SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLES

Supplementary Table 1. Genomes available in Gramene build 38
Supplementary Table 2. Ontology associations in Gramene build 38
Supplementary Table 3. Synteny analyses in Gramene build 38
Supplementary Table 4. Putative split gene events for 23 plant reference genomes in Gramene build 38
Supplementary Table 5. Variation data sets housed in the Ensembl variation module in Gramene build 38

Supplementary Table 6. BioCyc pathways databases in Gramene build 38

Supplementary Table 1. Genomes available in Gramene build 38

Species	Reference Genome Status	Assembly/Gene Space Annotation	Literature References	Status since Gramene build #31
Aegilops tauschii (goatgrass, wheat D-genome progenitor)	complete draft	GCA 000347335.1/2013-05-ENA	Jia et al (2013)	new
Arabidopsis lyrata	complete draft	v.1.0 /2008-12-Araly1.0	Hu et al (2011)	unchanged
Arabidopsis thaliana	complete draft	TAIR10/TAIR10	Arabidopsis Genome Initiative (2000)	updated
Brachypodium distachyon	complete draft	Brachy1.0/Barchy1.2	The International Brachypodium Initiative (2010)	unchanged
Brassica rapa (Chinese cabbage)	complete draft	IVFCAASv1/bra v1.01 SP2010 01	Wang et al (2011)	new
Chlamydomonas reinhardtii (green algae)	complete draft	GCA_000002595.2/2007-11-ENA	Merchant et al (2007)	new
Cyanidioschyzon merolae (red algae)	complete draft	ASM9120v1/2008-11-ENA	Matsuzaki et al (2004)	new
Glycine max (soybean)	complete draft	Glyma1.0/Glyma1.1	Schmutz et al (2010)	new
Hordeum vulgare (barley)	complete draft	030312v2/IBSC 1.0	The International Barley Genome Sequencing Consortium (2012)	new
Leersia perrieri (chr. 3s)	partial	454.pools.2012Feb/2012-10-CSHL	OGE/OMAP	new
Medicago truncatula	complete draft	GCA 000219495.1/2011-11-EnsemblPlants	Young et al (2011)	new
Musa acuminata (banana)	complete draft	MA1/2012-08-Cirad	D'Hont et al (2012)	new

Oryza barthii	complete draft, replaces previous chr. 3	OGE.2012Jul/2012-10-CSHL	OGE/OMAP	updated
Oryza brachyantha	complete draft, replaces previous chr. 3	Oryza brachyantha.v1.4b/OGEv1.4	OGE/OMAP	new
Oryza glaberrima	complete draft, replaces previous chr. 3	AGI1.1/2011-05-AGI	OGE/OMAP	updated
Oryza glumaepatula (chr. 3s)	partial	454.pools.2012Feb/2012-04-CSHL	OGE/OMAP	new
Oryza granulata (chr. 3s)	partial	454.pools.2012Feb/2012-10-CSHL	OGE/OMAP	new
Oryza longistaminata (chr. 3s)	partial	OGE.2012Jul/2012-10-CSHL	OGE/OMAP	new
Oryza meridionalis (chr. 3s)	partial	454.pools.2012Feb/2012-04-CSHL	OGE/OMAP	new
Oryza minuta BB (chr. 3s)	partial	BAC.Sanger.1.1 (May 2011)/CSHLv3.1	OGE/OMAP	new
Oryza minuta CC (chr. 3s)	partial	BAC.Sanger.1.1 (May 2011)/CSHLv3.1	OGE/OMAP	unchanged
Oryza nivara (chr. 3s)	partial	454.pools.1.1 (Jul 2010)/CSHL	OGE/OMAP	new
Oryza officinalis (chr. 3s)	partial	BAC.Sanger.1.1 (May 2011)/CSHLv3.1	OGE/OMAP	unchanged
Oryza punctata	complete draft, replaces previous chr. 3	OGE-v1.1-2013Mar/CSHL-2013Mar	OGE/OMAP	updated
Oryza rufipogon (chr. 3s)	partial	454.pools.1.1 (Jul 2010)/CSHL	OGE/OMAP	unchanged

Oryza sativa ssp. indica	complete draft	BGI 2005 Assembly/BGI GLEAN 2008 genes	Yu et al (2002); Zhao et al (2004)	unchanged
Oryza sativa ssp. japonica (rice)*	complete draft	MSU6/MSU6*	International Rice Genome Sequencing (2005)	unchanged*
Physcomitrella patens (moss)	complete draft	ASM242v1/2011-03-Phypa1.6	Rensing et al (2007)	new
Populus trichocarpa (poplar)	complete draft	JGI 2.0/2010-01-JGI	Tuskan et al (2006)	unchanged
Selaginella moellendorffii (spikemoss)	complete draft	v1.0/2011-05-ENA	Banks et al (2011)	new
Setaria italica (foxtail millet)	complete draft	JGIv2.0/JGIv2.1	Bennetzen et al (2012); Zhang G et al (2012)	new
Solanum lycopersicum (tomato)	complete draft	SL2.40/ITAG2.3	Tomato Genome Consortium (2012)	new
Solanum tuberosum (potato)	complete draft	3.0/SolTub 3.0	Potato Genome Sequencing Consortium (2011)	new
Sorghum bicolor	complete draft	Sorbi1/2007-12-JGI (Sbi1.4)	Paterson et al (2009)	unchanged
Triticum urartu (einkorn wheat, A-genome progenitor)	complete draft	GCA 000347455.1/2013-05-ENA	Ling et al (2013)	new
Vitis vinifera (grape)	complete draft	GGP 12X/2012-07-CRIBI	Jaillon et al (2007); Myles et al (2010)	unchanged
Zea mays (corn)	complete draft	B73 RefGen AGPv3/5b+	Schnable et al (2009)	updated

* IRGSP1/MSU7 will be made available in Gramene build#39

Supplementary Table 2. Ontology associations in Gramene build 38

Prefix	Ontology	Count
EO	plant environment ontology	501
GAZ	Gazetteer	456,894
GO	biological process	25,737
GO	cellular component	3,339
GO	molecular function	10,478
GRO	cereal plant growth stage	236
GR tax	gramene taxonomy	58,585
PO	plant anatomy	1,358
PO	plant structure development stage	340
SO	sequence ontology	2,244
TO	plant trait ontology	1,318

Supplementary Table 3. Synteny analyses in Gramene build 38

	Osj	Bd	Sb	Hv	At	Al	Vv	Pt	Sl	St
Oryza sativa japonica	-									
Brachypodium distachyon	YES	-								
Sorghum bicolor	YES	YES	-							
Hordeum vulgare	YES	YES	-	-						
Arabidopsis thaliana	-	-	-	-	-					
Arabidopsis lyrata	-	-	-	-	YES	-				
Vitis vinifera	-	-	-		YES	YES	-			
Poplar trichocarpa	-	-	-	-	YES	YES	YES	-		
Solanum lycopersicum	-	-	-	-	-	-	-	-	-	-
Solanum tuberosum	-	-	-	-	-	-	-	-	YES	-

Supplementary Table 4. Putative split gene events for 23 reference plant genomes in Gramene build 38*.

Species	# split gene predictions
Arabidopsis lyrata subsp. lyrata	195
Arabidopsis thaliana	43
Brachypodium distachyon	172
Brassica rapa subsp. Pekinensis	133
Chlamydomonas reinhardtii	137
Cyanidioschyzon merolae	8
Glycine max	456
Hordeum vulgare subsp. Vulgare	120
Medicago truncatula	588
Musa acuminata	1694
Oryza brachyantha	323
Oryza glaberrima	421
Oryza sativa Indica Group	711
Oryza sativa Japonica Group	622
Physcomitrella patens subsp. Patens	303
Populus trichocarpa	1870
Selaginella moellendorffii	232
Setaria italica	356
Solanum lycopersicum	1286
Solanum tuberosum	677
Sorghum bicolor	946
Vitis vinifera	794
Zea mays	1103

* Predicted split gene events for the present release are available
at ftp://ftp.gramene.org/pub/gramene/CURRENT_RELEASE/data/split_genes

Supplementary Table 5. Variation data sets housed in the Ensembl variation module in Gramene build 38

Rice	Variation data for *O. sativa japonica* and *indica* from NCBI dbSNP (http://www.ncbi.nlm.nih.gov/snp) 160,000 SNPs x 21 varieties (incl. Nipponbare ref.) from OryzaSNP, MSU6 McNally *et al* (2009). 1,311 SNPs x 395 accessions (Zhao et al, 2010)
Maize	Maize Diversity HapMap v1, 1.6 million SNPs x 27 NAM founder lines from Panzea, AGPv2 remapped to AGPv3. Gore et al (2009). 55 million SNPs & indels x 103 pre-domesticated and domesticated *Zea mays* varieties, including related *Tripsacum dactyloides* (Eastern gamagrass); HapMap v2 remapped to AGPv3 Chia *et al* (2012)
Arabidopsis	2010 Project SNP Discovery: 637,522 SNPs x 21 ecotypes, TAIR10. Clark et al (2007). 2010 Project 250K SNP chip genotypes v3.04, 214,000 SNPs x 1179 ecotypes, TAIR10 (Nordborg *et al*; http://www.gmi.oeaw.ac.at/research-groups/magnus-nordborg). 1001 Genomes/WTCHG SNPs from dbSNP, 18 ecotypes and 392 strains. Structural variants from Affy_250k, Perlegen_1M, WTCHG and 1001 Genomes Phenotype data added from a GWAS study of 107 phenotypes in 95 inbred lines (Atwell *et al* 2010)
Grape	71K grape SNPs. Myles *et al* (2011)
Oryza glaberrima	828K variants from the OGE/OMAP project (http://www.genome.arizona.edu/modules/publisher/item.php?itemid=7)
Sorghum	64.5K structural variations from a single study. Zheng *et al* (2011)
Barley	20.8M short variants (SNPs, indels, somatic mutations) from four cultivars, Barke, Bowman, Igri, Haruna Nijo and a wild barley (H. spontaneum). Mayer *et al* (2012)
Brachypodium	328K variants for *Brachypodium sylvaticum* from three populations from Corvallis (Oregon, USA), Spain and Greece, mapped to *Brachypodium distachyon* genome. Fox *et al* (2012)

Supplementary Table 6. BioCyc pathways databases in Gramene build 38

Name	Version	Species	Strain	Source	Literature Reference
RiceCyc	3.3	*Oryza sativa japonica*	Nipponbare	Gramene	Dharamawhardhana et al (2013)
MaizeCyc	2.1	*Zea mays*	B73	Gramene	Monaco et al (2012)
SorghumCyc	1.1	*Sorghum bicolor*	BTx623	Gramene	
BrachyCyc	2.0	*Brachypodium distachyon*	Bd21	Gramene	Fox et al (2013)
AraCyc	10.0	*Arabidopsis thaliana*	Columbia	Plant Metabolic Network	Mueller et al (2003)
PoplarCyc	5.0	*Populus trichocarpa* (and other *Populus* species and hybrids*)*	n/a	Plant Metabolic Network	Zhang et al (2010)
LycoCyc	2.0.1	*Solanum lycopersicum,* tomato	Heinz 1706	Sol Genomics Network	Bombarley et al (2011)
PotatoCyc	1.0.1	*Solanum tuberosum,* potato	n/a	Sol Genomics Network	
CoffeaCyc	1.1.1	*Coffea canephora,* coffee	n/a	Sol Genomics Network	
MedicCyc	1.0.1	*Medicago truncatula,* barrel clover	n/a	Noble Foundation	Urbanczyk-Wochniak and Sumner (2007)
PlantCyc	7.0	Plant Metabolic Pathway Database	n/a	Plant Metabolic Network	Zhang et al (2010)
MetaCyc	17.0	Reference Pathway Database	n/a	BioCyc	Caspi et al (2012); Altman et al (2013)
EcoCyc	17.0	Reference Pathway Database	K-12 MG1655	BioCyc	Keseler et al (2013)

		(*Escherichia coli*)			

SUPPLEMENTARY REFERENCES

Same as References 12-22 in manusccript file.

- Atwell, S., Huang, Y.S., Vilhjalmsson, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T. *et al.* (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature*, **465**, 627-631.
- Chia, J.M., Song, C., Bradbury, P.J., Costich, D., de Leon, N., Doebley, J., Elshire, R.J., Gaut, B., Geller, L., Glaubitz, J.C. *et al.* (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nat Genet*, **44**, 803-807.
- Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthmann, N., Hu, T.T., Fu, G., Hinds, D.A. *et al.* (2007) Common sequence polymorphisms shaping genetic diversity in Arabidopsis thaliana. *Science*, **317**, 338-342.
- Fox, S.E., Preece, J., Kimbrel, J.A., Marchini, G.L., Sage, A., Youens-Clark, K., Cruzan, M.B. and Jaiswal, P. (2012) Sequencing and De Novo Transcriptome Assembly of Brachypodium sylvaticum (Poaceae). *Applications in Plant Sciences*, **1**, 1200011.
- Gan, X., Stegle, O., Behr, J., Steffen, J.G., Drewe, P., Hildebrand, K.L., Lyngsoe, R., Schultheiss, S.J., Osborne, E.J., Sreedharan, V.T. *et al.* (2011) Multiple reference genomes and transcriptomes for Arabidopsis thaliana. *Nature*, **477**, 419-423.
- Gore, M.A., Chia, J.M., Elshire, R.J., Sun, Q., Ersoz, E.S., Hurwitz, B.L., Peiffer, J.A., McMullen, M.D., Grills, G.S., Ross-Ibarra, J. *et al.* (2009) A first-generation haplotype map of maize. *Science*, **326**, 1115-1117.
- Mayer, K.F., Waugh, R., Brown, J.W., Schulman, A., Langridge, P., Platzer, M., Fincher, G.B., Muehlbauer, G.J., Sato, K., Close, T.J. *et al.* (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature*, **491**, 711-716.
- McNally, K.L., Childs, K.L., Bohnert, R., Davidson, R.M., Zhao, K., Ulat, V.J., Zeller, G., Clark, R.M., Hoen, D.R., Bureau, T.E. *et al.* (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc Natl Acad Sci U S A*, **106**, 12273-12278.
- Myles, S., Boyko, A.R., Owens, C.L., Brown, P.J., Grassi, F., Aradhya, M.K., Prins, B., Reynolds, A., Chia, J.M., Ware, D. *et al.* (2011) Genetic structure and domestication history of the grape. *Proc Natl Acad Sci U S A*, **108**, 3530-3535.
- Zhao, W., Wang, J., He, X., Huang, X., Jiao, Y., Dai, M., Wei, S., Fu, J., Chen, Y., Ren, X. *et al.* (2004) BGI-RIS: an integrated information resource and comparative analysis workbench for rice genomics. *Nucleic Acids Res*, **32**, D377-382.
- Zheng, L.Y., Guo, X.S., He, B., Sun, L.J., Peng, Y., Dong, S.S., Liu, T.F., Jiang, S., Ramachandran, S., Liu, C.M. *et al.* (2011) Genome-wide patterns of genetic variation in sweet and grain sorghum (Sorghum bicolor). *Genome Biol*, **12**, R114.