

Apport du GBS chez les espèces fourragères

Philippe Barre

INRA Centre Poitou-Charentes

Unité de Recherche Pluridisciplinaire Prairies et Plantes Fourragères

URP3F



GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG

GGAGAG - AACAAAGAT - AAATT - AGCAGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGCAGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGCAGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGCAGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG
GGAGAG - AACAAAGAT - AAATT - AGT AGTGAAATCC - AAGC - A - TTACAG



Particularités biologiques

Forte allogamie, auto-incompatibilité et forte dépression de consanguinité

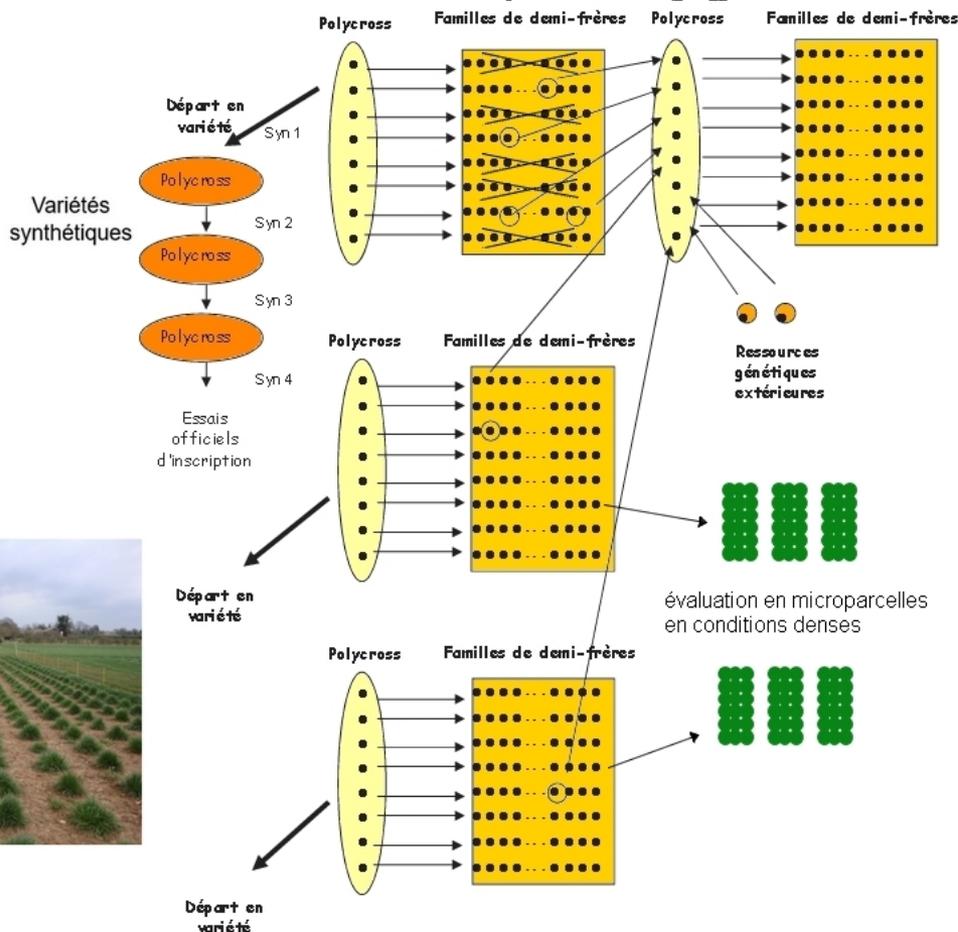
Très polymorphe:
15 pb/ SNP

Décroissance très rapide du déséquilibre de liaison

Tailles de génome :
- ray-grass anglais (2n=14) : 1C = 2,7 Gb
- luzerne (4X=32) : 1C = 0,9 Gb
référence: *Medicago truncatula* 1C = 0,5 Gb



Schéma de sélection classique du ray-grass



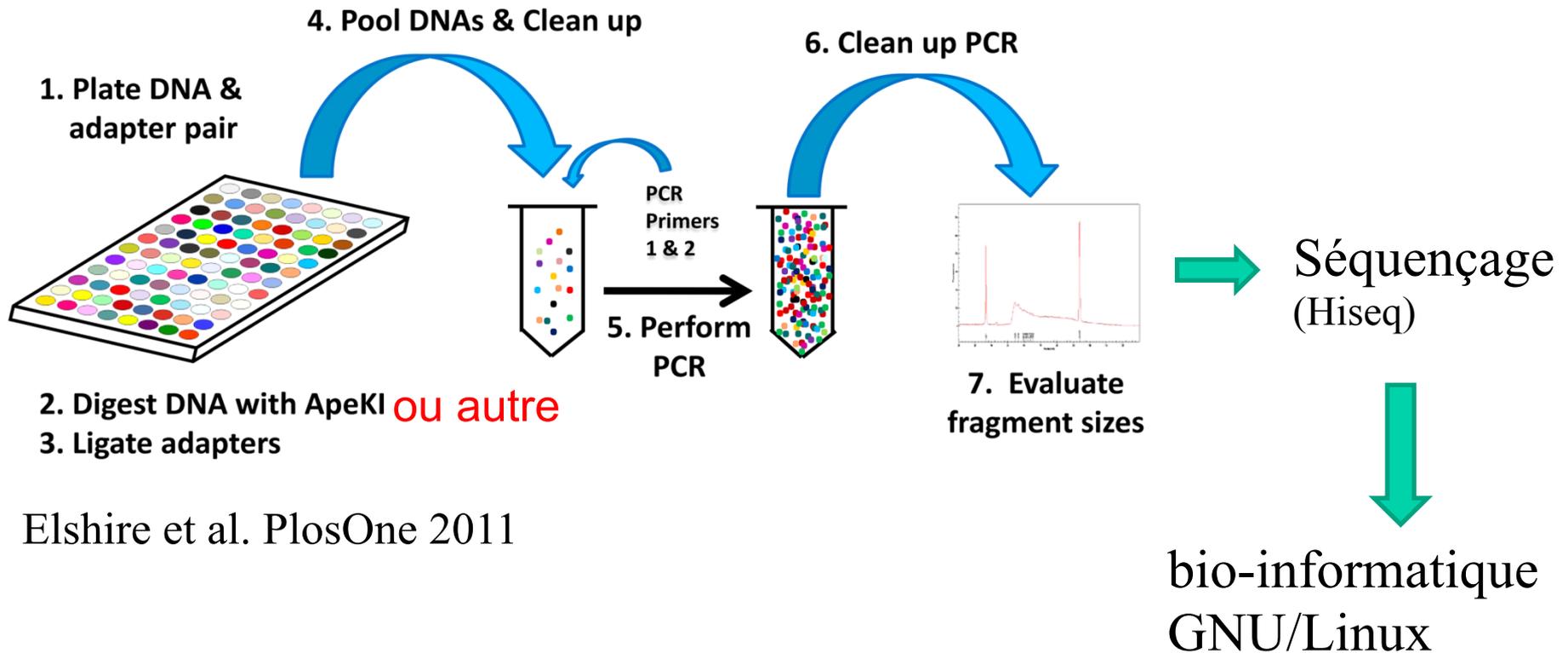
D'après Pauly (2012)

Choix de la méthode GBS/

génotypage SNP, capture de séquences

- Grand nombre de marqueurs (environ 10000)
- Pas besoin d'information de séquences préalable
- Evite les difficultés de définition d'amorces ou de sondes dans des régions très polymorphes
- Permet l'identification de nouveaux SNP
- Permet le génotypage de pool d'échantillons
- Peu onéreuse
- Expérience chez le RGA (T. Asp Univ. Aarhus)

Méthodes



Elshire et al. PlosOne 2011

- De-multiplexage
- Tri / qualité (30), sup. adaptateurs, longueur min. (60b) (Sickle, Scythe)
- Alignement sur séquence de référence (bwa, Picard, GATK)
- Détection des SNP biallélique (profondeur 10 ou 30 pools, MAF 5%) (GATK)
- Mise en forme des données (awk, python)

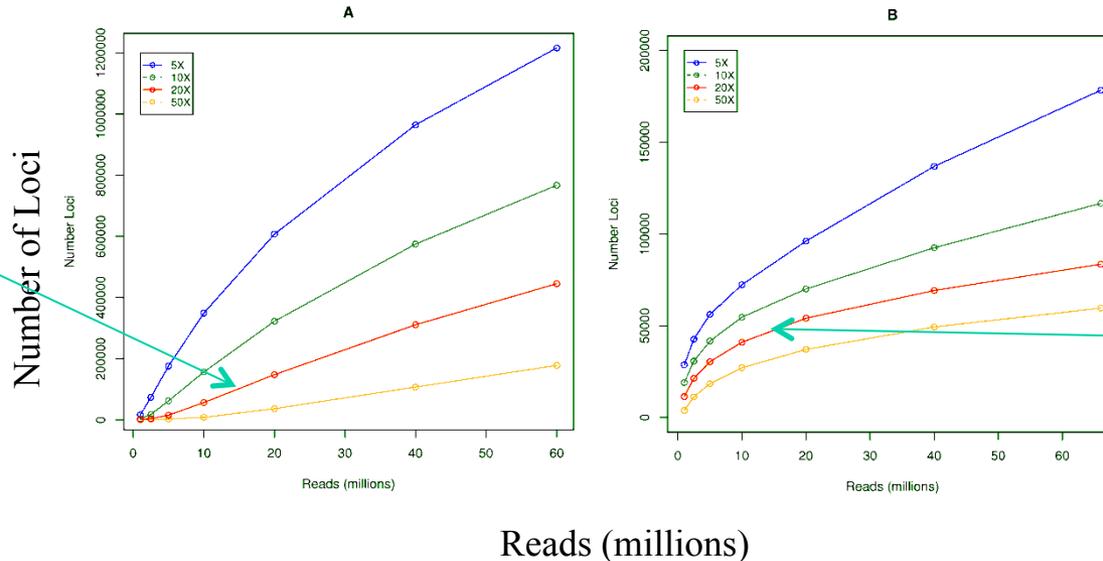
Cohérence entre l'enzyme de restriction et l'effort de séquençage

ApeKI (5 bp recognition site)

PstI (6 bp recognition site)

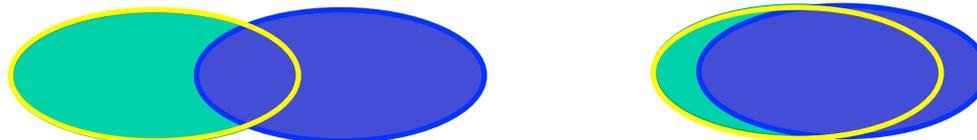
Lolium perenne

Byrne et al., PLOS, 2013



100000 marqueurs
Bcp de DM sur
l'ensemble de la
population

50000 marqueurs
Peu de DM sur
l'ensemble de la
population



Faire un test avec différentes enzymes de restriction pour définir le plateau : nb de loci / nb de séquences

Exemple: sélection génomique chez le ray-grass anglais

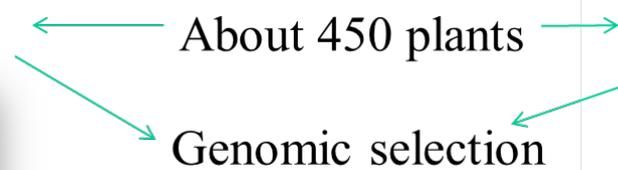


- Projet CropDL du métaprogramme SELGEN Gilles Charmet INRA Clermont
- URP3F: Sabrina Delaunay, Magali Caillaud, Camille Gréard ...
- EPGV: Marie-Christine Lepaslier, Aurélie Bérard ...
- Univ. d'Aarhus: Stephan Hentrup Istvan Nagy, Stephen Byrne, Torben Asp
- MIAT: Charles-Elie Rabier, Brigitte Mangin

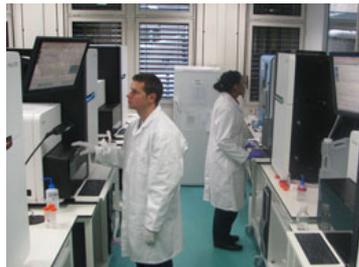
Population: Unlike (Gie GRASS)



Genotyping by GBS
with Pst1



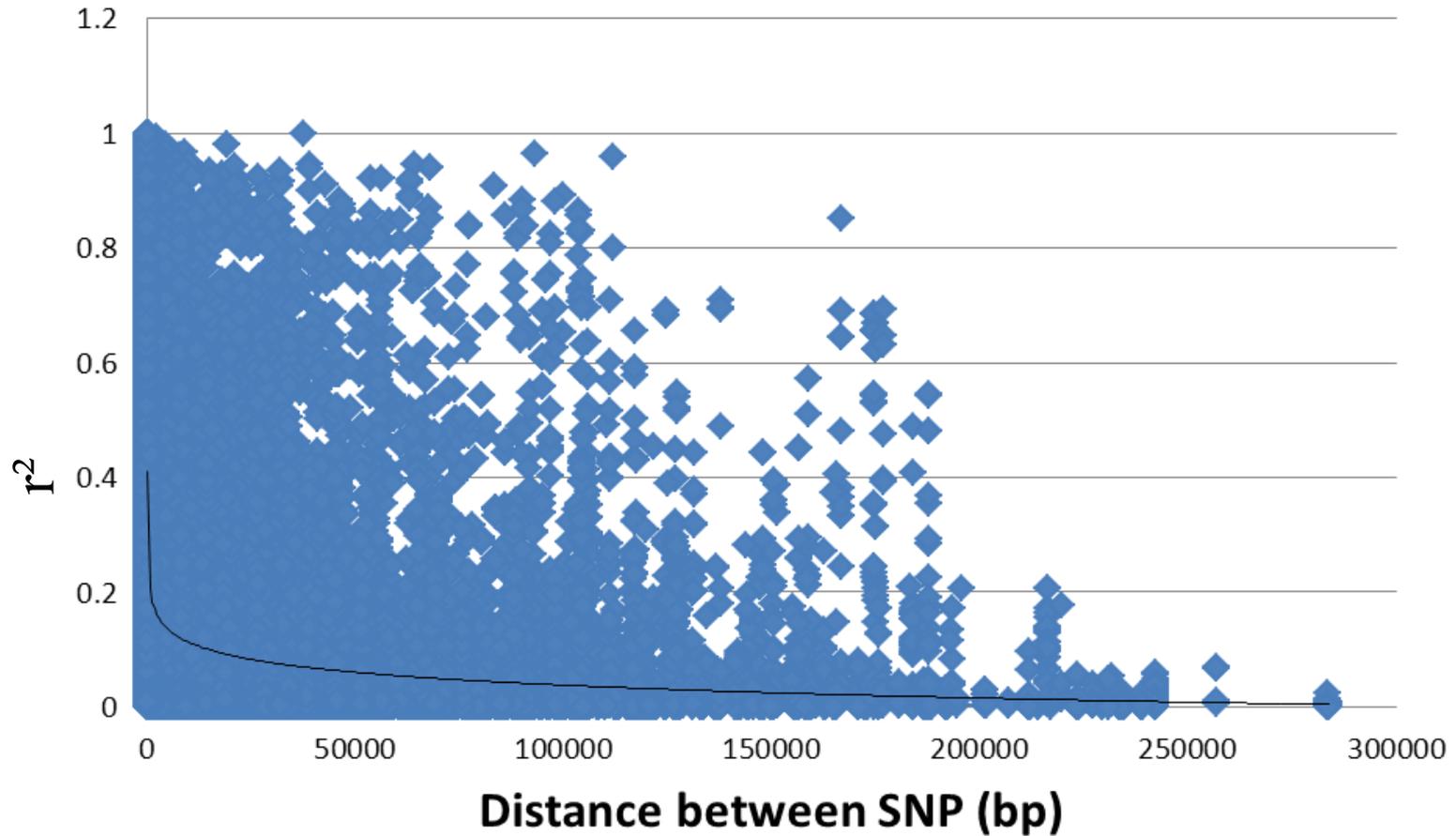
Phenotyping:
- Leaf length
- Plant height and heading date



17878 marqueurs
Répétables
40 % DM

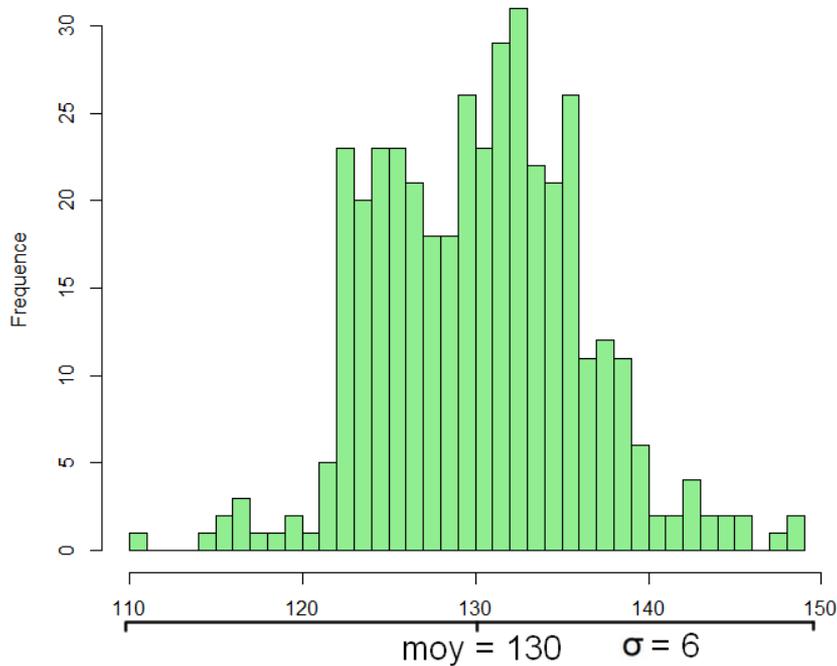
2 M reads / geno

Etendu du déséquilibre de liaison

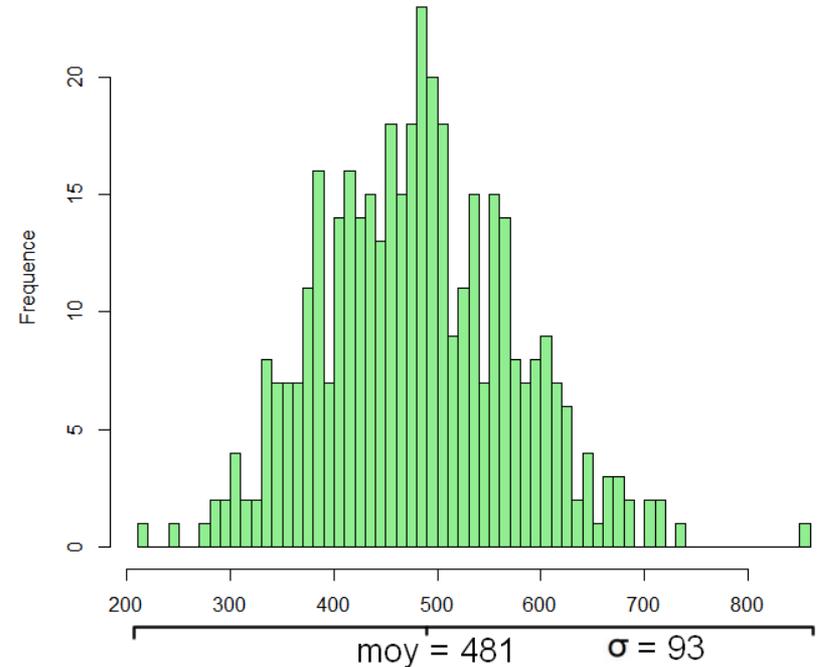


Données phénotypiques

Date d'épiaison



Longueur de feuille (mm)



	Date d'épiaison	Longueur de feuille
H ² sens large	0,96	0,59

Prédiction avec RR Blup

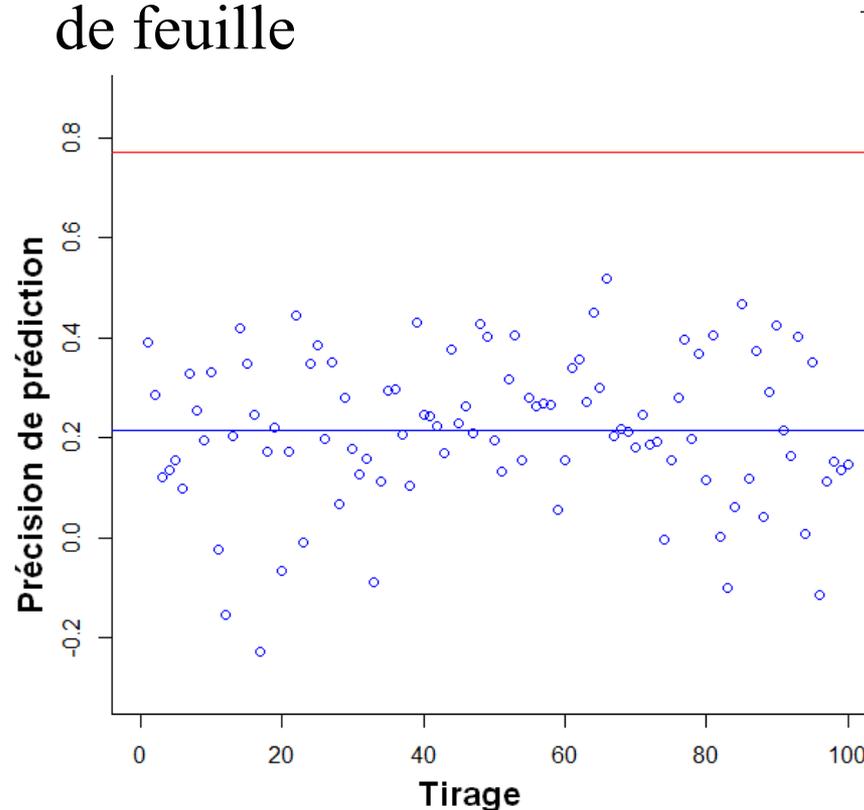
Précision de prédiction :
corrélation entre données prédites
et observées par validation
croisée (40/360 indiv)

Précision moyenne (100 tirages)

Date d'épiaison	0,40
Longueur de feuille	0,22

- Forte hétérogénéité entre tirages
- Précision variable entre caractères
- Amélioration :
 - augmenter le nombre de marqueurs
 - Diminuer le nombre de DM

Précision de prédiction pour
100 tirages pour la longueur
de feuille



Exemple: différenciation de variétés de luzerne



- Projet Amédiluze CASDAR Vincent Gensollen GEVES
- URP3F: Bernadette Julier, Sabrina Delaunay, Paola Lambroni ...
- GeT Plage: Olivier Bouchez, Clémence Genthon...
- Problématique: 21% des variétés déposées sont rejetées DHS alors que 97% sont acceptées VATE

Outil pour aider à la différenciation entre variétés :
marqueurs GBS ? Géotypage en pools ?



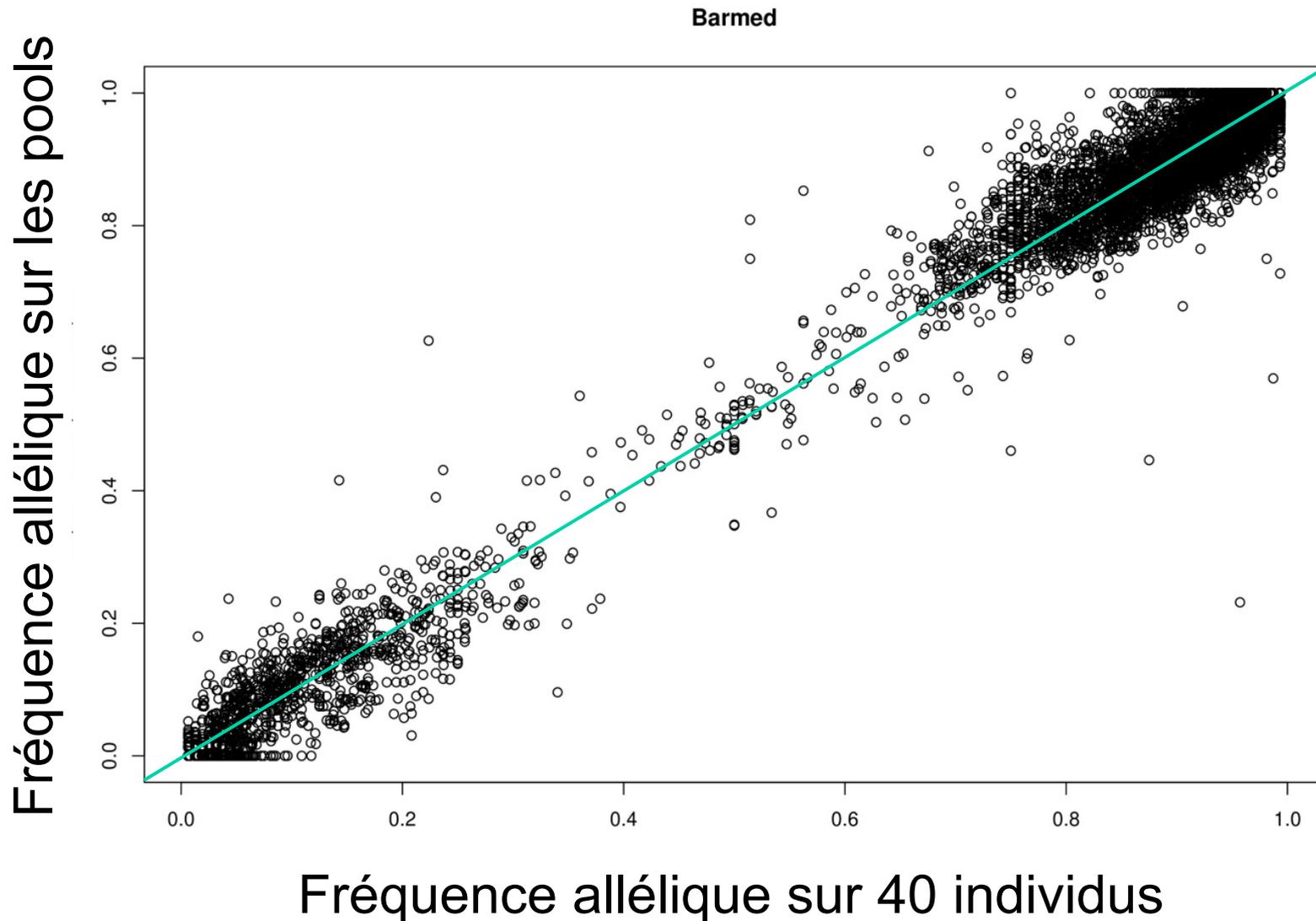
MINISTÈRE
DE L'AGRICULTURE
DE L'ALIMENTATION
DE LA PÊCHE
DE LA RURALITÉ
ET DE L'AMÉNAGEMENT
DU TERRITOIRE

*avec la contribution financière du
compte d'affectation spéciale
"Développement agricole et rural"*

Génotypage GBS

- Sur 3 variétés, représentées par 40 individus: GBS individuel
- Sur 20 variétés, représentées par 400 individus: 4 extractions d'ADN en pool de 100 par variété, GBS par pool
- Test Pst I / ApeKI : meilleurs résultats avec ApeKI
- Séquençage 8 lignes HiSeq 3000 : > 250 M reads / ligne
- 10 – 20 M reads / éch. → 3.7 – 4.6 M reads / éch. après alignement sur *M. truncatula*
- 35000 à 41000 SNP

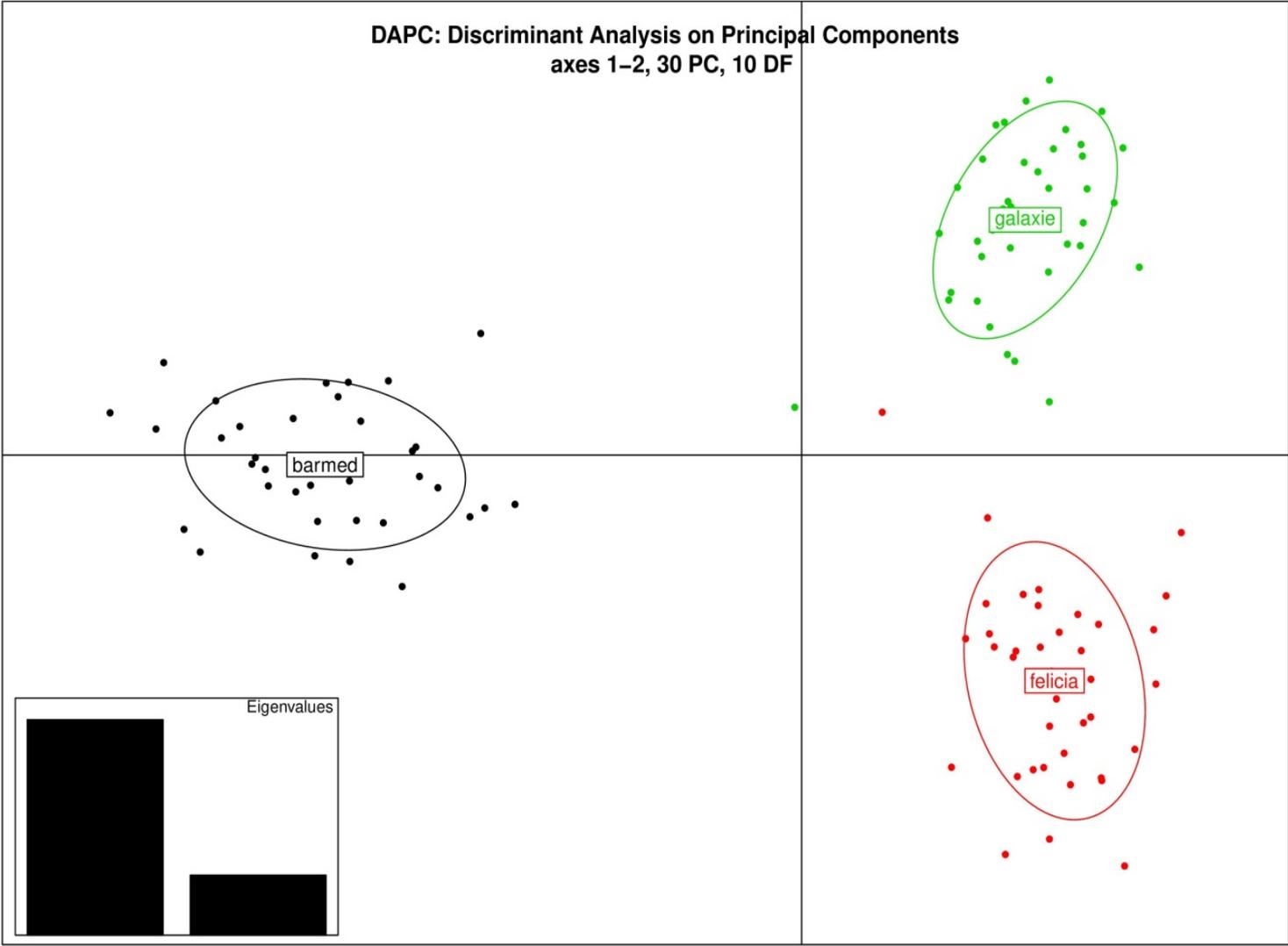
Correspondance entre génotypage sur les pools et génotypage sur 40 individus par variété



$$y = x$$

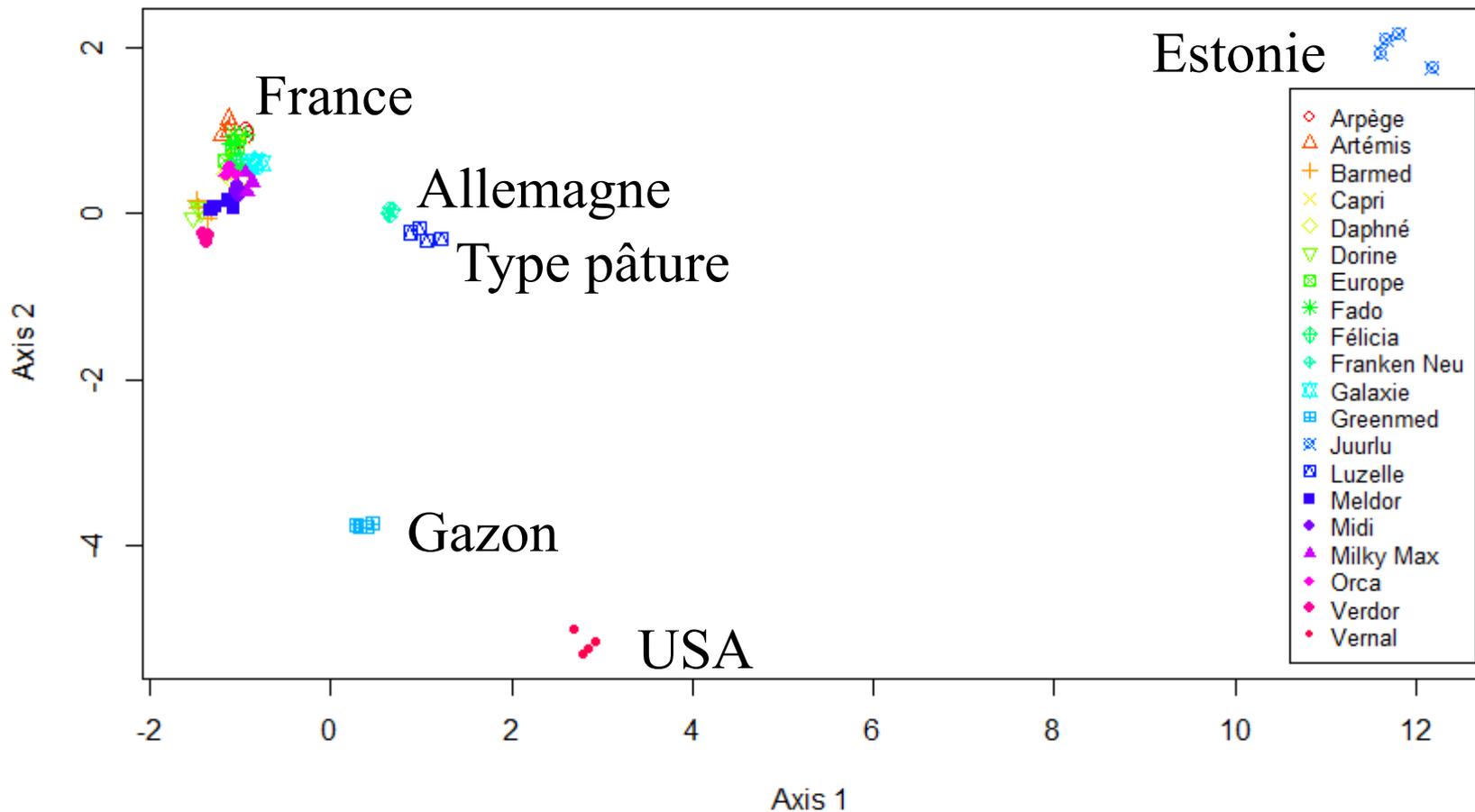
Distinction entre les 3 variétés dont 40 individus ont été génotypés individuellement

PCoA: Principal Coordinate Analysis



Distinction entre les 20 variétés représentées par 4 pools

PCA of data lucerne axes 1-2



Conclusions

- Le GBS permet d'obtenir un nombre important de marqueurs sans trop de DM si le séquençage est adapté à l'enzyme de restriction
- Temps d'obtention des données brutes : rapide mais nécessite du temps d'analyse info.
(pipeline)
- Coût dépend du nombre de marqueurs souhaités
- Bonne estimation des fréquences alléliques en pools: répétabilité et exactitude



Merci pour votre attention

