A Genome Assembly of the Barley 'Transformation Reference' Cultivar Golden Promise

Miriam Schreiber,* Martin Mascher,[†,‡] Jonathan Wright,[§] Sudharasan Padmarasu,[†] Axel Himmelbach,[†] Darren Heavens,[§] Linda Milne,** Bernardo J. Clavijo,[§] Nils Stein,[†,††] and Robbie Waugh*[,‡‡,§§,1]

*Cell and Molecular Sciences, **Information and Computational Sciences, [‡‡]Division of Plant Sciences, The James Hutton Institute, Invergowrie, Dundee, Scotland DD2 5DA, UK, [†]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany, [‡]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany, [§]Earlham Institute, Norwich Research Park, Norwich NR4 7UH, UK, [††]CiBreed - Center for Integrated Breeding Research, Georg-August University Göttingen, Department of Crop Sciences, Von Siebold Straße 8, 37075 Göttingen, Germany, and [§§]School of Agriculture and Wine, University of Adelaide, Plant Genome Building, Waite Campus, Urrbrae, Adelaide, South Australia

ORCID IDs: 0000-0002-3732-8822 (M.S.); 0000-0001-6373-6013 (M.M.); 0000-0002-7597-2774 (B.J.C.); 0000-0003-3011-8731 (N.S.); 0000-0003-1045-3065 (R.W.)

ABSTRACT Barley (*Hordeum vulgare*) is one of the most important crops worldwide and is also considered a research model for the large-genome small grain temperate cereals. Despite genomic resources improving all the time, they are limited for the *cv*. Golden Promise, the most efficient genotype for genetic transformation. We have developed a barley *cv*. Golden Promise reference assembly integrating Illumina paired-end reads, long mate-pair reads, Dovetail Chicago *in vitro* proximity ligation libraries and chromosome conformation capture sequencing (Hi-C) libraries into a contiguous reference assembly. The assembled genome of 7 chromosomes and 4.13Gb in size, has a super-scaffold N50 after Chicago libraries of 4.14Mb and contains only 2.2% gaps. Using BUSCO (benchmarking universal single copy orthologous genes) as evaluation the genome assembly contains 95.2% of complete and single copy genes from the plant database. A high-quality Golden Promise reference assembly will be useful and utilized by the whole barley research community but will prove particularly useful for CRISPR-Cas9 experiments.

Barley is a true diploid with 14 chromosomes (2n = 14). Its genome is around 5Gb in size and mainly consists of repetitive elements (International Barley Genome Sequencing Consortium 2012). Barley is and has been an important crop for thousands of years (Mascher *et al.* 2016). It was the fourth most produced cereal in 2016 worldwide (Faostat, http://www.fao.org/faostat/en/#home) and second most in the UK. While the majority of barley is used as feed, the most important market for 2-row spring barley is the whisky industry. An

iconic historical variety is the *cv*. Golden Promise which was used extensively for malting and whisky production and some distilleries still use it today. Golden Promise is a 2-row spring type which was mainly grown in Scotland in the 1970s and early 1980s and was identified as a semi-dwarf mutant after a gamma-ray treatment of the cultivar Maythorpe. In recent years, the main research interest in Golden Promise has come from its genetic transformability. Most barley transformations are successfully conducted using Golden Promise as it usually achieves the best shoot recovery from callus (Hensel *et al.* 2008). While many other cultivars have been tested and some successfully used, the transformation efficiency of Golden Promise is always superior (Murray *et al.* 2004; Ibrahim *et al.* 2010; Lim *et al.* 2018) With the rise of the CRISPR-Cas9 genome editing technology, a potential Golden Promise reference assembly has already sparked wide interest in the barley community. The use of CRISPR-Cas9 ideally requires a complete and correct reference assembly for the identification of target sites (Karkute *et al.* 2017). The Cas9 enzyme targets a position in the genome based on a sgRNA (single-guide RNA) followed by a PAM (protospacer-adjacent motif). The guide RNA is usually designed to be 20 bp long and

■ Table 1 Statistics for the different stages of the assembly process

	Contigs	Scaffolds	Dovetail	Hi-C
N50	22.4kb	86.67kb	4.14Mb	/
Number	786,696	245,820	128,283	8
Longest	352,153bp	1,540,019bp	22,832,123bp	612,216,794bp
Size	4.02Gb	4.11Gb	4.12Gb	4.13Gb

target-specific to avoid any off-target effects. The PAM region consists of three nucleotides "NGG" (Belhaj *et al.* 2013; Lawrenson *et al.* 2015). Any nucleotide variation between different cultivars can therefore cause problems with the CRISPR-Cas9 genome editing technology (Bortesi *et al.* 2016; Jaganathan *et al.* 2018). The time and cost involved in such increasingly common experiments highlights the value of a high-quality Golden Promise reference assembly.

MATERIALS AND METHODS

Contig construction and scaffolding

DNA extraction, library construction and sequencing: High molecular weight barley DNA was isolated from leaf material of 3-week old Golden Promise plants that had been kept in the dark for 48 hr to reduce starch levels. DNA was extracted using the GE Life Sciences Nucleon PhytoPure kit (GE Healthcare Life Sciences, Buckinghamshire, UK) according to the Manufacturers' instructions. Both paired-end and long mate-pair libraries were constructed and sequenced at the Earlham Institute by the Genomics Pipelines Group. A total of 2 μg of DNA was sheared targeting 1 kbp fragments on a Covaris-S2 (Covaris Brighton, UK), size selected on a Sage Science Blue Pippin 1.5% cassette (Sage Science, Beverly, USA) to remove DNA molecules <600bp, and amplification-free, paired-end libraries constructed using the Kapa Biosciences Hyper Prep Kit (Roche, New Jersey, USA). Long mate-pair libraries were constructed from 9 μg of DNA according to the protocol described in Heavens *et al.* (2015) based on the Illumina Nextera Long Mate Pair Kit (Illumina, San Diego, USA). Sequencing was performed on Illumina HiSeq 2500 instruments with a 2x250 bp read metric targeting >60x raw coverage of the amplification-free library and 30x coverage of a combination of different insert long mate-pair libraries with inserts sizes >7 kbp.

Contig and scaffold generation: Contigging was performed using the w2rap-contigger (Clavijo *et al.* 2017). Three mate-pair libraries were produced with insert sizes 6.5, 8 and 9.5kb and sequenced to generate approximately 284 million 2x250 bp reads. Mate-pair reads were processed and used to scaffold contigs as described in the w2rap pipeline (Clavijo *et al.* 2017; https://github.com/bioinfologics/w2rap). Scaffolds less than 500 bp were removed from the final assembly.

Chromosome conformation capture

Dovetail: Golden Promise 10-day old leaf material was sent to Dovetail Genomics (Santa Cruz, CA, USA) for the construction of Chicago libraries. Dovetail extracted high molecular weight DNA and conducted the library preparations. The Chicago libraries were sequenced on an Illumina HiSeqX (Illumina, San Diego, CA, USA) with 150bp paired-end reads. Using the scaffold assembly as input, the HiRise scaffolding pipeline was used to build super scaffolds (Putnam *et al.* 2016).

Hi-C: The Hi-C library construction from one week old seedlings of Golden Promise was performed as per protocol described in

Padmarasu *et al.* (2019) using DpnII for digestion of crosslinked chromatin. Sequencing of the Hi-C library was conducted on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA) with 101 bp paired-end reads. Super scaffolds from Dovetail were ordered and orientated to build the final pseudomolecule using the TRITEX assembly pipeline (Monat *et al.* 2019), with a detailed user guide available (https://tritexassembly.bitbucket.io).

Repeat and transcript annotation

The final assembly was analyzed for repetitive regions using Repeat-Masker (version 4.0.9) (Smit *et al.* 2013-2015) with the TREP Repeat library (trep-db_complete_Rel-16) (Wicker *et al.* 2002) and changing repetitive regions to lower case (-xsmall parameter) [repeat library downloaded from: http://botserv2.uzh.ch/kelldata/trep-db/downloadFiles.html]. The output of RepeatMask was condensed using the perl script "one-code-to-find-them-all" (Bailly-Bechet *et al.* 2014) with the parameters–strict and–unknown.

Transcript annotation was transferred from the BaRT transcriptome dataset (Rapazote-Flores *et al.* 2019) and the TRITEX gene annotation (Monat *et al.* 2019), using Gmap (version 2018-03-25) with the following parameters: -f 2 -n 1–min-trimmed-coverage = 0.8–min-identity = 0.9 (both files are available to download from figshare. BaRT: https://doi.org/10.6084/m9.figshare.9705278; TRITEX: https://doi.org/10.6084/m9.figshare.9705125).

Data validation and quality control

We used BUSCO with the plant dataset (embryophyta_odb9). For gene prediction BUSCO uses Augustus (Version 3.3) (Stanke *et al.* 2004; König *et al.* 2016). For the gene finding parameters in Augustus we set species to wheat and ran BUSCO in the genome mode (-m geno -sp wheat).

Data availability

Raw reads have been deposited to the NCBI sequence read archive. Bioproject: PRJNA533066 [SRA: Paired-end reads: SRR9291461, SRR9291462, SRR9291463, SRR9291464; Long mate-pair reads: SRR9266823, SRR9266824, SRR9266825, SRR9266826, SRR9266827, SRR9266828; Dovetail reads: SRR9202370, SRR9202371, SRR9202372, SRR9202373, SRR9202374; Hi-C data: SRR8922888]

The reference assembly is either available to download from figshare: https://doi.org/10.6084/m9.figshare.9332045 or through the European Nucleotide Archive (GCA_902500625).

RESULTS AND DISCUSSION

Genome assembly

Here we report a full-length Golden Promise genome assembly which was generated integrating short read sequencing and two chromosome conformation sequencing approaches. Approximately 624 million 2x250 bp paired reads were generated providing an estimated 62.4x coverage of the genome. 245,820 scaffolds were generated comprising 4.11 Gb of sequence with an N50 of 86.6kb. Gaps comprised only 1.6% of the scaffolds (Table 1). To generate full chromosome assembly, we utilized two different chromosome conformation captures. In a first
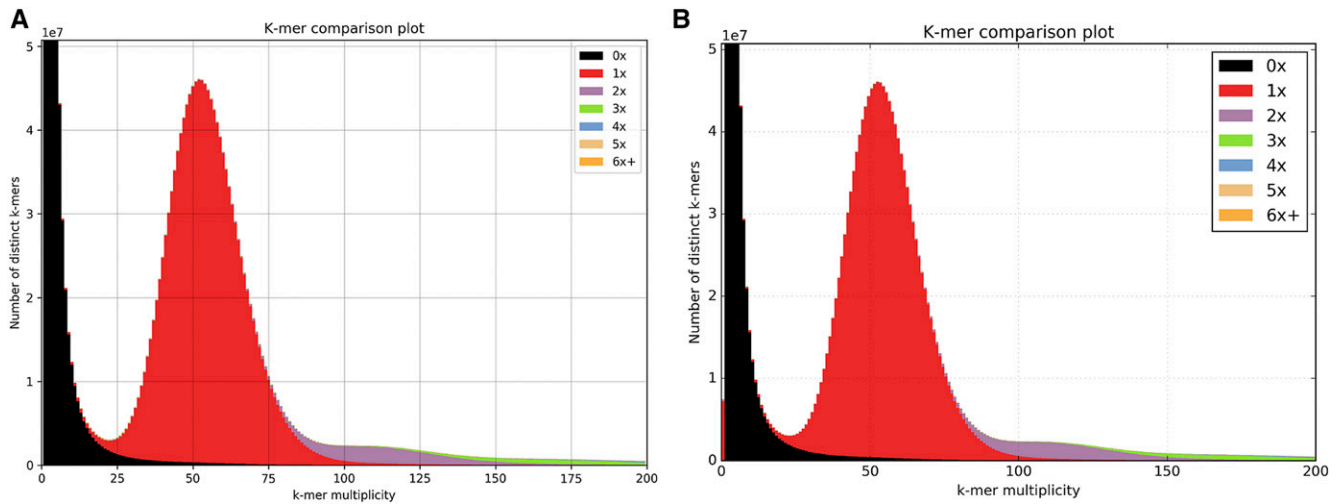
Figure 1 Spectra cn plots comparing k-mers from the paired-end reads to kmers in (a) the contig assembly and (b) the scaffold assembly.

step, we used Chicago Dovetail data which is generated by *in vitro* proximity ligation of large DNA fragments to increase the scaffold size and to correct false misjoins from the previous scaffolding. In the next step, we integrated Hi-C data which uses the native chromatin folding to increase the contiguity to full chromosome size. This resulted in a final assembly of 4.13Gb and 7 chromosomes plus an extra chromosome containing the unassigned scaffolds. We have provided the reference sequence as a blast and gmap searchable website for easy access: https://ics.hutton.ac.uk/gmapper/.

Completeness of the assembly

We used the spectra-cn function from the Kmer Analysis Toolkit (KAT) (Mapleson *et al.* 2017) to check for content inclusion in the contigs and scaffolds. KAT generates a k-mer frequency distribution from the paired-end reads and identifies how many times k-mers from each part of the distribution appear in the assembly being compared. It is assumed that with high coverage of paired-end reads, every part of the underlying genome has been sampled. Ideally, an assembly should contain all k-mers found in the reads (not including k-mers arising from sequencing errors) and no k-mers not present in the reads.

The spectra-cn plot in Figure 1a generated from the contigs shows sequencing errors (k-mer multiplicity <20) appearing in black as these are not included in the assembly. The majority of the content appears in a single red peak indicating sequence that appears once in the assembly. The black region under the main peak is very small indicating that most of this content from the reads is present in the assembly. The content that appears to the right of the main peak and is present twice or three times in the assembly represents repeats.

Scaffolds generally contain more miss-assemblies than contigs and this is reflected in the spectra-cn plot in Figure 1b generated from the scaffolds. The red bar at k-mer multiplicity 0 that is not present in the contigs spectra-cn plot reflects k-mers that appear in the scaffolds but do not appear in the reads. Approximately 7.2 million k-mers are represented in this region, less than 0.15% of the total.

Repetitive regions

The Golden Promise reference assembly was analyzed for repetitive regions using RepeatMasker with the TREP repeat library. This identified 73.2% (2.95 Gb) of the Golden Promise assembly as transposable elements (Table 2) with almost all from the class of

retroelements. The same analysis was also done for MorexV1 and MorexV2 showing that all three have very similar results (Table 2). Differences to the published results from MorexV1 and MorexV2 assembly (International Barley Genome Sequencing Consortium 2012; Mascher *et al.* 2017; Monat *et al.* 2019) are due to the different repeat libraries used.

Transcript annotation

For transcript annotation we transferred the latest barley annotation from MorexV2 onto the Golden Promise reference assembly. From a

■ **Table 2 Identified repetitive elements in the Golden Promise assembly. Values represent percentage coverage of the genome**

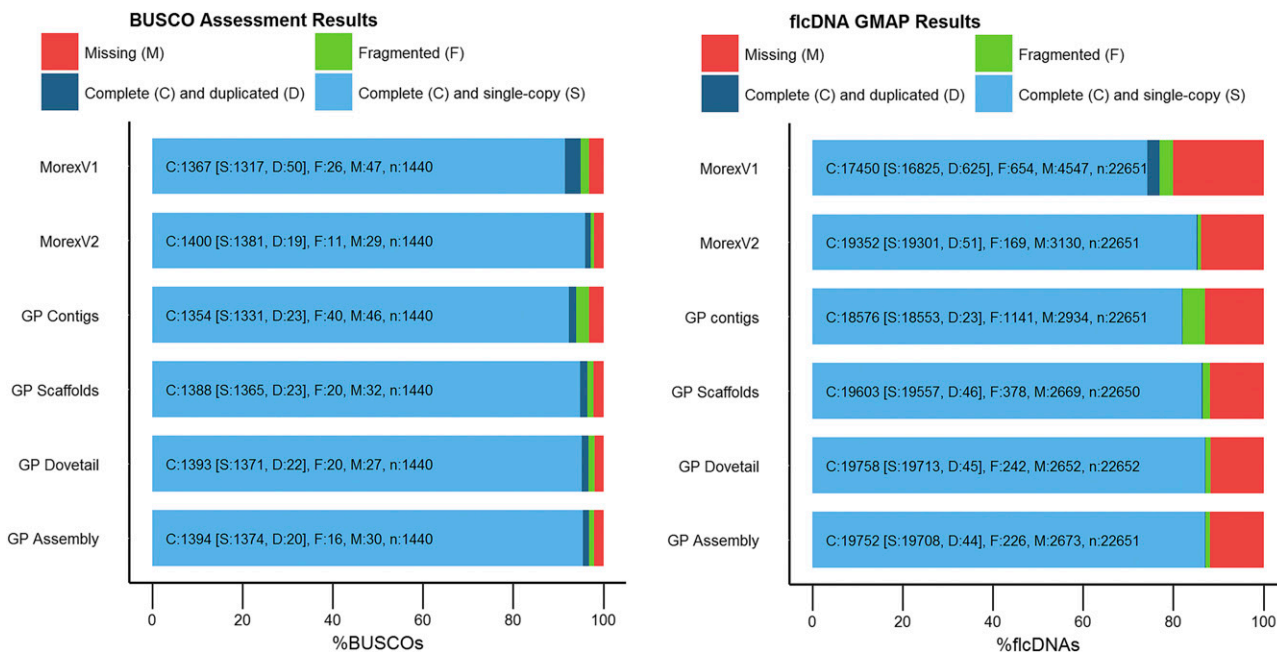|	Golden Promise	MorexV1	MorexV2
	72.88	70.65	74.93
Class I: Retroelement			
LTR Retrotransposon	63.16	62.25	64.25
LTR/Copia	19.87	21	20.94
LTR/Gypsy	42.97	40.93	42.99
Unclassified LTR	0.32	0.31	0.32
Non-LTR Retrotransposon			
LINE	0.25	0.24	0.24
SINE	0.03	0.03	0.03
Class II: DNA Transposon			
DNA Transposon Superfamily	8.25	7.39	8.97
CACTA superfamily (DTC)	7.77	6.92	8.49
hAT superfamily (DTA)	0.004	0.004	0.004
Mutator superfamily (DTM)	0.13	0.13	0.13
Tc1/Mariner superfamily (DTT)	0.2	0.19	0.2
Harbinger superfamily (DTH)	0.13	0.12	0.13
Unclassified (DTX)	0.02	0.02	0.02
MITE (DXX)	0.01	0.01	0.01
Helitron (DHH)	0.08	0.09	0.09
Unclassified Element (XXX)	0.46	0.3	0.74
Simple Sequence Repeats	0.63	0.36	0.59

Figure 2 Completeness assessment of the Golden Promise assembly in comparison to the previous steps of the assembly process and the published barley references MorexV1 and MorexV2 for both the BUSCO analysis (a) and the flcDNA mapping analysis (b).

total of 63,658 genes in MorexV2, 62,605 genes could be transferred onto Golden Promise. Among these genes 7.2% did not contain a valid start codon, 7.7% had a different nucleotide length and 5% had a premature stop codon in the gene. As some transcripts contained a combination of those errors, this still left 84% of correctly transferred transcripts.

Data validation and quality control

We used two approaches to evaluate the quality of the Golden Promise assembly based on gene content. The analysis was done for each of the steps along the assembly process. The first approach was done with BUSCO (Benchmarking Universal Single-Copy Orthologs, v3.0.2) (Simão *et al.* 2015; Waterhouse *et al.* 2018). It assesses the completeness of a genome by identifying conserved single-copy, orthologous genes. Even the contig stage had already more complete single copy genes, 92.4%, in comparison to the published barley assembly from the cultivar MorexV1 with 91.5% (Figure 2a). Throughout the assembly process this improved to 95.2% of complete and single copy genes in the final pseudomolecule. This is very close to the recently published MorexV2 assembly with 97.2% of single copy genes. As expected, the number of fragmented sequences decreased during the assembly process from 2.8% of fragmented genes to only 1.1% in the pseudomolecule.

The second approach used a flcDNA dataset which consists of 22,651 sequences generated from the cultivar Haruna Nijo (Sato *et al.* 2009; Matsumoto *et al.* 2011). These sequences were created from 12 different conditions and representing a good snapshot of the barley transcriptome. They can be used to identify the number of retained sequences in the Golden Promise pseudomolecule and give an impression on the segmentation of the pseudomolecule, highlighted by cDNAs which have been split within or across chromosomes. The 22,651 flcDNAs were mapped to the Golden Promise pseudomolecule using Gmap (version 2018-03-25; Wu and Watanabe 2005) with the following parameters: a minimum identity of 98% and a minimum trimmed coverage of 95%. The results for this dataset are very

similar to the BUSCO analysis. The contigs already contained 81.4% of complete and single copy genes in comparison to the 73% of the MorexV1 reference (Figure 2b). The final assembly contained 87.1% of complete and single copy genes, 14% more than the barley reference MorexV1 and around 400 genes more in comparison to MorexV2 accounting for a difference of 1.9%. Similar to the BUSCO analysis the number of duplicated complete genes and the number of fragmented genes is decreased in the Golden Promise assembly. Again, the overall comparison to MorexV2 shows very similar results emphasizing the high quality of both barley genomes.

CONCLUSION

Here, we presented such an assembly that is an improvement on the currently available barley reference from the cultivar MorexV1 (International Barley Genome Sequencing Consortium 2012; Mascher *et al.* 2017) and near-equivalent to the recently released MorexV2 (Monat *et al.* 2019). Importantly, it is a European 2-row cultivar, expanding barley genomic resources to European breeding material in contrast to the American 6-row cultivar Morex. The importance of having another genome assembly has already been demonstrated in the analysis of the highly divergent *Jekyll* genes (Radchuk *et al.* 2019). We anticipate it will benefit the whole barley research community but will be especially useful for groups working on CRISPR-Cas9.

LITERATURE CITED

Bailly-Bechet, M., A. Haudry, and E. Lerat, 2014   "One code to find them all": a perl tool to conveniently parse RepeatMasker output files. Mob. DNA 5: 13. https://doi.org/10.1186/1759-8753-5-13

Belhaj, K., A. Chaparro-Garcia, S. Kamoun, and V. Nekrasov, 2013   Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9: 39. https://doi.org/10.1186/1746-4811-9-39

Bortesi, L., C. Zhu, J. Zischewski, L. Perez, L. Bassie et al., 2016   Patterns of CRISPR/Cas9 activity in plants, animals and microbes. Plant Biotechnol. J. 14: 2203–2216. https://doi.org/10.1111/pbi.12634

Clavijo, B., G. Garcia Accinelli, J. Wright, D. Heavens, K. Barr et al., 2017   W2RAP: a pipeline for high quality, robust assemblies of large complex genomes from short read data. bioRxiv (Preprint posted February 22, 2017). https://doi.org/10.1101/110999

Heavens, D., G. G. Accinelli, B. Clavijo, and M. D. Clark, 2015   A method to simultaneously construct up to 12 differently sized Illumina Nextera long mate pair libraries with reduced DNA input, time, and cost. Biotechniques 59: 42–45. https://doi.org/10.2144/000114310

Hensel, G., V. Valkov, J. Middlefell-Williams, and J. Kumlehn, 2008   Efficient generation of transgenic barley: The way forward to modulate plant–microbe interactions. J. Plant Physiol. 165: 71–82. https://doi.org/10.1016/j.jplph.2007.06.015

Ibrahim, A. S., O. M. El-Shihy, and A. H. Fahmy, 2010   Highly efficient Agrobacterium tumefaciens-mediated transformation of elite Egyptian barley cultivars. American-Eurasian Journal of Sustainable Agriculture 4: 403–413.

International Barley Genome Sequencing Consortium, 2012   A physical, genetic and functional sequence assembly of the barley genome. Nature 491: 711–716. https://doi.org/10.1038/nature11543

Jaganathan, D., K. Ramasamy, G. Sellamuthu, S. Jayabalan, and G. Venkataraman, 2018   CRISPR for Crop Improvement: An Update Review. Front Plant Sci 9: 985. https://doi.org/10.3389/fpls.2018.00985

Karkute, S. G., A. K. Singh, O. P. Gupta, P. M. Singh, and B. Singh, 2017   CRISPR/Cas9 Mediated Genome Engineering for Improvement of Horticultural Crops. Front Plant Sci 8: 1635. https://doi.org/10.3389/fpls.2017.01635

König, S., L. W. Romoth, L. Gerischer, and M. Stanke, 2016   Simultaneous gene finding in multiple genomes. Bioinformatics 32: 3388–3395.

Lawrenson, T., O. Shorinola, N. Stacey, C. Li, L. Ostergaard et al., 2015   Induction of targeted, heritable mutations in barley and Brassica oleracea using RNA-guided Cas9 nuclease. Genome Biol. 16: 258. https://doi.org/10.1186/s13059-015-0826-7

Lim, W. L., H. M. Collins, R. R. Singh, N. A. J. Kibble, K. Yap et al., 2018   Method for hull-less barley transformation and manipulation of grain mixed-linkage beta-glucan. J. Integr. Plant Biol. 60: 382–396. https://doi.org/10.1111/jipb.12625

Mapleson, D., L. Venturini, G. Kaithakottil, and D. Swarbreck, 2017   Efficient and accurate detection of splice junctions from RNAseq with Portcullis. bioRxiv (Preprint posted November 10, 2017). https://doi.org/10.1101/217620

Mascher, M., H. Gundlach, A. Himmelbach, S. Beier, S. O. Twardziok et al., 2017   A chromosome conformation capture ordered sequence of the barley genome. Nature 544: 427–433. https://doi.org/10.1038/nature22043

Mascher, M., V. J. Schuenemann, U. Davidovich, N. Marom, A. Himmelbach et al., 2016   Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. Nat. Genet. 48: 1089–1093. https://doi.org/10.1038/ng.3611

Matsumoto, T., T. Tanaka, H. Sakai, N. Amano, H. Kanamori et al., 2011   Comprehensive Sequence Analysis of 24,783 Barley Full-Length cDNAs Derived from 12 Clone Libraries. Plant Physiol. 156: 20–28. https://doi.org/10.1104/pp.110.171579

Monat, C., S. Padmarasu, T. Lux, T. Wicker, H. Gundlach et al., 2019   TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. bioRxiv (Preprint posted May 9, 2019) https://doi.org/10.1101/631648

Murray, F., R. Brettell, P. Matthews, D. Bishop, and J. Jacobsen, 2004   Comparison of Agrobacterium-mediated transformation of four barley cultivars using the GFP and GUS reporter genes. Plant Cell Rep. 22: 397–402. https://doi.org/10.1007/s00299-003-0704-8

Padmarasu, S., A. Himmelbach, M. Mascher, and N. Stein, 2019   In Situ Hi-C for Plants: An Improved Method to Detect Long-Range Chromatin Interactions, pp. 441–472 in Plant Long Non-Coding RNAs: Methods and Protocols, edited by J. A. Chekanova and H.-L. V. Wang. Springer, New York.

Putnam, N. H., B. L. O'Connell, J. C. Stites, B. J. Rice, M. Blanchette et al., 2016   Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. Genome Res. 26: 342–350. https://doi.org/10.1101/gr.193474.115

Radchuk, V., R. Sharma, E. Potokina, R. Radchuk, D. Weier et al., 2019   The highly divergent Jekyll genes, required for sexual reproduction, are lineage specific for the related grass tribes Triticeae and Bromeae. Plant J. 98: 961–974.

Rapazote-Flores, P., M. Bayer, L. Milne, C.-D. Mayer, J. Fuller et al., 2019   BaRTv1.0: an improved barley reference transcript dataset to determine accurate changes in the barley transcriptome using RNA-seq. BMC Genomics 20: 968. https://doi.org/10.1186/s12864-019-6243-7

Sato, K., T. Shin-I, M. Seki, K. Shinozaki, H. Yoshida et al., 2009   Development of 5006 Full-Length CDNAs in Barley: A Tool for Accessing Cereal Genomics Resources. DNA Res. 16: 81–89. https://doi.org/10.1093/dnares/dsn034

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov, 2015   BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smit, A., R. Hubley, and P. Green, 2013–2015   RepeatMasker Open-4.0. http://www.repeatmasker.org/faq.html.

Stanke, M., R. Steinkamp, S. Waack, and B. Morgenstern, 2004   AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 32: W309–W312. https://doi.org/10.1093/nar/gkh379

Waterhouse, R. M., M. Seppey, F. A. Simão, M. Manni, P. Ioannidis et al., 2018   BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. Mol. Biol. Evol. 35: 543–548. https://doi.org/10.1093/molbev/msx319

Wicker, T., D. E. Matthews, and B. Keller, 2002   TREP: a database for Triticeae repetitive elements. Trends Plant Sci. 7: 561–562. https://doi.org/10.1016/S1360-1385(02)02372-5

Wu, T. D., and C. K. Watanabe, 2005   GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21: 1859–1875. https://doi.org/10.1093/bioinformatics/bti310

Communicating editor: S. Scofield