

Among living organisms, plants display a high level of genome complexity due to their large size, variations in polyploidy levels and high percentage of repetitive elements. In a current context of climate change, population growth and limited energy resources, increasing plant genomes knowledge is essential to better understand mechanisms that drive plant adaptation and evolution. Despite the Next Generation Sequencing revolution including the recent long read technologies, it remains challenging to obtain high quality assemblies at the genome scale.

The French Plant Genomic Resources Center (CNRGV) is dedicated to the analysis of plant genome complexity and has recently acquired an Irys System from BioNano Genomics. Using microfluidic and high molecular weight DNA molecules, this technology allows the rapid construction of optical whole genome maps. Such maps are very helpful for genomes analysis and comparison of structural variations among genotypes. We present how the hybrid scaffolding strategy that combines PacBio sequencing and BioNano Genomics optical mapping significantly improves the assembly of complex plant genomes.

Plant Genome Complexity

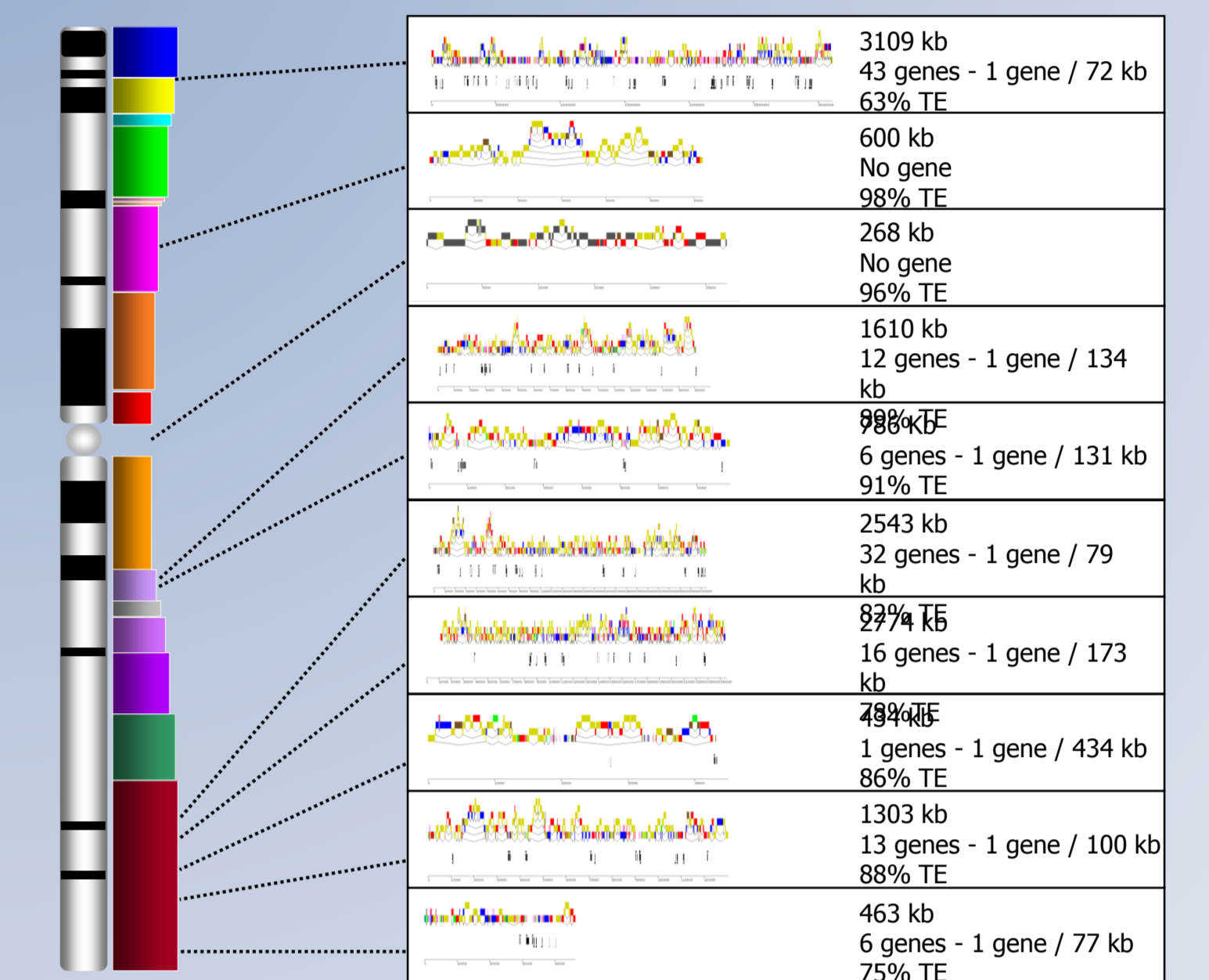


Very Large genome Size



Wheat: 17 Gb
90 % of TE - 6X

High percentage of transposable elements



Ex : Wheat Chinese Spring Chromosome 3B
(The picture was kindly provided by Etienne Paux)

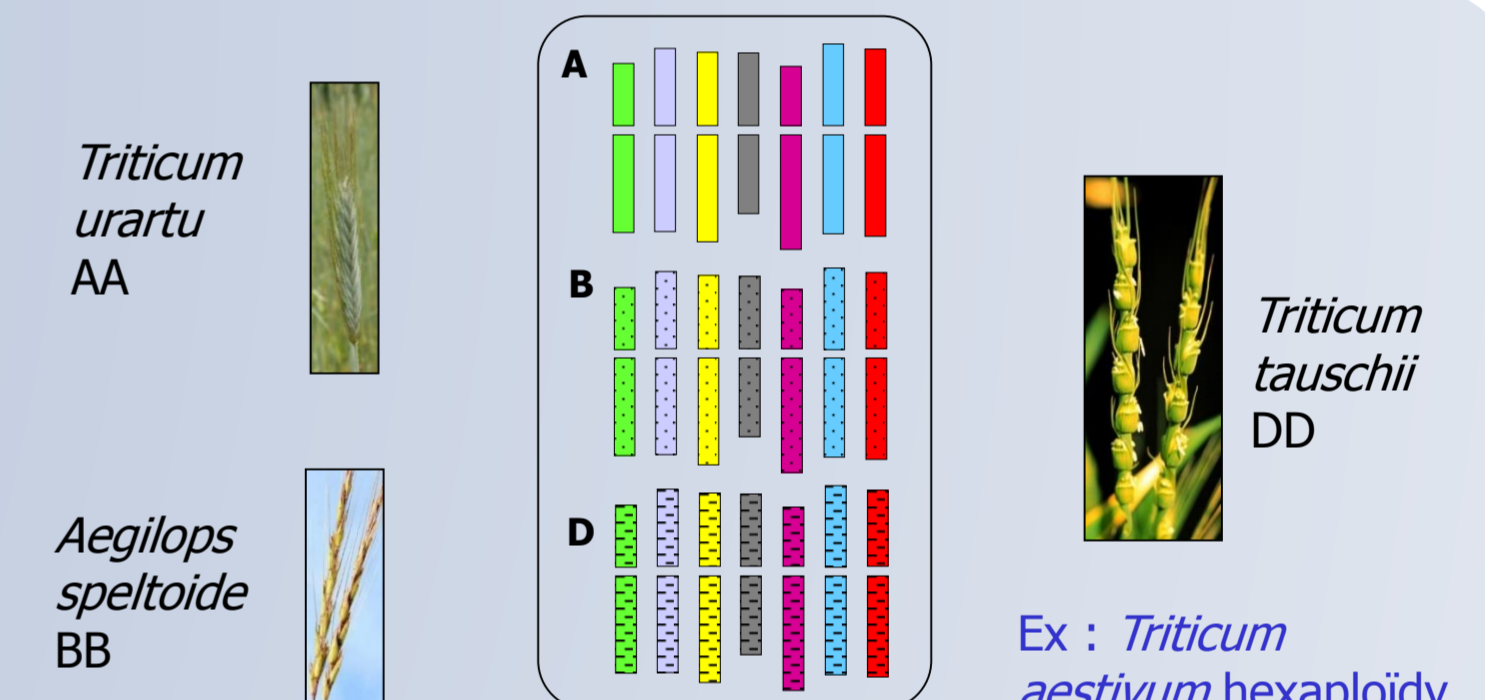
Combining technologies to decipher plant genome complexity

Polyploid genomes

Polyploidy is due to whole genome duplication :

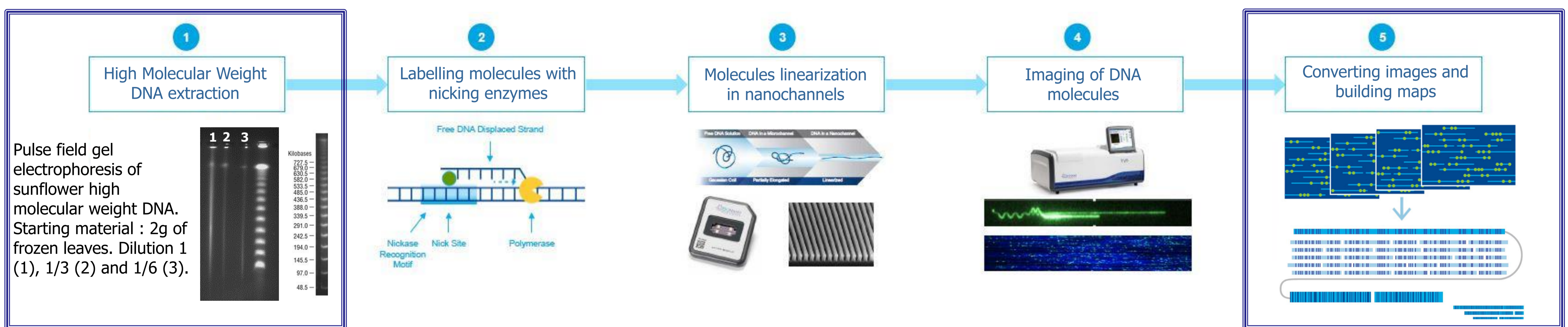
- by doubling a genome of one species (**autopolyploidy**)
- through hybridization of two related species and chromosome doubling (**allopolyploidy**).

Polyploidy is very common in plant species



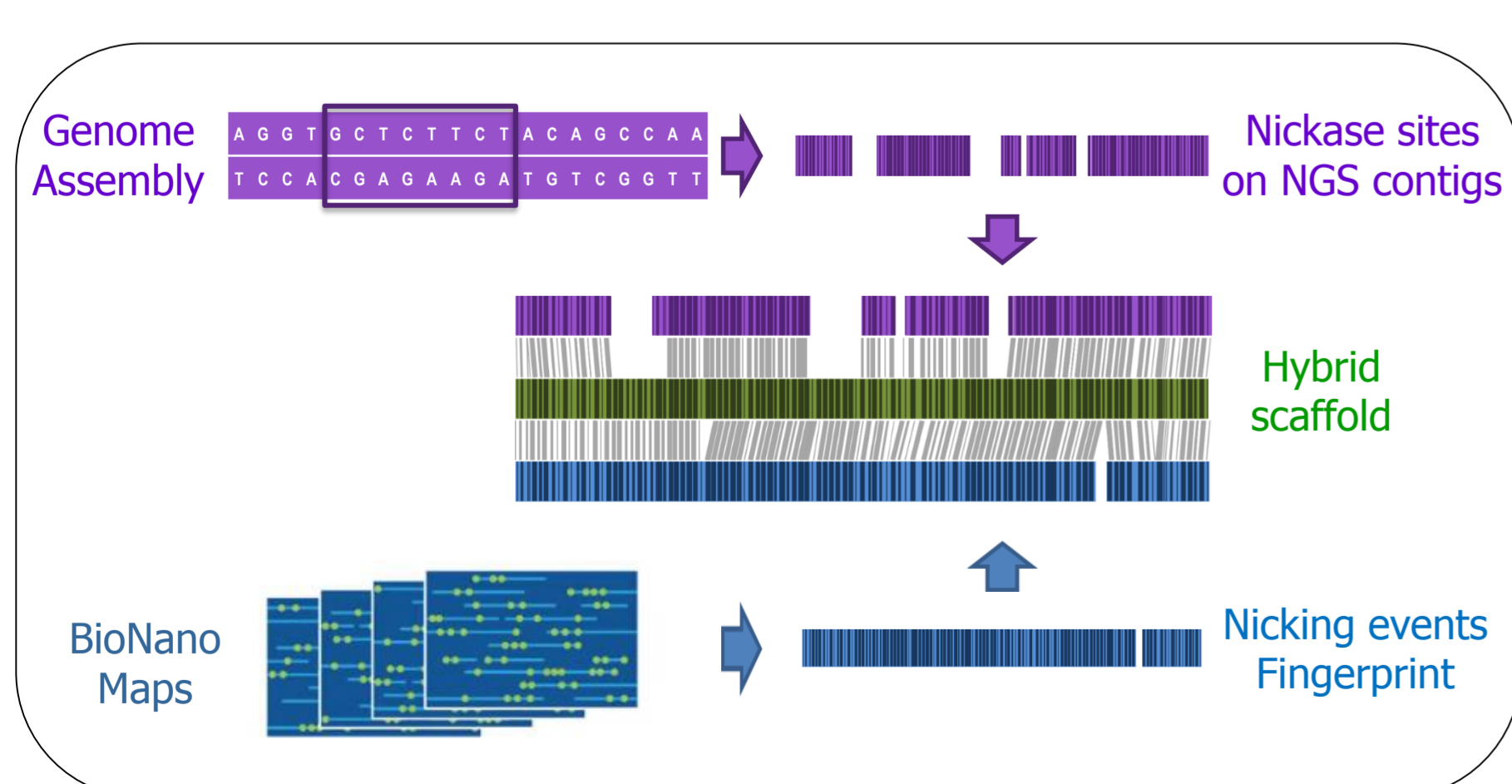
Building Whole Genome Optical Map

Irys System Workflow

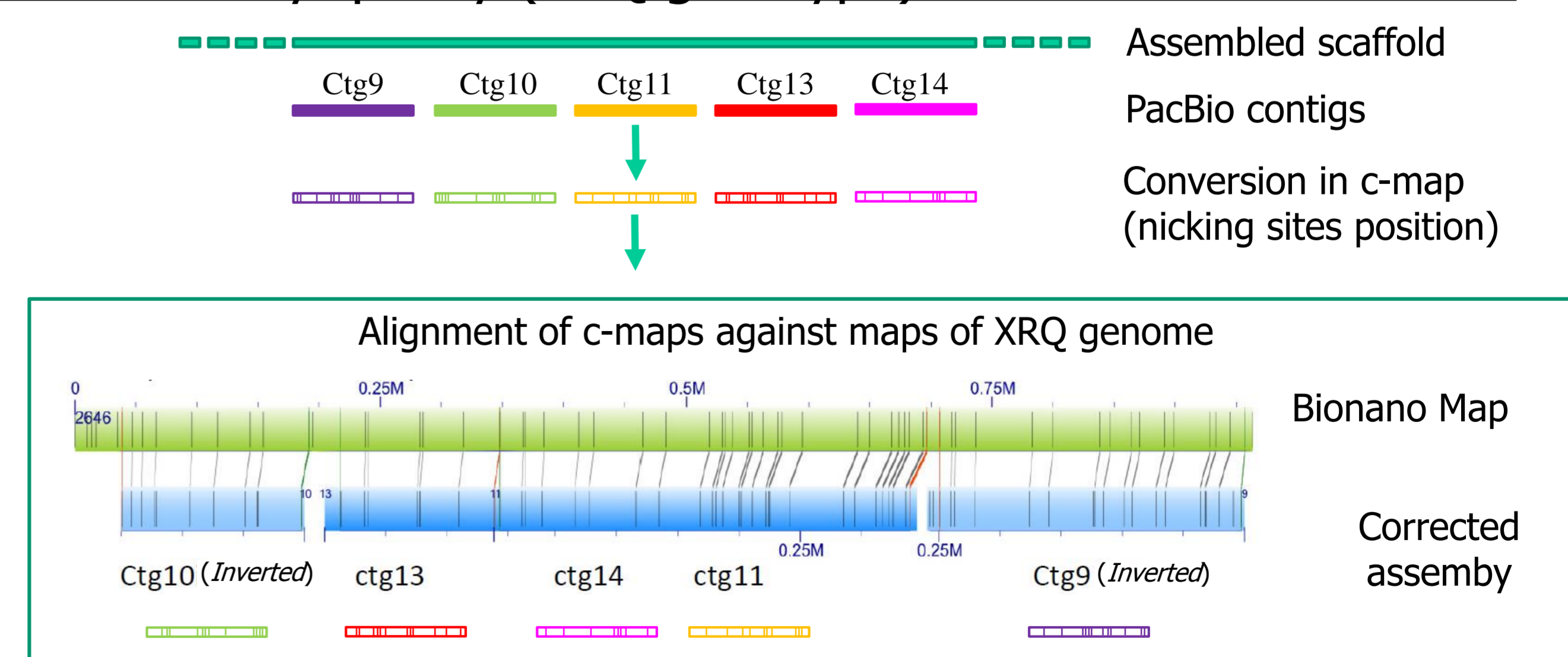


Improving genome assembly and correction of contig ordering through Hybrid Scaffolding

Hybrid scaffolding strategy → Improvement of the sunflower genome assembly quality (XRQ genotype) and conflicts resolution



	PacBio Assembly	Hybrid scaffold 2 Step (2 enzymes)
Number of Scaffolds	12318	1069
N50 length (Mb)	0.524	4.166
Max length (Mb)	3.35	24.670
Total length (Mb)	2930	2960
% genome coverage	81%	82%



Thanks to high quality DNA molecules, optical mapping uses physical reality to link and correct NGS scaffolds. It is now possible to improve assembly quality at the genome level and on targeted regions.