markers to deletion bins within the D genome of reference polyploid wheat (Chinese Spring) will allow for the construction of a physical marker scaffold that will aid in genome sequence completion and future mapping and cloning studies. To design an optimal marker array, we tested hybridization temperature, oligo length, and different statistical analysis methods. After determining appropriate marker design and experimental conditions we screened a pool of 206,486 repeat junction markers as well as 26,800 gene markers representing 6,700 genes. Screening results provided 46,224 markers in total that were selected for design of a final mapping array. These 46,224 markers are comprised of 30,900 repeat junction markers and 15,324 gene markers representing 6,330 genes. This final mapping array is being used to construct a high-resolution, physical map using D-genome deletion lines and radiation hybrid panels. Here we present our methods for design and analysis of the NimbleGen comparative genomic hybridization arrays constructed from *Ae. tauschii* repeat junction markers, the construction of a NimbleGen repeat junction array using selected markers, and its use in the development of a physical map for the D genome of hexaploid wheat.

Poster 10. Genome-wide characterization of transposable element repeat junctions in barley and their application in marker development for chromosome 3H.

Mona Mazaheri [1], Penny M.A. Kianian [1], Raed Seetan [2], Ajay Kumar [1], Anne M. Denton [2], Ruvini Ariyadasa [3], Nils Stein [3], Hana Simkova [4], Jaroslav Doelzel [4], Thomas Drader [5], Yong Q. Gu [5], and Shahryar F. Kianian [1].

[1] Department of Plant Science and [2] Department of Computer Sciences, North Dakota State University, Fargo, ND 58105, USA; [3] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany; [4] Institute of Experimental Botany, Olomouc, CZ-77200, Czech Republic; and [5] USDA–ARS, PWA, WRRC–GGD, Albany, CA 94710, USA.

Transposable elements (TEs) account for about 70% of the barley (*Hordeum vulgare* L.) genome. TEs move in the genome by inserting to new regions through copy-and-paste or cut-and-paste mechanisms. Insertion of TEs in DNA regions generates unique junctions between the TEs and the sequences in which they land. We investigated the uniqueness of these junctions throughout the barley genome and their potential application for marker development. The 10-Gb survey sequencing data of the seven barley chromosomes was searched with the 'RJPrimers' pipeline to estimate the frequency of repeat junctions (RJs). We found 988,750 RJs distributed evenly among the chromosomes with an average of 1 RJ per 10 Kb. Repeat junction markers (RJMs) for each chromosome were designed based on detected repeat junctions. Each RJM consists of one primer that spans unique insertion site of TE, whereas the second primer is designed from any region within 1 Kb of the junction. We randomly chose 36 RJMs from chromosome 3H to amplify five barley cultivars (Betzes, Golden Promise, Bowman, and Haruna Mugi) and two wild barley (*H. bulbosum* L.) accessions. Out of 36 RJMs, 28 (68%) amplified a single band and 21 (58%) were polymorphic. Based on scoring data of 21 RJMs, we were able to separate all tested barley accessions and cluster them into two main groups. The abundance of RJMs makes them an ideal marker for genome saturation. Further, the high level of polymorphism makes them ideal for molecular breeding applications.

Poster 11. Best practices for RNA-Seq differential expression analysis in barley.

Dharanya Sampath [1], Sarah Ayling [1], David Swarbreck [1], Melanie Febrer [1], Pete E. Hedley [2], Matthias Pfeifer [3], Burkhard Steuernagel [3], Klaus Mayer [3], Robbie Waugh [2], and Mario Caccamo [1].

[1] The Genome Analysis Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom; [2] James Hutton Institute, Invergowrie, Dundee, Scotland DD25DA, United Kingdom; [3] Munich Information Centre for Protein Sequences, Helmholtz Center Munich, 85764 Neuherberg, Germany.

Barley (*Hordeum vulgare*) being a member of the grass family is one of the most important large-genome cereals and a close relative of wheat. It has a complex diploid genome of 5.1 Gb and is being extensively used for genetic studies. As a part of the Barley Genome Sequencing Project, RNA-Seq experiments were performed using eight different tissues representing different developmental stages, including 4-day embryos, roots, and shoots from seedlings, young developing inflorescences (5-mm and 10–15 mm stages), at the six-leaf stage, and from the developing grain (5 and 15 days post anthesis). These datasets have been used to improve the barley gene model predictions and to detect differentially expressed transcripts. Here we discuss the results of the expression analysis that we implemented using different available

software packages. In the first instance, we processed the raw reads by filtering them for low quality and then proceed to map them to the latest barley genome reference using Bowtie (Langmead et al. 2009) and Tophat (Trapnell et al. 2009) to consider spliced reads. These mapped reads were used for a downstream expression analysis study using the Cufflinks packages (Cufflinks, Cuffcompare, and Cuffdiff) (Trapnell et al. 2010) and R bioconductor packages (DESeq, NOISeq, and edgeR) (Anders and Huber 2010; Tarazona et al. 2011; Robinson et al. 2010). A comparison based on the results of these approaches indicated differences in the sets of genes reported with significant fold changes. Identification of the biological function of the nonoverlapping genes produced by the different analyses has been undertaken to explore the impact of using any one of these analyses in isolation.

Anders S and Huber W. 2010. Differential expression analysis for sequence count data. Genome Biology 11:R106.

Langmeand B, Trapnell C, Pop M, and Salzberg S. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

Robinson MD, McCarthy DJ, and Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-40.

Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, and Conesa A. 2011. Differential expression in RNA-seq: a matter of depth. Genome Res 21:2213-2223.

Trapnell C, Pachter L, and Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105-1111.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, and Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotech 28:511-515.

Poster 12. A triticale (x Triticosecale Wittm.) reference transcriptome.

Michele Frick [1], Yong Xu [1], Carolyn Penniket [1], Laurian S. Robert [2], Linda J. Harris [2], Frances Tran [2], Danielle Schneiderman [2], Patrick J. Gulick [3], and André Laroche [1].

[1] Agriculture and Agri-Food Canada, Lethbridge Research Centre, Lethbridge, AB, Canada, T1J 4B1; [2] Agriculture and Agri-Food Canada, Eastern Cereal & Oilseed Research Centre, Ottawa, ON, Canada, K1A 0C6; and [3] Department of Biology, Concordia University, 7141 Sherbrooke W, Montréal, QC, Canada H4B 1R6.

Triticale produces high biomass and grain yields and is resistant to many abiotic and biotic stresses. Hexaploid triticale (AABBRR) has a large genome similar to that of common wheat (AABBDD), and possesses large numbers of repeated sequences, making it difficult to sequence. Little is known about coordination of gene expression of rye and wheat genomes in this interspecific hybrid but understanding this interaction is crucial for improving its agronomic traits. We have produced cDNAs from four different triticale lines/cultivars; AC Certa, Alta, line 797, and line 1308, from different tissues including leaf, stem, and reproductive tissues at different developmental stages or exposed to different abiotic treatments (salt and drought). Reads were cleaned, and we assembled 6.7 M 454 reads and 276. 3 M Illumina reads together and identified, using a stringent assembly protocol, a total of 72,218 contigs with N50 and N90 values of 1,658 and 588 bp, respectively. Of these 72,218 contigs, 50,524 transcripts appeared to be full length after comparison to the sequences obtained from the four monocot sequenced genomes (rice, maize, sorghum and Brachypodium). The noncoding sequences appear to represent long, noncoding cDNAs, short ORFs, and precursors of microRNAs. At least 31,230 of these transcripts appear to be translated. In addition, we determined that approximately 20% of the genes had transcripts with alternative splice variants. Furthermore, nearly 60% of these transcripts were electronically annotated. Compared to our previous results with development of a rye transcriptome, the transcripts corresponding to the R genome in triticale were segregated from the sequences originating from genome A and B. A detailed molecular characterization of the transcripts into different functional groups as well differential expression of the rye and wheat transcripts under control and stressed conditions will be presented. We also will report on the utilization of fractionated DNA to further extend the 5' region of transcripts. Therefore, using next generation sequencing to de novo assemble a comprehensive reference transcriptome for triticale was possible. This approach would be useful for species with large, unsequenced genomes.