



Rules&Tools

Tâche 2: modelling LD and estimating IBD

GABI: Didier Boichard, X1=Pilar Schneider/Julien Jacquin, H el ene
Gilbert

SAGA: Jean-Michel Elsen, Carole Moreno, Simon Teyss edre



Objectifs

Amélioration des méthodes de prise en compte du DL dans la cartographie

Sous tâche 1: Description du DL et estimations de proba IBD (1)

Méthodes d'identification des allèles QTL identiques

→ Similarité entre haplotypes

Exploration des tailles d'haplotypes ?

Sous tâche 2: Utilisation des informations pedigree pour l'estimation de l'identité

Relations de parentés connues entre les fondateurs génotypés

Sous tâche 3: Quelle approche spécifique aux croisements expérimentaux

LD intra race et entre race co-existent



Utilisation de l'origine raciale pour distinguer des haplotypes

→ deux ou trois générations génotypées
(croisements expérimentaux ...)

→ Carole

Sous tâches 1&2&3: approches

→ **Travail biblio**

→ **Développements algébriques/code**

→ **Tests des propriétés sur simulations**

→ **Applications aux données réelles**

→ Homogénéisation avec tâches 1 et 5

Pilar

→ Sous-tâches 1 et 2, janv 2010

→ Étape 1: investigation biblio
+ propriétés des méthodes



→ Étape 2 : comparaison sur simulations

→ Étape 3: comparaison sur un set de pedigree réels contrastés

Liens forts avec les tâches 1, 5 et 6, 3 et 4

Etape 2: analyse et comparaison des méthodes

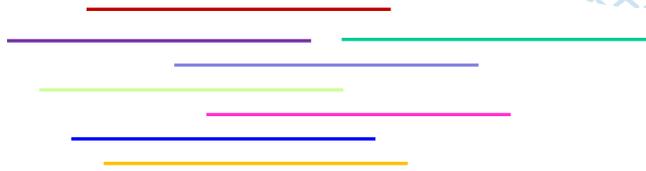
- Simulation de données
- Sélection de méthodes de calcul de $s(\text{IBD})$
- Comparaison des $s(\text{IBD})$ avec les statuts IBD connus
- Comparaison des $s(\text{IBD})$ pour la cartographie de QTL
- Utilisation sur données réelles

Simulation de données (1)

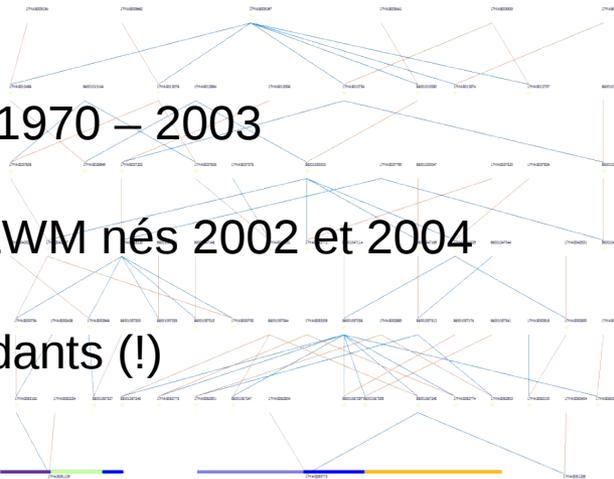
235 chromosomes LWF

(ou 1252 toutes races) reconstruits
dans le cadre de DELISUS

SSC18, 1252 SNP



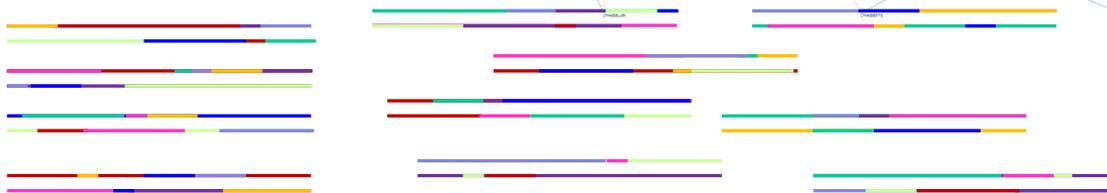
1593 ancêtres nés 1970 – 2003



PEDIG

206 males et 1786 femelles LWM nés 2002 et 2004

31813 descendants (!)



→ Statut IBD connu entre
régions simulées

Simulation de données (2)

SSC18, 1252 SNP



Description DL

SNP en fort DL

SNP en faible DL

avec ses voisines



$y = \frac{1}{2}$ polygénique père + $\frac{1}{2}$ polygénique mère + aléas méiose
+ effet génotype SNP + résiduelle

$h^2 = 0,25$

Effet QTL = 0,5 écart type génétique

Méthodes de calcul de s(IBD) (1)

→ Statut IBS : 0 ou 1

→ Mesure de similarité de Li and Jiang (2005) transformée en distance $d_{i,j}=(s_{i,i}-s_{i,j})/s_{i,i}$

$$s_{i,j} = \sum_{k=-l}^r w_1(x_k) I(h_i(k), h_j(k)) + \sum_{k=1}^{r'} w_2(x_k) + \sum_{k=-1}^{-l'} w_2(x_k),$$

nb allèles en commun
entre h_i et h_j

longueur à gauche et à droite des
allèles partagée entre h_i et h_j

Poids w_1 et w_2 = distance entre les marqueurs? DL entre les marqueurs D'?

Dans Li and Jiang (2005), mesure utilisée dans un algorithme de clusterisation DSCAN (Density Based Spatial Clustering of Applications with Noise)

→ Certains haplo clusterisés ensemble, autres = bruit

→ Q-score appliqué aux haplo clusterisés uniquement

m haplo dans cluster c , n haplo dans reste r

Tests par permutation

$$Q = \frac{(\mu_c - \mu_r) \sqrt{m+n-2}}{\sqrt{(\sigma_c^2 + \sigma_r^2) \left(\frac{1}{m} + \frac{1}{n} \right)}} \approx testT$$

Méthodes de calcul de s(IBD) (2)

- Statut IBS : 0 ou 1
- Mesure de similarité de Li et Jiang (2005) transformée en distance $d_{i,j} = (s_{i,i} - s_{i,j}) / s_{i,i}$
- Probabilités IBD de Meuwissen et Goddard (2001)

A Marqueur


$$P(\text{IBD}|\text{marker}) = P(\text{IBD}|S) = \frac{P(A = \text{IBD} \ \& \ S)}{P(A = \text{IBD} \ \& \ S) + P(A = \text{nonIBD} \ \& \ S)}$$

S = indicateur d'état IBS au marqueur

$$P(\phi = [1_1]) = f(c) = \frac{\exp[-2c]}{2N_e} \times \frac{1 - \exp\left[-T\left(2c + \frac{1}{2N_e}\right)\right]}{1 - \exp\left[-\left(2c + \frac{1}{2N_e}\right)\right]}$$

c = intervalle en Morgan

T = temps depuis la 1^{ère} génération

N_e = taille efficace de la population

→ généralisation 

→ intégration de la transmission au niveau du pedigree

Méthodes de calcul de s(IBD) (3)

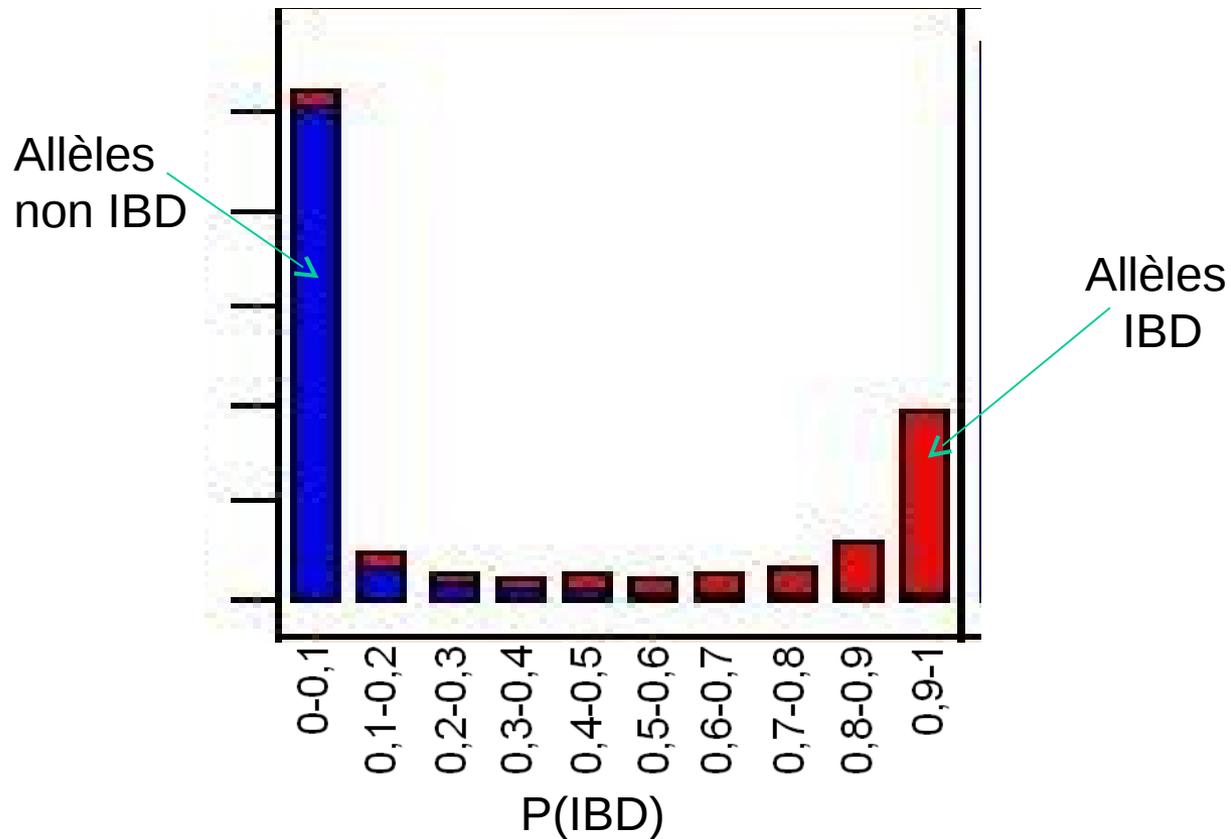
- Statut IBS : 0 ou 1
 - Mesure de similarité de Li et Jiang (2005) transformée en distance $d_{i,j} = (s_{i,i} - s_{j,j}) / s_{i,i}$
 - Probabilités IBD de Meuwissen et Goddard (2001)
 - Imputation et reconstruction de phases par HMM
- quantité à utiliser pour quantifier la ressemblance entre haplotypes pas encore claire

Méthodes de calcul de s(IBD) (3)

- Statut IBS : 0 ou 1
- Mesure de similarité de Li et Jiang (2005) transformée en distance $d_{i,j} = (s_{i,i} - s_{j,j}) / s_{i,i}$
- Probabilités IBD de Meuwissen et Goddard (2001)
- Imputation et reconstruction de phases par HMM
- PHASEBOOK (Druet and George, 2010)

Comparaison des $s(\text{IBD})$ avec les statuts IBD connus

→ 5 quantités à comparer aux statuts IBD entre allèles au QTL



Comparaison des s(IBD) pour la cartographie de QTL

modèle mixte avec effet polygénique, haplotype=effet aléatoire



Matrice de variance covariance entre haplotypes

efficacité du calcul
précision de cartographie
puissance de détection
taux de faux positif obtenu
robustesse des méthodes

Populations simulées

- étendue du pedigree disponible (avec/sans informations de typage)
- informations disponibles sur la population (banque d'haplotypes ?)
- présence du locus causal parmi les typages
- structure de DL dans la région

Données réelles

Voir tâche 5