

A phased diploid genome assembly of Italian ryegrass

Dario Copetti^a, Joel-Elias Kuon^b, Steven Yates^a, Roland Kölliker^a, and Bruno Studer^a

^a *Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, 8092 Zurich, Switzerland*

^b *Plant Biotechnology, Institute of Molecular Plant Biology, ETH Zurich, 8092 Zurich, Switzerland*

dario.copetti@usys.ethz.ch

Grasses of the genera *Lolium* and *Festuca* are the main feed sources for a sustainable livestock production due to their high palatability and biomass production. Since decades, their importance for the agriculture of temperate regions led to the development of new varieties through traditional breeding programs. However, newer crop improvement methods such as genomic selection could benefit from a high-quality reference genome assembly. In the past, attempts at producing genomic resources have struggled due to the complexity of the genome and the outcrossing nature of the species. We sequenced an individual of the *L. multiflorum* (Italian ryegrass) cv. Rabiosa, producing a highly-contiguous and complete assembly. Due to the high heterozygosity of the genotype, the resulting assembly was as large as the diploid genome, thus presenting the sequence of both alleles in separate collinear scaffolds. The generation of large-scale scaffolding datasets (i.e. chromosome conformation capture data and optical maps) allowed to phase sequences, reaching chromosome-level contiguity. The comparison of the two allelic sequences for a region showed an extensive amount of intergenic sequence variation, confirming that ryegrass genomes are highly dynamic. The high-quality genome assembly of the cv. Rabiosa is the first phased diploid assembly of a plant genome and provides a high-quality reference for expediting ryegrass breeding and studying the genome biology of outcrossing species.