

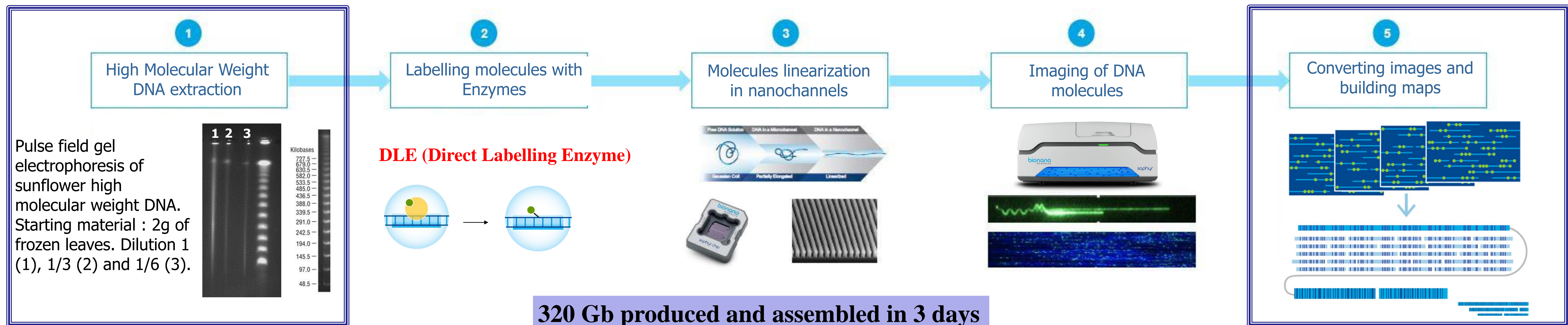
The French Plant Genomic Resources Center (CNRGV) is dedicated to the analysis of plant genome complexity. Indeed, among living organisms, plants display a high level of genome complexity due to their large size, variations in polyploidy levels and high percentage of repetitive elements. In a context of climate change, population growth and limited energy resources, increasing plant genomes knowledge is essential for a better understanding of mechanisms driving plant adaptation and evolution.

With the long reads sequencing technology, such as Pacific Bioscience that produces reads of several kilobases, the genome quality assembly was largely improved with a considerable contig number reduction and a better resolution of repeated sequences localization. However, it remains challenging to obtain scaffold as large as chromosomes.

At the CNRGV, in complement to different NGS technologies, we use whole genome optical maps to improve the genome assembly quality from several plant species. Using the new labelling chemistry and the Saphyr system from Bionano genomics, optical maps were produced to organize genome sequences from several species. This new approach allows us to work at the chromosome level and observe structural variation from kilobases to megabases. In this study, we produced optical maps from 6 sunflower genotypes (landrace and wild type relative). Here, we show the assembly improvement of the reference genome that goes from **0.5Mb to 175Mb scaffold N50**. In addition, we illustrate the comparison of the 6 genotypes at the chromosome level and highlight structural variation in specific genomic region of interest for the tomato genome with a combination of optical maps and either PacBio or 10X genomics sequencing technologies (the results of the combination of the 3 technologies are presented in another poster).

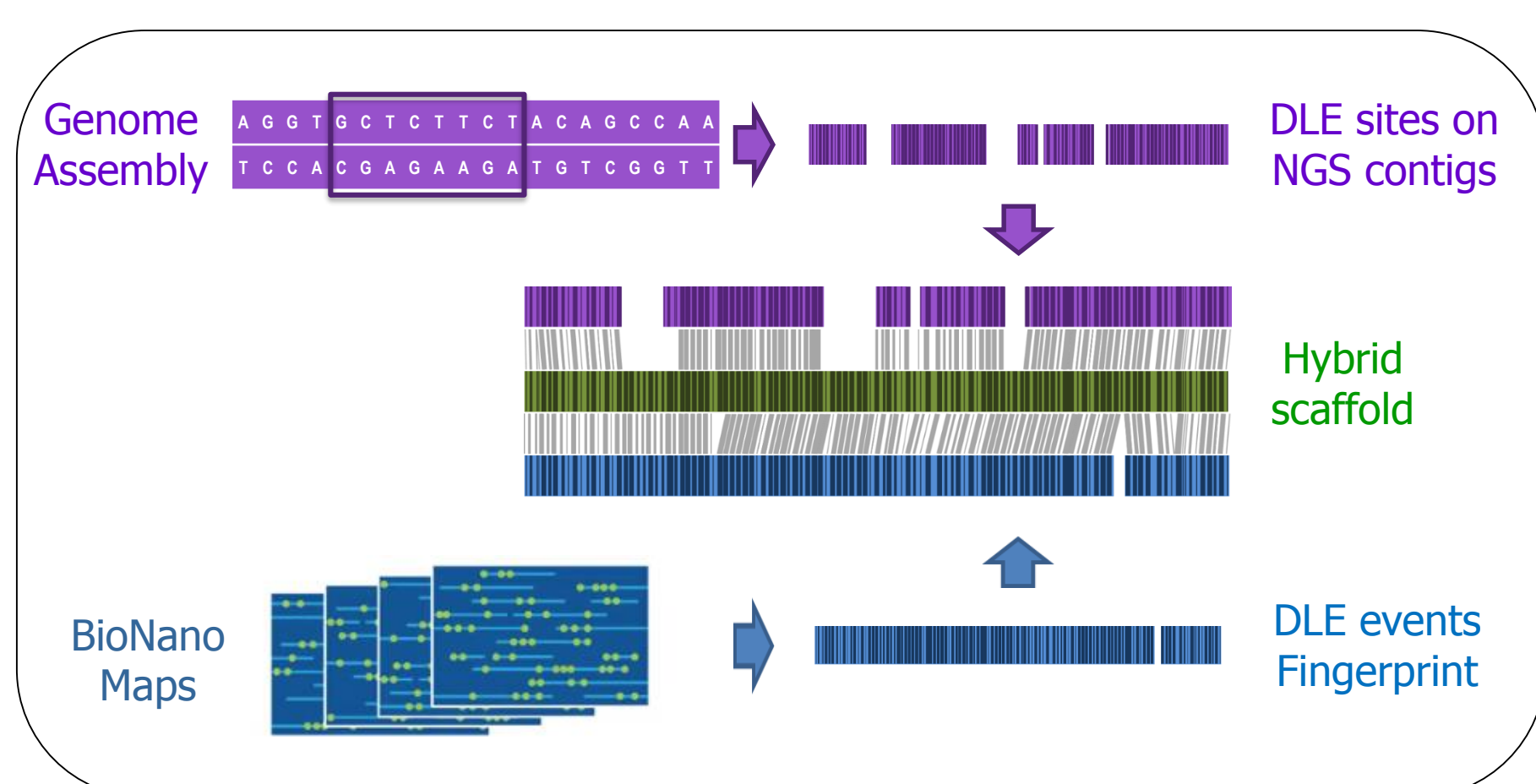
Building Whole Genome Optical Map

Optical map production Workflow



Improving genome assembly until the chromosome level

Hybrid scaffolding strategy → Genome assembly quality improvement of sunflower

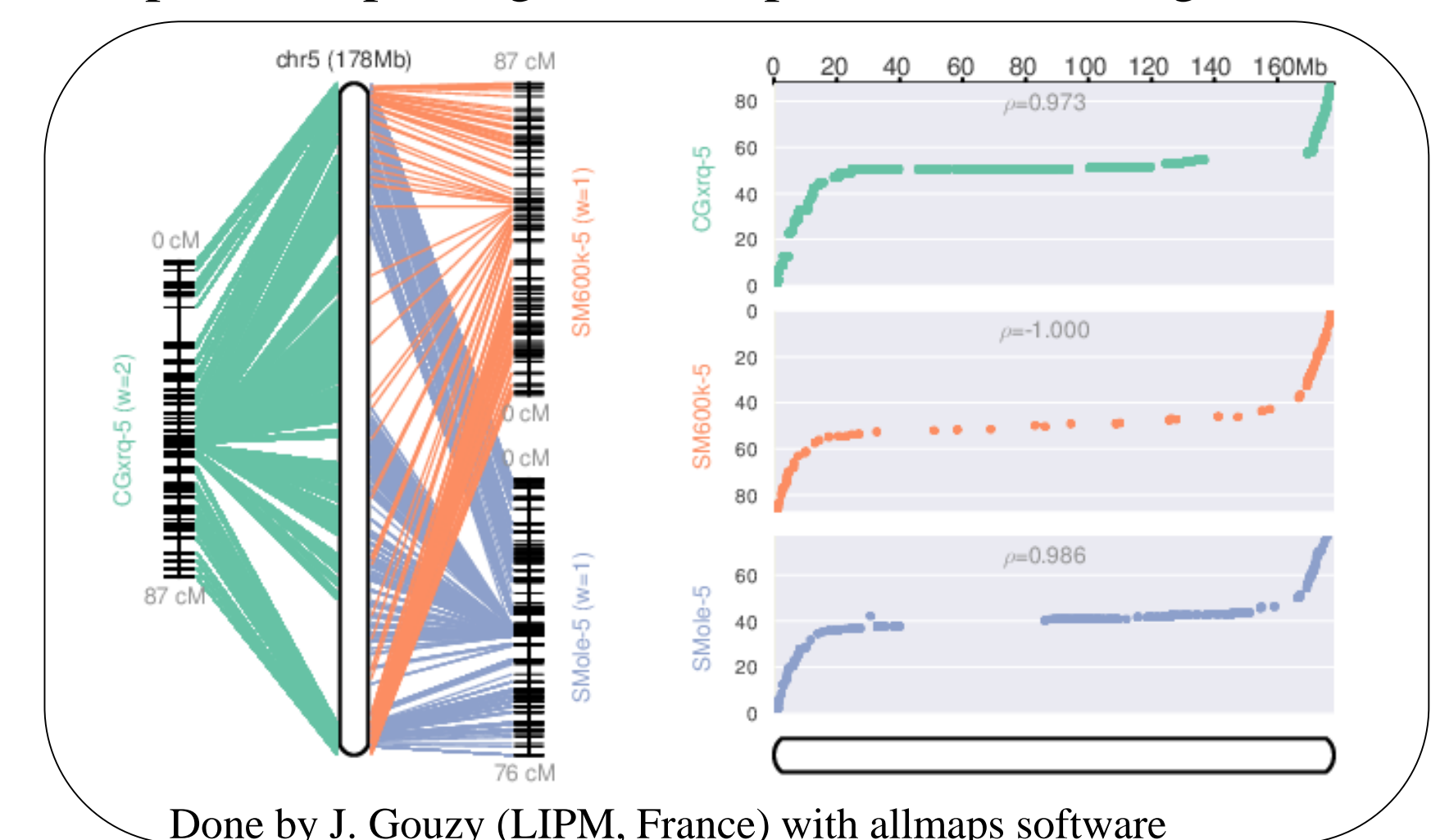


| | Statistic | Original BNG | Original sequence | Sequence used in hybrid scaffold | Hybrid scaffold | Hybrid A + leftover unscattered sequence |
|-------|-------------------|--------------|-------------------|----------------------------------|-----------------|--|
| DLE-1 | Number of maps | 69 | 11676 | 8738 | 25 | 5317 |
| | N50 (Mb) | 175.21 | 0.52 | 0.46 | 176.3 | 175.95 |
| | Total length (Mb) | 3057.67 | 2926.51 | 2792.45 (95.42%) | 3000.44 | 3134.36 |

350 fold

PacBio NGS contigs are organised and oriented according to the optical map during the hybrid scaffolding steps.

Optical map and genetic maps are 100 % congruent



The hybrid assembly produced 25 scaffolds with 17 that correspond to the 17 sunflower chromosomes.

Genomic comparison using optical maps only: sunflower example

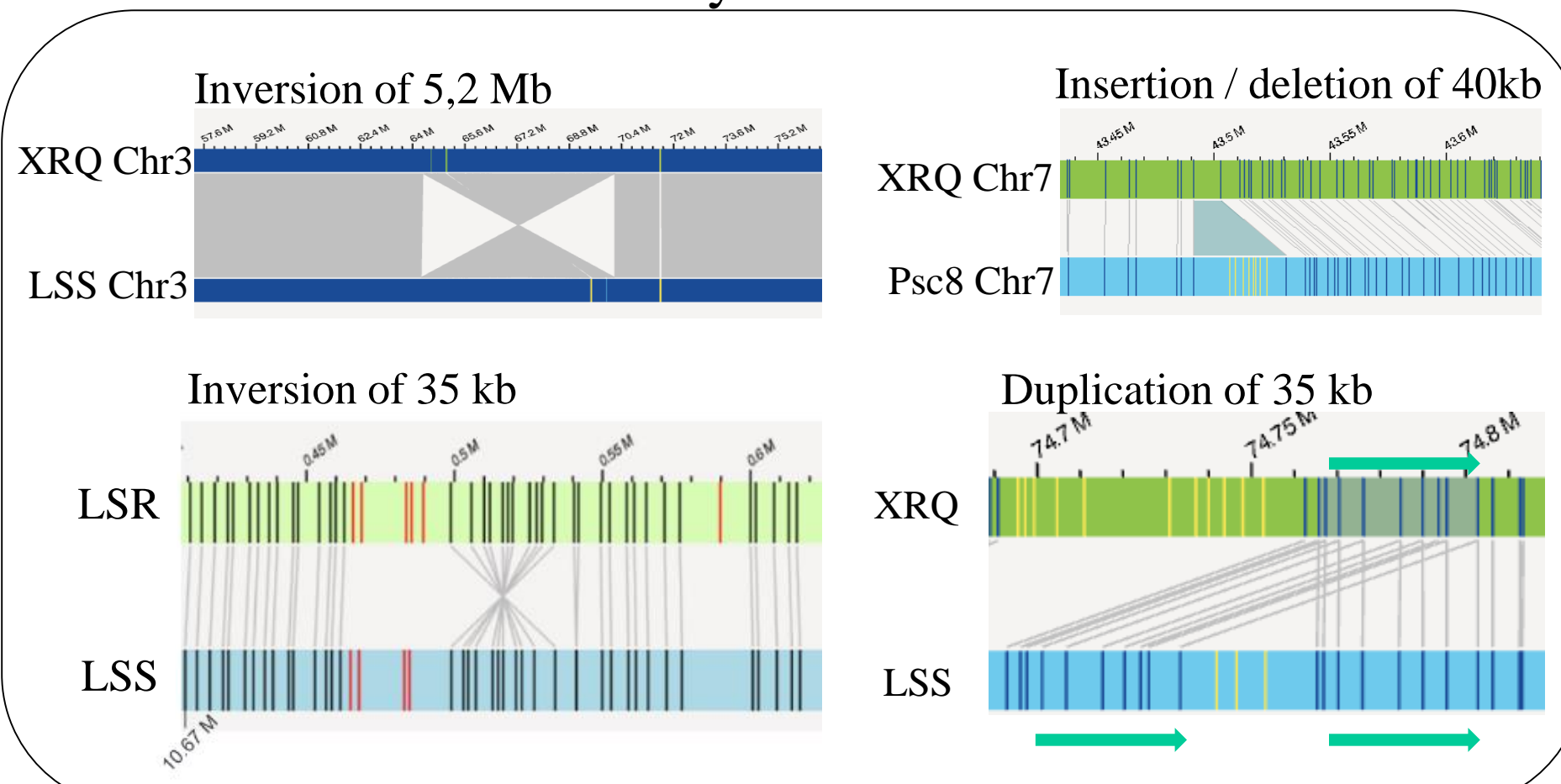
Sunflower genome : **diploid, 17 chromosomes - 3,2 Gb**

Optical map assembly of 6 sunflower genotypes

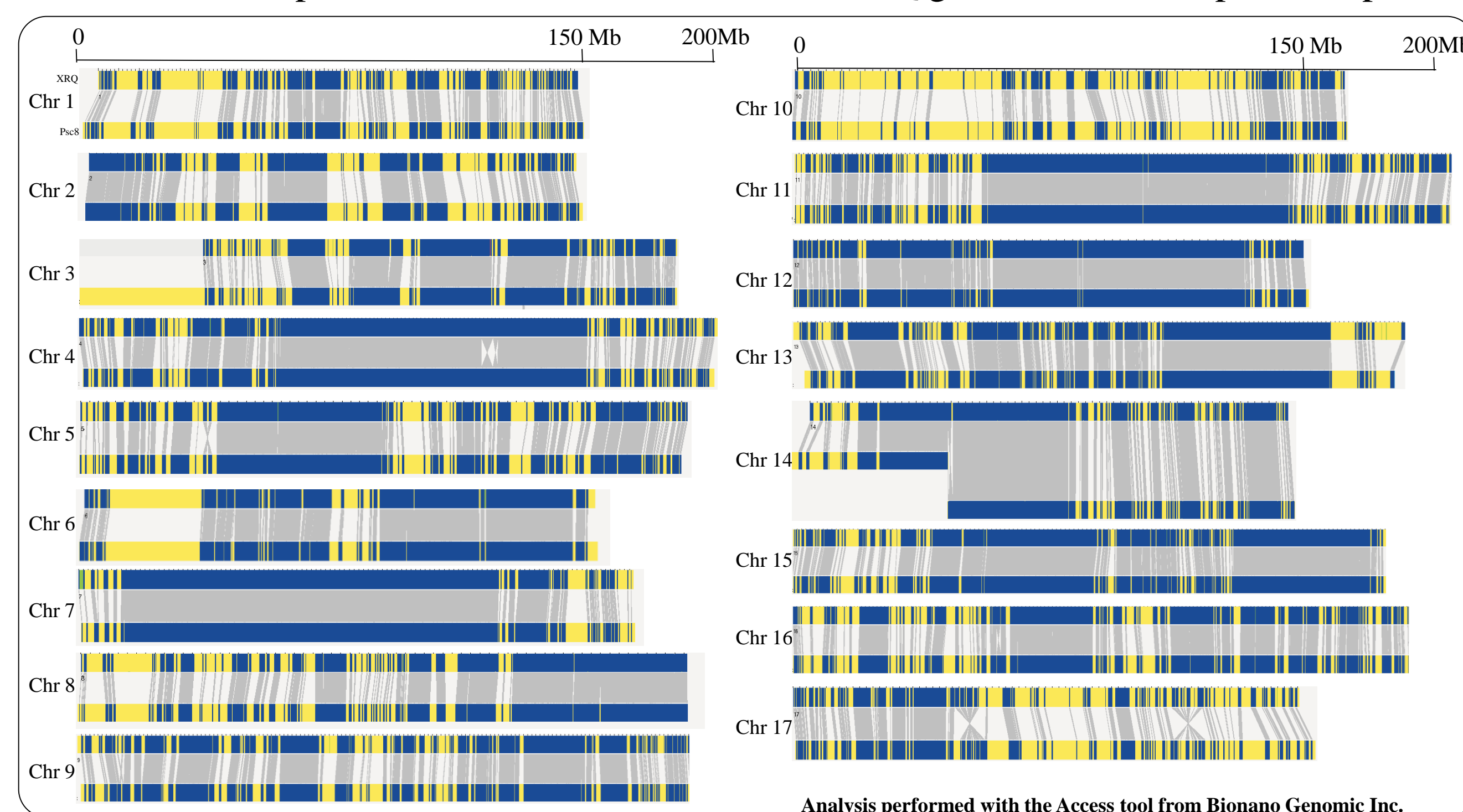
| Genotype | XRQ | LSS | LSR | Arikara | NB1 (Wild) | Psc8 |
|----------------------------------|------|------|------|---------|------------|------|
| Data collected (Gb) ¹ | 310 | 360 | 302 | 304 | 313 | 420 |
| Assembly size (Gb) | 3,06 | 3,05 | 3,04 | 4,8 | 5,43 | 3,03 |
| Genome map N50 (Mb) | 175 | 178 | 179 | 79 | 111 | 179 |
| Count | 69 | 32 | 32 | 304 | 118 | 23 |

¹Data produced in collaboration with Bionano Genomics Inc.

Structural variation analysis from Mb to Kb



Genomic comparison at the chromosome level: XRQ genome vs Psc8 optical maps



With optical maps that reach the chromosome size, it is possible to

- Compare and analyse the chromosome conservation between genotypes
- Analyse structural variation and link it to biological function or phenotypes
- Zoom on region of interest related to specific traits and know their genetic environment (conserved region or not, structural variation observed etc ...)

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 chromosome level and to compare genotypes only by comparing their optical maps.

