writing 99 of 122 words). Further contributors: Example Editor (fixture) 19 % (writing 23 of 122 words).

Data curation	Jens Allmer
Software	Jens Allmer
Writing – original draft	Jens Allmer
Writing – review & editing	Example Editor (fixture)

This article rests on

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