

A candle in the dark: Reference genome assembly for rye highlights the importance of data visualisation and manual editing

M. Timothy Rabanus-Wallace¹, Nils Stein¹, Martin Mascher¹, Andreas Graner¹, Andreas Boerner¹, Andreas Houben¹, Uwe Scholtz¹, Sudharsan Padmarasu¹, Axel Himmelbach¹, Ines Walde¹, Susanne Koenig¹, Manuela Knauff¹, Curtis Pozniak², Andy Sharpe², David Konkin², Faouzi Bekkaoui², Jana Cizkova³, Hana Simkova³, Jaroslav Dolezel³, Alex Hastie⁴, Heidrun Gundlach⁵, Manuel Spannagl⁵, Klaus Mayer⁵, Eva Bauer⁶, Anthony Hall⁷, Cristobal Uauy⁸, Frank Ordon⁹, Berndt Hackauf⁹, Viktor Korzun¹⁰, Joachim Fromme¹¹, André Laroche¹², Stefan Stojakowski¹³, Hanna Bolibok-Brągoszewska¹⁴, Monika Rakoczy-Trojanowska¹⁴, Alan Schulman¹⁵, Xuefeng Ma¹⁶, Jessie Poland¹⁷, Hikmet Budak¹⁸, Vijay Tiwari¹⁹, Lynne Reuber²⁰, Beat Keller²¹, Hakan Ozkan²²

¹ Leibniz-Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

² College of Agriculture and Bioresources, University of Saskatoon, Canada

³ Centre of Plant Structural and Functional Genomics, Institute of Experimental Botany CAS, Czech Republic

⁴ Bionano Genomics, San Diego, USA

⁵ Plant Genome and Systems Biology, Helmholtz Zentrum München, Germany

⁶ Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt, Technische Universität München, Germany

⁷ Earlham Institute, Norwich Research Park, UK

⁸ John Innes Centre, Norwich Research Park, UK

⁹ Institute for Resistance Research and Stress Tolerance, Julius Kühn-Institut, Germany

¹⁰ KWS LOCHOW, Einbeck, Germany

¹¹ Hybro Saatzucht, Schenkenberg, Germany

¹² Agriculture and Agri-Food, Ottawa, Canada

¹³ Department of Genetics, Plant Breeding and Biotechnology, West Pomeranian University of Technology, Szczecin, Poland

¹⁴ Department of Plant Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warsaw, Poland

¹⁵ University of Helsinki, Helsinki, Finland

¹⁶ Noble Research Institute, Ardmore, USA

¹⁷ Department of Agronomy, Kansas State University, Manhattan, USA

¹⁸ Plant Sciences & Plant Pathology, Michigan State University, East Lansing, USA

¹⁹ Plant Science & Landscape Architecture, University of Maryland, Upper Marlboro, USA

²⁰ 2Blades Foundation, San Mateo, USA

²¹ Department of Plant and Microbial Biology, Zürich University, Zürich, Switzerland

²² Faculty of Agriculture, Cukurova University, Adana, Turkey

E-mail of corresponding author: wallace@ipk-gatersleben.de

The rye genome poses a major assembly challenge owing to its large size, repetitive content, and fixed heterozygosity. A suite of new technologies (such as molecule-linked reads, chromosome-conformation-capture reads, and optical maps) help to overcome many of these challenges, but integrating the data from very diverse sources of data to produce an assembly is difficult to automate with optimal results. Several years of work by an international consortium of institutions has produced a new reference quality

genome for rye, which was completed following the philosophy that the results of automated procedures are best taken as suggestions, to be carefully refined by a human curator with access to an array of intuitive visualisations. Such close curation can markedly increase the quality of a genome assembly, and visually-intuitive representations of a genome assembly (and its relationship to the underlying data), are similarly valuable for those using the genome for downstream applications.