

Jeudi 20 décembre 2018



Université Claude Bernard



Caractérisation de la diversité génétique du gène PRDM9 dans les populations humaines

Mélanie Bonhomme
Arnaud Duvermy
Loïc Guille
Rémi Séraphin

Maître d'ouvrage :
Pr Laurent Duret, directeur de recherche du CNRS
UMR CNRS 5558 - LBBE "Biométrie et Biologie Évolutive"

Sommaire

- I. Contexte biologique
- II. HLAsan
- III. Méthode sur les jonctions
- IV. Conclusion

I. Contexte biologique

- PRDM9 code pour une protéine à doigts de zinc
- Détermine les points chauds de recombinaison homologue
- Gène qui évolue le plus rapidement parmi les gènes de mammifères
- Très fort polymorphisme

➡ Etudier la diversité génétique du gène PRDM9 dans de nouvelles population grâce au projet 1000Genomes

➡ Développer procédure pour déterminer les allèles d'un individu

PRDM9

Localisation : chromosome 5

Longueur : 85011 pb

Nombre d'exons : 11

Exon 11 est constitué de motifs de 84 paires de bases répétées en tandem

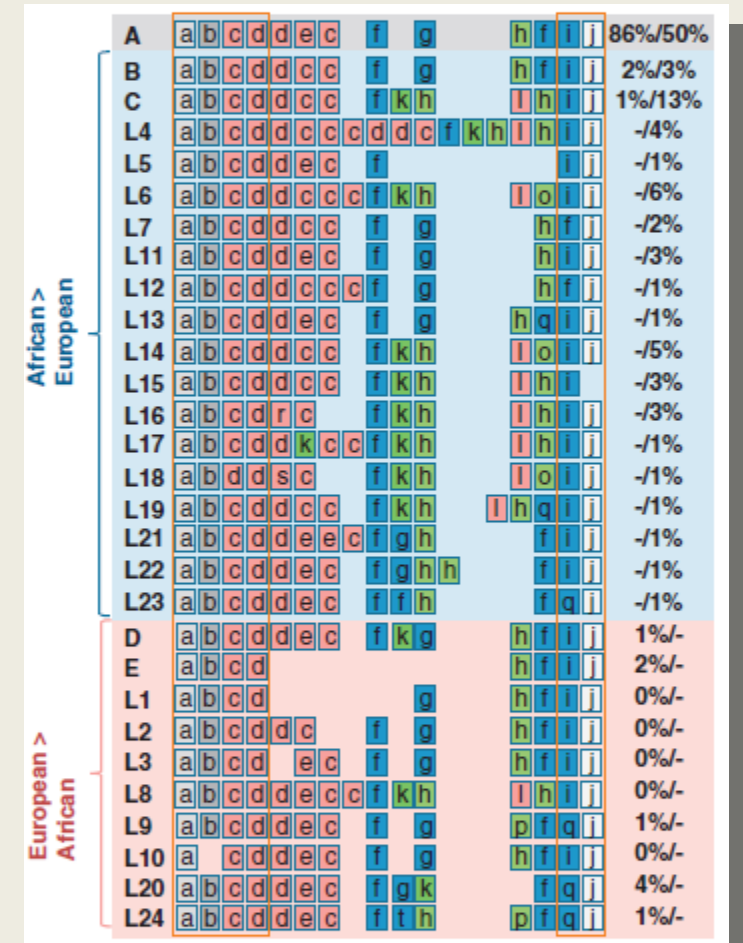
Exemple :

motif A

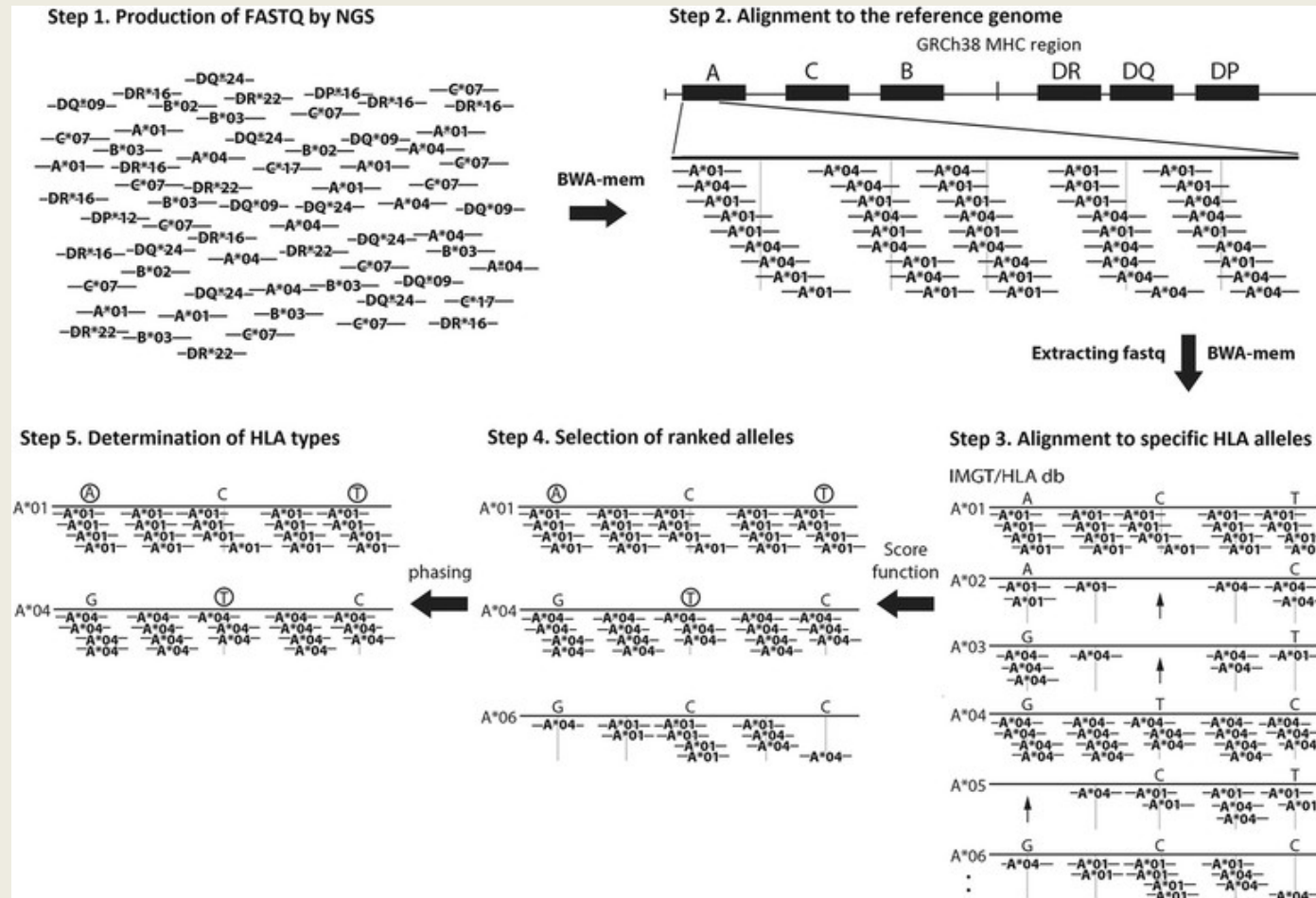
« TGTGGACAAGGTTTCAGTGTTAAATCAGATGTTATTACACACCAAAGGAC
ACATACAGGGGAGAAGCTCTACGTCTGCAGGGAG »

motif B

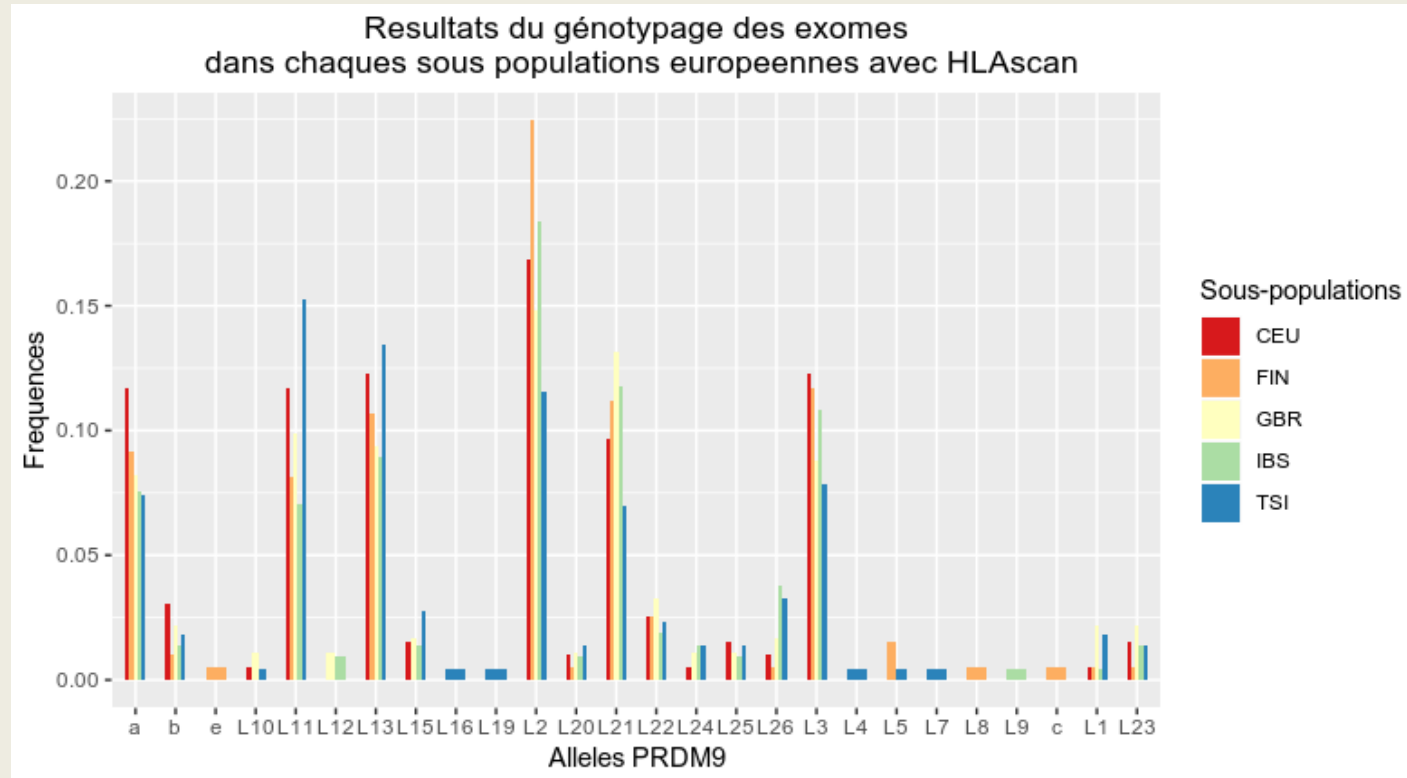
« TGTGGGCGGGGCTTTAGCTGGAAGTCACACCTCCTCATTACACAGAGG
ATACACACAGGGGAGAAGCCCTATGTCTGCAGGGAG »



Logiciel HLAscan



Résultats HLAscan



Problèmes données low coverage du projet 1000 Génomes
Sur de génotypes connus on retrouve .. Erreurs

Méthode sur les jonctions

1. Création des profils :

- Base de donnée pour toutes les jonctions existantes
- Génération de reads de 100 nucléotides de long pour chaque génotype
- Alignement exact des reads sur les jonctions
- Comptage du nombre de reads pour chaque jonction

```
1 Jonctions/jonctionsCF.fa 6711
2 Jonctions/jonctionsHF.fa 6544
3 Jonctions/jonctionsDD.fa 6136
4 Jonctions/jonctionsCD.fa 5954
5 Jonctions/jonctionsEC.fa 5832
6 Jonctions/jonctionsBC.fa 5695
7 Jonctions/jonctionsPF.fa 5096
8 Jonctions/jonctionsDE.fa 4875
9 Jonctions/jonctionsKF.fa 4673
10 Jonctions/jonctionsFG.fa 4673
11 Jonctions/jonctionsFI.fa 4667
12 Jonctions/jonctionsAB.fa 4658
13 Jonctions/jonctionsGH.fa 4580
```

Exemple de fichier de sortie pour le profil AA

Méthode sur les jonctions

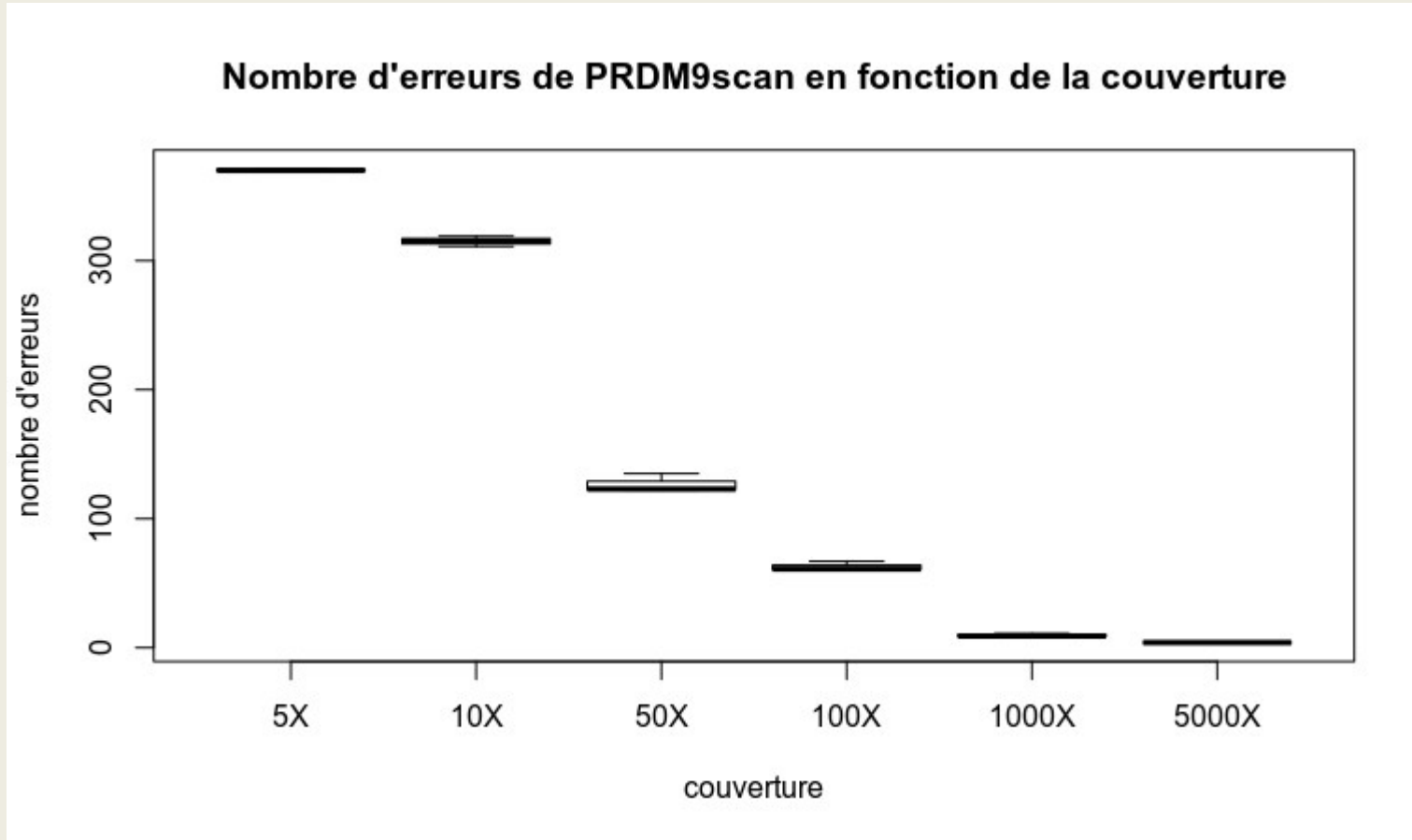
2. Détermination du génotype d'un individu

- Génération de reads pour individu X
- Alignement exact des reads sur les jonction
- Comptage du nombre de reads pour chaque jonctions
- Comparaison aux différents profils grâce à une distance :

$$distance = \sum \left| \frac{\text{nombre de reads sur jonction } y \text{ de l'individu } X}{\text{nombre de reads total alignés de l'individu } X} - \frac{\text{nombre de reads sur jonction } y \text{ du profil}}{\text{combre de reads total alignées sur le profil}} \right|$$

- La distance la plus petite correspond au profil le plus probable

Résultats méthode jonctions



Conclusion

☾ Test de 2 méthodes

→ Méthode sur les jonctions efficace à des couvertures élevées

→ Piste d'amélioration : fixer un seuil pour les profils

→ Données 1000 génomes inutilisables pour le moment à cause de la méthode de capture et de la trop faible profondeur