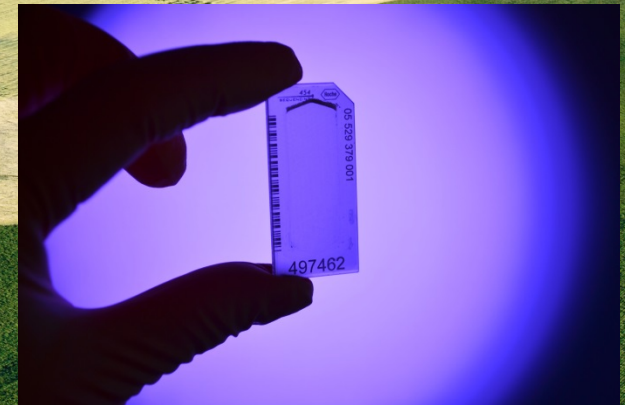
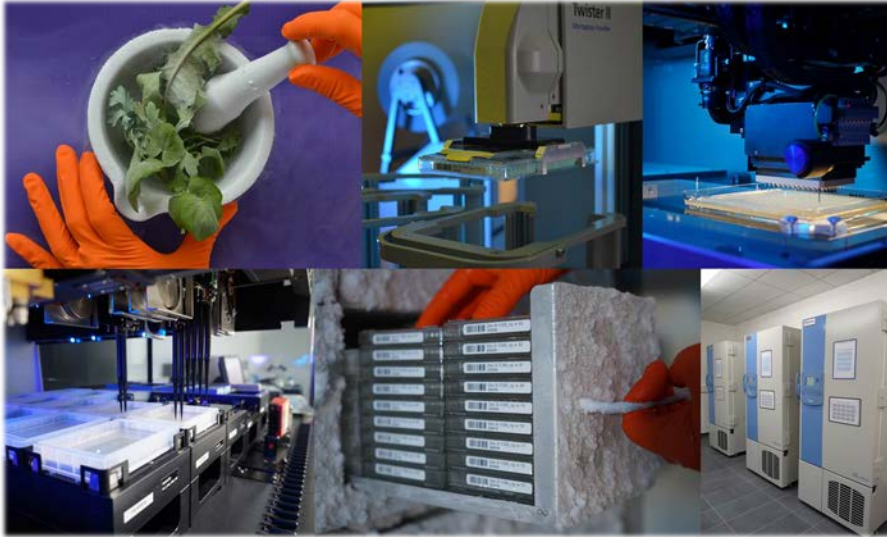


**“From the genome to the field :
how to improve the isolation of genomic regions
of interest for plant breeding.”**



The French Plant Genomic Center

Created in 2004 by INRA (French National Institute for Agricultural Research)



ISO 9001:2008
October 2005

• 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 for studying plant genomes
- ⇒ Develop innovative solutions
- ⇒ Develop genomic projects in collaboration
- ⇒ Host scientists in the frame of collaborations

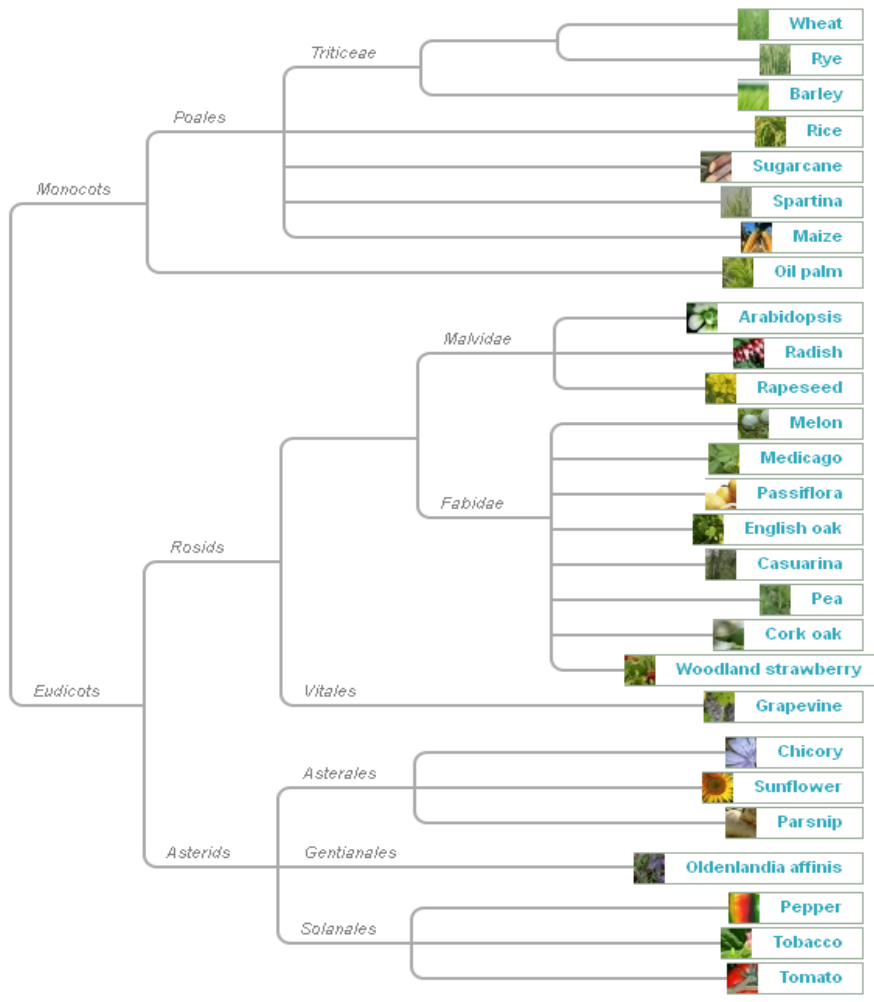
Genomic Libraries available at CNRGV

Library

- > Acyrthosiphon pisum
- > Aphanomyces
- > Arabidopsis
- > Barley
- > Casuarina
- > Chicory
- > Common bean
- > Cork oak
- > English oak
- > Grapevine
- > Hyposoter didymator
- > Maize
- > Medicago
- > Melon
- > Oil palm
- > Oldenlandia affinis
- > Parsnip
- > Passiflora
- > Pea
- > Pepper
- > Radish
- > Ralstonia solanacearum
- > Rapeseed
- > Rice
- > Rye
- > Spartina
- > Spodoptera frugiperda
- > Sugarcane
- > Sunflower
- > Tobacco
- > Tomato

Library

In order to command clones, an on-line ordering system procedure is available on our web site. You have to click on the detailed file associated with each collection. Each biological material sending, require signing a specific **MTA (Material Transfer Agreement)**.



myCNRGV

E-mail:

Password:

Create your account

Retrieve your account



- 157 Genomic Libraries
- 15 M unique clones



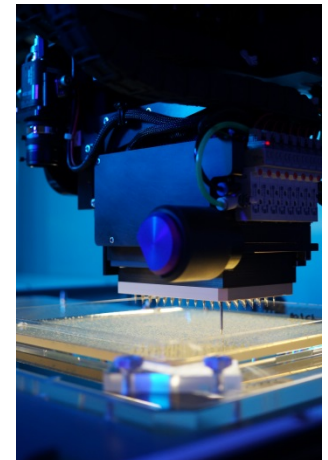
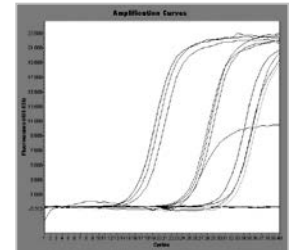
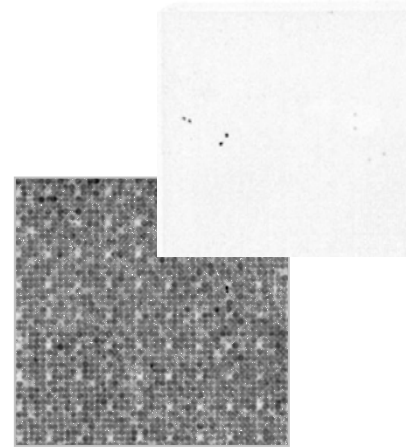
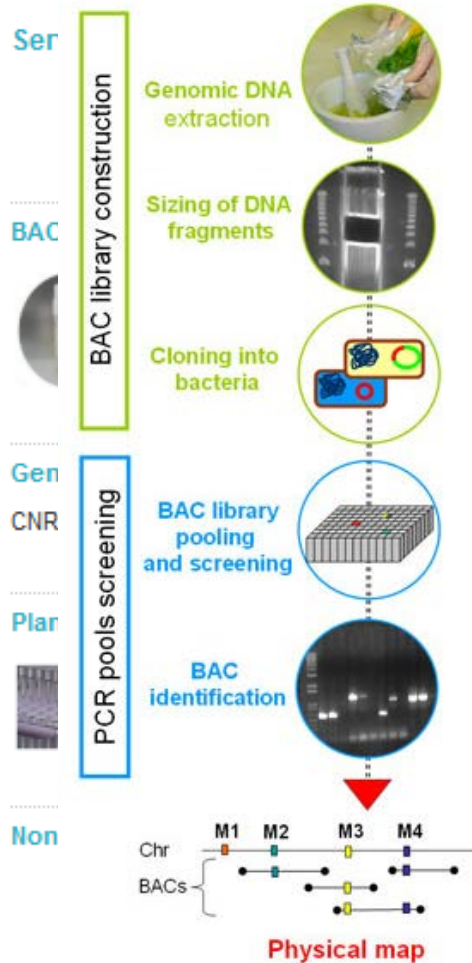
Services at CNRGV

Home Communication Library **Services** Projects

Home > Services

Services

- > BAC, cDNA, ORFeome clone distribution
- > Genomic tools
- > Plant BAC library construction
- > Non gridded BAC library
- > Macroarray production
- > Macroarray screening
- > DNA Pool production
- > PCR Pool screening
- > BAC clone characterization
- > Rearranging
- > Automated colony Picking
- > Others services



Interactions with laboratories around the world



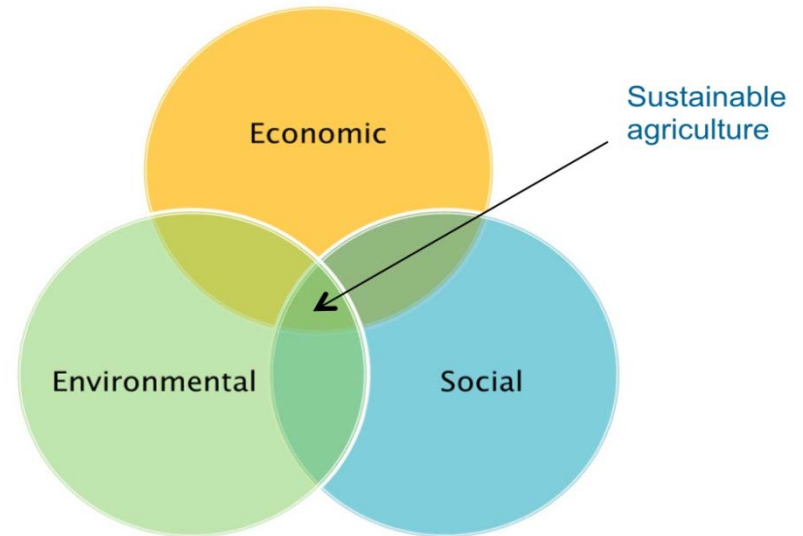
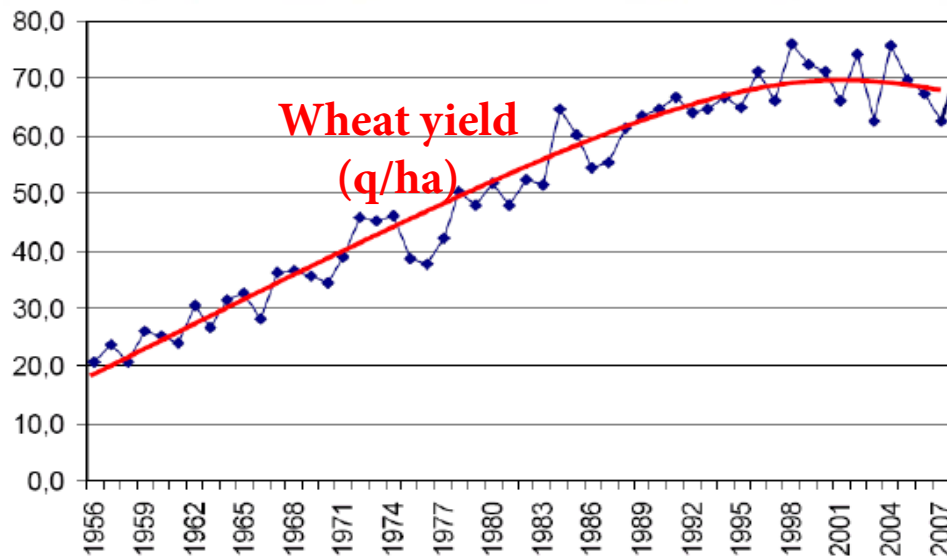
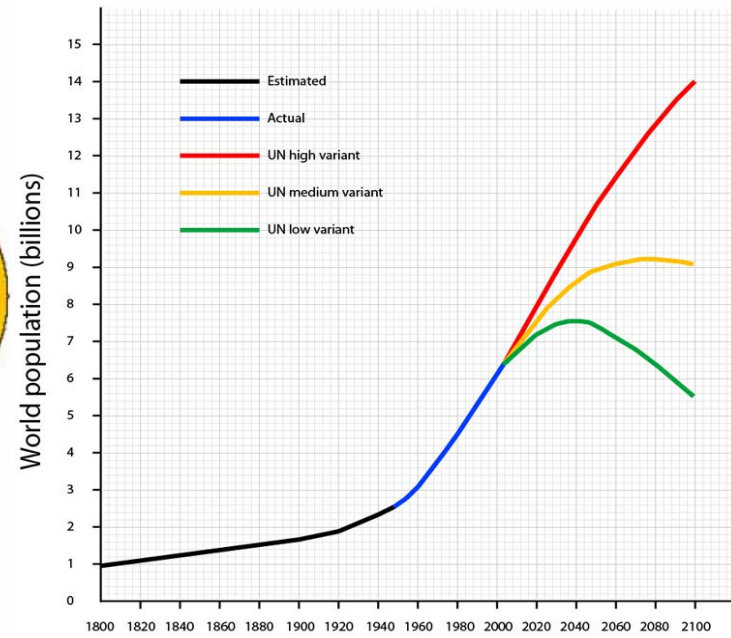
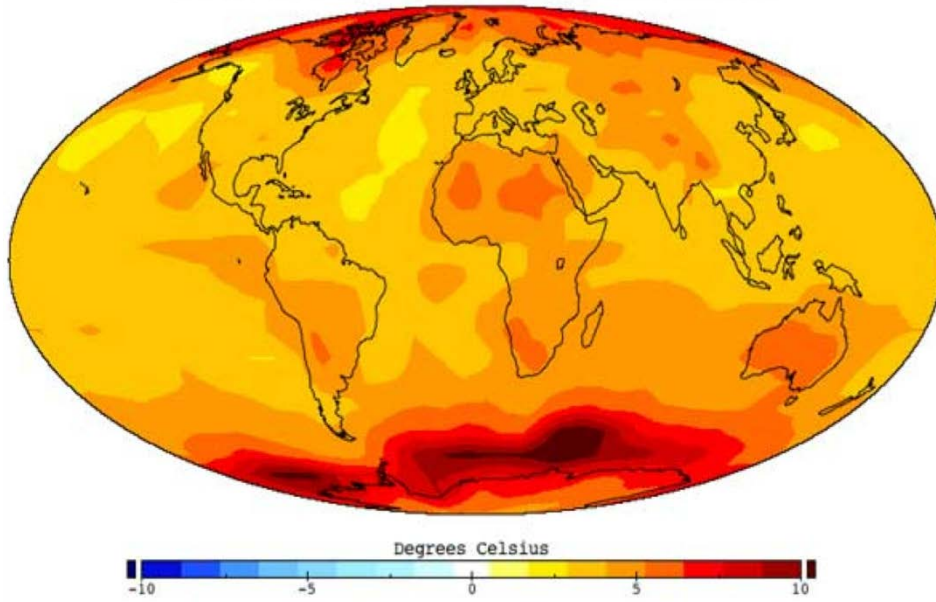
➤ **Since January 2011 : 2 617 240 clones distributed**

International Collaborations

- Member of the International Wheat Genome Sequencing Consortium (IWGSC) and reference centre for wheat genomic resources
- Member of the International Barley Sequencing Consortium (IBSC) and reference centre for barley genomic resources
- Reference center for *Medicago truncatula* genomic resources
- Involvement in the sunflower sequencing project with Canada and USA (Genomics of Sunflower - HELIAGENE)
- Involvement in the sugarcane sequencing project (Sugarcane Genome Sequencing Initiative)
- Involvement in the oil palm sequencing project

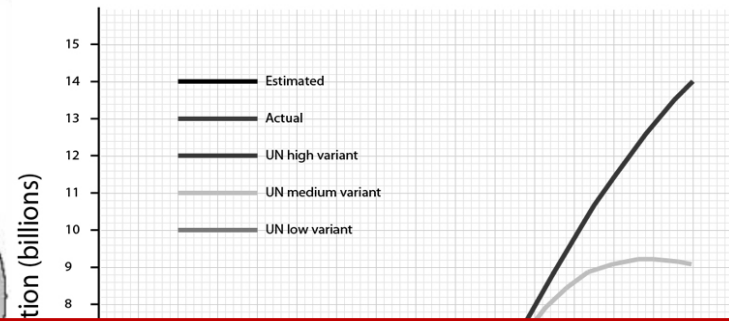
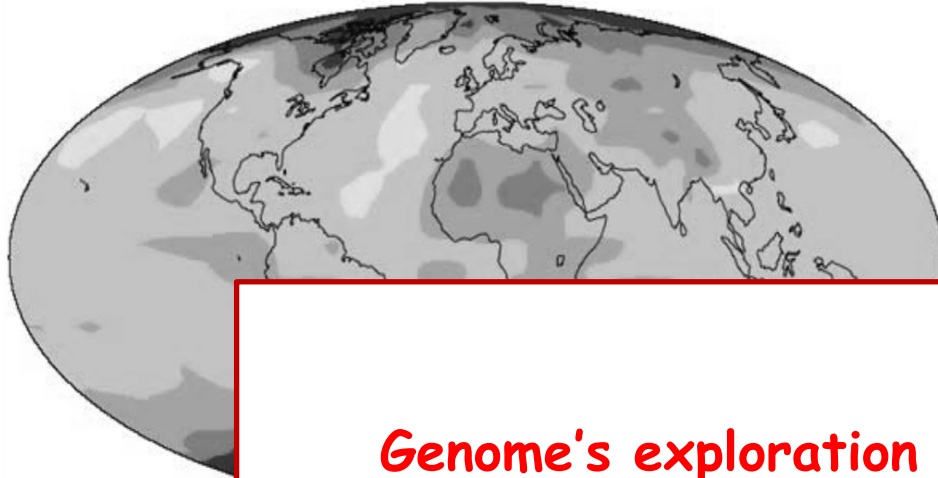
Agriculture must face multiple challenges: Global warming effects, Population growth, Erosion of genetic progress, Consumer expectations

Surface Air Temperature Increase 1960 to 2060

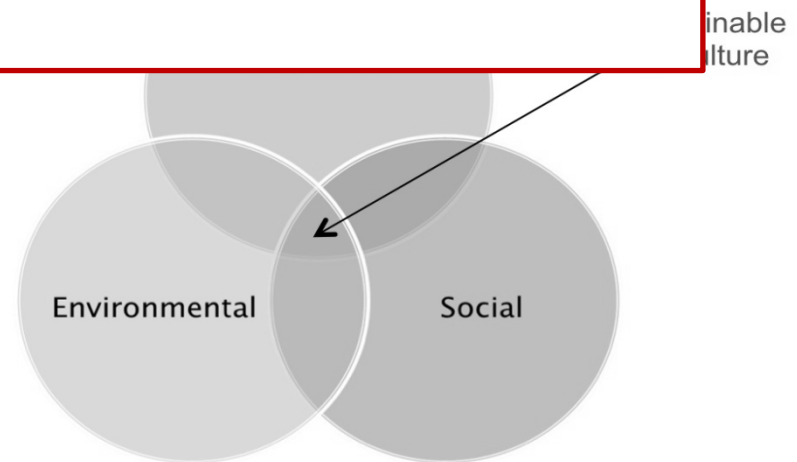
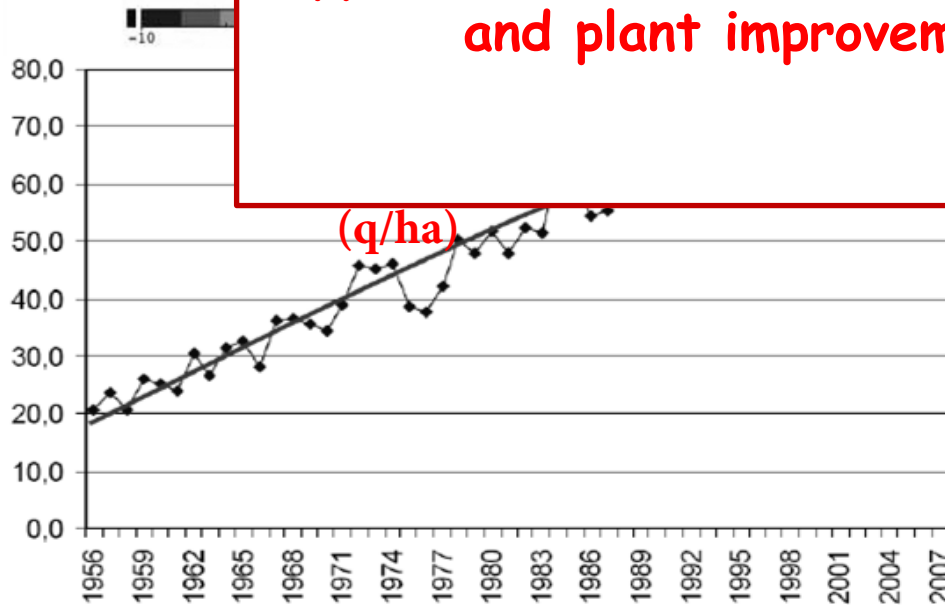


Agriculture must face multiple challenges: Global warming effects, Population growth, Erosion of genetic progress, Consumer expectations

Surface Air Temperature Increase 1960 to 2060



Genome's exploration is one of the strategic approaches for better understanding plant evolution and plant improvement and adaptation



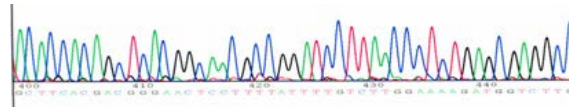
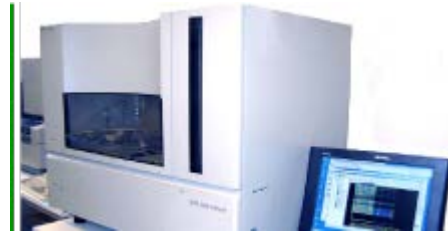
What do we expect from genomics?

From the plant



The objectives

- Genes inventory
- Genes Localisation



- Sequencing
- Function

The applications

Construction of elites genotypes by:

- producing molecular markers
- Identifying candidate genes for QTL analysis
- Identifying favorable alleles

Various targets for crop improvement



➤ Yield potential and yield stability

- ✓ Photosynthesis efficiency
- ✓ Harvest Index
- ✓ Reduced inputs (fertilizers, pesticides..)



➤ Adaptation to climate change

- ✓ Avoidance
- ✓ Tolerance (Drought, heat, cold..)
- ✓ Post stress recovery...



➤ Durable resistance to biotic stress

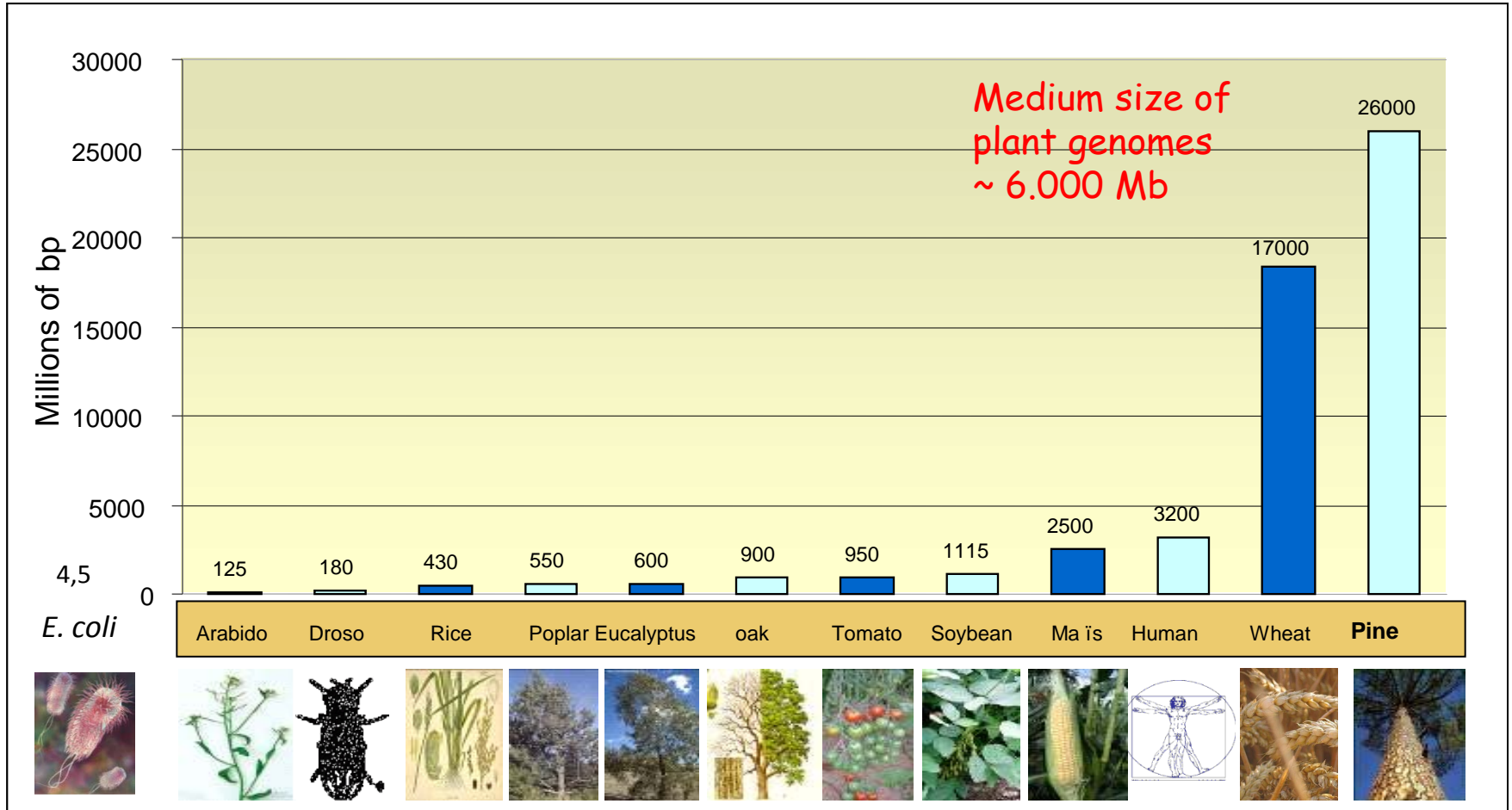
- ✓ Usual suspects (virus, fungi)
- ✓ new pests and diseases
- ✓ invasive species



➤ Quality of grain and coproducts

- ✓ Grain protein content, malting quality, foam stability
- ✓ Starch, straw
- ✓ Food safety -Allergenicity-mycotoxin contamination..

Genome exhibits high levels of diversity



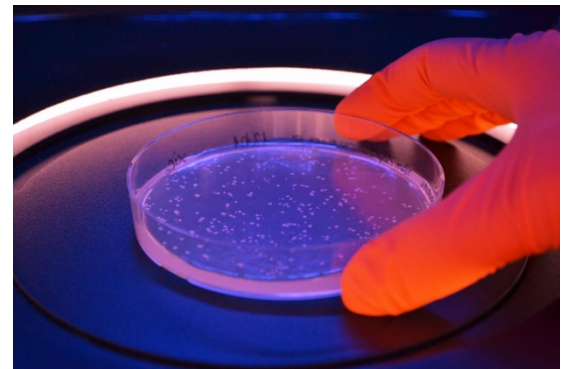
The size of a genome doesn't reflect the complexity of the organism

- Repeat elements (>80 % in maize, barley, sunflower, wheat etc ...)
- Polyploidy

Sequencing strategies for complex genomes

Two main sequencing strategies :

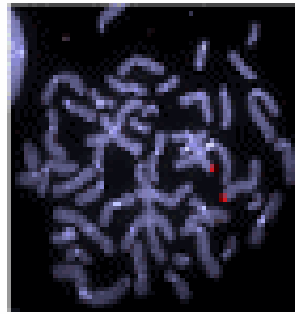
- **Using genomic libraries (BAC)**
- **Whole Genome Shotgun**



Genome sequencing strategies

BAC by BAC

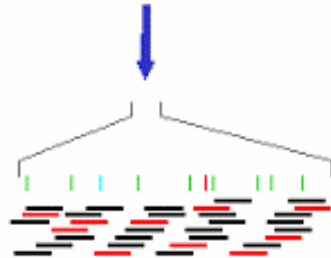
SHOTGUN



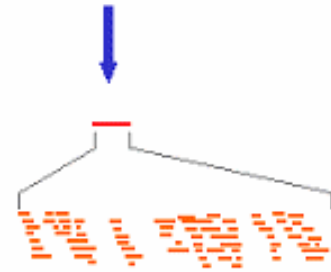
1- Construction of BAC libraries



2- Physical map



3- **MTP** selection



4- BACs sequencing



1- Genome fragmentation



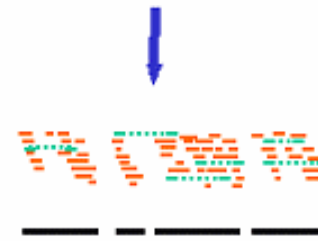
2- paired end Sequencing



3- Sequences Assembling by contigs

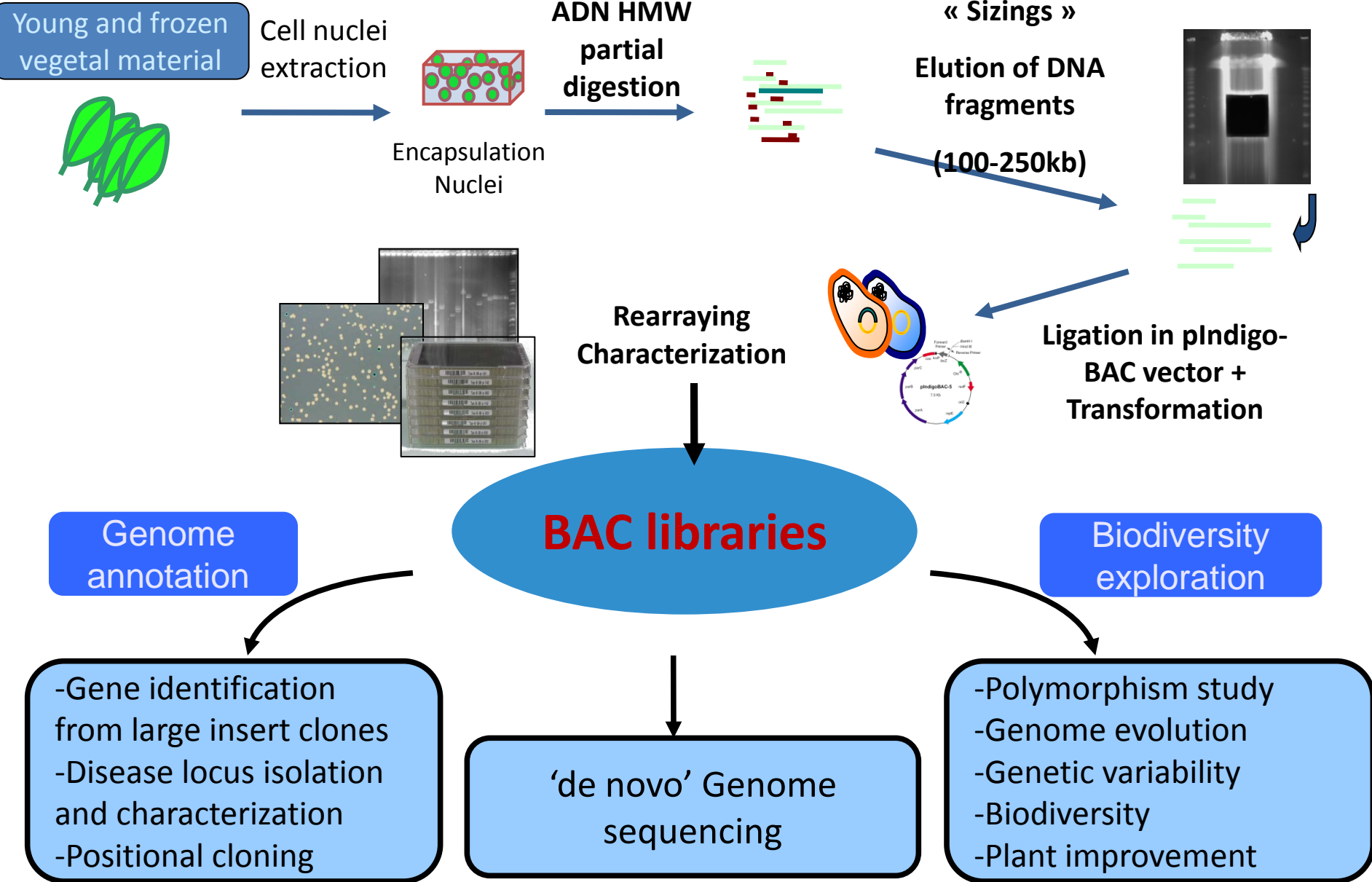


4- Combined data to assemble contigs into scaffolds



BAC libraries :

the starting point for various scientific questions



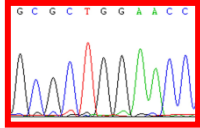
Plant Genomes available

Name	Genome size (Mb)	Ploïdie	Sequencing strategy-coverage	Year
<i>Arabidopsis thaliana</i>	125	Diploïde	BAC – 15x	2000
Rice (<i>Oryza sativa ssp. Japonica</i>)	389	Diploïde	BAC - 10 x	2002
<i>Sorghum bicolor</i>	770	Diploïde	WGS – 8,5x	2009
Grapevine (Pinot noir cv PN40024)	487	Dihaploïde	WGS – 8,5x	2007
Soybean	1 100	Diploïde	WGS – 8x	2010
Brachypodium	300	Diploïde	WGS – 8x	2010
Poplar	485	Diploïde	WGS – 7,5x	2006
Maize (B73)	2600	Diploïde	BAC – 6x	2009
Cuncumbere	367	Diploïde	WGS – 72x (NGS)	2009
<i>Medicago truncatula</i>	500	Diploïde	BAC	2011
<i>Lotus Japonicus</i>	472	Diploïde	BAC – 8,4x	2008
Tomato	950	Diploïde	BAC + NGS	2012
Potato	840	Dihaploïde	WGS – 70x (NGS)	2011
Melon	450	Dihaploïde	454 (13,5x)	2012
Strawberry	240	Diploïde	WGS – ×39 (NGS)	2011
Apple	742	Dihaploïde	WGS – 16,9 x (NGS)	2011
Papaya	372	Diploïde	WGS – 3x	2008
Peach	220	Dihaploïde	WGS – 7x	2013
Barley	5 100	Diploïde	BAC + WGS	2012

Wheat in progress

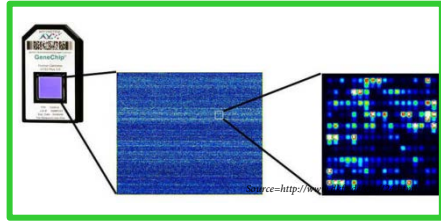
The Evolution of Sequencing Technologies

Automated Sanger Sequencing



First DNA sequencer launched by Applied Biosystem.
Output : 1000 bases per day.

High Throughput Genotyping



Affymetrix released the first commercial SNP array targeting about 1,500 human SNPs.

First Plant Genome Sequenced / Draft sequence Human genome



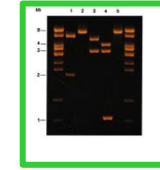
Arabidopsis Thaliana

Launch of the Genome Sequencer 2G system by 454 Life technologies based on pyrosequencing technology.
Output : 20 millions bases per run.

High Throughput Sequencing Technologies

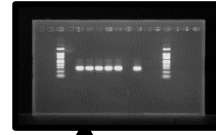


RFLP Genotyping



Kary Mullis et al., developed PCR.

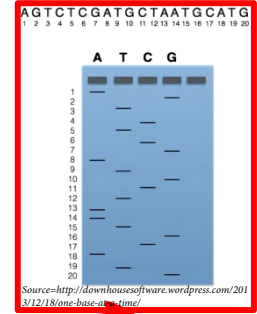
PCR



White et al., described the first polymorphic RFLP marker and proposed methods for mapping the human genome

Sanger Sequencing

Era of DNA sequencing begins : Maxam/Gilbert and Sanger described sequencing techniques.



DNA double helix Description



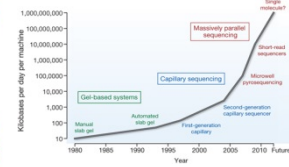
High quality X-ray crystallography data by R. Franklin and M. Wilkins allowed Watson and Crick to elucidate double helix structure of DNA.

The Next Generation ?



Single molecule sequencers like the MinION will provide us more and more bases in less time.

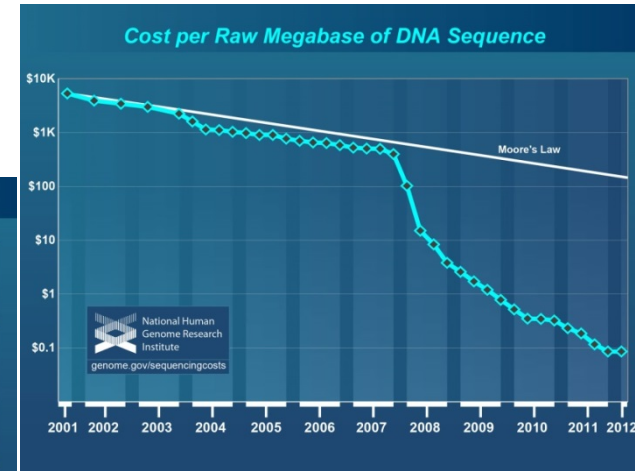
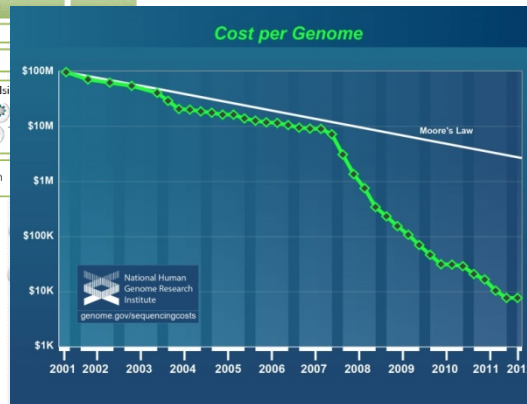
Will it be possible to sequence a whole genome in a USB key?



The Next-generation DNA sequencing technologies

- Complex plant genomes sequencing projects became possible (many billions of bases per day for hundreds or thousands of dollars per gigabase instead of millions or billions of dollars per gigabase)

Société	Roche			Illumina			Life Technologies				
Plateforme	GS Junior	464	MiSeq	HiSeq 1000	HiSeq 2000	Genome Analyzer IIx	Ion Torrent PGM	SOLID 4	SOLID 6600	SOLID 6600xl	
Technologie	Titanium	FLX Titanium	FLX + Titanium				Chip 314 Chip 316				
Acides nucléiques (matrice)											
Ligation adaptateurs											
Méthode d'amplification	PCR en émulsion			« Bridge PCR »			PCR en émulsion				
Méthode de séquençage	Synthèse (Pyroséquençage)			Synthèse			Ligation				
Durée de séquençage/run	10h	10h 20h	26h	8jrs	8jrs	14jrs	2h	12jrs			
Capacité (Mb) séquençage/run	50	500 900	1500	100000	200000	95000	>10 >100 >1000	70000			
Taille moyenne des reads	400	400 700	150+150	100+100	100+100	150+150	100 >100 >100	50+35			
Coût (\$) /run	1100	6200	750	10000	20000	11500	500 750 950	8150			
Coût machine + annexes (K\$)	110+25	500+30	125	560	690	250	50+20	480+55			
Exactitude de séquençage (%)	99	99	99,9	99,9	99,9	99,9	99	99,95	99,95	99,99	



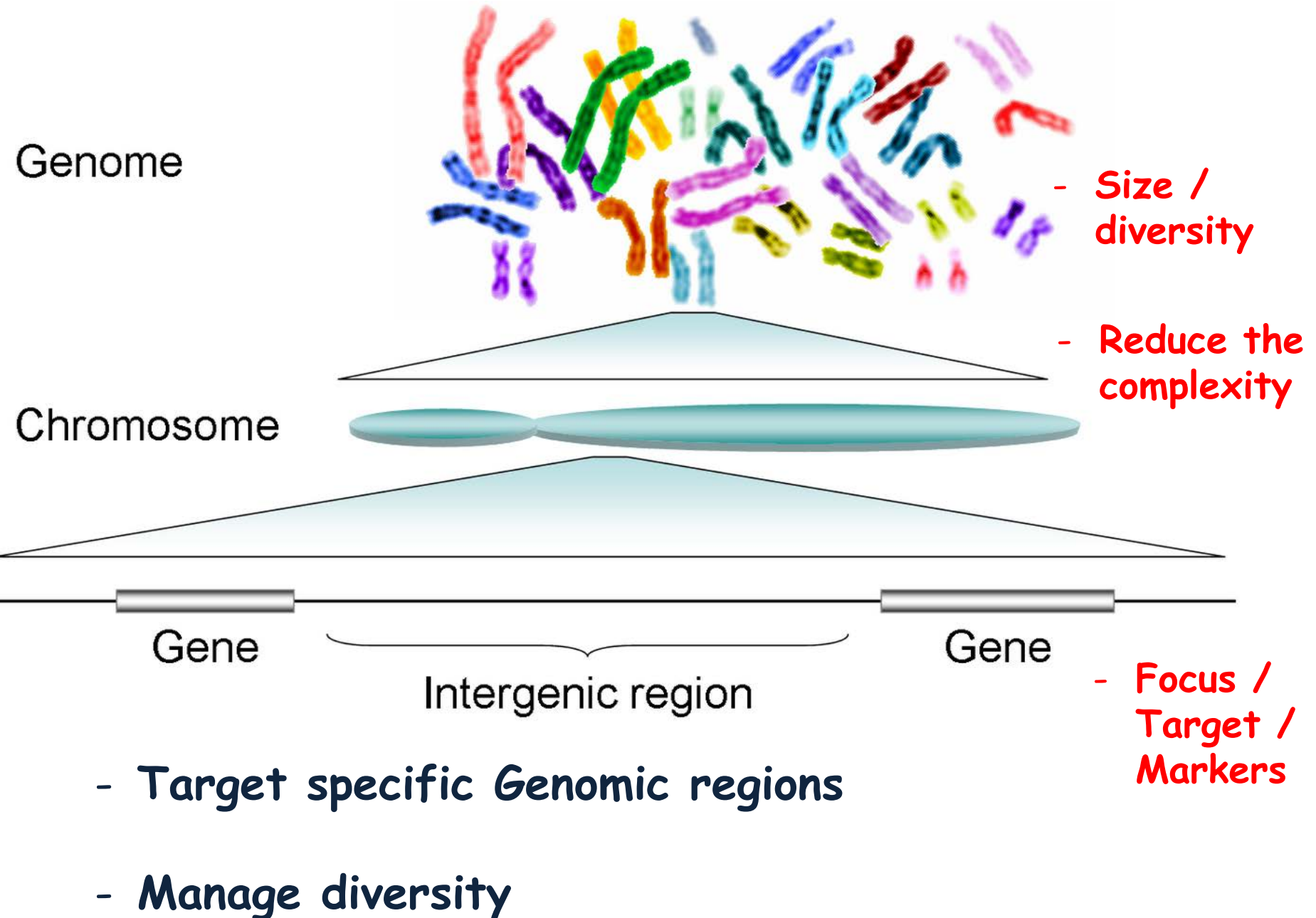
but de novo assembling of plant genomes remains challenging

- Size of the reads
- Repeated elements / Gene families are difficult to assemble and may collapse into a mosaic
- How to assemble these genomes accurately?

-> Despite the progress made with the NGS technologies we still don't have enough reference plant genomes with high confident data (false conclusions ?)

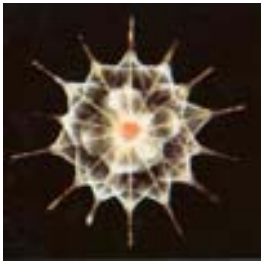
-> third-generation sequencing technologies ?

The challenges / The expectations





Diversity of the living world

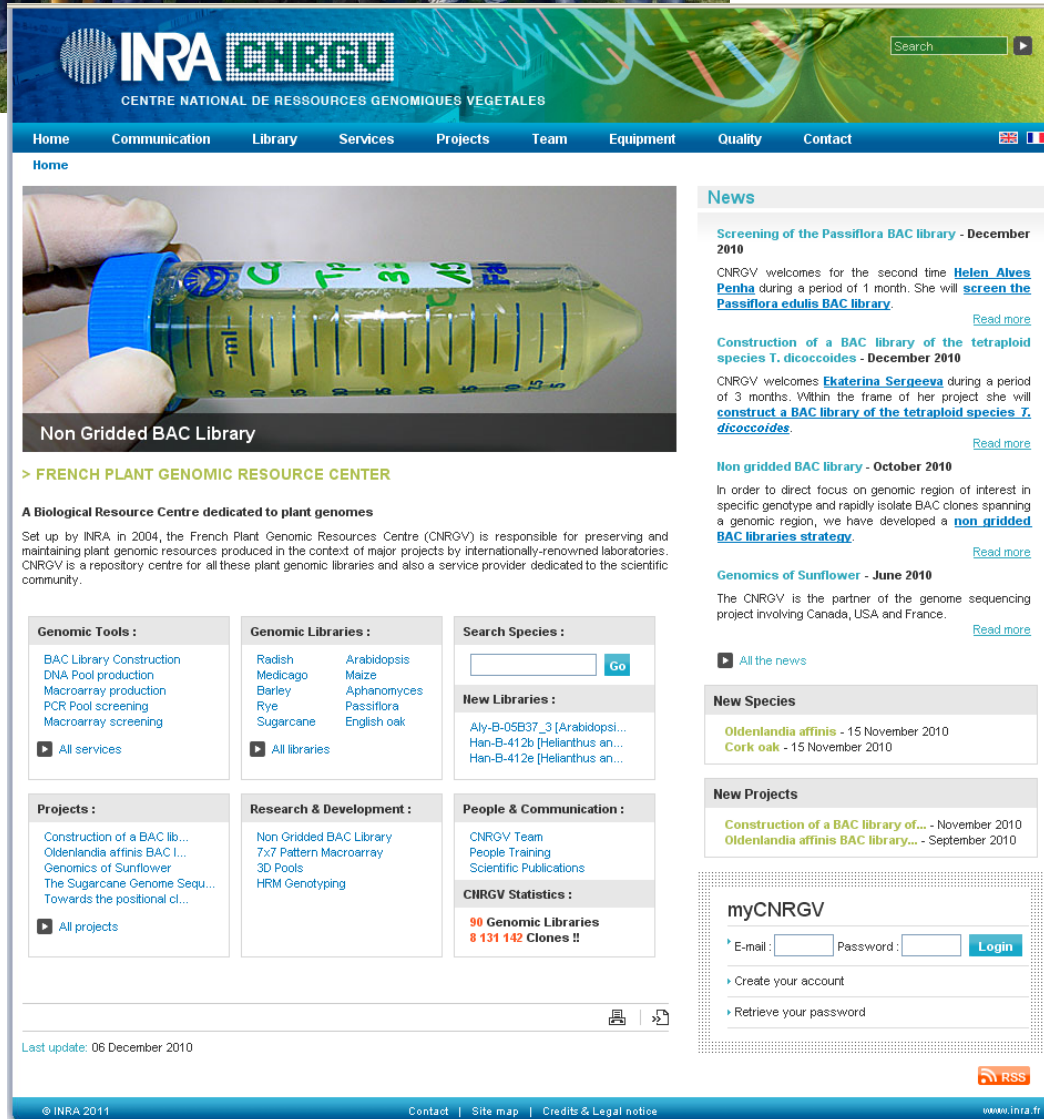



DNA : shared by all living organisms



Thank you for your attention

Arnaud BELLEC
Sonia VAUTRIN
Genséric BEYDON
William MARANDE
Joëlle FOURMENT
Elisa PRAT
Nadine GAUTIER
Nathalie RODDE
Nadège ARNAL
Audrey COURTIAL
Stéphane CAUET
David PUJOL
Nicolas THERON
Laetitia HOARAU
Hélène BERGES



The screenshot shows the homepage of the Centre National de Ressources Génomiques Végétales (CNRGV). The header features the INRA and CNRGV logos, a search bar, and a navigation menu with links to Home, Communication, Library, Services, Projects, Team, Equipment, Quality, and Contact. The main content area is divided into several sections:

- Home:** A featured image of a hand holding a blue-tipped pipette tip with a yellow liquid, labeled "Non Gridded BAC Library".
- News:** A list of recent news items, including "Screening of the Passiflora BAC library - December 2010", "Construction of a BAC library of the tetraploid species *T. dicoccoides* - December 2010", "Non gridded BAC library - October 2010", and "Genomics of Sunflower - June 2010". Each item includes a brief description and a "Read more" link.
- Genomic Tools:** A list of services such as BAC Library Construction, DNA Pool production, and Macroarray production.
- Genomic Libraries:** A list of species for which libraries are available, including Radish, Arabidopsis, Medicago, Maize, Barley, Aphanomyces, Rye, Passiflora, and Sugarcane, English oak.
- Search Species:** A search box with a "Go" button.
- New Libraries:** A list of newly added libraries, such as Aly-B-05B37_3 [Arabidopsi...], Han-B-412b [Helianthus an...], and Han-B-412e [Helianthus an...].
- New Species:** A list of new species, including Oldenlandia affinis and Cork oak.
- New Projects:** A list of new projects, such as Construction of a BAC library of... and Oldenlandia affinis BAC library....
- Projects:** A list of ongoing projects, including Construction of a BAC lib..., Oldenlandia affinis BAC l..., Genomics of Sunflower, The Sugarcane Genome Sequ..., and Towards the positional cl....
- Research & Development:** A list of research and development activities, including Non Gridded BAC Library, 7x7 Pattern Macroarray, 3D Pools, and HRM Genotyping.
- People & Communication:** A list of communication activities, including CNRGV Team, People Training, and Scientific Publications.
- myCNRGV:** A login section with fields for E-mail and Password, a "Login" button, and links to "Create your account" and "Retrieve your password".

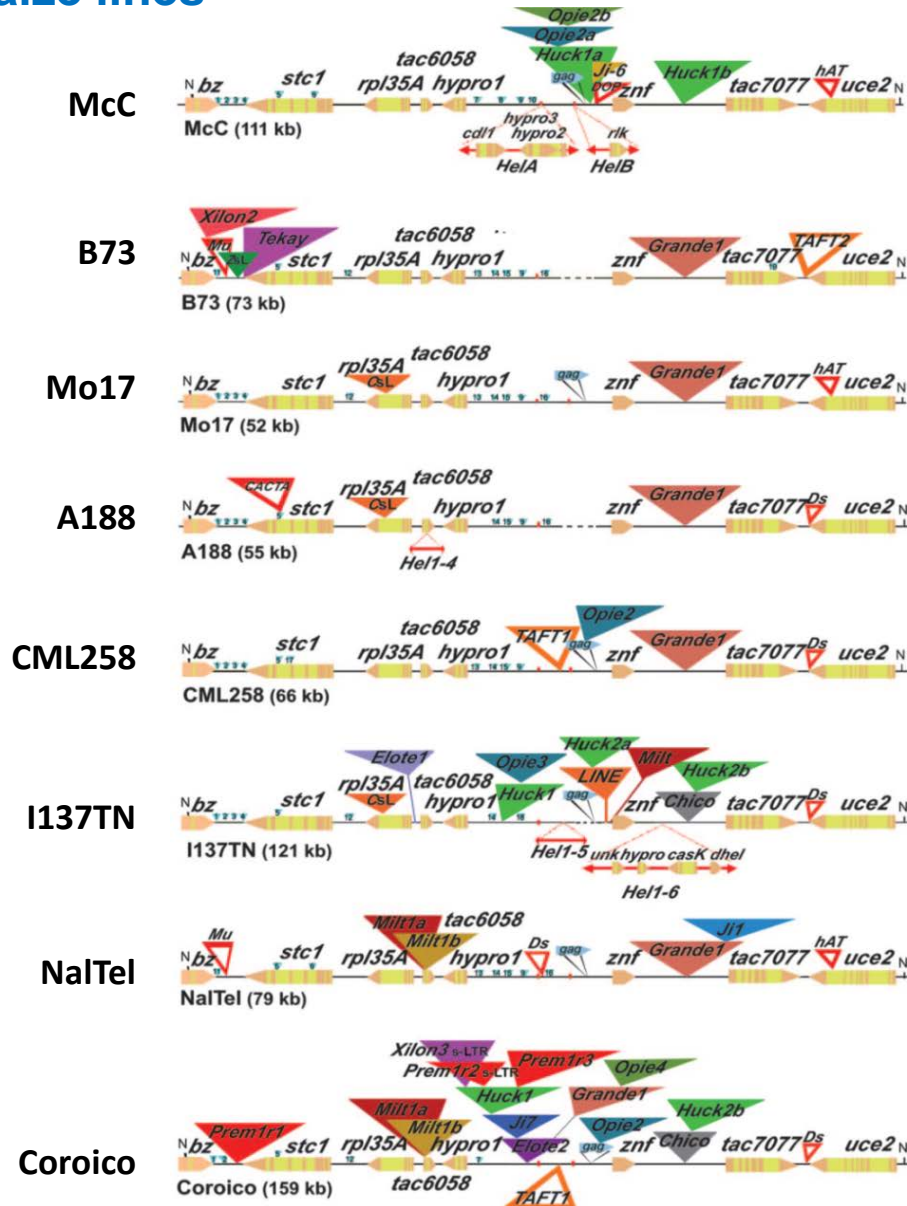
At the bottom of the page, there is a footer with the text "© INRA 2011", "Contact | Site map | Credits & Legal notice", and "www.inra.fr".



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

COMPARATIVE ANALYSIS OF A GENIC REGION: THE *BRONZE* LOCUS

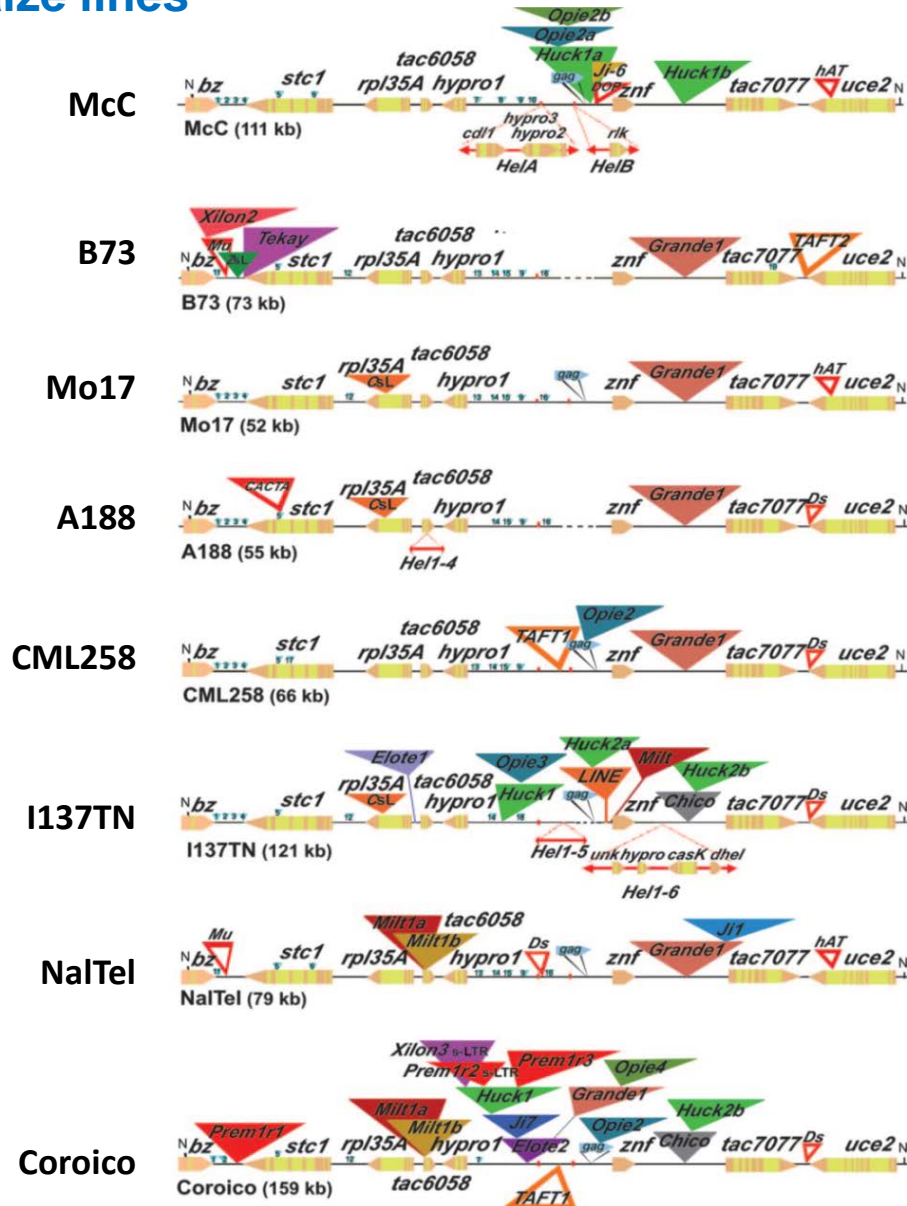
Maize lines



84% sequence shared
between genotypes

COMPARATIVE ANALYSIS OF A GENIC REGION: THE *BRONZE* LOCUS

Maize lines



Only 25% sequence shared between genotypes