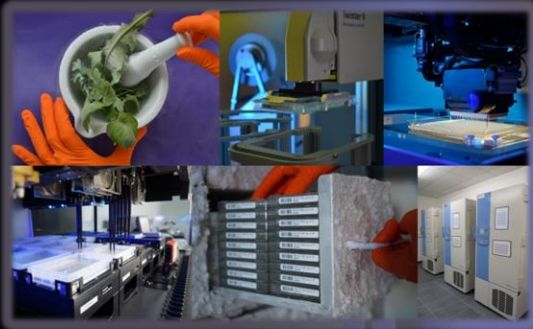




**LA RÉGION** OCCITANIE  
Pyrénées-Méditerranée

# Toward a better understanding of plant genomes: from a global organization to region of interest

Céline CHANTRY-DARMON



# CNRGV

## The French Plant Genomic Center

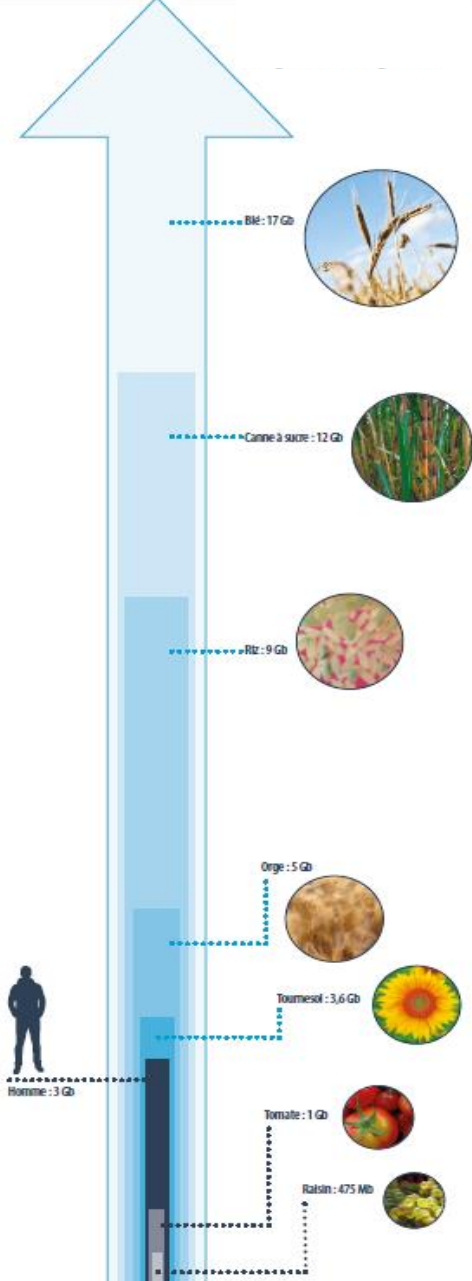


- Created in 2004 by INRA
- Depository of genomic libraries for the scientific community
  - BAC libraries
- A dedicated structure to assist plant genomic programs
  - Distribute the genomic resources at the international level
  - Provide high quality research material and efficient tools and services
  - Develop genomic projects in collaboration
  - Host scientists
  - Develop innovative solutions



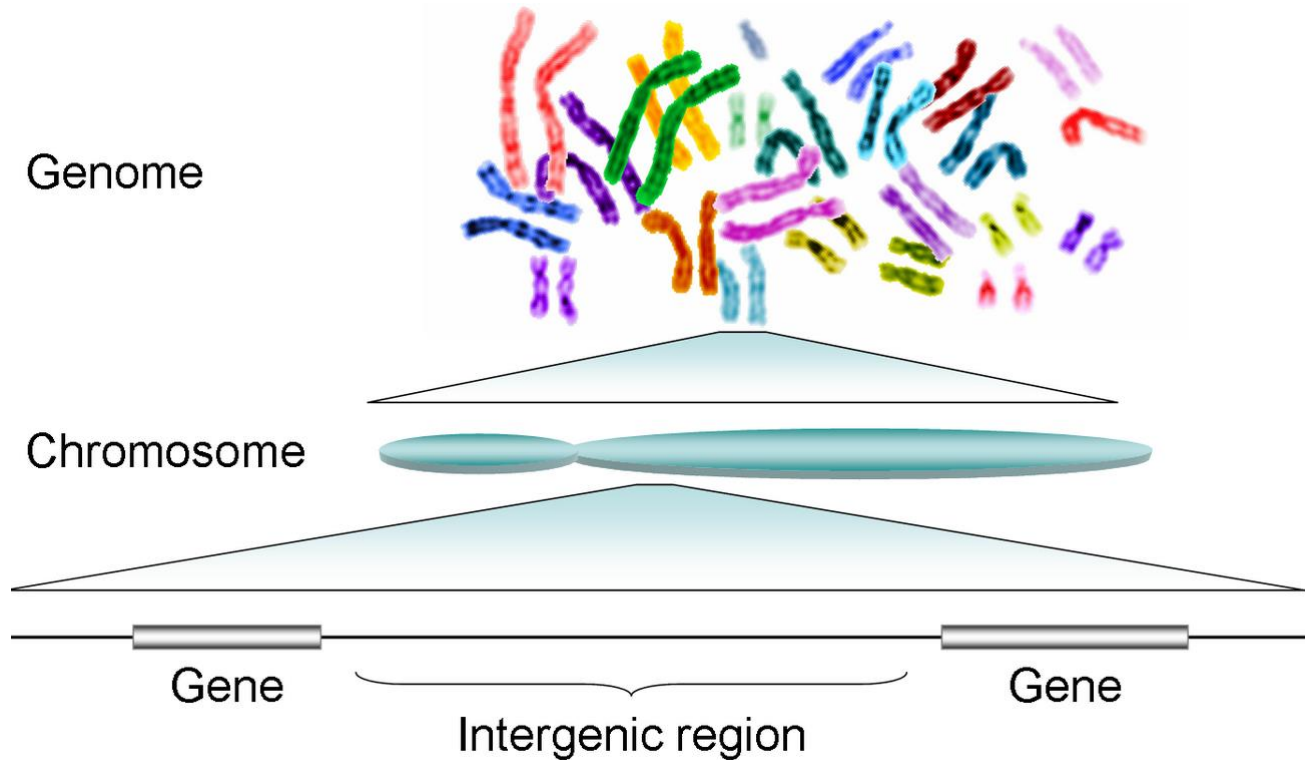
ISO 9001:2008  
Octobre 2005

# The plant genomes complexity



- Genomes size from 130Mb (*Arabidopsis thaliana*) to 150Gb (*Paris japonica*)
- High percentage of transposable elements (>80% in maize, barley, sunflower, wheat...)
- Polyploid genomes (bread wheat 6X, strawberry 8X...)

# The challenge

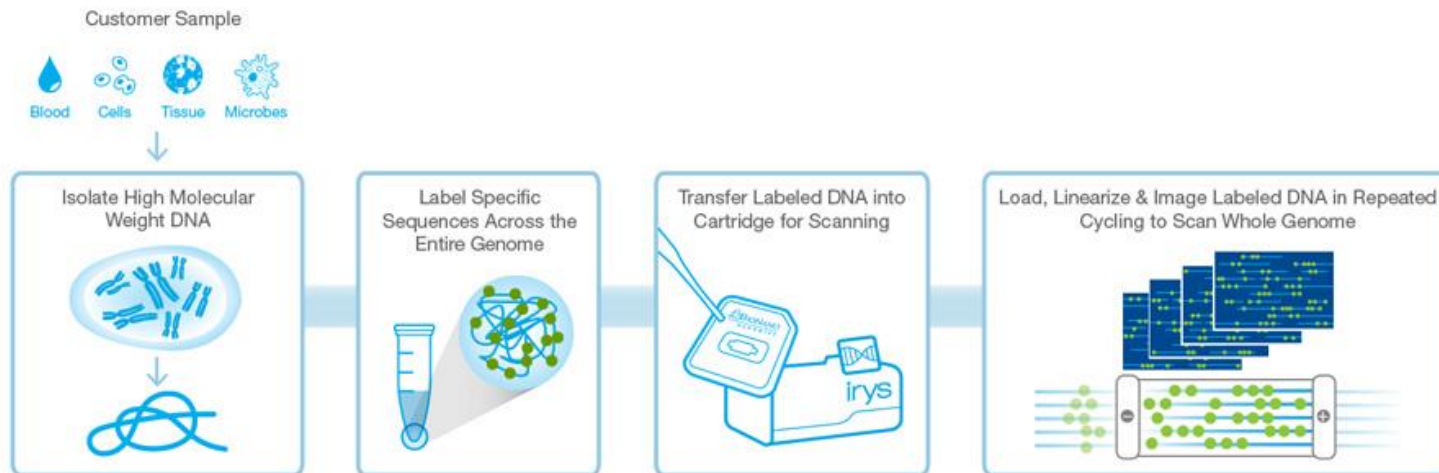


- **Manage genome size and diversity**
- **Decrease genome complexity**
- **Target genomic region of interest**

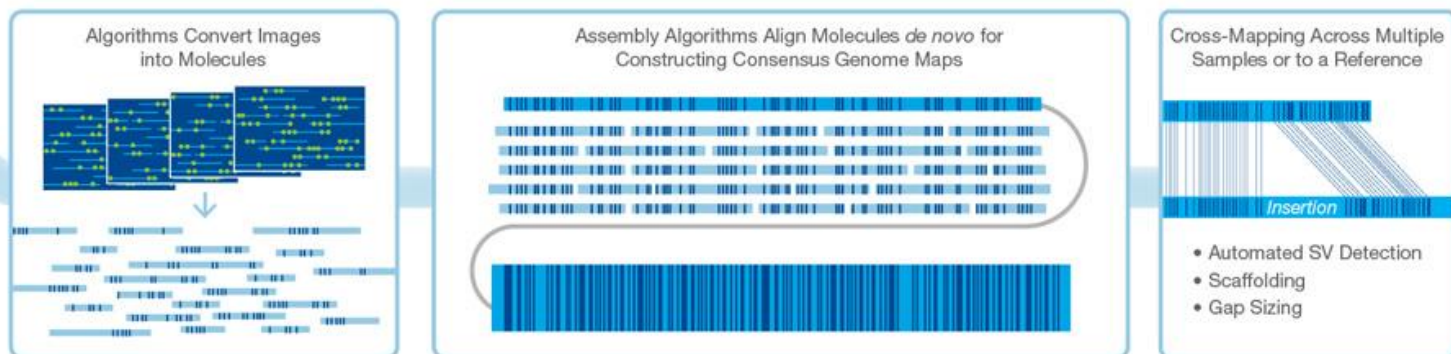


# **PART I: A BETTER UNDERSTANDING OF GLOBAL GENOME ORGANIZATION**

# The BioNano Technology



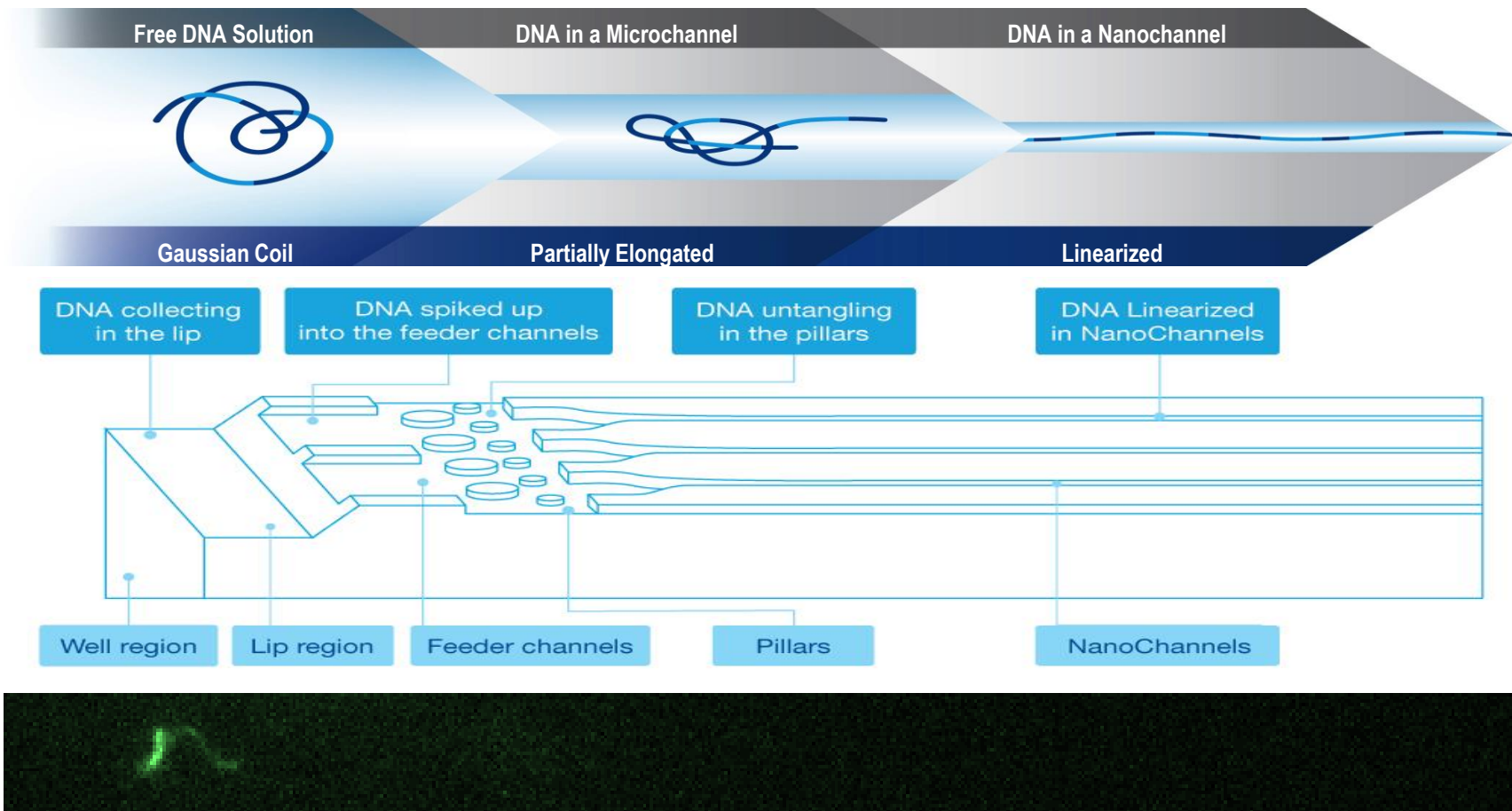
High-Throughput, High-Resolution Imaging Gives Contiguous Reads up to Mb Length



➤ 50Gb data generated per flowcell (=> 100Gb / chip)



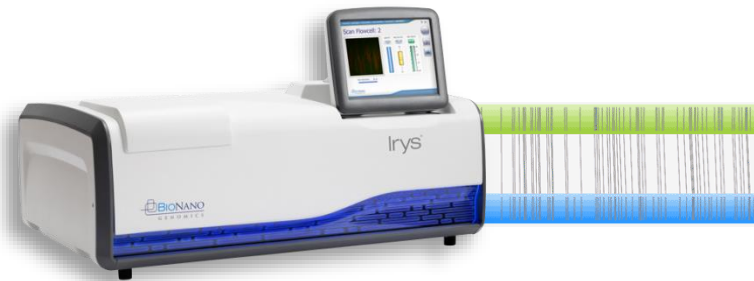
# HMW DNA molecules in nanochannels



➤ HMW DNA molecules from 100kb to >2Mb



# The Next Generation Mapping applications



**For Genome Finishing,  
the maps serve as a scaffold:**

- Sequencing contigs are converted in silico into molecular barcodes by highlighting the same sequence motifs
- These sequencing based barcodes are then aligned to the BioNano maps

**For SV discovery/detection, compare to a reference or gold standard, looking for changes in the patterns:**

- Shifts in barcode patterns reveal insertion (addition), deletion (subtraction), inversion (re-orientation, translocation of genome segments)



# Sunflower Genome Finishing

- Species: *Helianthus annuus*
- 3.6 Gb
- 2n=34 chromosomes
- Genome sequence >100X PacBio

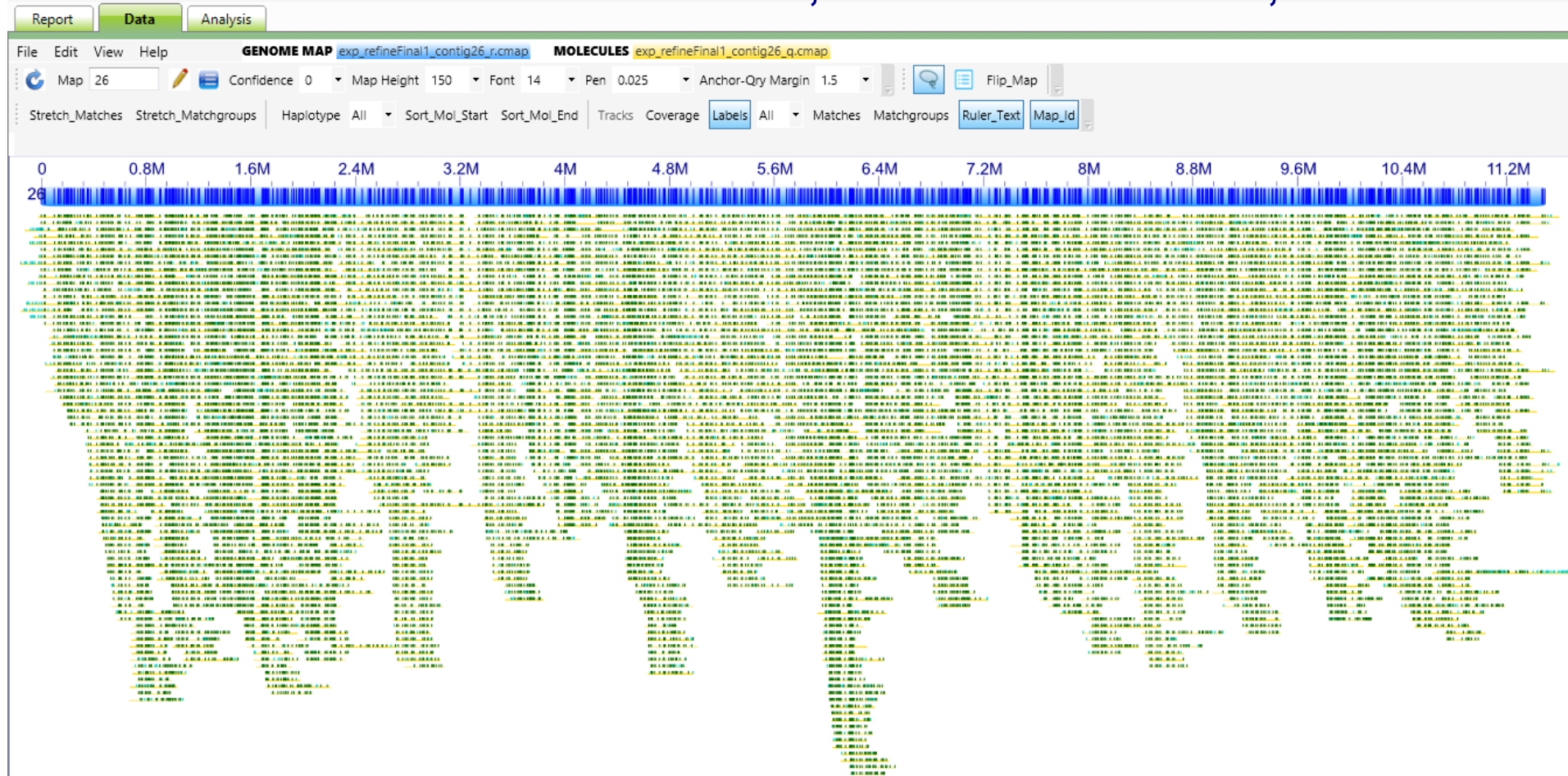
# contigs	LEN Max	N50 BP	#>N50	MEDIAN	BP
12 318	3,35 Mb	524 kb	1 684	120 kb	2,93

=> 80% of the genome inside contigs

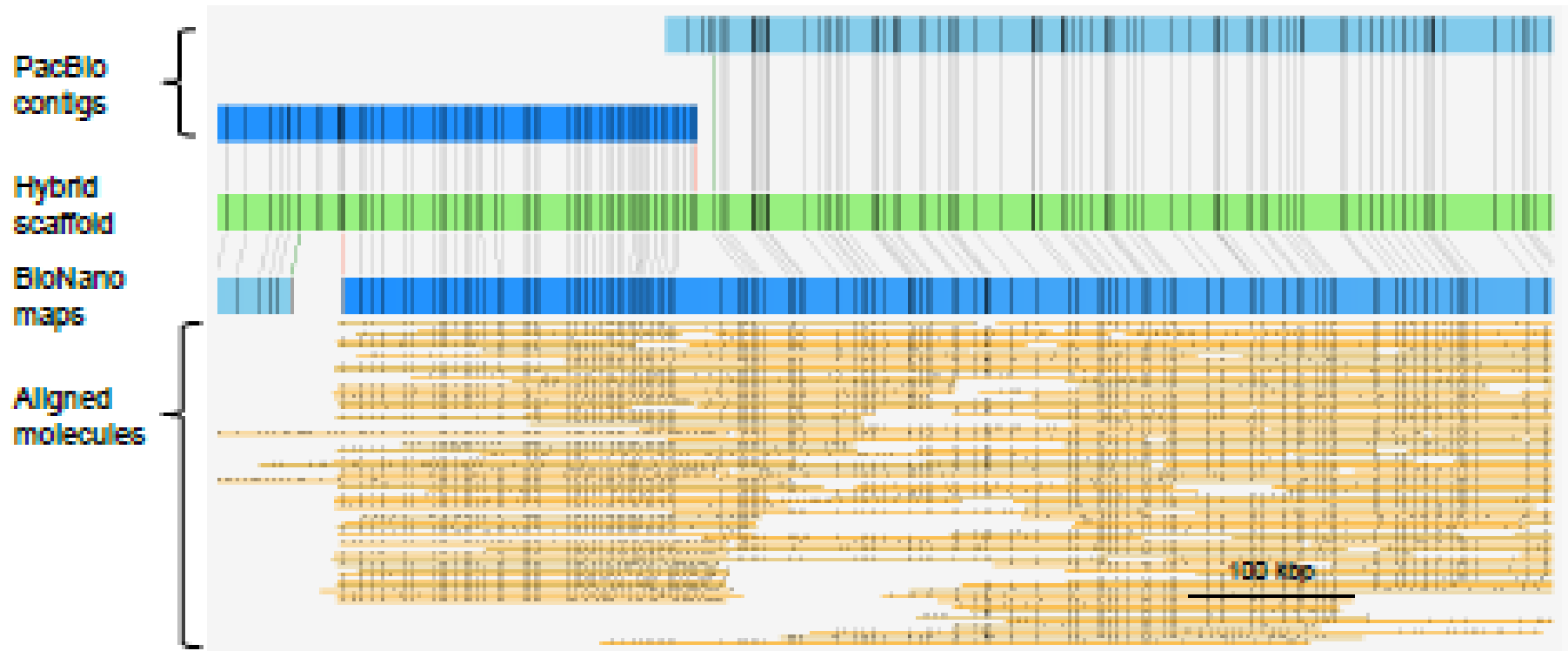
Gouzy et al., 2016

# Optical map of the sunflower genome

=> 175X de couverture, molecules de 150kb à 2,3Mb



# Hybrid scaffolding assembly



# Sunflower genome enhancement

	PacBio Assembly	BioNano Assembly	Hybrid scaffold
Count	12318	2228	1430
Median length (Mb)	0.120	0.999	1.442
N50 length (Mb)	0.524	1.979	2.87
Max length (Mb)	3.35	11.49	17.45
Total length (Mb)	2930	3191	2922
% genome	81%	88%	81%

More than 5 fold  
increase



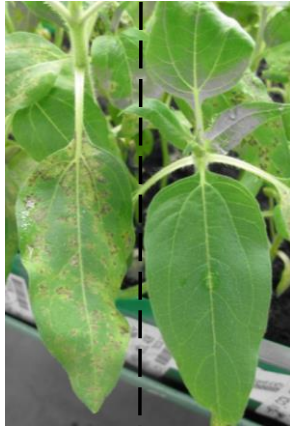
# **PART II: FOCUSING ON SPECIFIC REGIONS WITH NON-GRIDDED BAC LIBRARY**

# QRM1 QTL in Sunflower



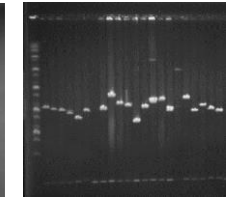
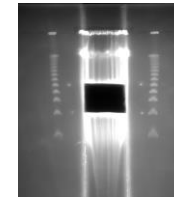
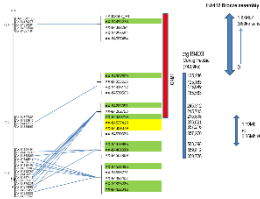
## Phenotypic analysis (232 recombinant plants)

Susceptible | Resistant



- QTL controlling resistance to downy mildew
  - Strong effect on LG10
  - Explain 65% of the phenotypic variability
  - 2 lines:
    - Susceptible (PSC8)
    - Resistant (XRQ)
  - *In silico* physical mapping
- ⇒ reduction of the genetic gap to 0.4 cM

# Non-gridded BAC library construction

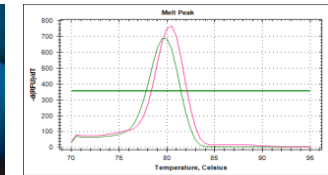
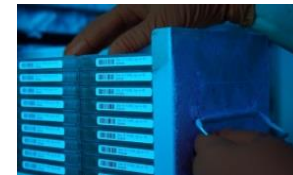


- Establishment of genetic map
- Markers definition on the QMR1 locus

- BAC library construction for PSC8
- Pooling of the transformants



- BAC-Pool Sequencing
- 35 to 100 x coverage
- PacBio Technology :
  - MTP of BACs : 1 contig

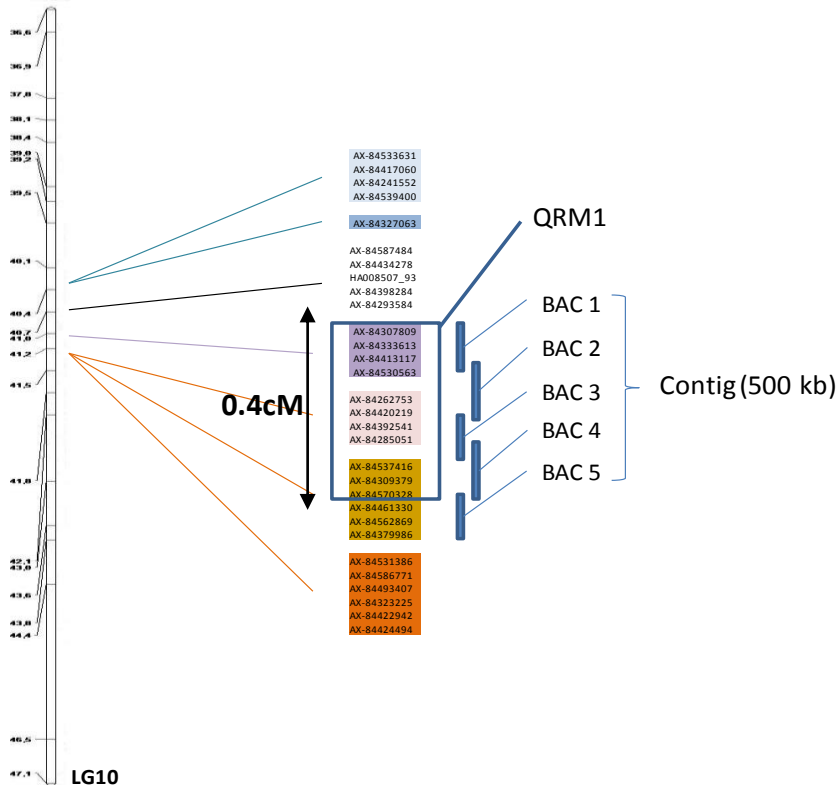


- Screening of pools with specific markers
- Identification of positive BAC clones
- BAC clone characterization



Method derived from Isidor et al., 2005

# Map-based cloning of QRM1



- Sequencing of the 5 BAC clones
- 1 contig of 500kb
- Identification of 2 candidate genes:
  - A MAP Kinase Interacting Kinase
  - A Proteinase inhibitor





## **PART III: FOCUSING ON SPECIFIC REGION WITH SEQUENCE CAPTURE**

ARTICLE

Received 11 Feb 2015 | Accepted 16 Jul 2015 | Published 1 Sep 2015

DOI: 10.1038/ncomms9101

OPEN

# Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters

Wenjun Jiang<sup>1</sup>, Xuejin Zhao<sup>2</sup>, Tslil Gabrieli<sup>3</sup>, Chunbo Lou<sup>2</sup>, Yuval Ebenstein<sup>3</sup> & Ting F. Zhu<sup>1</sup>

PROTOCOL

## Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments

Wenjun Jiang & Ting F Zhu

School of Life Sciences, Center for Synthetic and Systems Biology, Ministry of Education Key Laboratory of Bioinformatics, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Tsinghua University, Beijing, China. Correspondence should be addressed to W.J. (jiangwj12@mails.tsinghua.edu.cn) or T.F.Z. (tzhu@biomed.tsinghua.edu.cn).

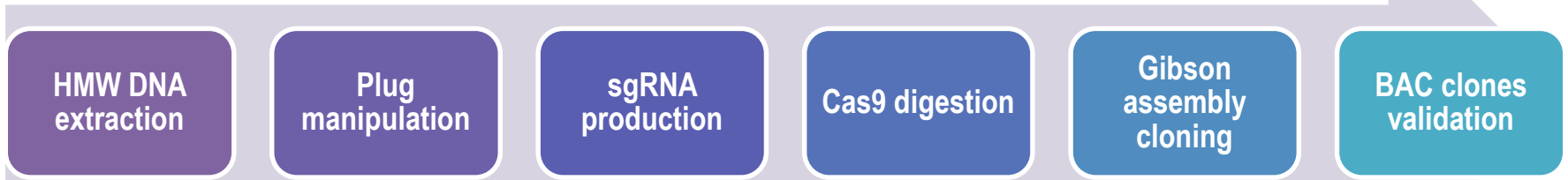
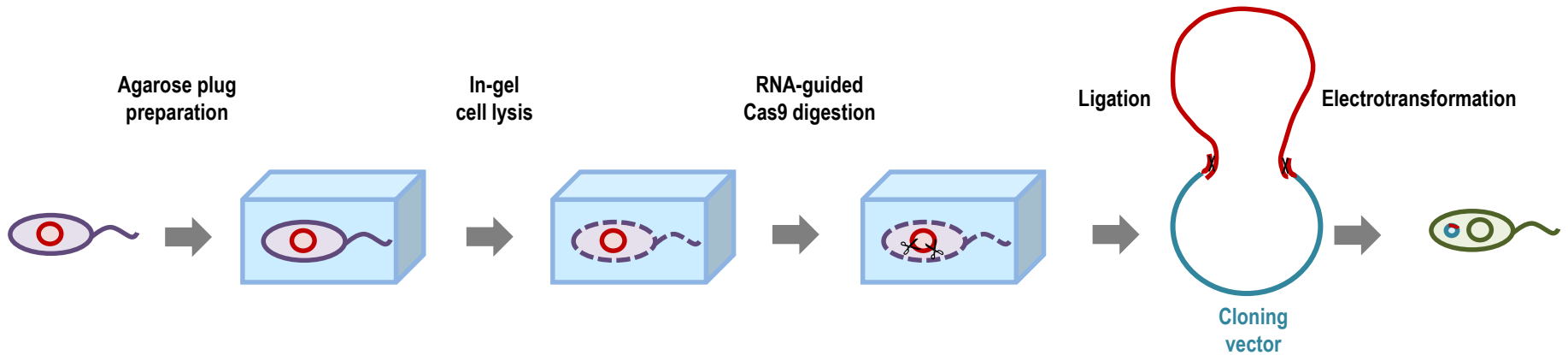
Published online 21 April 2016; doi:10.1038/nprot.2016.055

# CATCH: Cas9-Assisted Targeting of Chromosome segments

- **One step targeting cloning of large genes cluster:**
  - Targeted
  - Fast
  - Specific
- **Based on CRISPR/Cas9 technique**
  - => programmable molecular scissors**

*Jiang and Zhu, 2016*

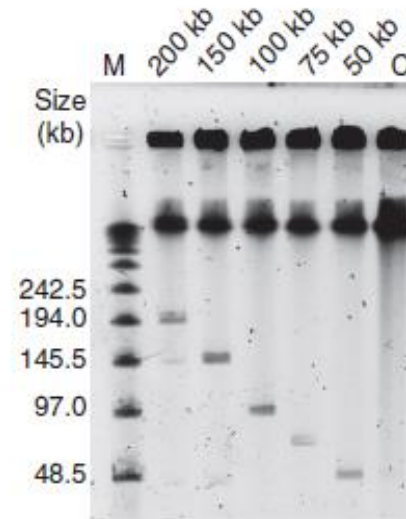
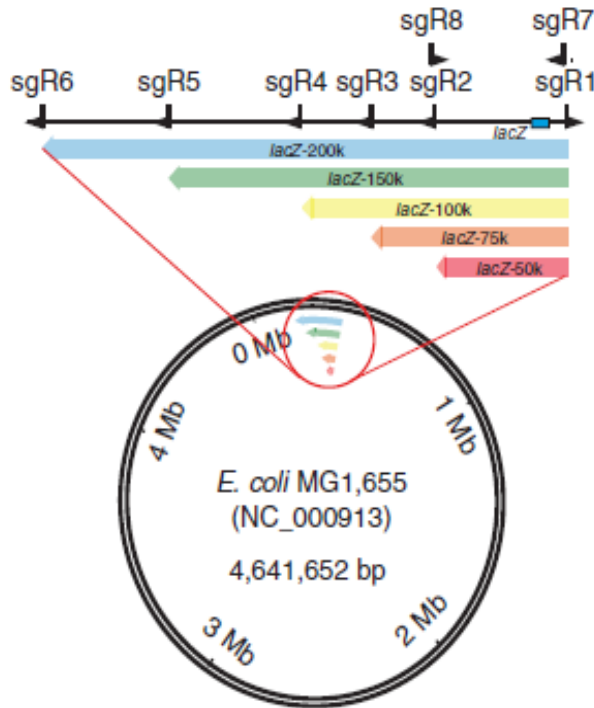
# General workflow



Jiang and Zhu, 2016

# sgRNA design and results

- Isolation of DNA fragment of several size : from 50 to 200kb



- In various bacteria
- Limitations in cloning efficiencies with long DNA fragments

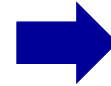
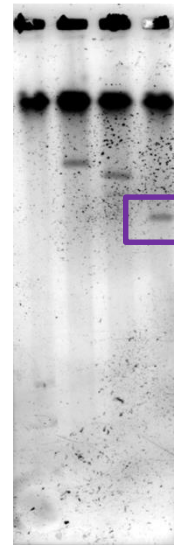
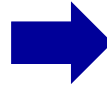
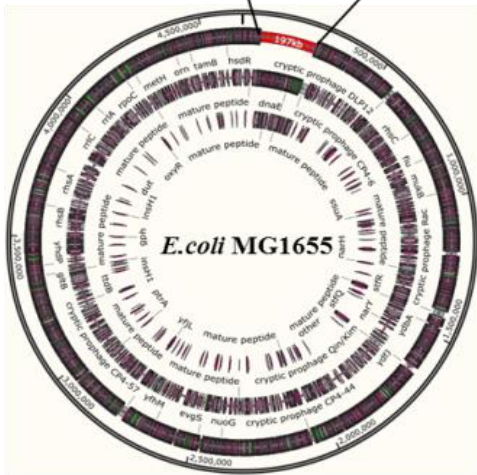
Jiang and Zhu, 2016

# Direct sequencing

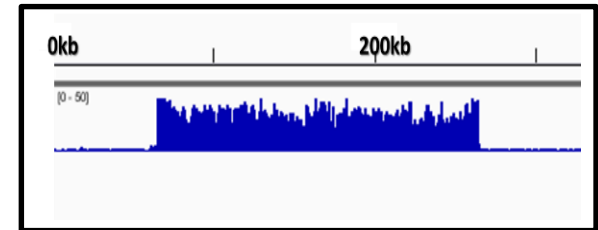
guide RNA1 binding site

197kb

guide RNA2 binding site



Illumina Hiseq2000 deep-sequencing



*Tsilil Gabrieli and Yuval Ebenstein, 2016*

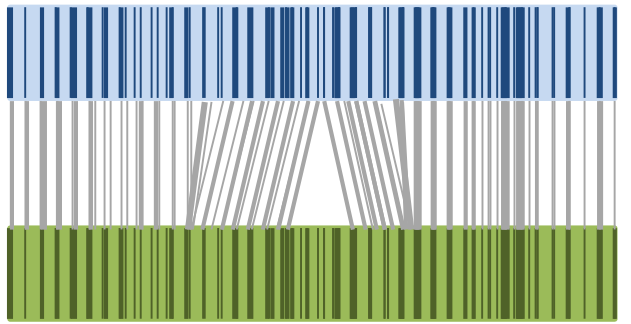
# Choice of a model plant genome

- ***Medicago truncatula***:
  - Small genome size ~ 500 Mb
  - Diploid
  - Complete genome sequence
  - BAC Minimal Tiling Path (MTP)
  - BAC libraries
  - Plants available at LIPM



# I have a dream...

Insertion / Deletion



- The straightest way from phenotype to gene
- From the global to the specific
- Integration of the different tools

GAACTTGCT....ATGGCTTAATCGCC



# Acknowledgements



@CNRGV



PLANT GENOMIC CENTER

<http://cnrgv.toulouse.inra.fr/>

**Hélène BERGES**

Nadège ARNAL

Arnaud BELLEC

Genséric BEYDON

Caroline CALLOT

Stéphane CAUET

Céline CHANTRY-DARMON

Joëlle FOURMENT

Nadine GAUTIER

Laetitia HOARAU

Céline JEZIORSKI

William MARANDE

Elisa PRAT

David PUJOL

Nathalie RODDE

Roseana RODRIGUES

Sonia VAUTRIN



Laboratoire Interactions Plantes Micro-organismes



UNE CULTURE POUR LE FUTUR

Jérôme GOUZY  
Nicolas LANGLADE  
Stéphane MUNOS  
Sébastien CARRERE

Sandra MOREAU  
Pascal GAMAS  
Frédéric DEBELLE



John BAETEN  
Kees-Jan FRANCOIJS



UNION EUROPÉENNE

**LA RÉGION OCCITANIE**  
Pyrénées - Méditerranée