

Supplementary Material

Supplementary Table S1. Genomes available in build 47

Supplementary Table S2. Counts of putative contiguous gene split models in 39 plant reference genomes in build 47

Supplementary Table S3. Ontology associations in build 47

Supplementary Table S4. (B)LASTz-chain-net pairwise whole-genome alignments available in build 47

Supplementary Table S5. Synteny analyses in build 47

Supplementary Table S6. Variation data sets housed in the Ensembl variation module in build 47

Supplementary Table S1. Genomes available in build 47

Species	Assembly/Gene Space Annotation	Variation
<i>Aegilops tauschii</i> (goatgrass, wheat D-genome progenitor)	ASM34733v1/GCA_000347335.1	
<i>Amborella trichopoda</i>	GCA_000471905.1 (AMTR1.0)/2014-01-AGD	
<i>Arabidopsis lyrata</i>	Araly1.0/Araly1.0	
<i>Arabidopsis thaliana</i>	TAIR10/TAIR10	SNPs & Structural
<i>Brachypodium distachyon</i>	Brachy1.0/Barchy1.2	SNPs
<i>Brassica oleracea</i>	GCA_000092065.1/ENA	
<i>Brassica rapa</i> (Chinese cabbage)	IVFCAASv1/bra_v1.01_SP2010_01	
<i>Chlamydomonas reinhardtii</i> (green algae)	v3.0/2007-11-ENA	
<i>Cyanidioschyzon merolae</i> (red algae)	ASM9120v1/2008-11-ENA	
<i>Glycine max</i> (soybean)	Glyma1.0/Glyma1.1	
<i>Hordeum vulgare</i> (barley)	030312v2/IBSC_1.0	SNPs
<i>Leersia perrieri</i>	GCA_000325765.3/OGE	
<i>Medicago truncatula</i>	MedtrA17_3.5/2011-11-EnsemblPlants	
<i>Musa acuminata</i> (banana)	MA1/2012-08-Cirad	

<i>Oryza barthii</i> (complete; replaces previous chr. 3s)	GCA_000182155.1/OGE Maker (Aug 2013)	
<i>Oryza brachyantha</i> (complete; replaces previous chr. 3s)	[GCA_000231095.2] OGEv1.4b/OGEv1.4	
<i>Oryza glaberrima</i> (African rice)	[GCA_000147395.1] AGI1.1 (May 2011)/2011-05-AGI (MIPS)	SNPs
<i>Oryza glumaepatula</i> (complete; replaces previous chr. 3s)	GCA_000576495.1/OGE Maker (Aug 2013)	
<i>Oryza longistaminata</i> (complete; replaces previous chr. 3s)	v0117-2013Aug/OGE	
<i>Oryza meridionalis</i> (complete; replaces previous chr. 3s)	GCA_000338895.1/OGE Maker (Aug 2013)	
<i>Oryza nivara</i> (complete; replaces previous chr. 3s)	GCA_000576065.1/OGE Maker (Aug 2013)	
<i>Oryza punctata</i> (complete; replaces previous chr. 3s)	GCA_000573905.1/OGE Maker (Aug 2013)	
<i>Oryza rufipogon</i> (complete; replaces previous chr. 3s)	PRJEB4137/OGE	
<i>Oryza sativa</i> ssp. <i>indica</i> (Asian rice)	ASM465v1/2010-07-BGI	SNPs
<i>Oryza sativa</i> ssp. <i>japonica</i> (Asian rice)	IRGSP-1.0/MSU 7.0	SNPs
<i>Ostreococcus lucimarinus</i>	v2.0 GCA_000092065.1/JGI	
<i>Physcomitrella patens</i> (moss)	ASM242v1/2011-03-Phypa1.6	
<i>Prunus persica</i> (peach)	Prupe1_0 / GCA_000346465.1/Prupe1_0 / Peach v1.0	
<i>Populus trichocarpa</i> (poplar)	JGI 2.0/2010-01-JGI	
<i>Selaginella moellendorffii</i> (spikemoss)	v1.0/2011-05-ENA	
<i>Setaria italica</i> (foxtail millet)	JGIv2.0/JGIv2.1	
<i>Solanum lycopersicum</i> (tomato)	SL2.50/ITAG2.3	SNPs
<i>Solanum tuberosum</i> (potato)	v3.0/SolTub_3.0	

<i>Sorghum bicolor</i>	Sorbi1/2007-12-JGI	SNPs & Structural
<i>Theobroma cacao</i> (cocoa)	GCA_000403535.1/ENA	
<i>Triticum aestivum</i> (bread wheat)	IWGSP1/IWGSP1 (MIPS)	SNPs
<i>Triticum urartu</i> (einkorn wheat, A-genome progenitor)	GCA_000347455.1/GCA_000347455.1 (ENA, Apr 2013)	
<i>Vitis vinifera</i> (grape)	IGGP 12x/2012-07-CRIBI	SNPs
<i>Zea mays</i> (corn)	B73_RefGen_v3/5b+	SNPs

Supplementary Table S2. Counts of putative contiguous gene split models (frequently associated with annotation artifacts) in 39 plant reference genomes in build 47

Species	Putative contiguous split gene models
<i>Aegilops tauschii</i>	740
<i>Amborella trichopoda</i>	994
<i>Arabidopsis lyrata</i> ssp. <i>lyrata</i>	180
<i>Arabidopsis thaliana</i>	37
<i>Brachypodium distachyon</i>	158
<i>Brassica oleracea</i>	911
<i>Brassica rapa</i> ssp. <i>Pekinensis</i>	151
<i>Chlamydomonas reinhardtii</i>	141
<i>Cyanidioschyzon merolae</i>	8
<i>Glycine max</i>	390
<i>Hordeum vulgare</i> ssp. <i>Vulgare</i>	108
<i>Leersia perrieri</i>	189
<i>Medicago truncatula</i>	394
<i>Musa acuminata</i>	1624
<i>Oryza barthii</i>	338

<i>Oryza brachyantha</i>	302
<i>Oryza glaberrima</i>	518
<i>Oryza glumaepatula</i>	394
<i>Oryza longistaminata</i>	291
<i>Oryza meridionalis</i>	272
<i>Oryza nivara</i>	310
<i>Oryza punctata</i>	282
<i>Oryza rufipogon</i>	300
<i>Oryza sativa</i> ssp. <i>indica</i>	796
<i>Oryza sativa</i> ssp. <i>japonica</i>	636
<i>Ostreococcus lucimarinus</i>	26
<i>Physcomitrella patens</i> subsp. <i>Patens</i>	276
<i>Populus trichocarpa</i>	1651
<i>Prunus persica</i>	448
<i>Selaginella moellendorffii</i>	215
<i>Setaria italica</i>	332
<i>Solanum lycopersicum</i>	1216
<i>Solanum tuberosum</i>	653
<i>Sorghum bicolor</i>	953
<i>Theobroma cacao</i>	113
<i>Triticum aestivum</i>	1234
<i>Triticum urartu</i>	674
<i>Vitis vinifera</i>	725
<i>Zea mays</i>	844

Supplementary Table S3. Ontology associations in build 47

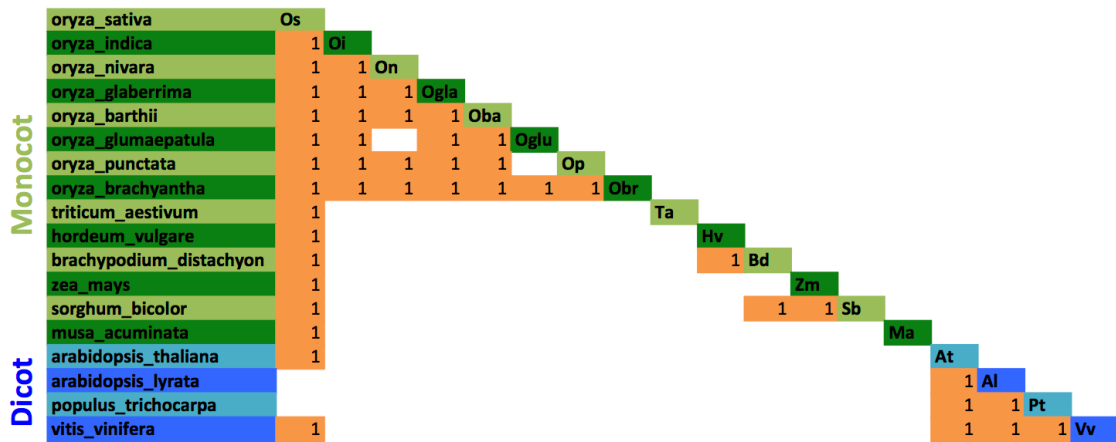
Prefix	Species	Total Count
GO	<i>Aegilops tauschii</i>	4233
	<i>Arabidopsis lyrata</i>	5397
	<i>Arabidopsis thaliana</i>	5463
	<i>Brachypodium distachyon</i>	4483
	<i>Brassica rapa</i>	4712
	<i>Chlamydomonas reinhardtii</i>	3666
	<i>Cyanidioschyzon merolae</i>	2169
	<i>Glycine max</i>	4779
	<i>Hordeum vulgare</i>	4714
	<i>Medicago truncatula</i>	4580
	<i>Musa acuminata</i>	4511
	<i>Oryza brachyantha</i>	4837
	<i>Oryza glaberrima</i>	4745
	<i>Oryza sativa ssp. indica</i>	5171
	<i>Oryza sativa ssp. japonica</i>	5115
	<i>Physcomitrella patens</i>	4218
	<i>Populus trichocarpa</i>	5088

	<i>Selaginella moellendorffii</i>	4267
	<i>Setaria italica</i>	4852
	<i>Solanum lycopersicum</i>	4804
	<i>Solanum tuberosum</i>	4568
	<i>Sorghum bicolor</i>	5000
	<i>Triticum aestivum</i>	4388
	<i>Triticum urartu</i>	4257
	<i>Vitis vinifera</i>	5116
	<i>Zea mays</i>	4681
PO	<i>Arabidopsis thaliana</i>	594
	<i>Brachypodium distachyon</i>	3
	<i>Physcomitrella patens</i>	4
	<i>Solanum lycopersicum</i>	75
	<i>Vitis vinifera</i>	23
	<i>Zea mays</i>	56

Supplementary Table S4. (B)LASTz-chain-net pairwise whole-genome alignments available in build 47

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
Monocot																																							
1	oryza_sativa	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	oryza_rufipogon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	oryza_indica	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	oryza_nivara	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	oryza_glaberrima	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	oryza_barthii	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	oryza_meridionalis	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	oryza_longistaminata	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	oryza_glumaepatula	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	oryza_punctata	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	oryza_brachyantha	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	leersia_perrieri	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	triticum_aestivum	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	triticum_urartu	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	aegilops_tauschii	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	hordeum_vulgare	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	brachypodium_distachyon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	zea_mays	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	sorghum_bicolor	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	setaria_italica	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	musa_acuminata	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	arabidopsis_thaliana	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	arabidopsis_lyrata	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	brassica_oleracea	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	brassica_rapa	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	glycine_max	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	medicago_truncatula	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	prunus_persica	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	populus_trichocarpa	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	theobroma_cacao	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	vitis_vinifera	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	solanum_lycopersicum	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	solanum_tuberosum	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
34	amborella_trichopoda	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
35	selaginella_moellendorffii	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
36	physcomitrella_patens	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
37	ostreococcus_lucimarinus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
38	chlamydomonas_reinhardtii	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
39	cyanidioschyzon_merolae	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Supplementary Table S5. Synteny analyses in build 47



Supplementary Table S6. Variation data sets housed in the Ensembl variation module in build 47

Species	Variants	Source	Studies
<i>Arabidopsis thaliana</i>	14,234,197	250K SNPs x 1179 accessions	(Clark <i>et al</i> , 2007, Atwell <i>et al</i> , 2010, Gan <i>et al</i> , 2011)
	SV: 13,667	1001 genomes project: 411 resequenced accessions	
<i>Brachypodium distachyon</i>	327,988	3 accessions of <i>Brachypodium sylvaticum</i>	(Fox <i>et al</i> , 2013)
<i>Hordeum vulgare</i>	19,474,539	Resequencing 4 accessions plus wild barely & RNA-Seq (Illumina HiSeq 2000) from embryos of 9 cultivars and Morex	(International Barley Genome Sequencing <i>et al</i> , 2012)
		~5M from 90 Morex x Barke	(Fox <i>et al</i> , 2014)
		~6M from 84	

		Oregon Wolfe	
<i>Oryza glaberrima</i>	7,172,036	Resequenced 20 accessions African rice & wild progenitor <i>O. barthii</i>	Oryza Genome Evolution project
<i>Oryza sativa</i> ssp. <i>japonica</i>	5,512,746	160K SNPs x 20 accessions	(McNally <i>et al</i> , 2009)
		1311 SNPs x 395 accessions	(Zhao <i>et al</i> , 2010)
		~4M BGI Japonica vs Indica SNPs	NCBI dbSNP
<i>Oryza sativa</i> ssp. <i>indica</i>	5,418,373	~4M BGI Japonica vs Indica SNPs	NCBI dbSNP
<i>Solanum lycopersicum</i>	71,156,450	WGS of 84 tomato accessions & related species	(Tomato Genome Sequencing <i>et al</i> , 2014)
<i>Sorghum bicolor</i>	SV: 32,246	Genomic Variants archive (dGVA)	(Zheng <i>et al</i> , 2011)
	265,000	SAP, MCC, RS	(Morris <i>et al</i> , 2013)
	6,578,420		(Mace <i>et al</i> , 2013)
<i>Triticum aestivum</i>	12,220,924	~1.57 M SNPs from the Wheat HapMap. ~725,000 non-redundant SNP from CerealsDB (Axiom, iSelect and KASP array sets) & ~10.3M inter-homoeologous variants may be visualized or retrieved through the genome browser	(Jordan <i>et al</i> , 2015); CerealsDB
<i>Vitis vinifera</i>	457,404	Resequencing USDA germplasm collection	(Myles <i>et al</i> , 2010)
<i>Zea mays</i>	51,162,745	HapMap1: NAM	(Gore <i>et al</i> , 2009)

		founder lines	
		HapMap2: pre-domesticated & domesticated lines	(Chia <i>et al</i> , 2012)
		Panzea 2.7 GBS	Panzea.org

Supplementary References

Atwell, S., et al. (2010). "Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines." Nature **465**(7298): 627-631.

Chia, J. M., et al. (2012). "Maize HapMap2 identifies extant variation from a genome in flux." Nat Genet **44**(7): 803-807.

Clark, R. M., et al. (2007). "Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*." Science **317**(5836): 338-342.

Fox, S. E., et al. (2014). "De novo transcriptome assembly and analyses of gene expression during photomorphogenesis in diploid wheat *Triticum monococcum*." PLoS One **9**(5): e96855.

Fox, S. E., J. Preece, J. A. Kimbrel, G. L. Marchini, A. Sage, K. Youens-Clark, M. B. Cruzan and P. Jaiswal (2013). "Sequencing and De Novo Transcriptome Assembly of *Brachypodium sylvaticum* (Poaceae). ." Applications in Plant Sciences **1**(3): 1200011.

Gan, X., et al. (2011). "Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*." Nature **477**(7365): 419-423.

Gore, M. A., et al. (2009). "A first-generation haplotype map of maize." Science **326**(5956): 1115-1117.

International Barley Genome Sequencing, C., et al. (2012). "A physical, genetic and functional sequence assembly of the barley genome." Nature **491**(7426): 711-716.

Jordan, K. W., et al. (2015). "A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes." Genome Biol **16**: 48.

Mace, E. S., et al. (2013). "Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum." Nat Commun **4**: 2320.

McNally, K. L., et al. (2009). "Genomewide SNP variation reveals relationships among landraces and modern varieties of rice." Proc Natl Acad Sci U S A **106**(30): 12273-12278.

Morris, G. P., et al. (2013). "Population genomic and genome-wide association studies of agroclimatic traits in sorghum." Proc Natl Acad Sci U S A **110**(2): 453-458.

Myles, S., J. M. Chia, B. Hurwitz, C. Simon, G. Y. Zhong, E. Buckler and D. Ware (2010). "Rapid genomic characterization of the genus vitis." PLoS One **5**(1): e8219.

Tomato Genome Sequencing, C., et al. (2014). "Exploring genetic variation in the tomato (*Solanum section Lycopersicon*) clade by whole-genome sequencing." Plant J **80**(1): 136-148.

Zhao, K., et al. (2010). "Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome." PLoS One **5**(5): e10780.

Zheng, L. Y., et al. (2011). "Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*)." Genome Biol **12**(11): R114.