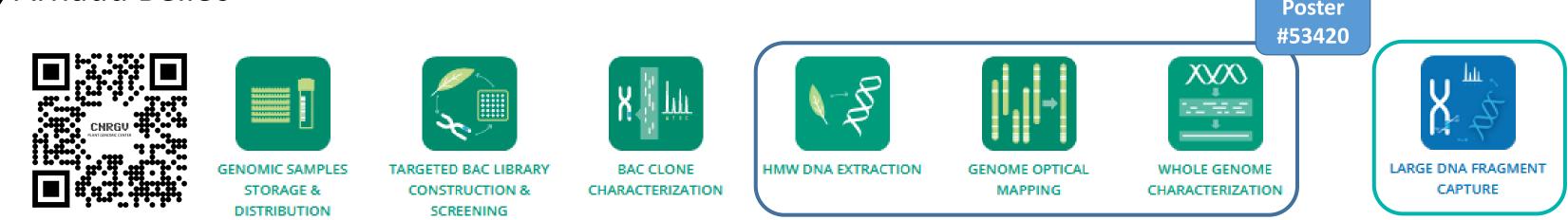
Genetic Diversity Analysis using **Targeted sequencing :** Tools and Strategies

William Marande, Caroline Callot, Stephane Cauet, Charlotte Cravero, Anthony Theron, Nathalie Rodde, Robin Schwartz, Nadine Gautier, Nadège Arnal, Isabelle Dufau, David Pujol, Sonia Vautrin, Arnaud Bellec



Gaining access to plant genome assemblies enhances our comprehension of the diversity within plant species. The large cost reduction of a genome sequence production enables the access to biodiversity, but the exploration of intra-species variability within a region of interest and even more for large genome is still costly. Therefore the utilization of targeted enrichment methods stands out as a strategic approach for biodiversity analysis of larger panel. This method offers precise and reliable information to link a genomic region to a trait of interest carried by a specific genotype.

The french genomic center (**CNRGV**), offers **two new strategies** for efficiently **targeting large regions of interest in complex plant genomes**. The first one is based on a microfluidic-based system, Xdrop from **Samplix®**, that allows the enrichment of targeted long DNA molecules starting from only a few nanograms and one PCR marker. We explore this technology by screening a diversity of sunflowers for a large locus (150kb) implicated in the resistance of broomrape.

In a second hand, we investigate the potential of **Twist BIOSCIENCE** target enrichment solution. A panel of 90 probes was designed to cover the locus KTR 2/3 (size of 63 kb) involved in the resistance of anthracnose (*Colletotrichum lindemuthianum*) on common bean.



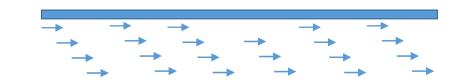
Region of interest ~80 kb

PCR specific marker

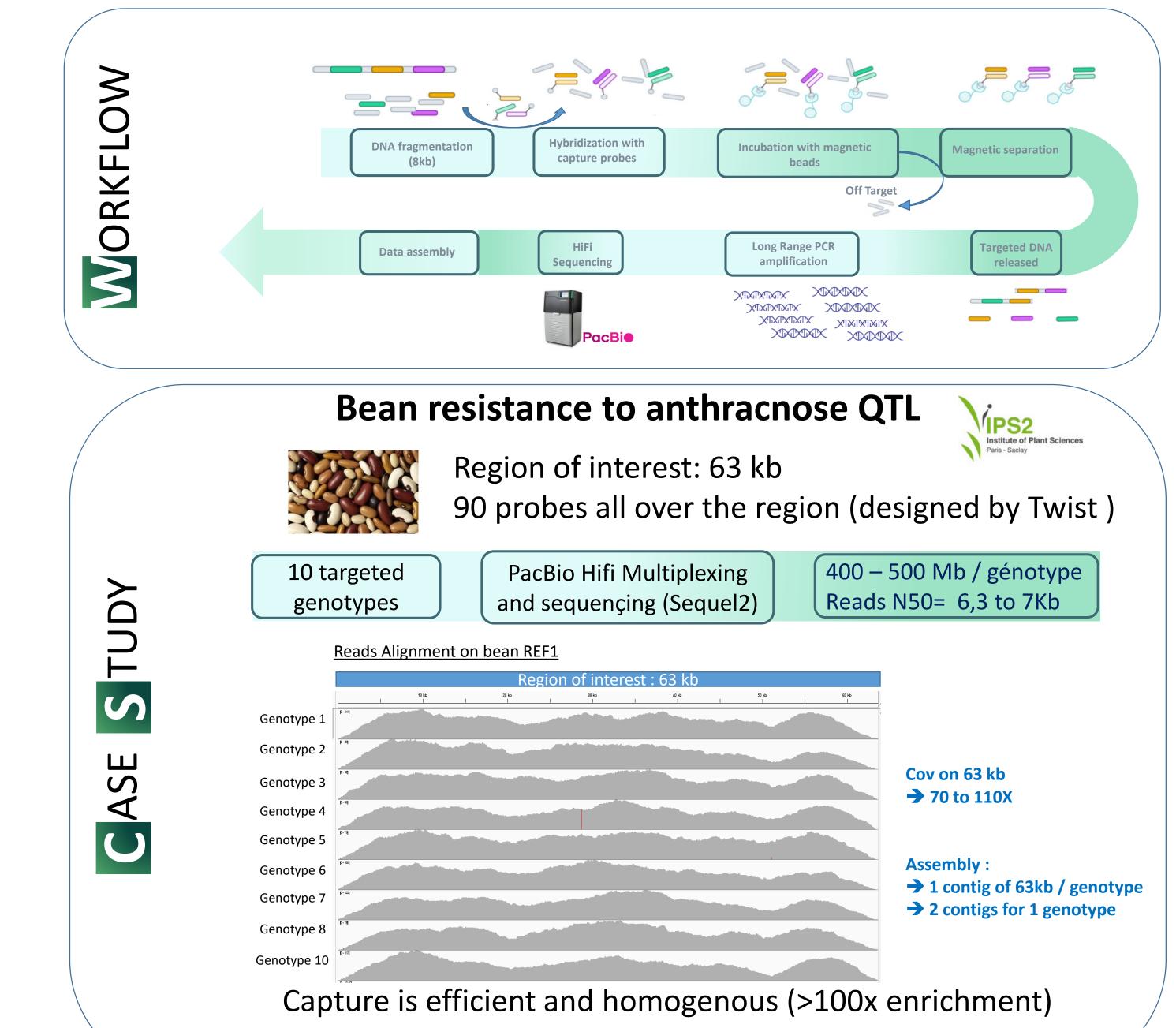
- Project that requires little knowledge
 1 specific PCR marker conserved in diversity
- 100ng HMW DNA
- Region variability is not a limitation
- Possibility of multiplexing / 30 samples

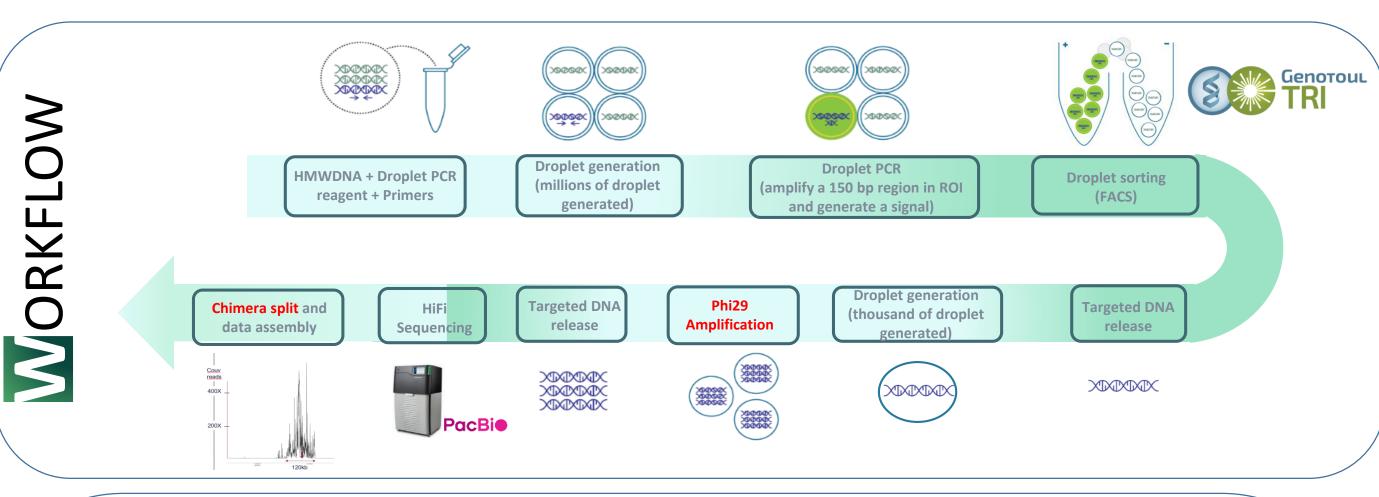


Region of interest ~ 50kb to Mbs

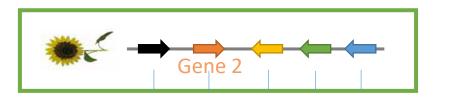


- Project that requires a reference genome
 Probes are designed all over the region of interest
 200ng HMW DNA
- Large structural variation may be a limitation
- Possibility of multiplexing / 96 samples









Region of interest: 150 kb Gene 2 primers used for the capture

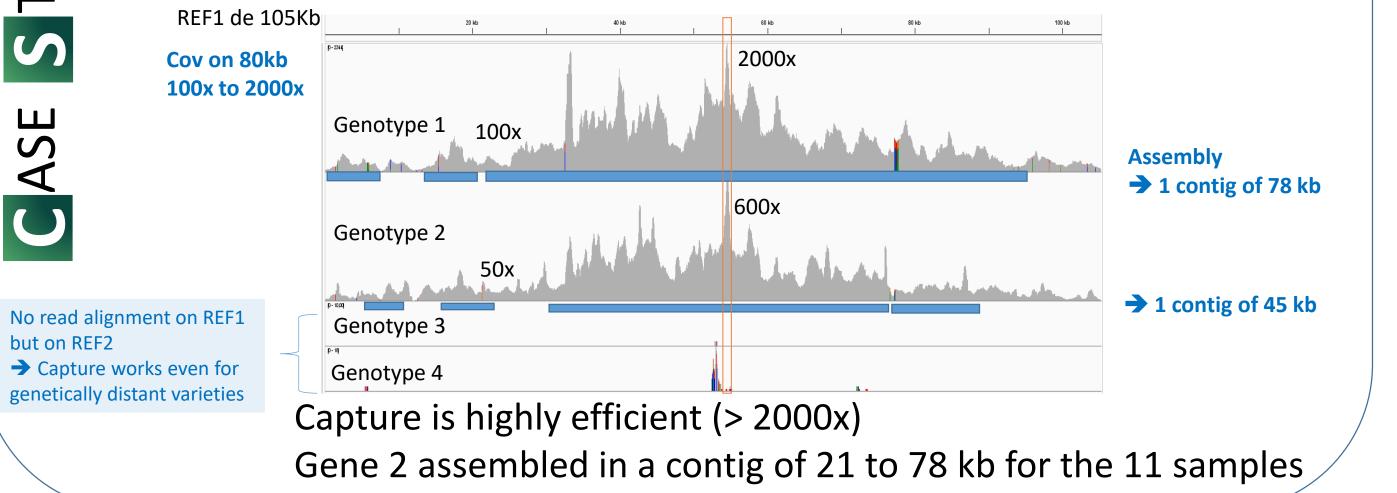
2,2 Gb to 4,5 Gb / genotype

Reads N50: 6,5 to 8,6 Kb

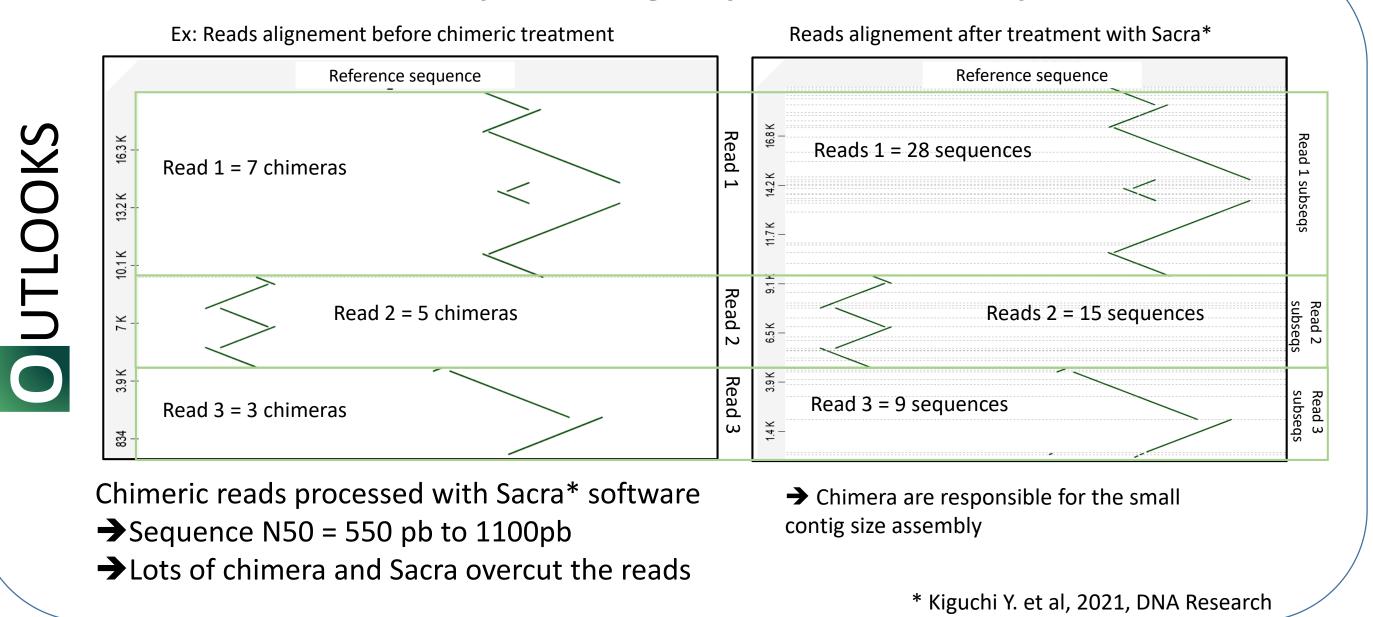


Reads Alignment on sunflower REF1 Gene2

-UDΥ







The region of interest assembled in 1 contig for 9 genotypes

- Sequencing yield improvement
- Cleaning test before sequencing in progress (sequencing production is better)
- Method ready for more complex projects
- Maize project with 10 genotypes and 2 regions of 150 kb
- Wheat project with 20 genotypes and 1 region of 700 kb

CONCLUSION

KS

Ο

0

The presented capture methods offer efficient solutions to target large genomic region of interest in complex plant genomes

What's Next?



Adaptive Sampling: Targeted sequencing method based on a software controlled enrichement. No need for upfront sample preparation.

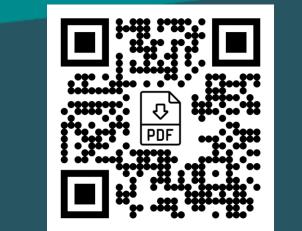
P2 solo from Oxford Nanopore Technologies



Provide an accurate and reliable genomic information for the region of interest
 Allow a rapid comparison of a region of interest between several genotypes

Acknowledgements and fundings

The authors would like to thank everyone who helped generate and analyze data for these poster
Y. Merlet from Twist Bioscience and C. Nyholm Andersen from Samplix for the experiment design
C. Claudel from INNOLEA and V. Geffroy from IPS2 for the genomic sequence information
The Samplix approach was partially funded by "Région Occitanie"





24 Chemin de Borde Rouge Auzeville - CS 52627 31326 Castanet tolosan cedex FRANCE Tél : 05.61.28.52.53 https://cnrgv.toulouse.inrae.fr https://twitter.com/CNRGV