**Deciphering the Faba Bean Genome: High-Quality Assemblies and Resolution of Haplotypes**

Kübra Arslan1, Silvia F. Zanini1, Amine Abbadi2, Olaf Sass3, Björn Usadel4,5, Agnieszka A. Golicz1

*Agnieszka.Golicz@agrar.uni-giessen.de*

1 Institute of Agronomy and Plant Breeding I, Justus Liebig University Gießen, Germany

2 NPZ Innovation GmbH, Holtsee, Germany

3 Nordeutsche Pflanzenzucht Hans-Georg Lembke, Germany

4 Faculty of Mathematics and Natural sciences, Institute for Biological Data Science,CEPLAS, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

5 IBG-4: Bioinformatics, Institute of Bio- and Geosciences, BioSC, CEPLAS, Forschungszentrum Jülich, 52428 Jülich, Germany

Faba bean (*Vicia faba)* is an agriculturally important crop plant with a large diploid genome. Its haploid size is estimated to comprise 13 billion base pairs (2n=12), with one chromosome larger than the entire human genome. Faba bean has also one of the highest percentage of repeats among crop plant genomes.

While assembling genomes representing completely homozygous lines is conceptually most straight forward, in real world breeding material a range of different levels of heterozygosity can be observed. This feature can make faba bean genomes more challenging to assemble, with haplotype resolved assemblies requiring further validation.

In our faba bean pangenome project we used Pac-Bio HiFi sequencing (30X) and assembled genomes from several diverse faba bean lines representing breeding material and spanning a range of heterozygosity.

By assessing heterozygosity and repeat percentage, we examine the challenge of haplotig purging in this large genome, ensuring the accurate separation of haplotypes in our assemblies with different approaches. We are optimizing our designed pipeline by testing it on a single individual serving as an example for hybrid synthetic variety fababean genome (Vertigo) to reveal the quality of haplotype resolution without the need for additional sequencing data.

Through this work, we anticipate achieving high-quality fababean genome assemblies and enhancing haplotype resolution in large repetitive diploid plant genomes.