

# The challenges while building linkage maps for diploid and polyploid species with genotyping-by-sequencing markers

Cristiane Hayumi Taniguti<sup>1</sup>, Jeekin Lau<sup>1</sup>, Antonio Augusto Franco Garcia<sup>2</sup>, Marcelo Mollinari<sup>3</sup>, David Byrne<sup>1</sup>, Oscar Riera-Lizarazu<sup>1</sup>

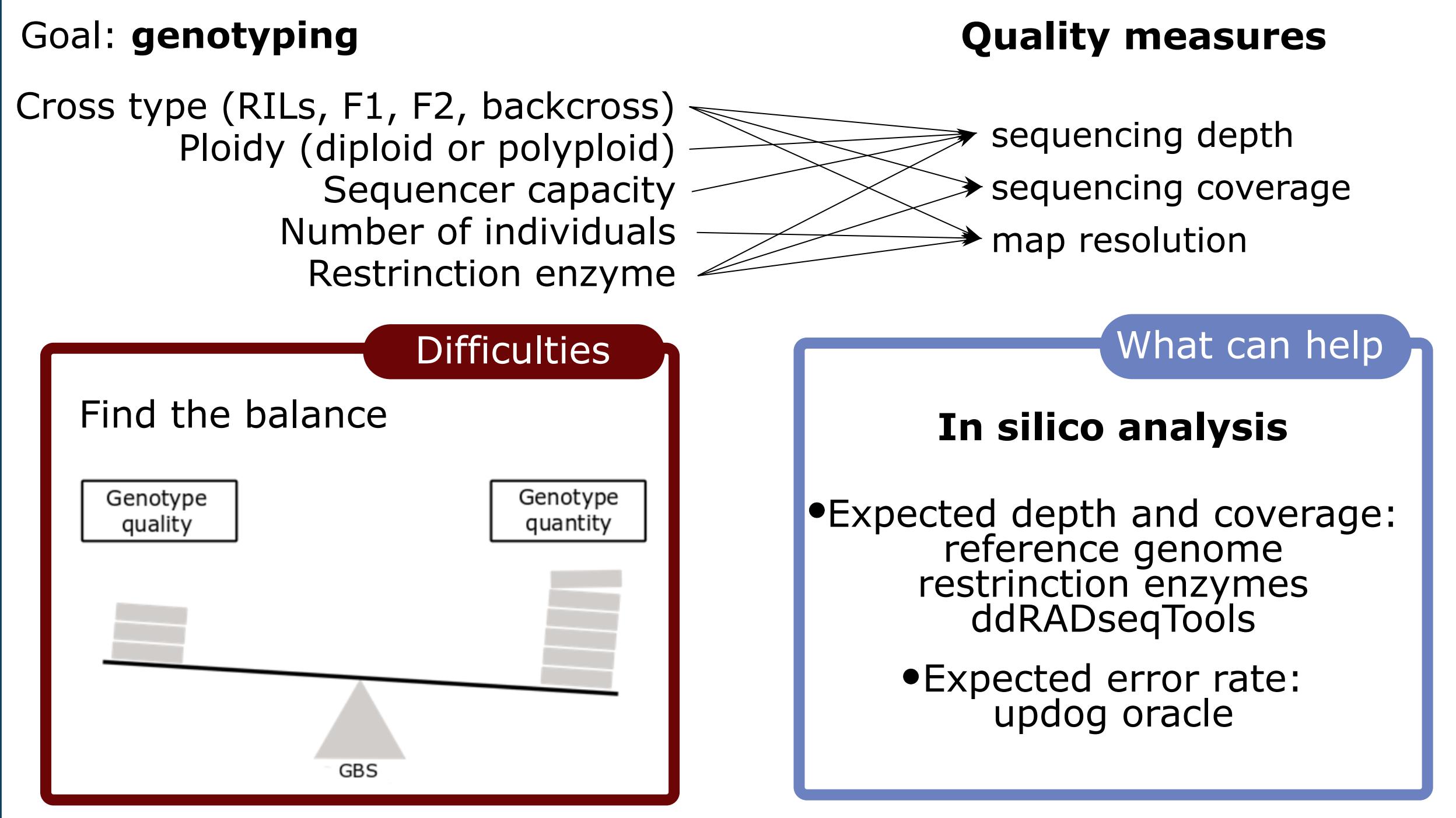
1 – Department of Horticultural Sciences, Texas A&M University, College Station, TX, USA; 2 – Department of Genetics, University of São Paulo, SP, Piracicaba, Brazil; 3 – Bioinformatics Research Center, Department of Horticultural Sciences, North Carolina State University, Raleigh, NC, USA



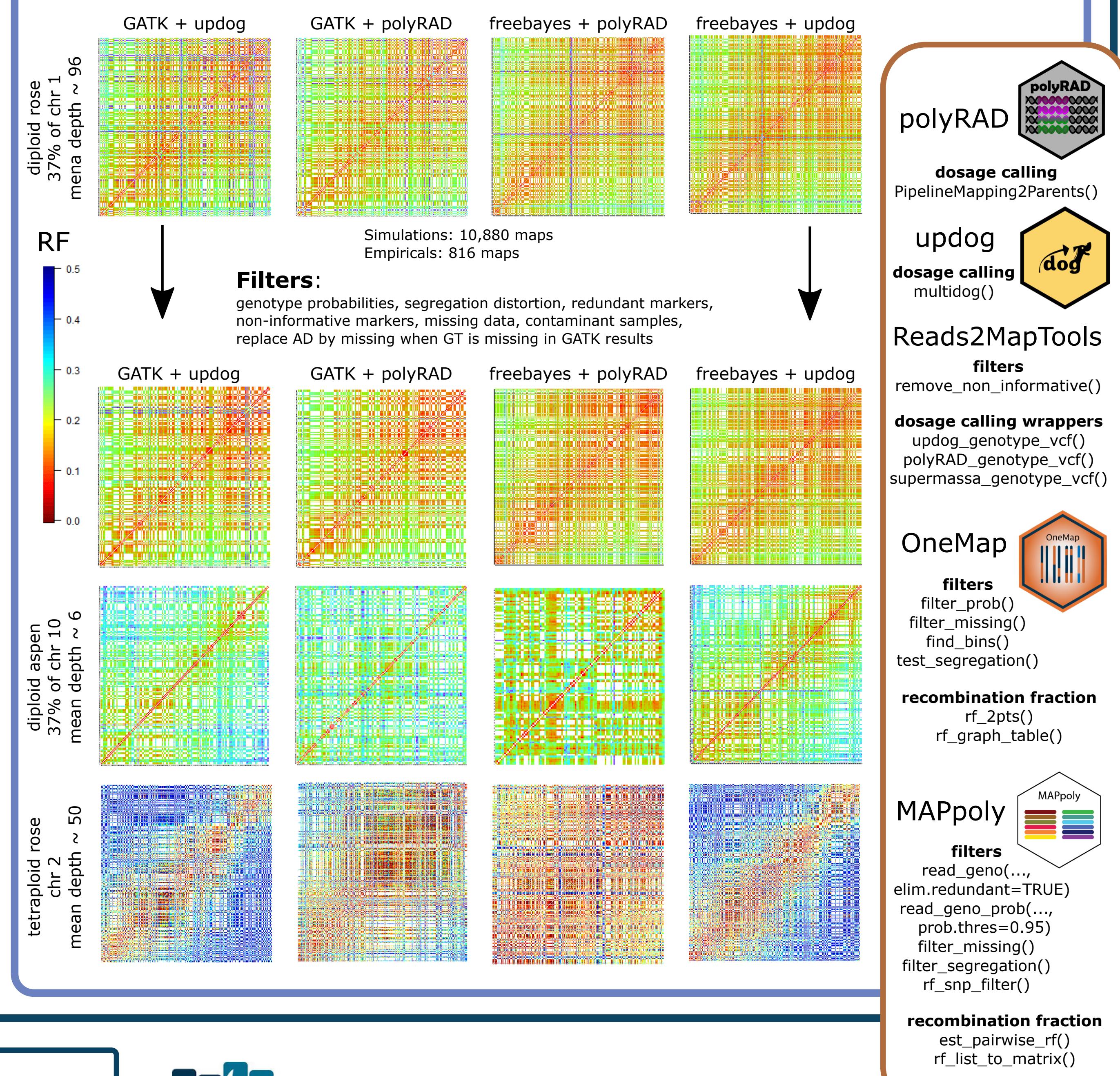
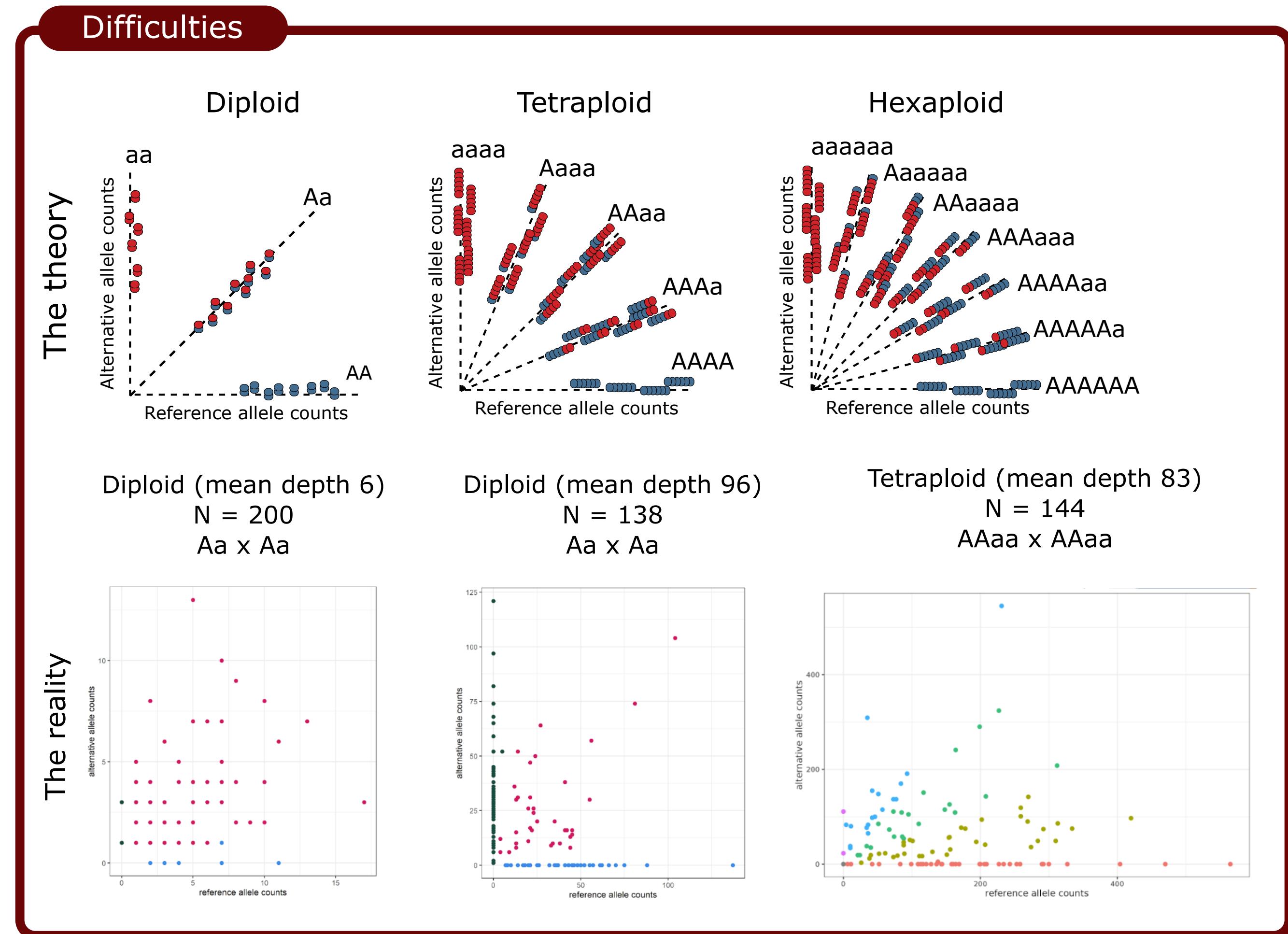
## Context

Linkage maps are important tools to explore the genetic complexity of populations. Their common applications are for population haplotyping, quantitative trait loci (QTL) studies, and assisting reference genome assemblies. The genetic markers and the statistical genetic methods to build linkage maps improved over the years. Currently, thousands of molecular markers are available through high-throughput genotyping. There are also diverse open-source software available able to perform the analysis in different mapping populations design, in inbred and outbred crosses, and in different ploidies. This scenario brings new challenges while building the linkage map. The higher amount of data requires higher computational capacity and software optimization. Furthermore, markers coming from sequencing technologies such as Genotyping-by-Sequencing (GBS) are more prone to errors which bring difficulties in the grouping, ordering, and estimation of genetic distances of the linkage map. The consequences of genotyping errors can be even harder to overcome in highly heterozygous and polyploid populations for which it is required also to estimate the linkage phase and alleles dosages. We present here some examples of linkage map building in diploid and polyploid outcrossing roses highlighting the difficulties in each step and some suggestions on how to overcome them.

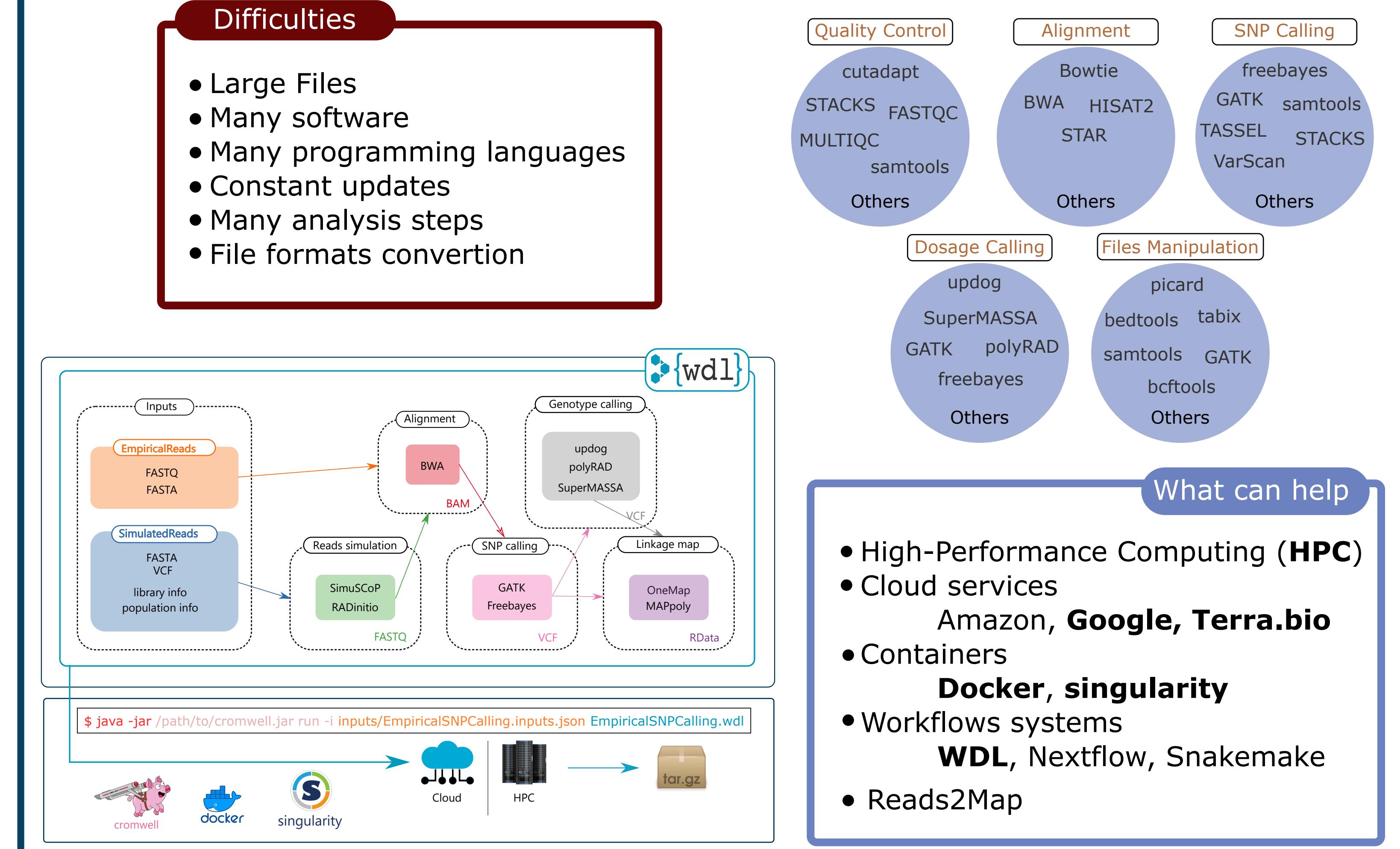
## Experimental design



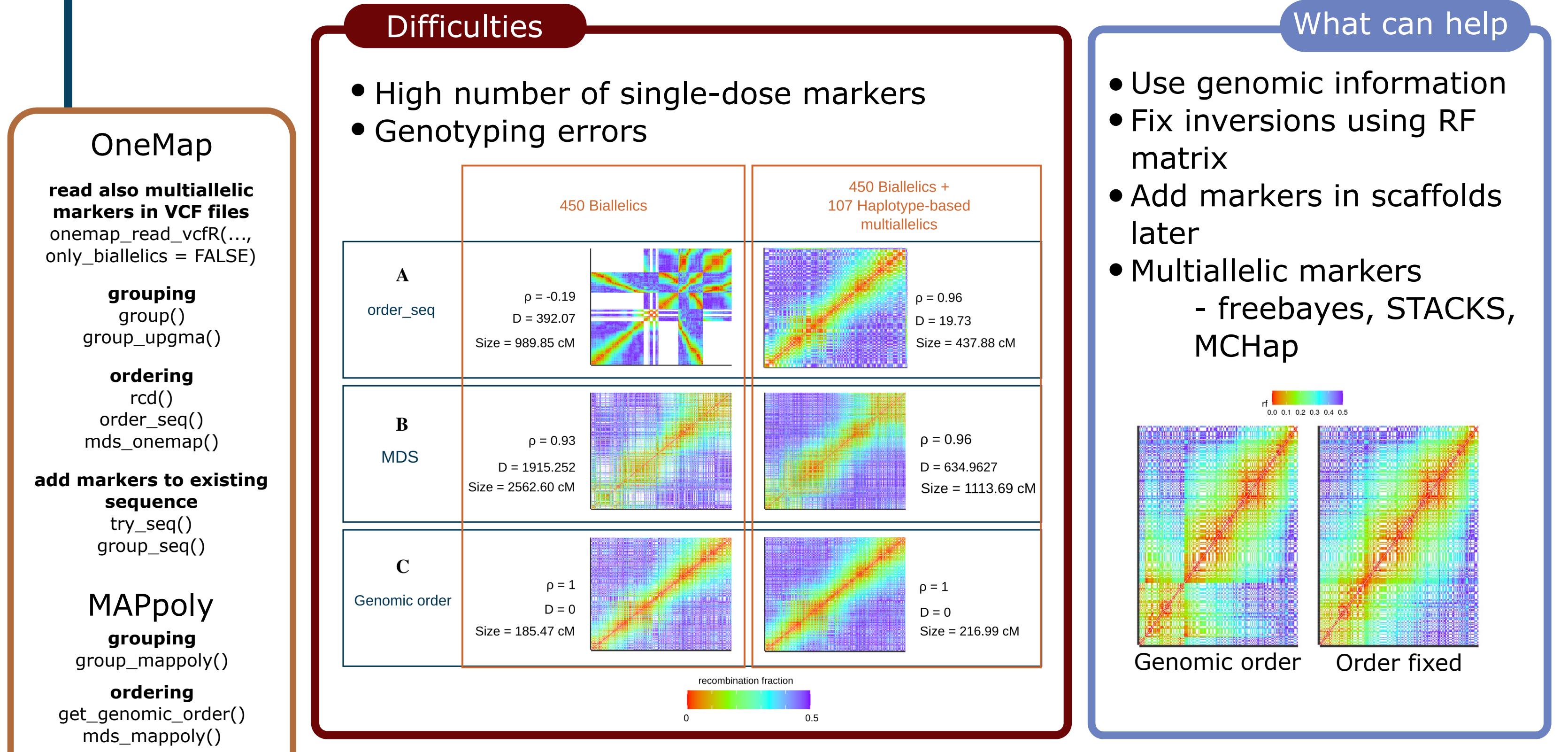
## Dosage calling



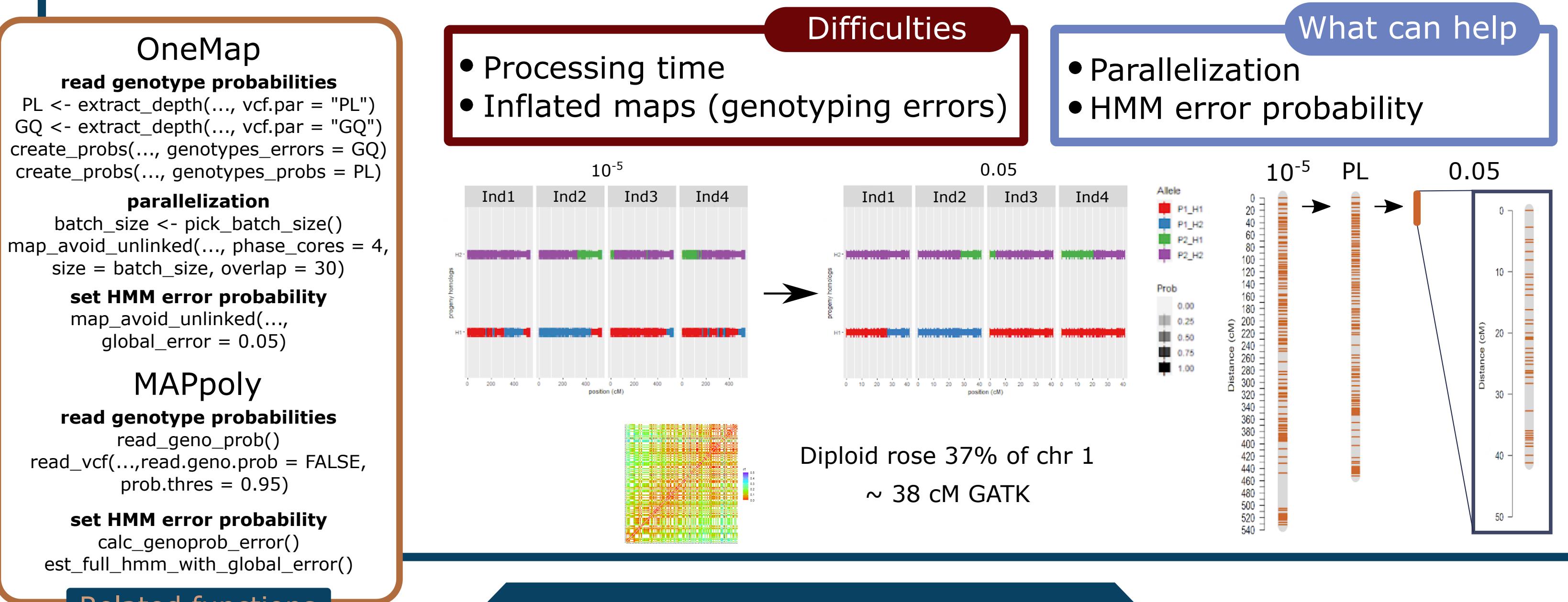
## SNP calling



## Grouping and Ordering



## Genetic Distances and Haplotypes



## Mining Results

