

Sources and versions of genomic data used in this work.

Organism	Source ¹	Version	Web reference
<i>Arabidopsis thaliana</i>	TAIR	TAIR6_pep_20051108	www.arabidopsis.org/
<i>Caenorhabditis elegans</i>	WormBase	WS150	www.wormbase.org/
<i>Chlamydomonas reinhardtii</i>	JGI	v3.0	www.chlamy.org/
<i>Cryptosporidium parvum</i>	cryptoDB	3.3	cryptodb.org/cryptodb/
<i>Cyanidioschyzon merolae</i>	C.merolae genome project	-	merolae.biol.s.u-tokyo.ac.jp/
<i>Dictyostelium discoideum</i>	dictyBase	Primary	dictybase.org/
<i>Drosophila melanogaster</i>	ENSEMBL	BDGP4.2	www.ebi.ac.uk/ensembl/
<i>Entamoeba histolytica</i>	geneDB	Proteins_17102005	www.genedb.org/
<i>Giardia lamblia</i>	GiardiaDB	-	www.mbl.edu/Giardia/
<i>Homo sapiens</i>	ENSEMBL	NCBI 36	www.ebi.ac.uk/ensembl/
<i>Leishmania major</i>	geneDB	v5	www.genedb.org/
<i>Oryza sativa</i>	TIGR	v4.0	www.tigr.org/tdb/e2k1/osa1/
<i>Ostreococcus lucimarinus</i>	JGI	v2.0	www.jgi.doe.gov/
<i>Phaeodactylum tricornutum</i>	JGI	v2.0	www.jgi.doe.gov/
<i>Phytophthora sojae</i>	JGI	v1.1	www.jgi.doe.gov/
<i>Plasmodium falciparum</i>	geneDB	3D7 v2.1.1	www.genedb.org/
<i>Populus trichocarpa</i>	JGI	v1.1	www.jgi.doe.gov/
<i>Saccharomyces cerevisiae</i>	ENSEMBL	SGD1	www.ebi.ac.uk/ensembl/
<i>Schizosaccharomyces pombe</i>	geneDB		www.genedb.org/
<i>Takifugu rubripes</i>	ENSEMBL	4.0	www.ebi.ac.uk/ensembl/
<i>Tetrahymena thermophila</i>	TGD	Predictions_Aug_2004	www.ciliate.org/
<i>Thalassiosira pseudonana</i>	JGI	v1	www.jgi.doe.gov/
<i>Toxoplasma gondii</i>	toxodb	v4.0	www.toxodb.org/
<i>Trypanosoma brucei</i>	geneDB	v4	www.genedb.org/

¹ nb: Source of data does not necessarily correlate with sequencing centre(s). See individual projects for full acknowledgements of participants.