



Contents

- [1 Genome assembly and annotation V1.2](#)
- [2 Downloads genome annotation version 1.6](#)
 - ◆ [2.1 Gene structure annotation](#)
 - ◇ [2.1.1 GFF3](#)
 - ◇ [2.1.2 FASTA](#)
 - ◆ [2.2 Functional annotation - genonaut dumps](#)
 - ◇ [2.2.1 Latest Release 2012.3](#)
 - ◇ [2.2.2 Master branch - Nightly-build](#)
 - ◇ [2.2.3 Additional functional annotations](#)
- [3 Mapping to old annotation releases](#)
- [4 Organellar genomes](#)
- [5 Version history](#)

Genome assembly and annotation V1.2

If you use this data, please cite: Lang, D., A.D. Zimmer, S.A. Rensing, R. Reski (2008): Exploring plant biodiversity: the Physcomitrella genome and beyond. Trends in Plant Science 13, 542-549.

[V1.2.1 filtered assembly](#)

1,985 [filtered](#) genomic scaffolds based on the V1.0 JGI assembly.

[V1.2.2 filtered gene model transcripts](#)

27,960 [filtered](#), [annotated](#) transcripts based on the V1.1 JGI gene model selections.

[V1.2.2 filtered gene model proteins](#)

27,960 [filtered](#), [annotated](#) proteins based on the V1.1 JGI gene model selections.

[V1.2.2 GFF3](#)

27,960 [filtered](#) gene structures in [GFF3 format](#).

[Dbxrefs V1.2](#): database cross references to NCBI Genbank and NCBI Gene

Files are compressed with gzip. Sequences are provided as FASTA format.

Downloads genome annotation version 1.6

Please cite the following paper if you use the V1.6 genome annotation for your work:

Zimmer, A.D., D. Lang, K. Buchta, S. Rombauts, T. Nishiyama, M. Hasebe, Y. van de Peer, S.A. Rensing, R. Reski (2013): Reannotation and extended community resources of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. [BMC Genomics 14, 498.](#)

In addition, if you use any of the released files please provide this site and the release as a reference:

current: V1.6 2012.3 <https://www.cosmoss.org>https://www.cosmoss.org/physcome_project/wiki/Downloads

Gene structure annotation

GFF3

V1.6 protein coding GFF3

32,275 gene structures of 38,357 protein coding transcripts in GFF3 format.

V1.6 non-protein coding GFF3

non-protein coding genes/regions (798 rRNA, 432 tRNA, 229 miRNA, 213 snRNA, 6 SRP) in GFF3 format.

FASTA

- [38,357 mRNA transcript sequences](#)
- [38357 protein sequences](#)
- [18180 5'UTR sequences](#)
- [38357 CDS sequences](#)
- [19,041 3'UTR sequences](#)

Functional annotation - genonaut dumps

The community-curated genonaut annotation (descriptions, gene names, protein names, GO) is dumped regularly and available as release and master branch in a bitbucket repository. The master branch will soon be configured to reflect nightly dumps.

Latest Release 2012.3

[cosmoss.genonaut.gene_name.txt](#)

gene name tab-delimited table (including annotator information)

[cosmoss.genonaut.protein_name.txt](#)

protein name tab-delimited table (including annotator information)

[cosmoss.genonaut.aliases.txt](#)

Downloads

[gene aliases tab-delimited table \(including annotator information\)](#)
[cosmoss.genonaut.description.txt](#)
description lines tab-delimited table (including annotator information)
[cosmoss.genonaut.annot](#)
Transcript/Protein-wise GO annotation in [BLAST2GO](#) annotation format
[cosmoss.genonaut.gaf2](#)
Transcript/Protein-wise GO annotation in GO Annotation File Format 2.0 ([GAF2.0](#))
[cosmoss.genonaut.map](#)
Locus-wise GO annotation in [topGO](#) input format
[cosmoss.genonaut.txt](#)
Full text annotation genonaut database including author information

Master branch - Nightly-build

If you use these include **Nightly-built downloaded on XX-XX-XXXX** in your methods section.

[cosmoss.genonaut.gene_name.txt](#)
gene name tab-delimited table (including annotator information)
[cosmoss.genonaut.protein_name.txt](#)
protein name tab-delimited table (including annotator information)
[cosmoss.genonaut.aliases.txt](#)
gene aliases tab-delimited table (including annotator information)
[cosmoss.genonaut.description.txt](#)
description lines tab-delimited table (including annotator information)
[cosmoss.genonaut.descriptions.txt](#)
description lines tab-delimited table
[cosmoss.genonaut.annot](#)
Transcript/Protein-wise GO annotation [BLAST2GO](#) annotation format
[cosmoss.genonaut.gaf2](#)
Transcript/Protein-wise GO annotation in GO Annotation File Format 2.0 ([GAF2.0](#))
[cosmoss.genonaut.map](#)
Locus-wise GO annotation in [topGO](#) input format
[cosmoss.genonaut.txt](#)
Full text annotation genonaut database including author information

Additional functional annotations

[cosmoss.TAPScan.csv](#)
TAPScan Transcription factor classification Lang et al. 2010 [Genome Biology and Evolution 2, 488-503](#)
[cosmoss.mapman.txt](#)
[MapMan](#) annotation D. Lang (unpublished)

Mapping to old annotation releases

[Conversion of *P. patens* gene accessions numbers \(Excel xls file\)](#)

Organelle genomes

mitochondrial genome annotation

NC_007945.1 mitochondrial genome] encoded mRNAs, tRNAs, rRNAs in GFF3 format.

plastid genome annotation

NC_005087.1 plastid genome encoded mRNAs, tRNAs, rRNAs in GFF3 format.

plastid encoded proteins

85 proteins NC_005087.1

mitochondrial encoded proteins

42 proteins NC_007945.1

Version history

--Lang 14:18, 9 October 2012 (UTC) Updated V1.6 Genome Annotation Release

--Lang 05:35, 27 April 2010 (UTC) organelle proteins

--AndZim 07:27, 1 April 2010 (UTC) optimized LocusIDs (CGI)

--AndZim 10:47, 8 October 2009 (UTC) removed additional bacterial and human contaminations.

--Lang 16:05, 16 September 2009 (UTC) changed description lines to GenBank status as of today.

--Lang 14:22, 7 October 2009 (UTC) added GFF3