

Development of a CRISPR-Cas9 Large DNA Fragment Targeting Strategy for Plant Genome Understanding

Céline CHANTRY DARMON¹, Carine SATGÉ¹, Nicolas MILON², Arnaud BELLEC¹, William MARANDE¹, Stéphane CAUET¹, Caroline CALLOT¹, Pauline DURIEZ³, Stéphane MUNOS³, Aurélien BANCAUD² and Hélène BERGES¹

¹ French Plant Genomic Center CNRGV – INRA, 24 Chemin de Borde Rouge, 31326 Castanet-Tolosan, France

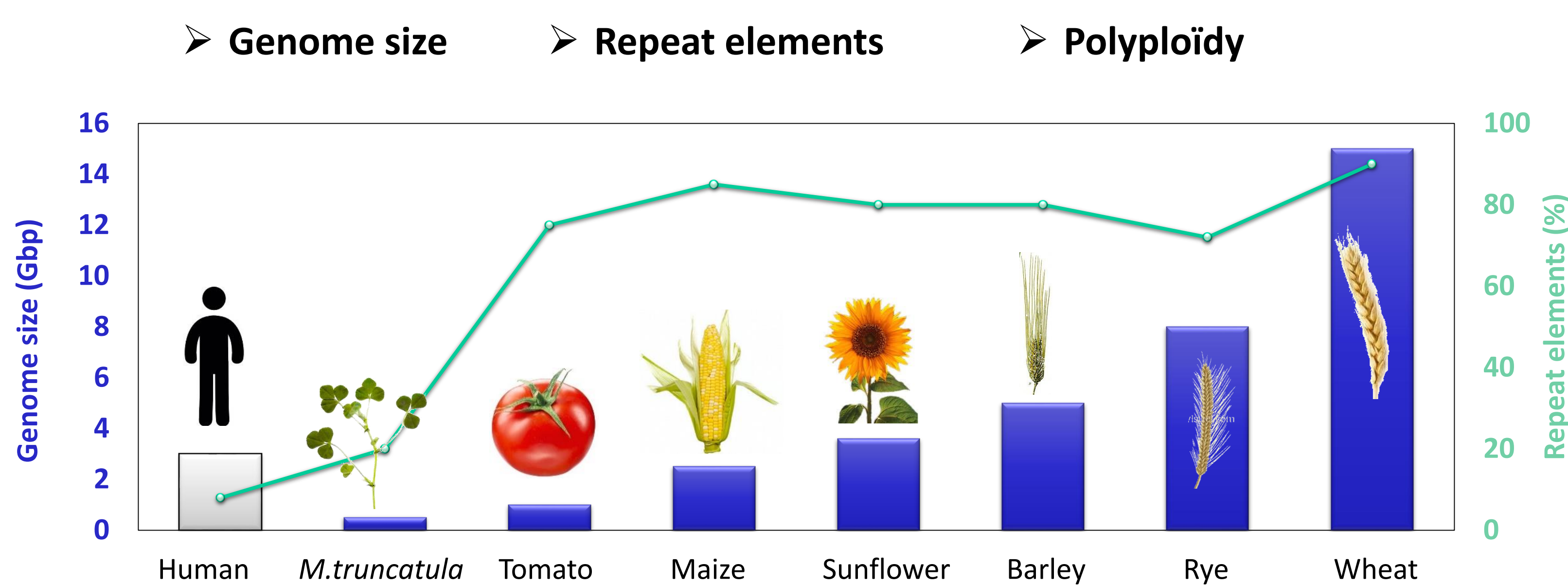
² LAAS, CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France

³ Laboratoire Plante Micro-organismes INRA/CNRS - LIPM, 24 Chemin de Borde Rouge, 31326 Castanet-Tolosan, France

Genome exploration is one of the strategic approaches of choice to better understand how plants resist, adapt and evolve. However, despite the sequencing of long reads with Third Generation Sequencing technologies, the study of numerous plant genomes remains challenging due to their complexity in terms of size, polyploidy or high percentage of repetitive elements. Moreover, in many cases, only a specific region is of interest. In these cases, the cloning of long DNA fragments is a challenge and whole genome sequencing results in higher costs and a huge amount of data to analyze. New strategies for targeting large regions of interest are really needed.

Here, we report a new sequence capture approach for large DNA fragments. Our protocol is based on the CATCH method (Cas9-Assisted Targeting of CHromosomal segments) previously described by Jiang et al., 2015 and Jiang and Zhu, 2016. This method is based on the endonuclease function of the CRISPR/Cas9 system to cut a region of interest in genomic DNA. We used this technique to capture and to sequence a 200 kbp genomic region of interest from the *Medicago truncatula* genome.

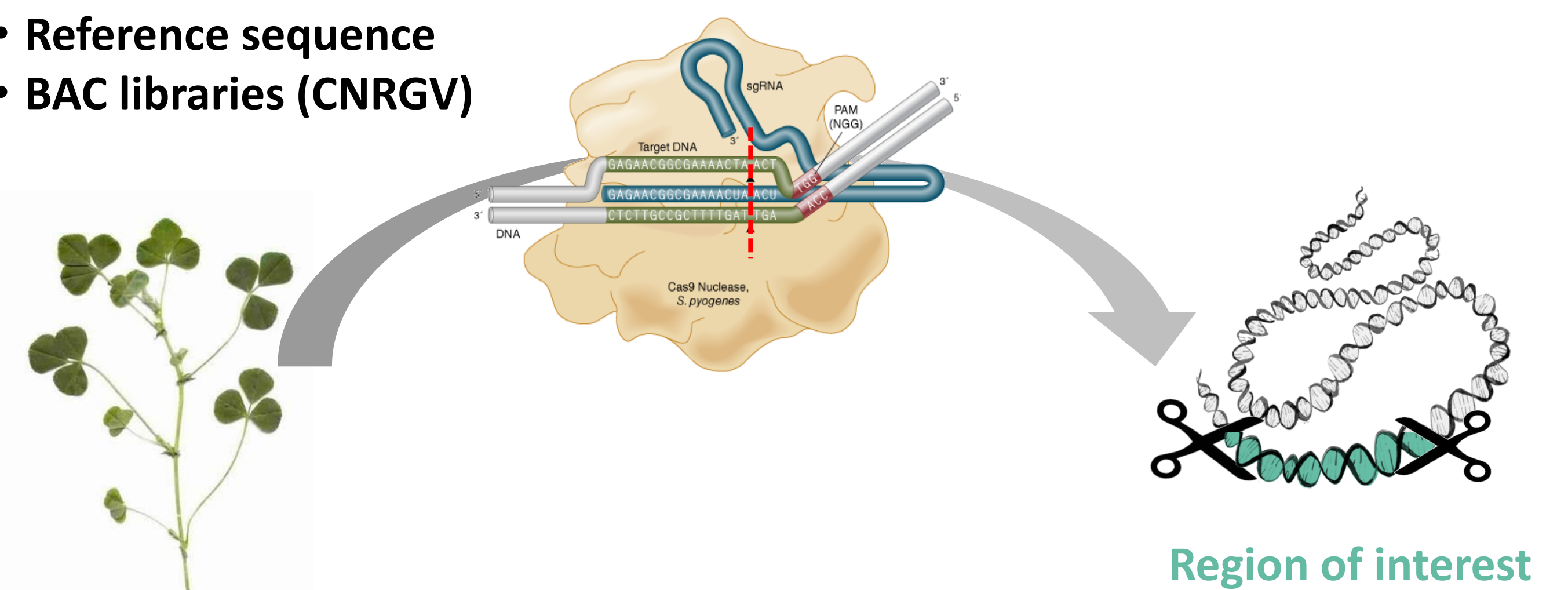
Complexity Of Plant Genomes



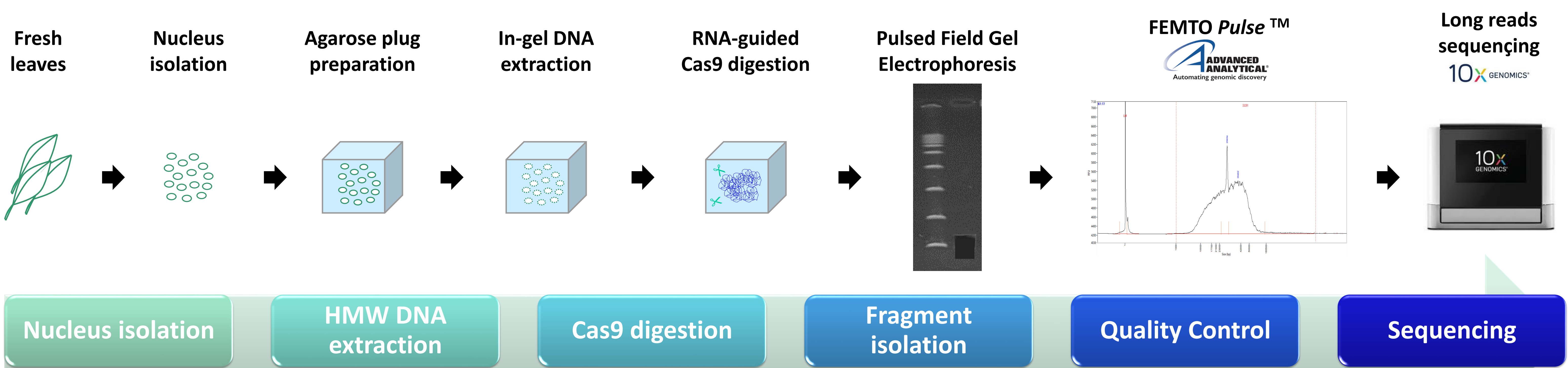
The CRISPR-Cas9 System To Target A Genomic Region Of Interest

Medicago truncatula:

- Small genome 450 Mbp
- Reference sequence
- BAC libraries (CNRGV)

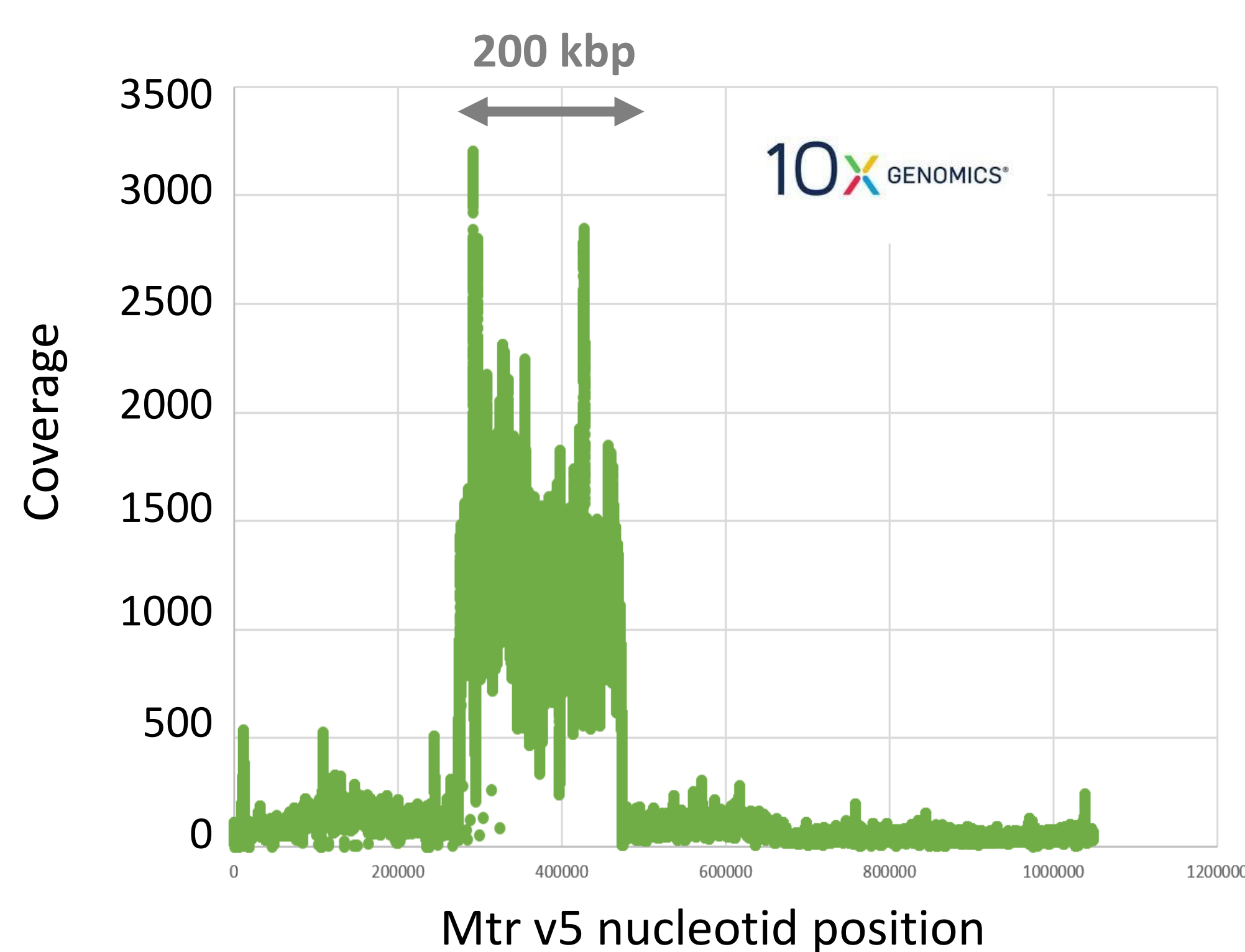


General Workflow



Sequencing Analysis of A 200 kbp Genomic Region

Alignment of the reads on the *M. truncatula* chromosome 5 sequence



Sequencing statistics	
Total number of reads	158 801 802
Mapped reads to Mtr v5 genome (%)	96.3
Mapped reads to the 200 kbp target (%)	1.2
Mean depth	2 983
Enrichment factor	68 fold



With the CATCH technique, we are able to isolate a 200 kbp region of interest from the *M. truncatula* genome. With an appropriate sequencing technology, we obtained an adequate coverage, a high sequencing depth and an enrichment factor of 68 fold for the entire region of interest. The assembling step is in progress.

We are currently testing this method on the **sunflower** which is an important crop specie. We target a **120 kbp** region rich in **repeat elements**. This region is known to be involved in the resistance to a parasitic plant. We have already **isolated and sequenced the region of interest**. Mapping and assembling steps are in progress.