



HAL
open science

Quelle place pour la sélection génomique chez les plantes pérennes ?

Catherine Bastien, David Cros, Patrice P. This

► To cite this version:

Catherine Bastien, David Cros, Patrice P. This. Quelle place pour la sélection génomique chez les plantes pérennes ?. Ecole chercheur "Sélection génomique", Sep 2013, Bruz, France. hal-02810865

HAL Id: hal-02810865

<https://hal.inrae.fr/hal-02810865v1>

Submitted on 6 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Quelle place pour la sélection génomique chez les plantes pérennes ?

Catherine Bastien¹, David Cros², Patrice This³

- Particularités des plantes pérennes
 - Exemple 1: **le palmier à huile**
 - Exemple 2: **la vigne**
 - Exemple 3: **le peuplier**

¹ INRA-AGPF Orléans, ² CIRAD Montpellier, ³ INRA-UMR AGAP Montpellier



Points communs des plantes pérennes



Pommier



Pêcher



Cerisier



Palmier
à huile



Vigne



Eucalyptus



Pin
maritime



Peuplier

$$\text{Gain génétique} / \text{unité temps} = \frac{\text{Précision de prédiction} \times \text{Intensité de sélection} \times \text{Variabilité génétique}}{\text{Longueur cycle}}$$

- Des **cycles de sélection longs** (maturité sexuelle + durée d'évaluation)
- Un **encombrement spatial** limitant les capacités de phénotypage
- Une **sélection multi-caractères** pour des environnements diversifiés (**GxE**)

Points communs des plantes pérennes



Pommier

2010



Pêcher

2010



Cerisier



Palmier
à huile

2013



Vigne

2007



Eucalyptus

2011



Pin
maritime
en cours

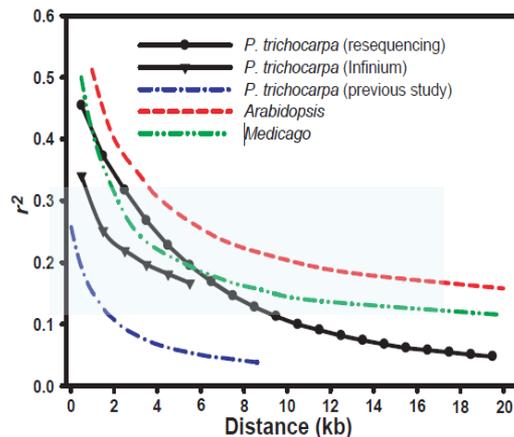


Peuplier

2006

- Des **génomés de référence** et d'importants développements de SNP
- Mais un **DL** qui chute rapidement à courte distance ($r^2 < 0,2$ à partir de 1 à 4 kb suivant espèce)

Populus trichocarpa
Slavov et al. 2012



Densité de
marqueurs **élevée**

Taille efficace
réduite

Population
d'entraînement **grande**

Précision de
prédiction

Quelle place pour la sélection génomique chez les plantes pérennes ?

Exemple du palmier à huile

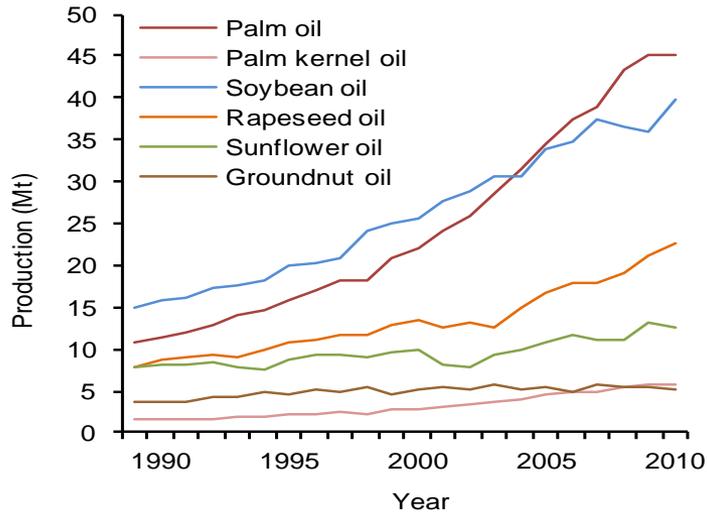
David Cros

CIRAD, Unit « Genetic improvement and adaptation of Mediterranean and tropical plants », Montpellier



Le palmier à huile

1ère plante oléagineuse au monde (production >55 Mt)

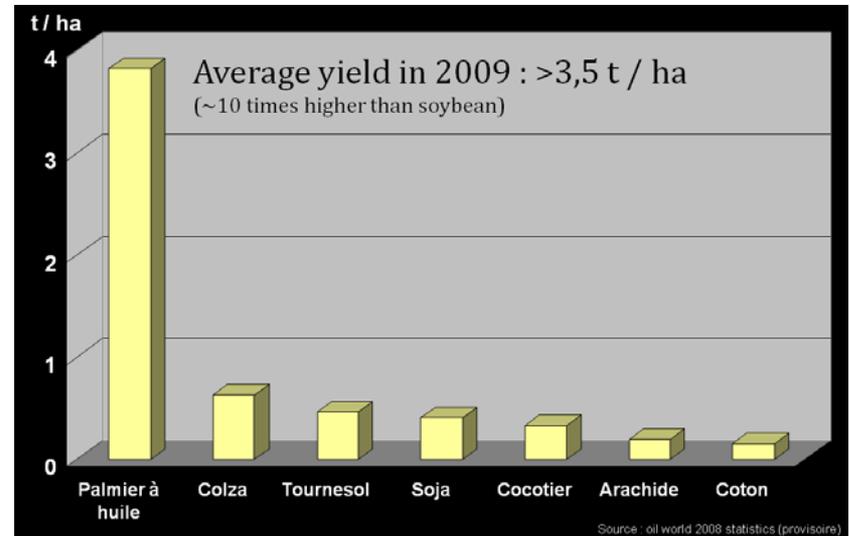


Cultivé dans toute la zone tropicale



80% pour l'alimentation humaine
(margarine, friture, etc.)

20% pour l'industrie
(savonnerie, cosmétiques, etc.)



Meilleur rendement chez les plantes oléagineuses

Caractéristiques d'importance pour l'amélioration

Le palmier à huile (*Elaeis guineensis*) :

➤ **diploïde** ($2n = 32$)



➤ **allogame**

(monoïque avec cycles males et femelles: « dioïque temporelle »)

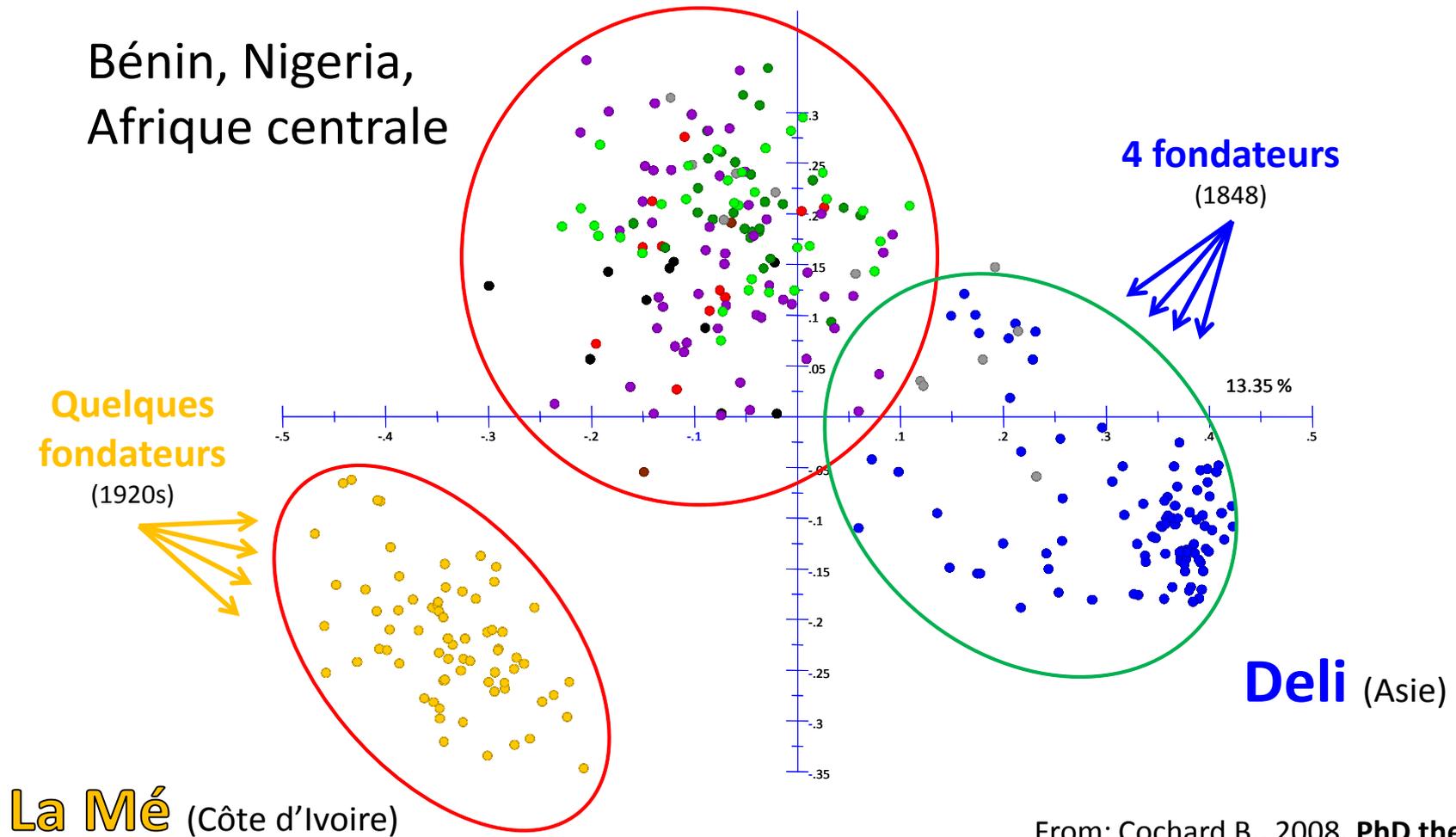


➤ **multiplication végétative difficile**



➤ **pollinisation artificielle facile**

Diversité génétique chez *E. guineensis*



ACP sur 318 *E. guineensis* (14 marqueurs SSR)

La Sélection Récurrente Réciproque pour le rendement

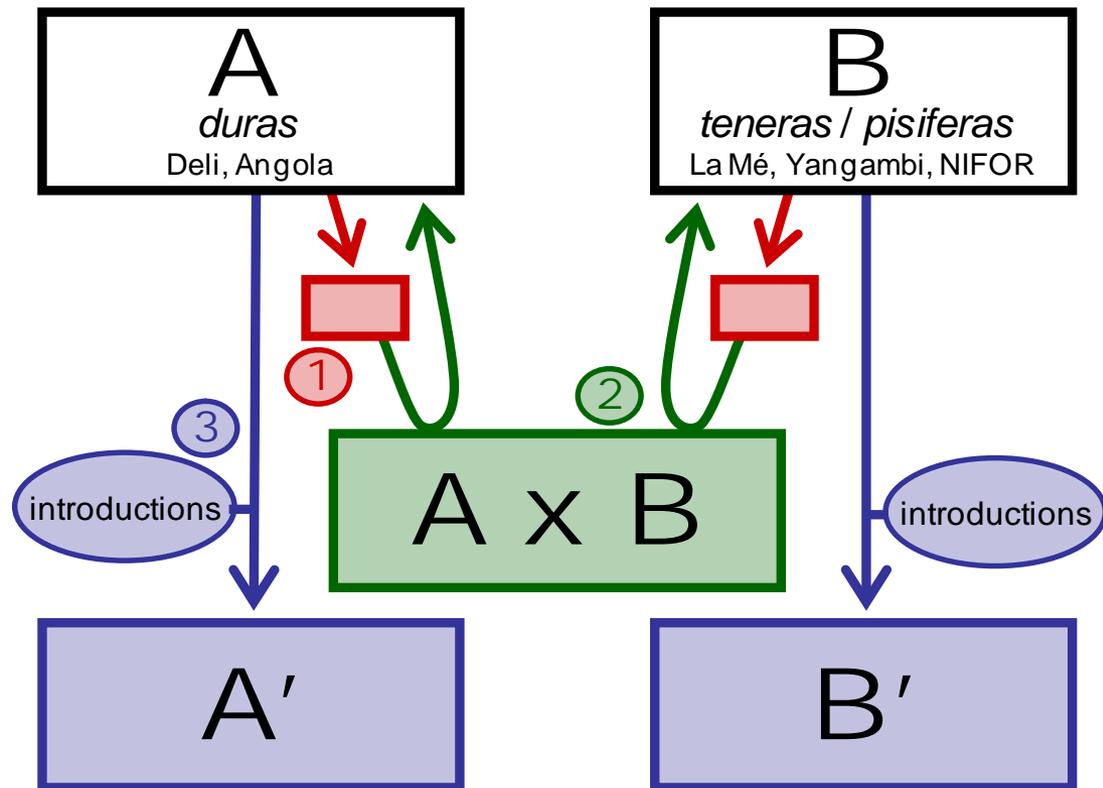
↪ groupes hétérotiques A et B (basés sur PM et NR)

↪ SRR (1957, adapté du maïs) :

① Les populations A et B de départ sont soumises à une **sélection sur les caractères les plus hérissables** (%PF, %OP).

② Les individus A et B sélectionnés sont croisés entre eux et évalués dans des essais génétiques. La **sélection finale** est faite **sur les BLUPs** pour tous les caractères (%PF, %OP, %FR, PM, NR).

③ Les individus sélectionnés sont **autofécondés ou recombinaés** au sein de chaque groupe pour former 2 populations A et B améliorées qui serviront de point de départ au cycle suivant et à produire des semences commerciales.



Connaissances sur le génome

Cartes génétiques
(~ 1700 to 2250 cM)

**Séquence publiée en juillet 2013 (MPOB),
population AVROS :**
26X, 1.8-Gb, +34000 gènes (Singh et al 2013)

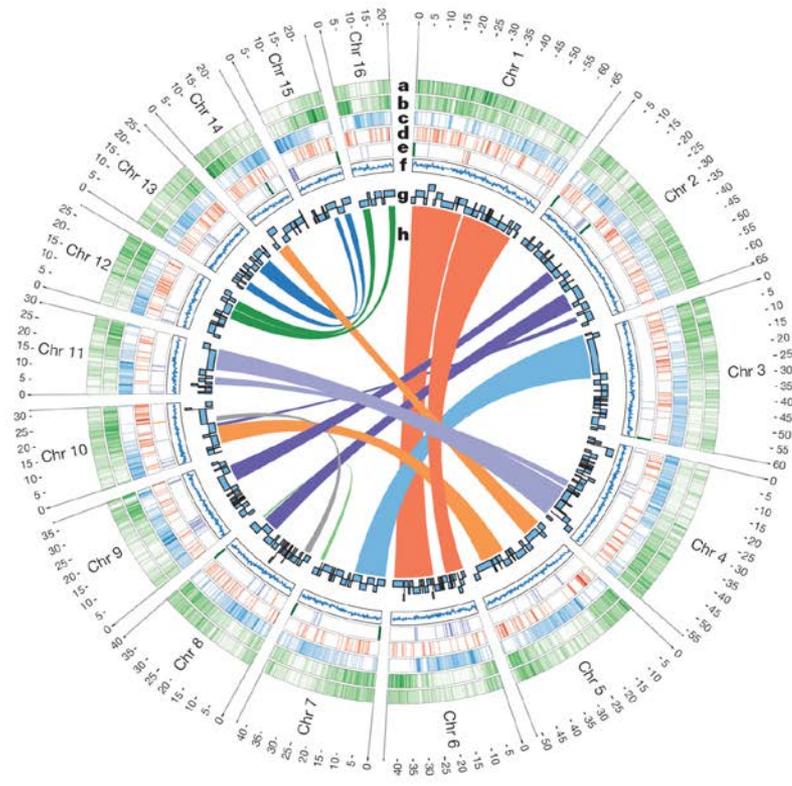
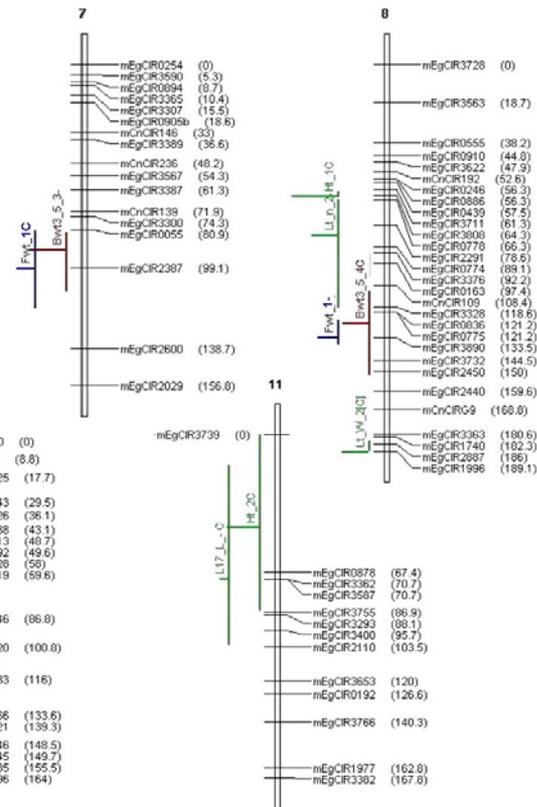


Figure 1 | The chromosomes of oil palm. *E. guineensis* has 16 chromosome pairs, ordered by size, which correspond to 16 linkage groups identified by genetic mapping (Supplementary Table 7). Tracks displayed are: a, gene density; b, methyl-filtered read density; c, retroelement density; d, simple sequence repeats; e, low copy number repetitive elements, including telomere repeat TTTAGG (green), 5S rRNA (orange) and pericentromeric repeats (purple); f, regional G-C content (range 0.3–0.45); g, genetically mapped scaffolds from the P5-build; and h, segmental duplications. Densities for telomere repeats are exaggerated for visual clarity.

QTLs (rendement, développement végétatif, qualité de l'huile)

Séquençage en cours (CIRAD), population Deli



Ressources génomiques & utilisation actuelle

- +300 **SSR** dont SSR d'EST
- **SNP intragéniques** en développement
- 2014 : NGS → marqueurs **Dart + SNP**

- **16 SSR** utilisés en routine pour vérifier la **légitimité**

Perspectives d'application de la SG

Premiers résultats prometteurs...

- 1 **simulation** (Wong et Bernardo, 2008)
- 1 étude sur **données réelles** en cours de publication

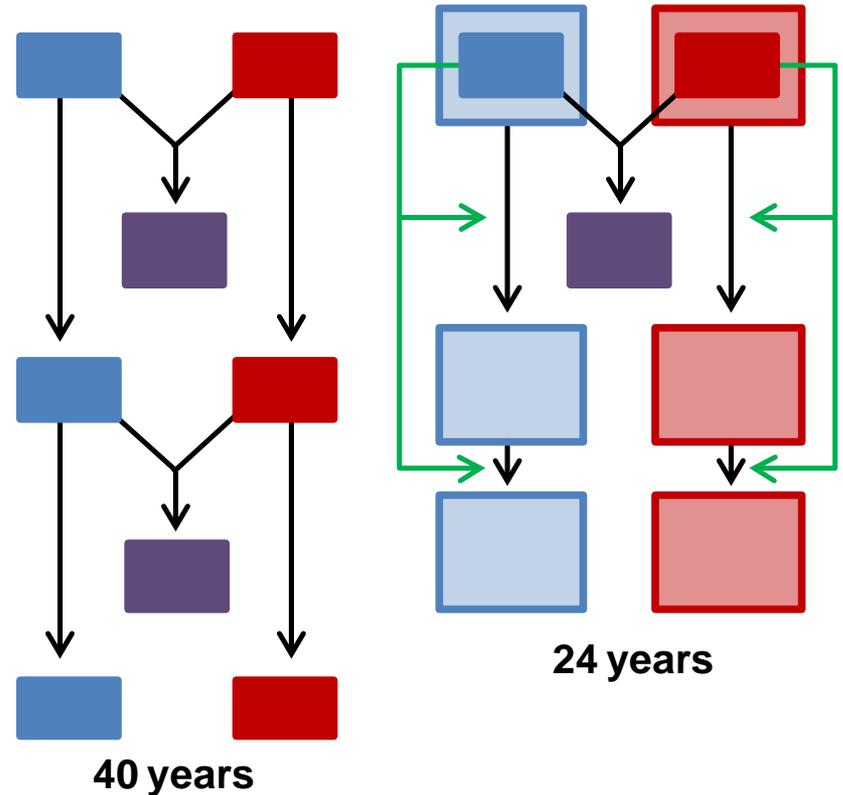
GS:

$$\frac{\text{intensity} * \text{accuracy} * \sigma_a}{\text{generation interval}} = \text{Genetic gain per year}$$

if accuracy > 0.6

with accuracy = $r(\text{TBV}, \text{EBV})$

(accuracy with RRS ~ 0.98)



→ Genomic selection

■ ■ Individuals progeny-tested (RRS) or progeny-tested and genotyped (GS)

■ ■ Genotyped individuals

Spécificités de la sélection génomique chez le palmier

Point de départ...

- intervalle de génération long
- tests en croisement lourds → peu d'individus testés
- faible taille efficace des populations

Place de la SG...

+ réduction possible si précision suffisante

+ augmentation intensité sélection

- petites populations d'apprentissage

+ DL fort limite marqueurs et taille des populations d'apprentissage

Quelle place pour la sélection génomique chez les plantes pérennes ?

Exemple de la vigne

Patrice This

³ INRA-UMR AGAP Montpellier



La vigne et le vin

- 800 000 ha de vigne = 3% de la SAU
- 15 % de la valeur de la production agricole
- 87000 exploitations
- Image forte de la France
- Solde balance commerciale positive (9.5 Mrd € 2012, 2^{ème} place après aéronautique)
- Vigne à vin ou à raisins de table (1% surface)
- Vin AOC, IGP, table
- Principalement *V. vinifera*



Caractéristiques d'importance pour l'amélioration

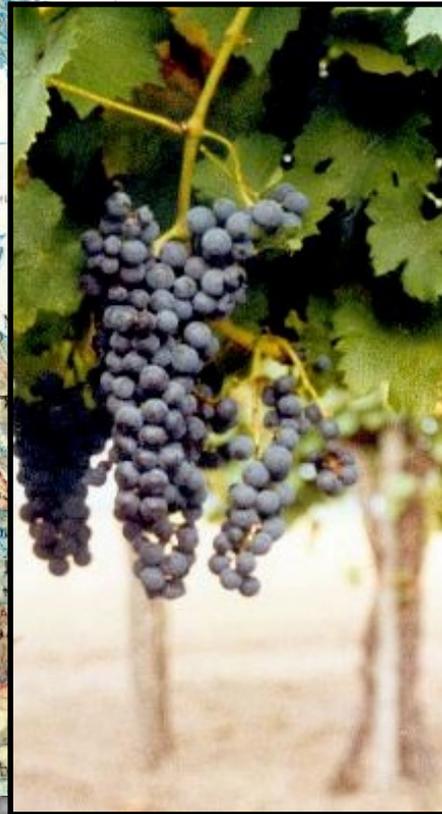
La vigne (*Vitis vinifera*) :

- **diploïde** ($2n = 38$)
- **allogame**
(hermaphrodisme)
- **multiplication végétative très facile, existence d'une variabilité clonale**
- **pollinisation artificielle délicate**
car besoin de castration manuelle
- **cycle de 3 à 4 ans (graine à graine) en condition normale, peut s'accélérer**
- **Genre *Vitis*, espèces interfertiles, origine**
- **Amérique, Asie et Europe (*V. vinifera*)**



Historique d'amélioration de la vigne

- La vigne en 6000 ans
- et de



Vignes sauvages

Vignes Cultivé

Les enjeux de la viticulture française et de la création variétale

Réduire voire supprimer la consommation de fongicides anti-mildiou et anti-oïdium

- 35000 tonnes de matières actives pour 3 % de la SAU
- Coût de plus 220 millions d'euros par an
- Conséquences sur l'environnement, sur la santé humaine, ...
- Apparition de souches de pathogènes résistantes au fongicides

Développer des moyens de lutte vis à vis des agents pathogènes difficiles à contrôler par les méthodes disponibles

- Pertes de rendement dépassant 70 % en cas d'attaque sévère
- 30 % du vignoble touché, soit 240 000 sur 800 000 ha
- Nécessité de rechercher des alternatives à la lutte chimique : impasse technique et réglementaire

Maintenir la qualité et la typicité des vins dans un contexte de changement climatique

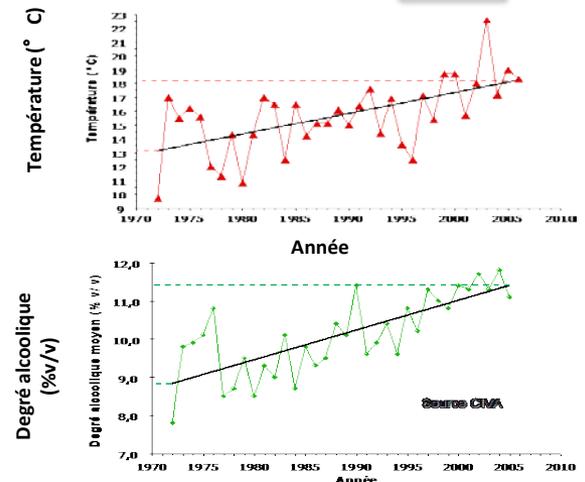
- Elévation de la température pendant la maturation des raisins : 5 °C
- Avancement de la date des vendanges : 4 semaines
- Augmentation du degré alcoolique : 2 degrés
- ...



chardonnay avec protection phytosanitaire

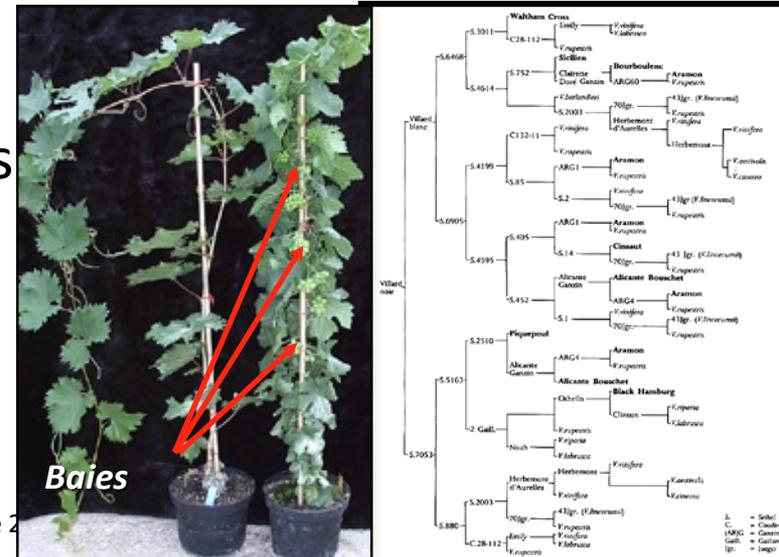


chardonnay sans protection phytosanitaire



La création variétale Vigne

- Publique en France, quelques rares privés en Europe et dans le monde
- Durée de cycle de sélection /évaluation jusqu'à 20 ans
- Différents schémas de sélection
 - les plus courants = croisements biparentaux et sélection massale
 - quelques schémas d'introgression de résistance (7 générations)
 - schémas anciens de croisement inter-spécifique avec large généalogie
- Schémas du futur
 - Pyramidage à l'aide de populations de génotypes Dwarfs
 - Introduction de la SG dans la sélection multilocale



Connaissances sur le génome

Cartes génétiques
(~ 1700 cM)

Séquence publiée en 2007

Pinot PN 40024 (12x) , PN Entav –INRA 115 (6,5x):
12X, 585 Mb, +26500 gènes (Jaillon et al 2007)



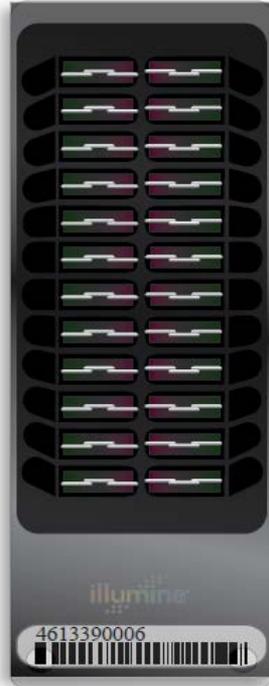
Grape Genome Browser (12X)

Genoscope Home || Vitis Project | B
Contact | Vitis (8X)

Search by position or name

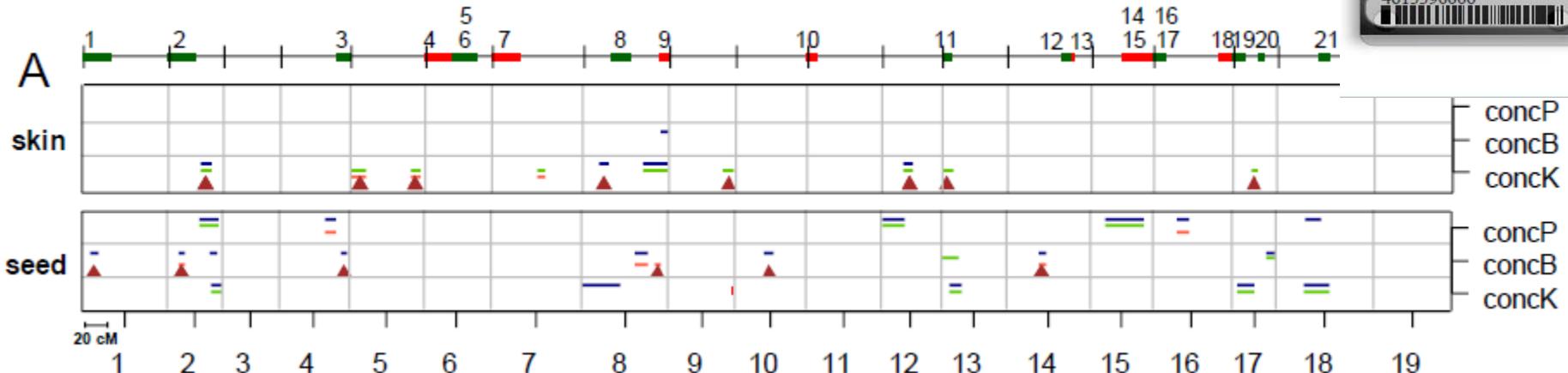
Ressources génomiques

- SSRs (plusieurs centaines)
- SNPs
 - 18k illumina OPA
 - Reséquençage et GBS



QTLs (résistances,
qualité, stress abiotiques)

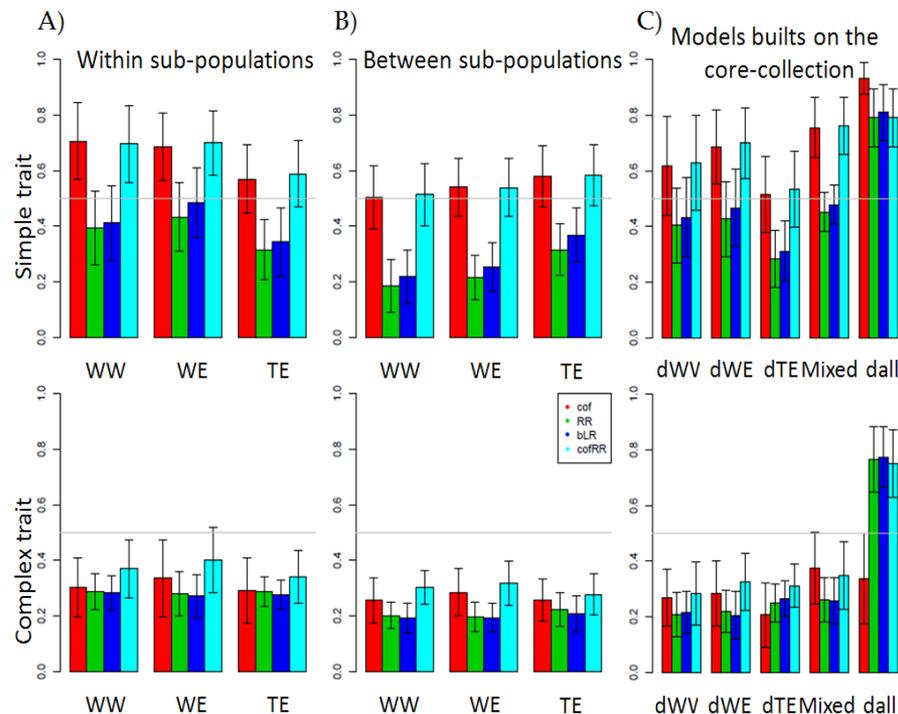
name, locus, or
allowed



Perspectives d'application de la SG

Premiers résultats prometteurs...

- **1 simulation** (Fodor et al, soumis)
 - Etude effet composition population entraînement
 - Etude différentes méthodes de calcul de la prédiction du phénotype
 - Test nombre de marqueurs nécessaires
- **1 étude sur données réelles** en cours d'analyse



Spécificités de la sélection génomique chez la Vigne

Point de départ...

- intervalle de génération long
- Données sur génome
- Interactions G x E
- Variabilité clonale utilisée en sélection clonale

Place de la SG...

+ réduction possible si précision suffisante et permet de réaliser des cycles sans phénotypages

+ intégration des données génomiques dans les analyses SG

+ prédictions des comportements dans différents environnements

+ SG est-elle utilisable dans un contexte de variabilité somatique ?

Quelle place pour la sélection génomique chez les plantes pérennes ?

Exemple du peuplier

Catherine Bastien

INRA-AGPF, Orléans



Importance des peupleraies cultivées en France / Europe

- essence ligneuse la plus productive en climat tempéré: $\sim 15 \text{ m}^3/\text{ha}/\text{an}$
- des **rotations relativement courtes (12-20 ans)** limitant les risques
- production d'un bois homogène à haute valeur ajoutée industrielle (déroulage, biocarburants) en plantations **monoclonales**



~ 800 000 ha dont 236 000 ha
en France



~ 4,8 M m³/an, deuxième
essence feuillue

Une création variétale très majoritairement
publique en Europe

Registre européen des variétés cultivées:

- **145** cultivars (clones) – 85% hyb. intersp.
- **10** variétés multiclones

mais...

seulement moins de 15 clones très utilisés par pays

Une grande inertie au renouvellement variétal

Caractéristiques d'importance pour l'amélioration

- diploïde ($2n = 38$)
- allogame, dioïque, intervalle graine à graine 4-8 ans



- Domestication très limitée et importante diversité génétique entre et à l'intérieur des espèces



P. nigra



P. deltoides



P. trichocarpa

- La multiplication végétative est facile et la foresterie clonale a été acceptée depuis longtemps



- De nombreux avantages à l'hybridation interspécifique: complémentarité (enracinement/résistances aux ravageurs, QB) hétérosis en F1 et F2 (+35% - +73% pour la croissance)
- Des effets d'interactions G x E marqués

Schémas de sélection

Genetic Resources



P. deltoides
n=500
♂/♀

P. nigra
n=1200
♂/♀

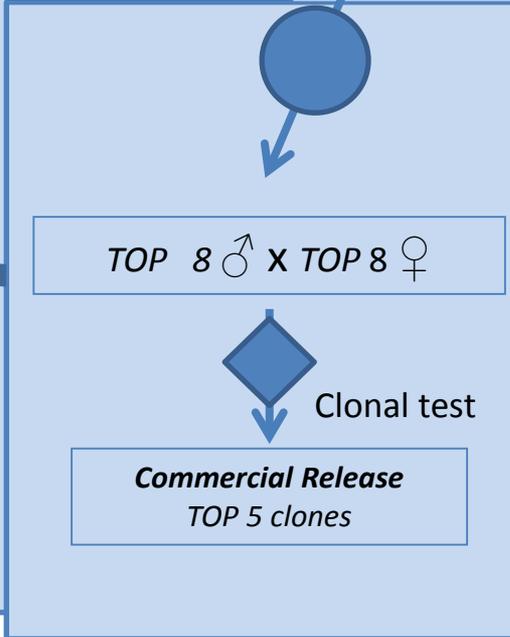
P. trichocarpa
n=600
♂/♀

Clonal test

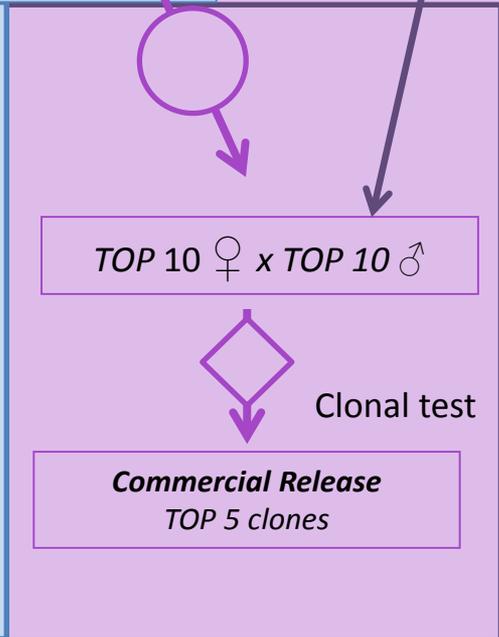


Breeding population
n=120
♂/♀

Intraspecific SR



Interspecific SRR



**2013 :
3 générations**

Schémas de sélection

Une sélection multicritère par étapes pour concilier **intensité de sélection** et **coût de phénotypage**



20-25 crosses/5000 seeds 2500 seedlings



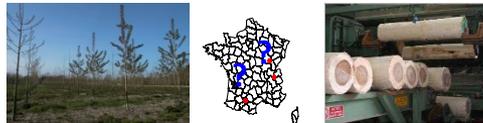
800 clones



800 clones



50 clones



25 clones



5 clones

Durée d'évaluation = 15-20 ans

Coûts de phénotypage **élevés** en plantation et pour les **propriétés du bois**

Interaction Clone x Environnement élevée

Mais peu de corrélations défavorables entre caractères

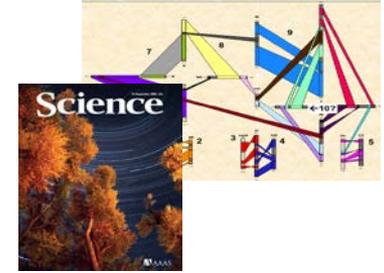


Connaissances sur le génome

- Un génome de référence séquencé et annoté chez *P. trichocarpa*: $n=19$, ~ 480 Mb

Populus: A Model System for Plant Biology

Annu. Rev. Plant Biol. 2007. 58:435–58



- Plusieurs cartes génétiques sur les différentes espèces : 1800 cM- 2800 cM
- Des QTL qui expliquent rarement plus de 10% de V_G sauf pour les résistances aux maladies
- Identification de SNP:

P. trichocarpa (Slavov et al. 2012)

~ 7,000,000 SNP

P. nigra (Giacomello et al. 2012)

~ 3,300,000 SNP

P. deltoides
en cours

- Puces de génotypage : (études d'association en pop. naturelles)

P. trichocarpa:

Slavov et al 2012, Geraldès et al 2013

Infinium Illumina **34 K**

P. nigra

(Noveltree, Energy-Poplar)

Infinium Illumina **10,5K**



aujourd'hui

5 – 17 SNP /cM

Perspectives d'application de la SG

- Déception vis-à-vis de la SAM : un modèle plus proche du modèle de Fisher qui séduit
- Pas de résultats encore disponibles chez **le peuplier** mais des résultats encourageants sur **Eucalyptus** (schéma et contexte génétique proches)

Tree Genetics & Genomes (2013) 9:37–51
DOI 10.1007/s11295-012-0528-1

ORIGINAL PAPER

Efficiency of genomic selection with models including dominance effect in the context of *Eucalyptus* breeding

Marie Denis · Jean-Marc Bouvet

Research

New Phytologist

Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees

Marcos D. V. Resende^{1,2}, Márcio F. R. Resende Jr.¹, Carolina P. Smanoloni^{3,4}, Cesar D. Petrolli^{5,1}, Alexandre A. Missiaggia¹, Aurelio M. Aguiar⁶, Jupiter M. Abad¹, Elizabeth K. Takahashi⁷, Antonio M. Rosado⁸, Danielle A. Faria⁹, Georgios J. Pappas Jr.^{10,11}, Andrzej Kilian¹² and Dario Grattapaglia^{1,2}
¹EMBRAPA Forest Research, Colombo, PR, 81531-000, Brazil; ²Universidade Federal de Viçosa – Viçosa, MG, 36570-000, Brazil; ³EMBRAPA Genetic Resources and Biotechnology – EPqB, 70770-910, Brasília, DF, Brazil; ⁴Universidade de Brasília – Campus Darcy Ribeiro Brasília, DF, 70910-900, Brazil; ⁵IBRIMA Galvão S.A., Rod. Aracruz/Brejo do Riacho, km 25, Aracruz, ES, 29107-000, Brazil; ⁶CENIPA Colômbia Tupy Brasília S.A. Belo Horizonte, MG, 31156-000, Brazil; ⁷Universidade Católica de Brasília – UCB, Vité mod. B, Brasília, DF, 70790-000, Brazil; ⁸INIA T – Forestry Area Technology, P.O.B. 7116, Tucumán, A.C.T., Avenida 2000

- Une situation expérimentale en cours d'évaluation chez *P. nigra*

	Objectifs	Population d'entraînement	Population de validation
N	Précision de la prédiction génomique G_{intra} , BV_{intra} Multi-trait Apparemment ↳ Cycle ↗ intensité	$N_e = 15$ $N = 1200$ Phénotypage acquis Génotypage 50% 10K	apparemment variable à tester Schéma multi-population
DxN	Précision de la prédiction génomique G_{inter} , BV_{inter} Non additivité ↳ Cycle ↗ intensité		$N_{e \text{ nigra}} = 10$ $N = 600$ Phénotypage acquis

Spécificités de la sélection génomique chez le peuplier

Point de départ...

- intervalle de génération long
- Nbx caractères complexes peu héritable
- Diversité issue de ≠ populations
- Forte variabilité clonale surtout en xt interspécifiques

Place de la SG...

+ réduction possible si précision suffisante et recombinaison sans phénotypage
+ de réactivité vis-à-vis de nouvelles demandes

+ améliorer la prédiction si densité en marqueurs suffisante
+ optimiser le phénotypage et la S multi-caractères

+ si précision suffisante en modèle multi-populations (origine DL)

+ augmentation de l'intensité de sélection
+ meilleure gestion de la consanguinité
+ si prédiction inclut les effets non-additifs

Conclusions - plantes pérennes

- Une **grande attente** pour la **SG** après les espoirs déçus de la **SAM**
- Principaux **verrous / questions / limites** actuels

1- Définir des populations d'entraînement

taille efficace, apparemment avec pop. de sélection, multi-populations

2- Continuer à optimiser le géotypage

Développement de SNP, imputation, géotypage par re-séquençage, coût

3- Travailler sur les modèles de prédiction

Prise en compte de GxE

Inclure si disponible l'information sur QTL connus

Inclure les effets non additifs

Prédiction pure-bred / prédiction cross-bred

Prédiction multi-populations

Mise à jour des prédictions au fil des générations