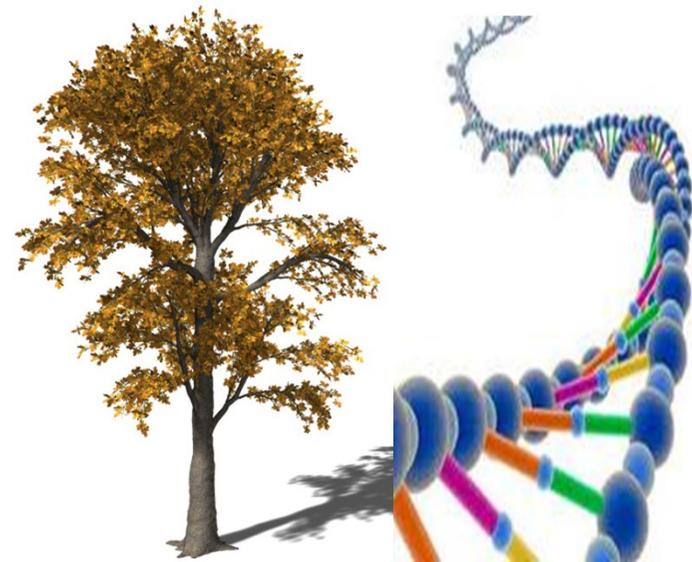


Décryptage du génome du chêne pour mieux comprendre les bases génétiques de d'adaptation des arbres forestiers à leur environnement

Christophe
Plomion



Récentes avancées en génomique des arbres forestiers

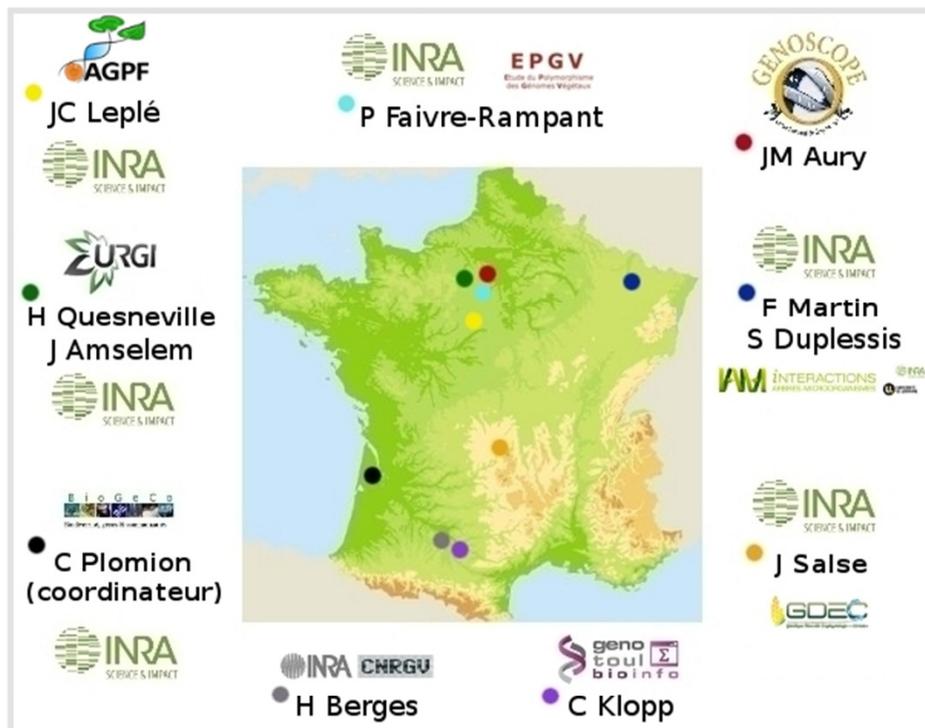
Académie d'Agriculture

section 2 « forêts et filière bois » & section 6 « sciences de la vie »



22 milliards de séquences
= 4,6 triards de nucléotides
= 2 770 mètres d'ADN de chêne

Partners



- Who** : INRA Bordeaux – Christophe Plomion, Catherine Bodénès, Christophe Boury , François Ehrenmann, Antoine Kremer, Céline Lalanne, Grégoire Le Provost, Thibault Leroy, Isabelle Lesur, Franck Salin

What : Linkage Mapping, QTL detection and association mapping, Bioinformatics and Databasing, Evolutionary genetics (speciation, local adaptation), Functional genomics (bud phenology, drought stress, root hypoxia), Web site
- Who** : INRA Versailles-Grignon (URGI) – Joëlle Amselem, Tina Alaëitabar, Nicolas Francillonne, Célia Michotey, Delphine Steinbach, Hadi Quesneville

What : Bioinformatics (structural & functional annotation), Web interface, Databasing
- Who** : INRA Versailles-Grignon – Nathalie Boudet, Aurélie Canaguier, Cécile Guichard, Isabelle Le Clainche, Patricia Faivre-Rampant

What : BAC lib./clone storage/BAC end sequencing
- Who** : CEA Evry – Jean-Marc Aury, Valérie Barbe, Caroline Belser, Arnaud Couloux, Corinne Da Silva, Barbara Estrada-Mairey , Stéphanie Fouteau, Karine Labadie, Jonathan Mercier, Patrick Wincker

What : Sequencing, resequencing, Bioinformatics (transcriptome & genome assembly)
- Who** : INRA Nancy – Sébastien Duplessis, Francis Martin, Annegret Kohler, Clémence Marchal

What : Annotation jamboree and annotation of biotic interaction related genes
- Who** : INRA Clermont-Ferrand – Florent Murat, Jérôme Salse

What : Comparative genomics, Genome evolution

- Who** : INRA Toulouse Genotoul – Christine Gaspin, Christophe Klopp

What : Non-coding RNAs
- Who** : INRA Toulouse CNRGV – Hélène Bergès

What : BAC library construction, clone storage
- Who** : INRA Orléans AGPF – Jean-Charles Leplé

What : Non-coding RNAs

Decoding the oak genome: public release of sequence data, assembly, annotation and publication strategies

CHRISTOPHE PLOMION,^{*,†1} JEAN-MARC AURY,^{‡1} JOËLLE AMSELEM,^{§1} TINA ALAEITABAR,[§] VALÉRIE BARBE,[‡] CAROLINE BELSER,[‡] HÉLÈNE BERGÈS,[¶] CATHERINE BODÈNÈS,^{*,†} NATHALIE BOUDET,^{**} CHRISTOPHE BOURY,^{*,†} AURÉLIE CANAGUIER,^{**} ARNAUD COULOUX,[‡] CORINNE DA SILVA,[‡] SÉBASTIEN DUPLESSIS,^{††} FRANÇOIS EHRENMANN,^{*,†} BARBARA ESTRADA-MAIREY,[‡] STÉPHANIE FOUTEAU,[‡] NICOLAS FRANCILLONNE,[§] CHRISTINE GASPIN,^{‡‡} CÉCILE GUICHARD,^{**} CHRISTOPHE KLOPP,^{‡‡} KARINE LABADIE,[‡] CÉLINE LALANNE,^{*,†} ISABELLE LE CLAINCHE,^{**} JEAN-CHARLES LEPLÉ,^{§§} GRÉGOIRE LE PROVOST,^{*,†} THIBAUT LEROY,^{*,†} ISABELLE LESUR,^{*,†} FRANCIS MARTIN,^{††} JONATHAN MERCIER,[‡] CÉLIA MICHOTÉY,[§] FLORENT MURAT,^{¶¶} FRANCK SALIN,^{*,†} DELPHINE STEINBACH,[§] PATRICIA FAIVRE-RAMPANT,^{**} PATRICK WINCKER,^{‡***†††} JÉRÔME SALSE,^{¶¶} HADI QUESNEVILLE[§] and ANTOINE KREMER^{*,†}



accords internationaux des Bermudes (1998) et de Fort Lauderdale (2003), déclaration de Toronto (2009)

www.oakgenome.fr

Search

OAK GENOME SEQUENCING

Sequencing of the oak genome and identification of genes that matter for forest tree adaptation

Home OAK genome in numbers Data Links Other resources Partners Publications Contact Funding Gallery Outreach

Home

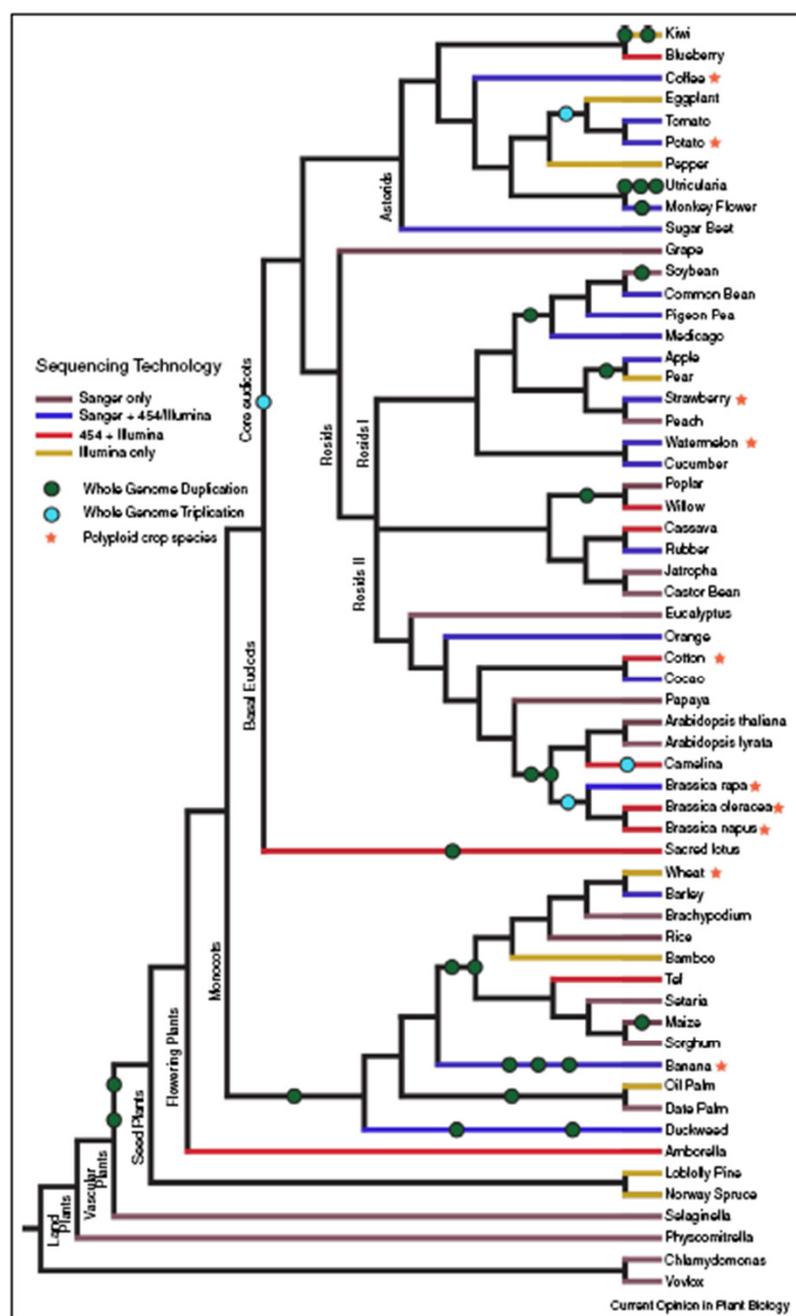
Work on the genome of a species of oak tree native from Europe (*Quercus robur*, the pedunculate oak) began in January 2012. It was conducted by a **French consortium** led by Christophe Plomion at INRA Bordeaux. Three years later, this genome has been sequenced, assembled and annotated. The consortium characterized around 50,000 genes and estimated that 50% of the 1.5 billion base pairs of this genome was made of repetitive elements.

This first oak genome assembled so far, provides a foundation to study the biology and evolution of these species. In particular, it will allow to identify genes essential for the adaptation of these long-lived organisms to their environment, genes involved in the symbiotic relationships between the mycelia of truffles and its root, genes responsible for the biosynthesis of wood extractives such as tannins and whisky-lactone that gives their flavor and taste to alcohol and wine.

BioGeCo
Biodiversité, génomes & communautés



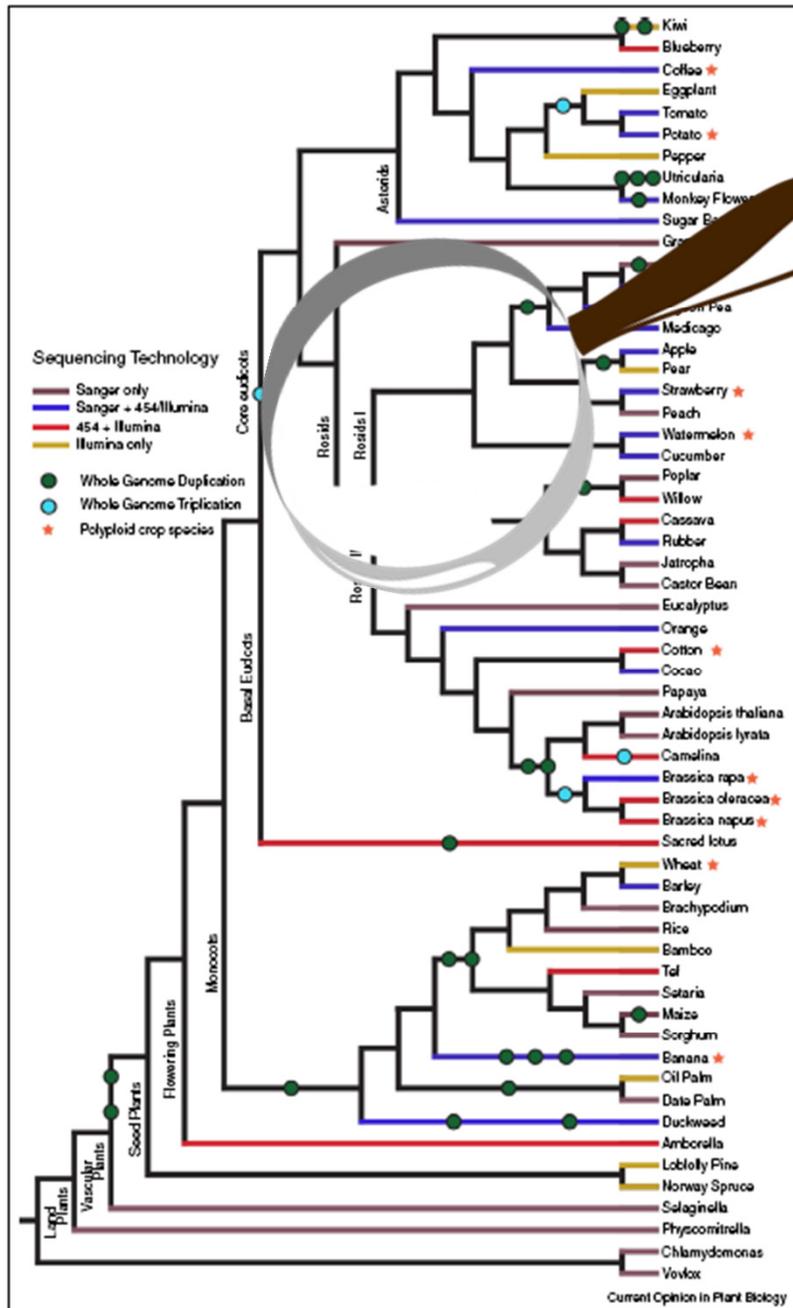
www.oakgenome.fr



Progress, challenges and the future of crop genomes

Todd P Michael¹ and Robert VanBuren²

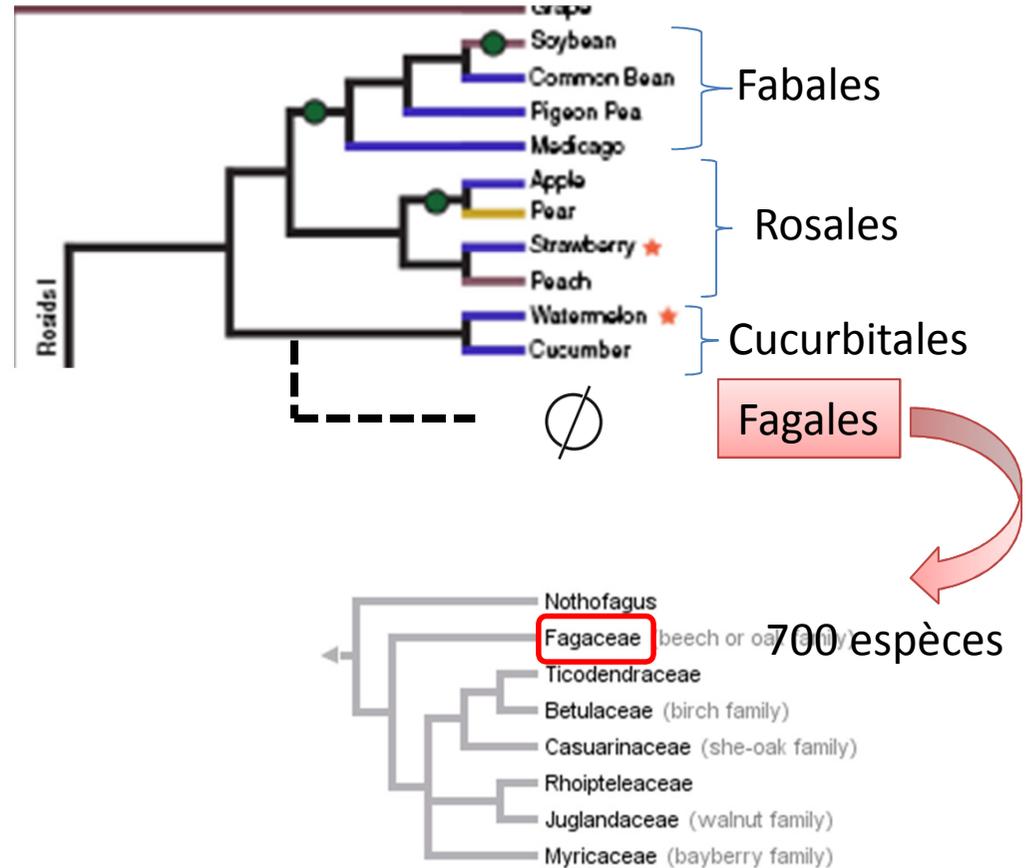
Current Opinion in
Plant Biology



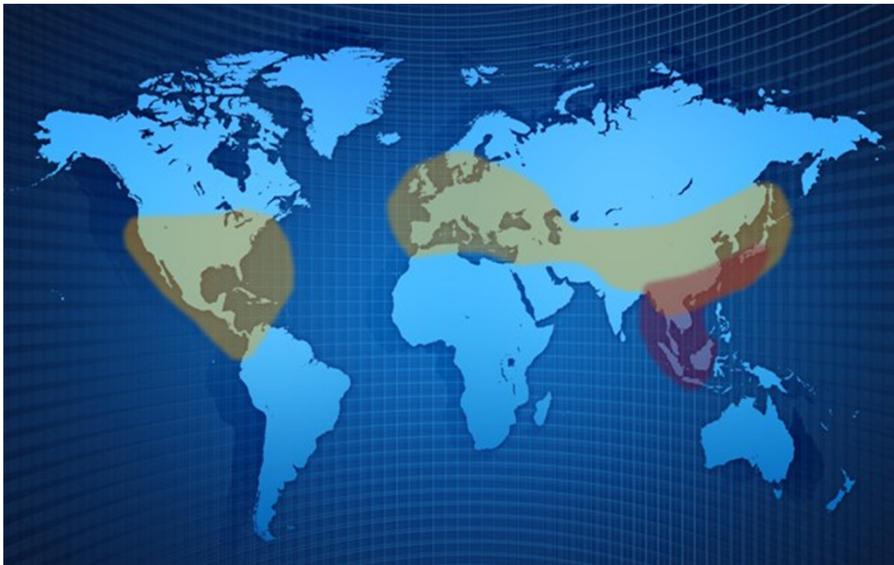
Progress, challenges and the future of crop genomes

Todd P Michael¹ and Robert VanBuren²

Current Opinion in
Plant Biology



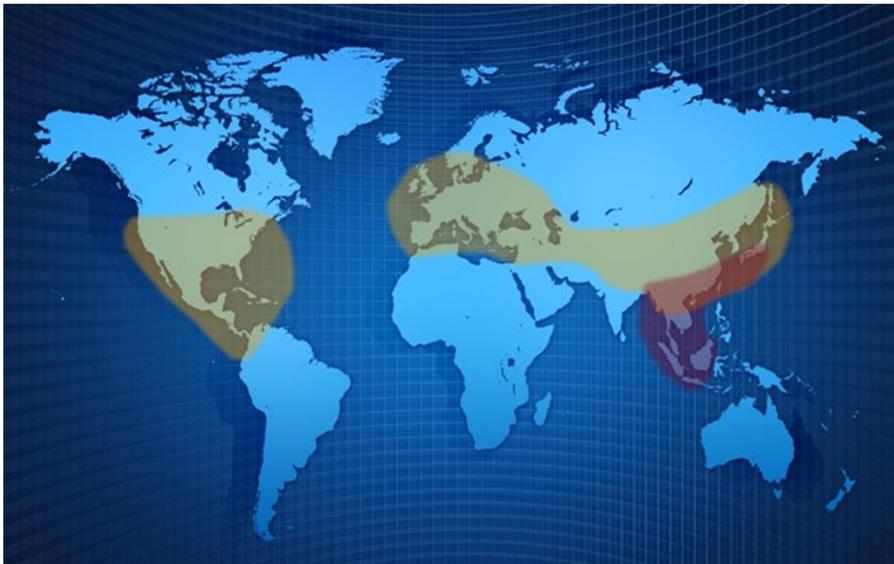
Le genre *Quercus*



Distribution géographique

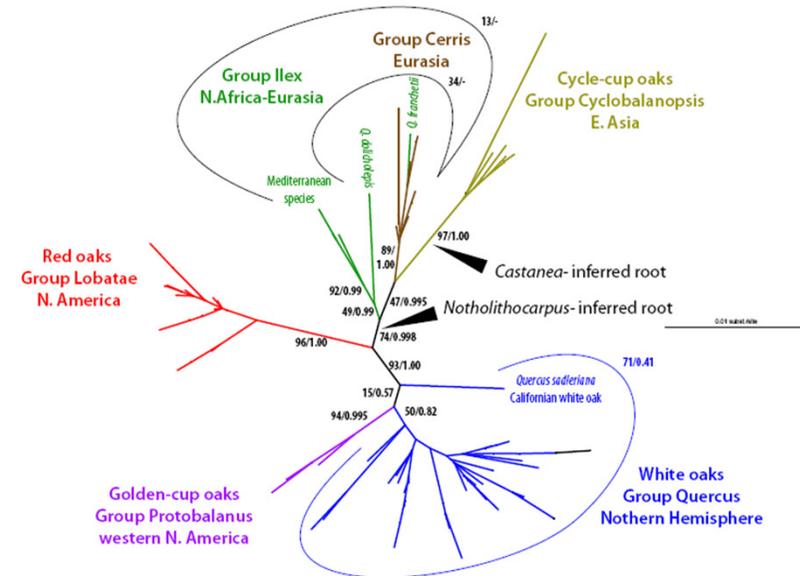
➤ ~ 600 espèces

Le genre *Quercus*



Distribution géographique

- ~ 600 espèces



Phylogénie

Hubert *et al* 2014 Systematics & Biodiversity, 1-19

- Spéciation : 50-55 MA
- Chêne blancs européen < 5 MA

Chêne blancs européens



Objectifs

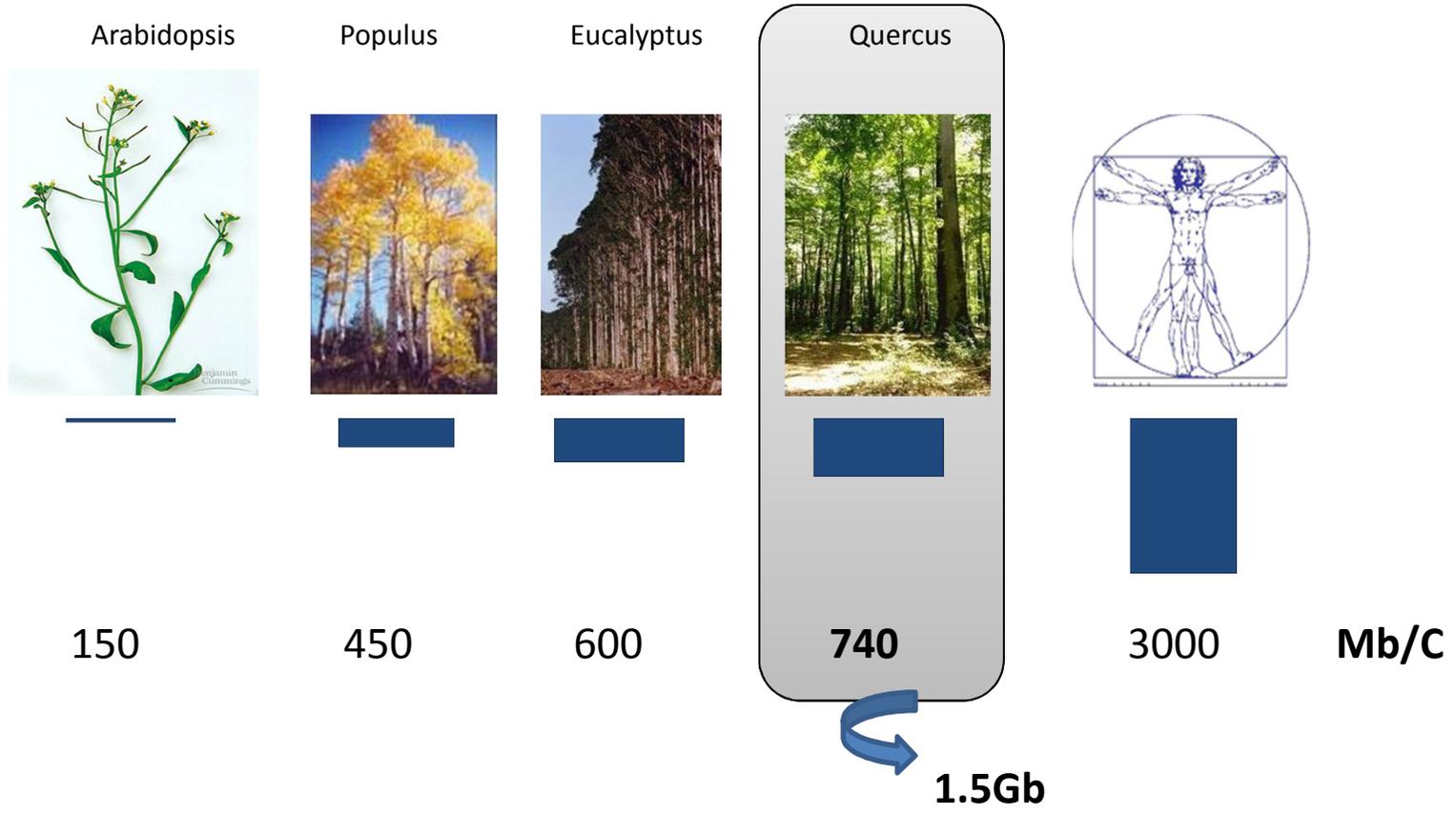
1. Etablir une **séquence de référence du génome** du chêne
2. Comprendre sa **structure**, son **évolution** et son **fonctionnement**
3. Comprendre les **bases moléculaires des différences adaptatives**

Objectifs

1. Etablir une **séquence de référence du génome** du chêne
2. Comprendre sa structure, son évolution et son fonctionnement
3. Comprendre les bases moléculaires des différences adaptatives

La complexité du génome du chêne : sa taille physique

Quercus robur : $2n=2x=24$

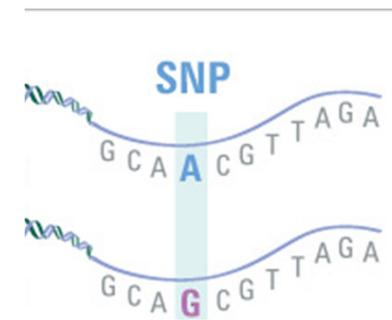


La complexité du génome du chêne : son hétérozygotie

✚ Au niveau le plus élémentaire :

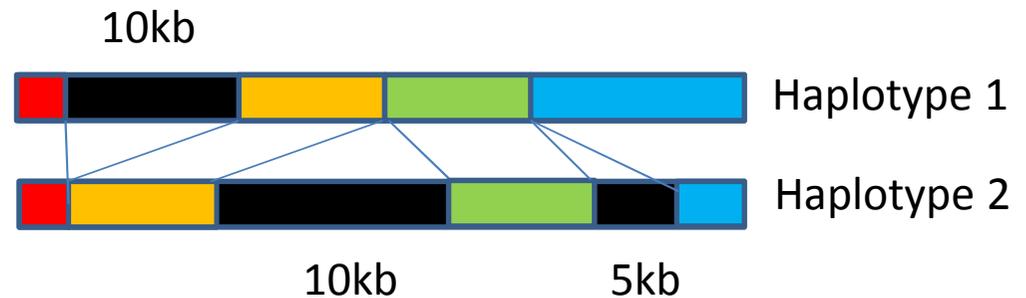
➤ 1 SNP/45 bp estimé par le reséquençage de 546 régions exoniques

➤ 1 SNP / 40 bp dans des régions intergéniques de 3 BAC alléliques séquencés



✚ Au niveau de la macrostructure :

➤ De grandes insertions/délétions d'éléments transposables [black box] dans 3 BAC alléliques séquencés



Stratégie

- Couvrir les deux haplotypes $2 \times 750 \text{ Mb} = 1.5 \text{ G}$
 - $> 100 \times$
- Produire des longues séquences pour mieux séparer les deux haplotypes
 - En 2012: Sanger (600 pb), Pyroséquençage (450b) et Illumina chevauchants (170b)
 - En 2014: Moleculo (lecture synthétiques, 5kb)
- Séquencer les extrémités de longs fragments (ponts)
 - Banques de 150kb (BAC), 3-5kb, 5-8kb, 8-11kb
- Assembler une version diploïde V2n
- Fusionner les deux H en une version haploïde V1n
- Aligner V1n sur une carte génétique → 12 pseudo-molécules

Résultat

	V2n (1.5 Gb)	V1n (750 Mb)
Taille assemblée	1.4Gb	814Mb
Nb de scaff	8 827	1 409
N50 scaff	0.8G	1.3Mb

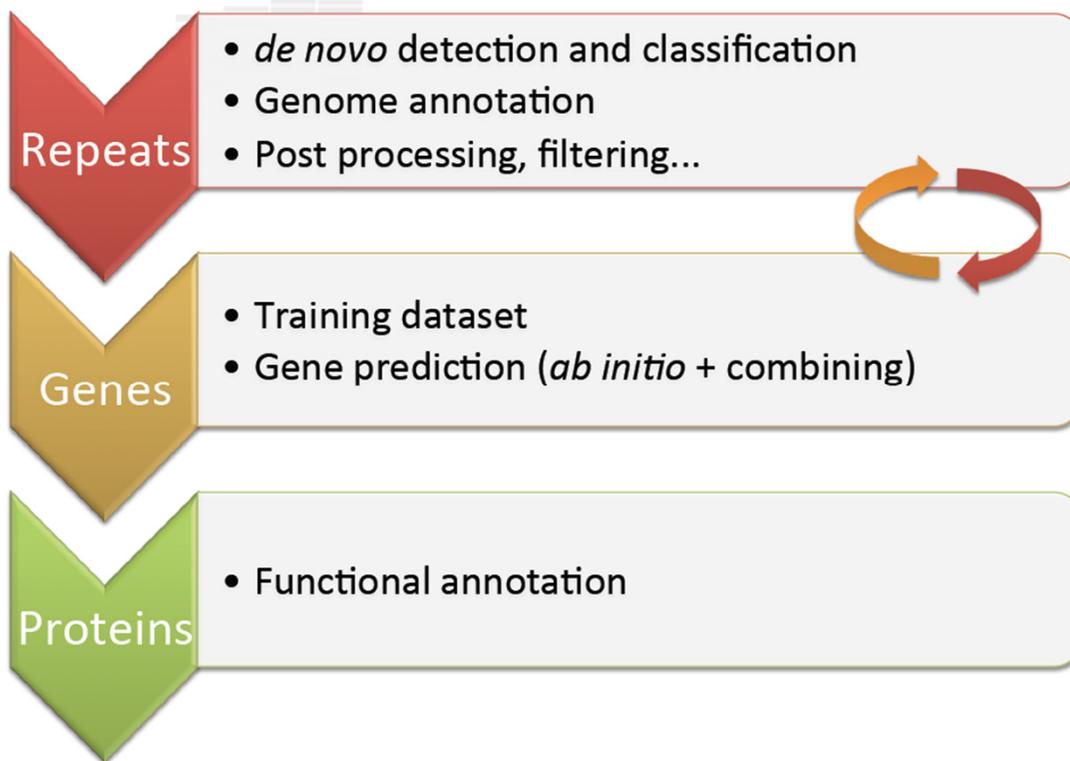


N50: la moitié de la taille du génome est contenue ds des scaff d'au moins 1.3 Mb

Objectifs

1. Etablir une séquence de référence du génome du chêne
2. Comprendre sa **structure**, son **évolution** et son **fonctionnement**
3. Comprendre les bases moléculaires des différences adaptatives

Annotation



Eléments transposables
→ 50.6% du génome

➤ 29.5k gènes (2.8kb – CDS 1kb)

Navigation dans le génome du chêne

The screenshot displays the Quercus Robur V2 genome browser interface. The top navigation bar includes a dropdown menu for 'Quercus Robur V2', menu items for 'File', 'View', and 'Help', and a 'Login' link. Below the navigation bar is a coordinate scale from 0 to 450,000. A search bar contains the text 'qrob_v2_scaffold_0983' and 'qrob_v2_scaffold_0983:1..134093 (134.05)'. The main area shows several tracks:

- Chunks**: A track showing genomic chunks.
- Repeat_Annotation**: A track showing repeat annotations with red and black bars.
- Gene annotation V2.2 (transcript)**: A track showing gene models with yellow arrows.
- Marker**: A track showing marker positions.
- QTLs peak**: A track showing QTL peaks.
- Haplome_v2.3**: A track showing haplotype information, with a specific haplotype 'Qrob_H2.3_Sc0000269:637328..758082' highlighted.
- QTLs CI**: A track showing QTL confidence intervals, with several entries like 'Bourran1_2003_QTL2_CI_Bud_burst_A4'.

On the left side, there is an 'Available Tracks' panel with a search box and a list of track categories and their counts:

- Chunks**: 2 (checked)
- 1-Repeats**: 2 (checked)
- 2-Prediction**: 6 (checked)
- 3-Transcriptome**: 4
- 4-Protein**: 7
- 5-RNA-seq(BAM)**: 24
- 6-Chloro and mito**: 2
- 7-Marker and QTL**: 3 (checked)
- BAC**: 2
- Haplome**: 1 (checked)
- V1 on V2**: 4

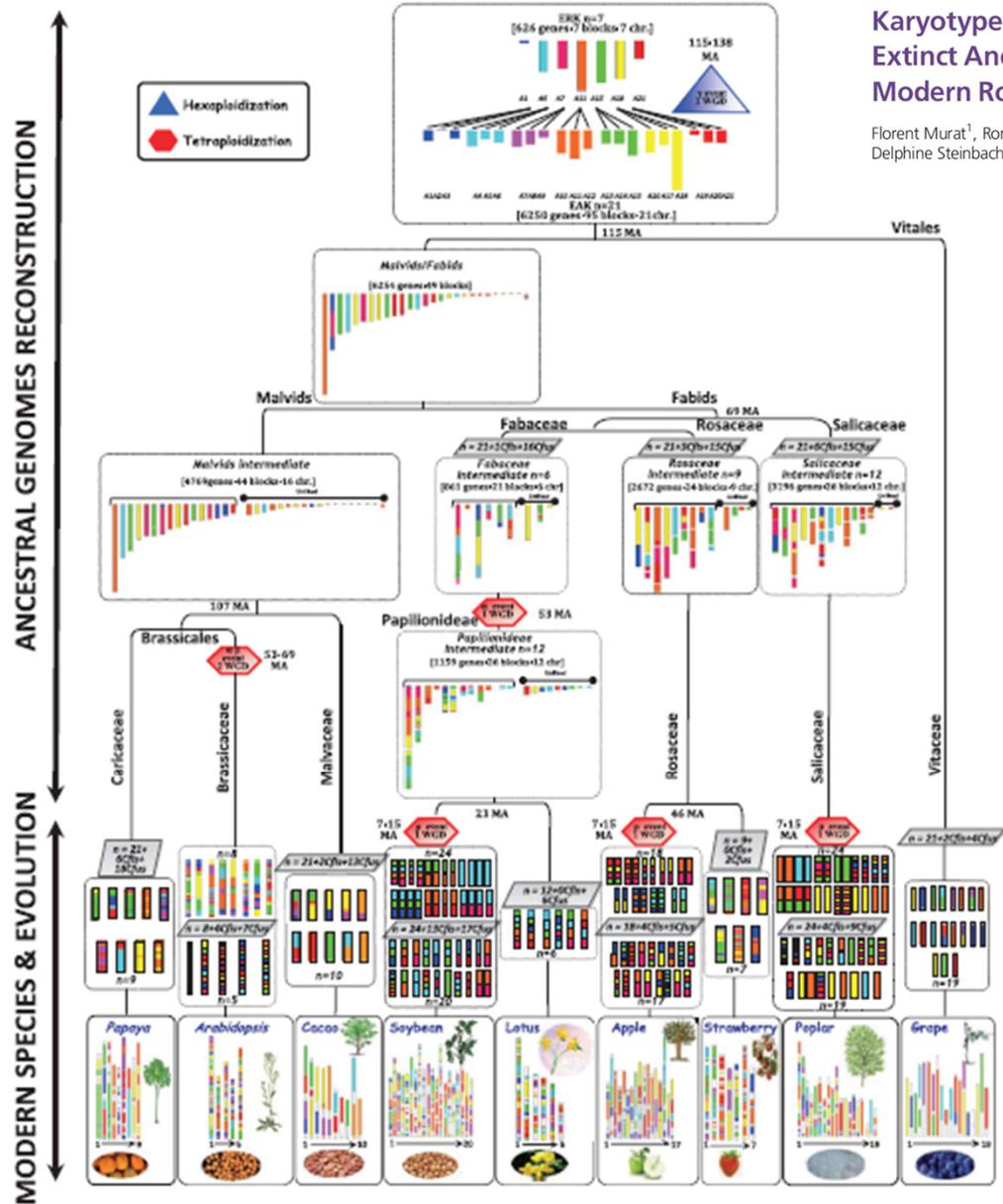
Objectifs

1. Etablir une séquence de référence du génome du chêne

2. Comprendre sa **structure**, son **évolution** et son **fonctionnement**

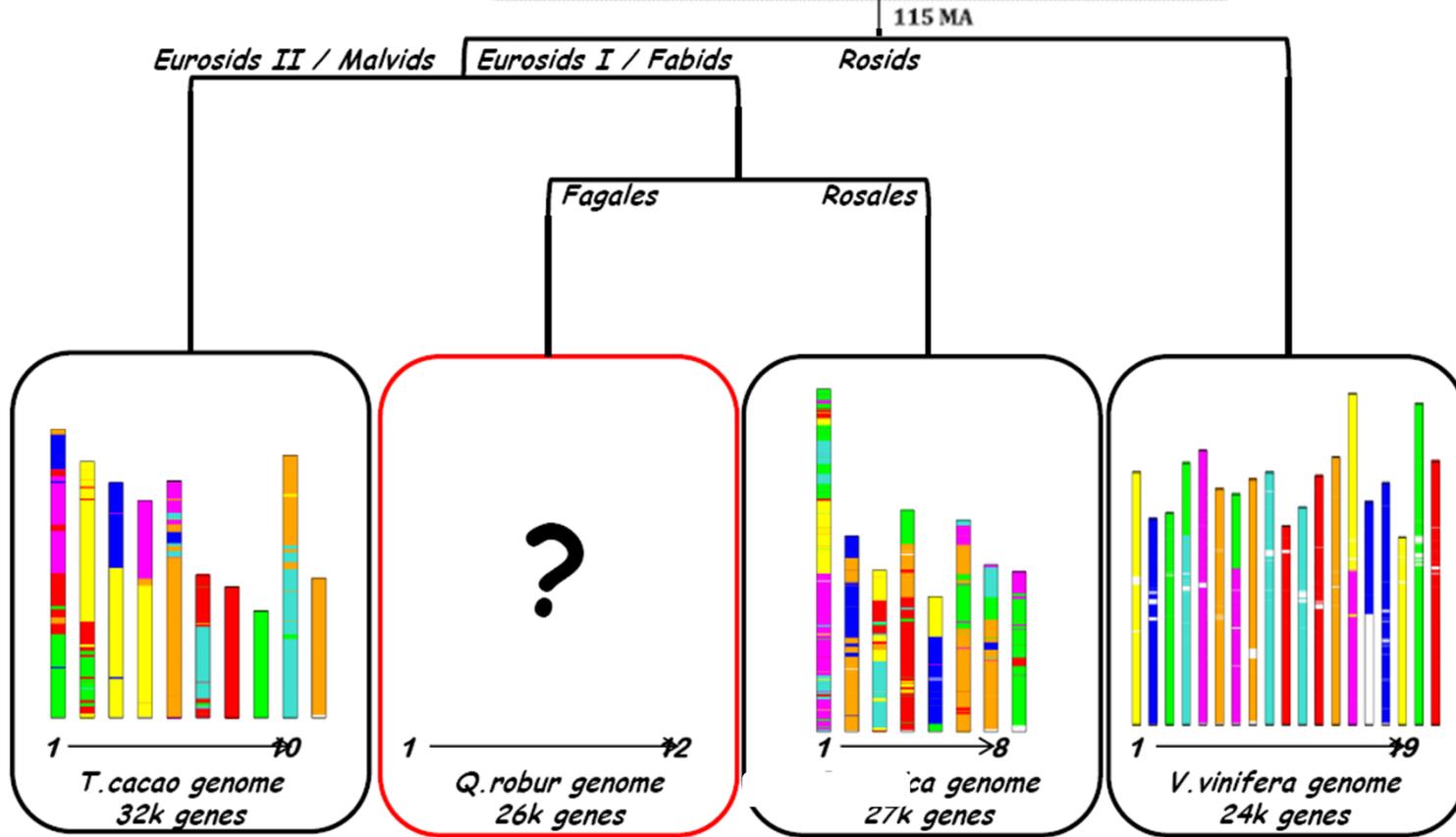
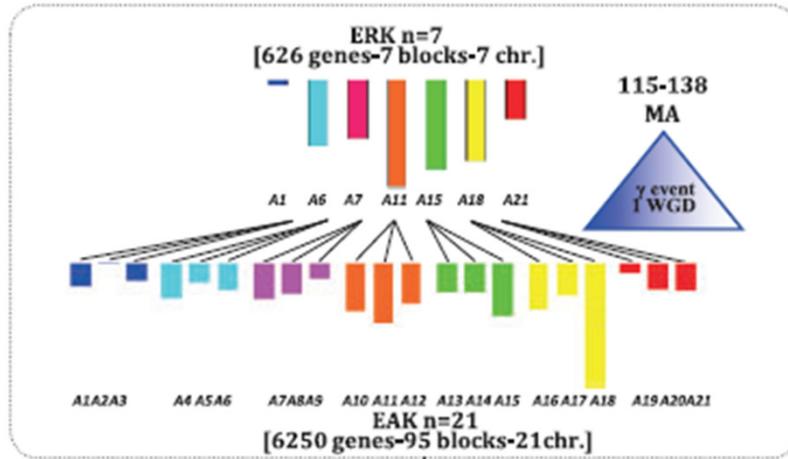


3. Comprendre les bases moléculaires des différences adaptatives



Karyotype and Gene Order Evolution from Reconstructed Extinct Ancestors Highlight Contrasts in Genome Plasticity of Modern Rosid Crops

Florent Murat¹, Rongzhi Zhang¹, Sébastien Guizard¹, Haris Gavranović², Raphael Flores³, Delphine Steinbach³, Hadi Quesneville³, Eric Tannier⁴, and Jérôme Salse^{1,*}



Objectifs

1. Etablir une séquence de référence du génome du chêne
2. Comprendre sa **structure**, son **évolution** et son **fonctionnement**
3. Comprendre les bases moléculaires des différences adaptatives

Approches sans *a priori*:

- Le répertoire des gènes du chêne est-il différent de celui des autres plantes séquencées ?
- Quels sont les familles de gènes en expansion chez le chêne?
- La caractéristique » ARBRE vs HERBACEE » se traduit-elle par un répertoire particulier de gènes ?

Approches avec *a priori*:

- annotation expertisée de certaines familles de gènes/fonctions cibles

Objectifs

1. Etablir une séquence de référence du génome du chêne
2. Comprendre sa structure, son évolution et son fonctionnement
3. Comprendre les **bases moléculaires des différences adaptatives**

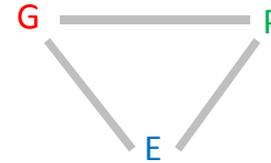
Génome



Transcriptome

Stratégie

- Intégrer trois sources d'information
- Balayer les processus à différentes échelles évolutives en tenant compte des forces évolutives
- Prendre en compte la contribution du polymorphisme de l'ADN et la régulation des gènes
- S'appuyer sur des dispositifs *in situ* et *ex situ*



➤ PLASTICITE ET HETEROGENEITE INTRA INDIVIDUEL
➤ ADAPTATION (synchronique/allochronique)
➤ SPECIATION

➤ MUTATION DERIVE SELECTION MIGRATION

