

Inférence de la démographie et de la sélection à partir de données génomiques temporelles

Mathieu Uhl^{1,2}, Paul Bunel^{1,3}, Miguel de Navascués¹
Simon Boitard¹, Bertrand Servin³

¹ INRAE, CBGP, Montpellier, France

² CNRS, CEFÉ, Montpellier, France

³ INRAE, GenPhySE, Toulouse, France

Semaine de restitution DevOCGen
Décembre 2025

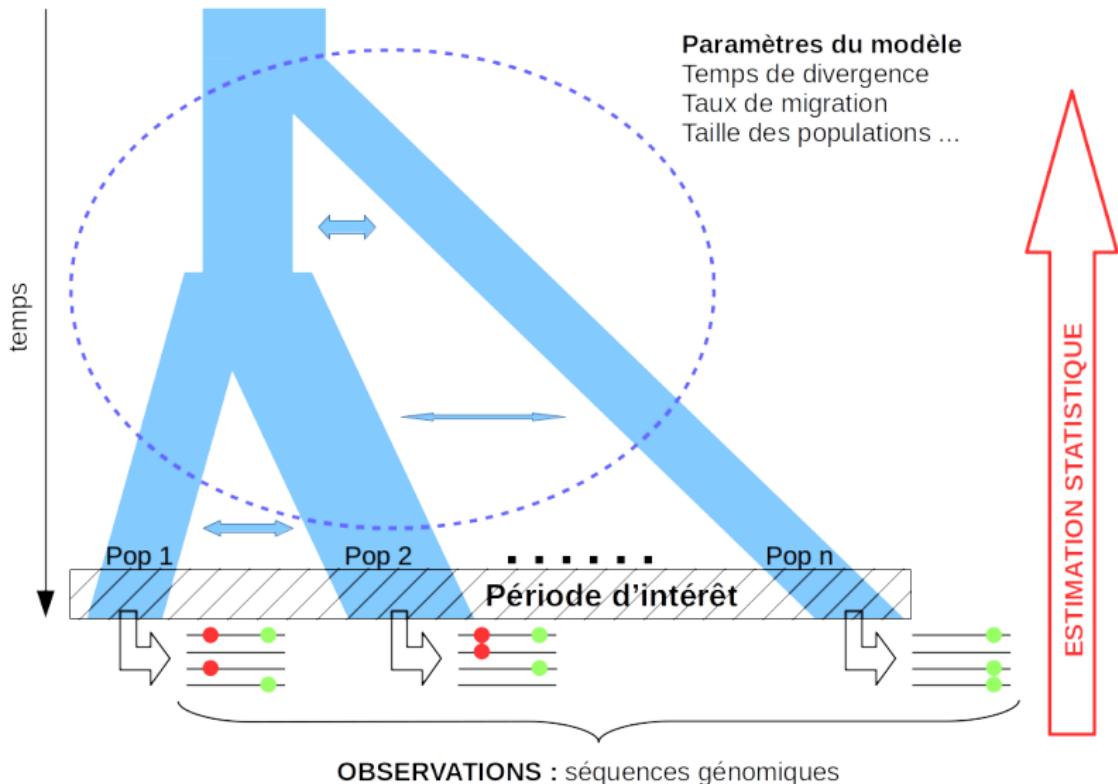
Outline

- 1 Contexte: intérêt des données génomiques temporelles?**
- 2 La méthode SelNeTime**

Plan

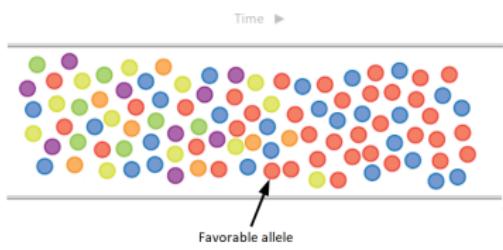
- 1 Contexte: intérêt des données génomiques temporelles?**
- 2 La méthode SelNeTime**

Estimation en génétique des populations

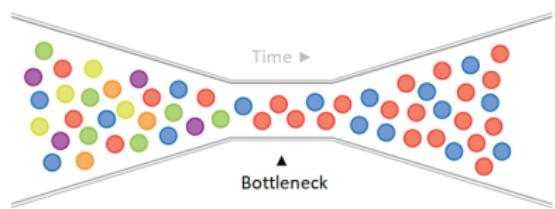


Effets confondants de la démographie et de la sélection

Sélection



Déclin démographique



Ignorer la démographie d'une population peut conduire à détecter de faux locus sous sélection

Données génomiques temporelles

Plusieurs contextes et échelles temporelles



Evolution
expérimentale



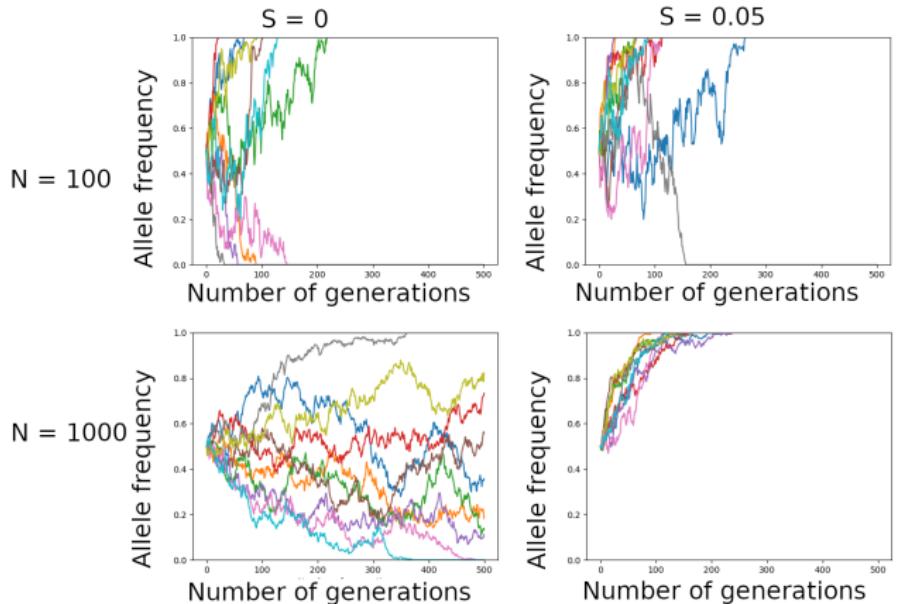
Suivi de
populations



ADN
ancien

Données génomiques temporelles

Les trajectoires de fréquences alléliques nous informent sur la démographie et la sélection.



Données génomiques temporelles

- Interviennent dans plusieurs contextes et échelles temporelles.
- Permettent un focus sur une période temporelle spécifique.
- Permettent (en principe!) de distinguer les effets démographiques et adaptatifs caractérisant cette période.

Plan

- 1 Contexte: intérêt des données génomiques temporelles?**
- 2 La méthode SelNeTime**

En résumé

- **Package python**

<https://forge.inrae.fr/genetic-time-series/selnetime>

- **Analyse en deux temps des données génomiques temporelles :**

- 1 Estime N en utilisant tous les locus et en supposant $s = 0$ (pas de sélection)
- 2 Estime s à chaque locus sachant \hat{N} .

- **Simule** des données génomiques temporelles.

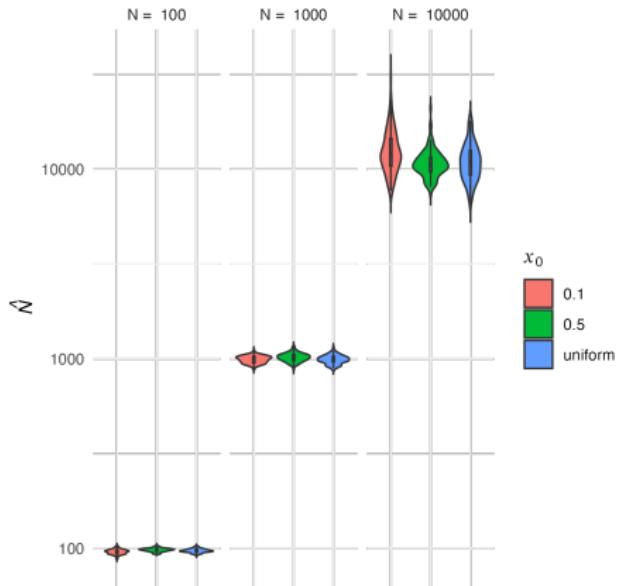
- **Tutoriel**

<https://forge.inrae.fr/genetic-time-series/selnetime-tutorial>

- **Preprint** recommandé par PCI Math Comp Biol:

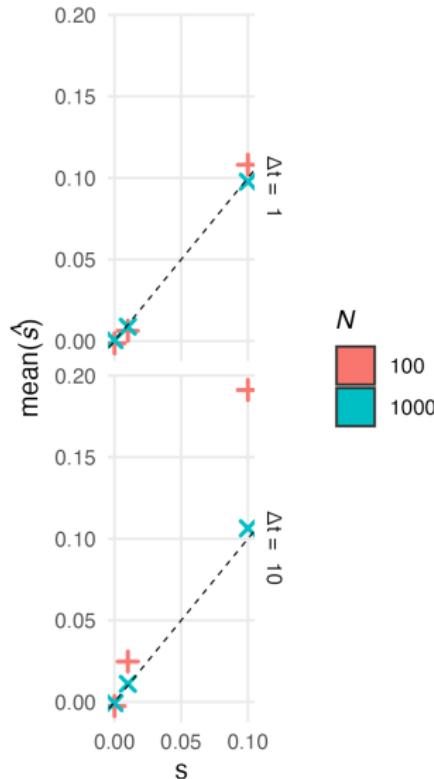
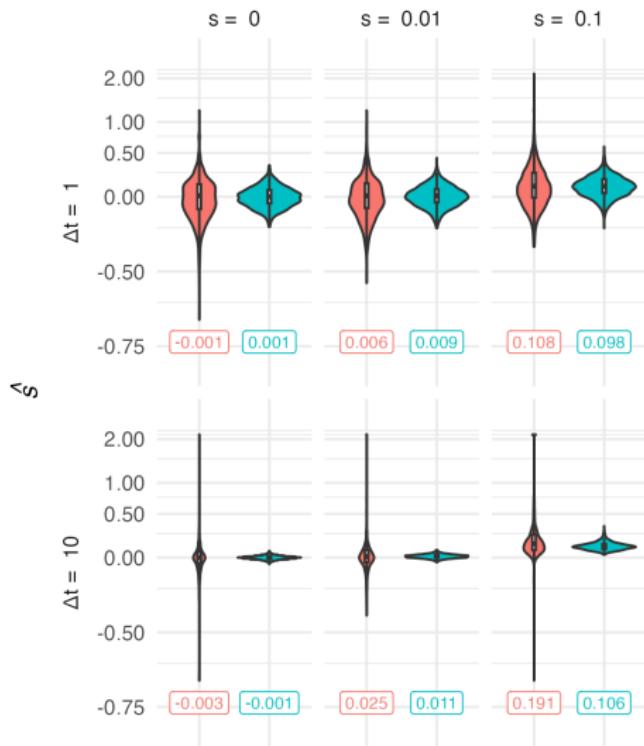
<https://mcb.peercommunityin.org/articles/rec?id=406>

Estimation de N sans sélection

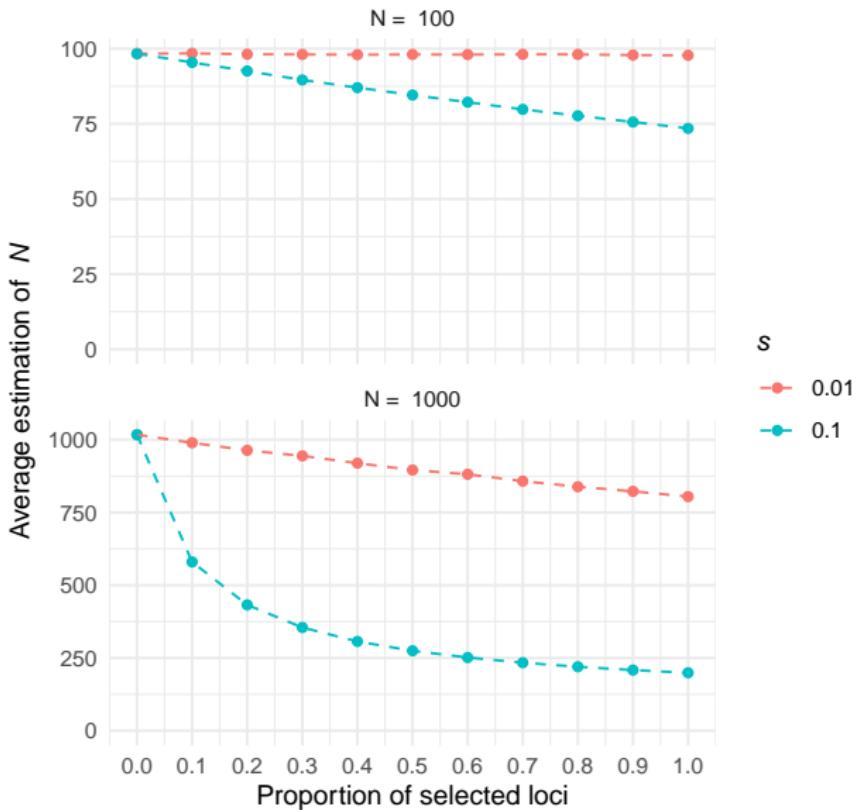


- Un jeu de données simulé = 15 diploïdes échantillonés 10 fois avec un écart de 10 générations, 1000 locus
- 100 jeu de données simulés pour chaque scénario.

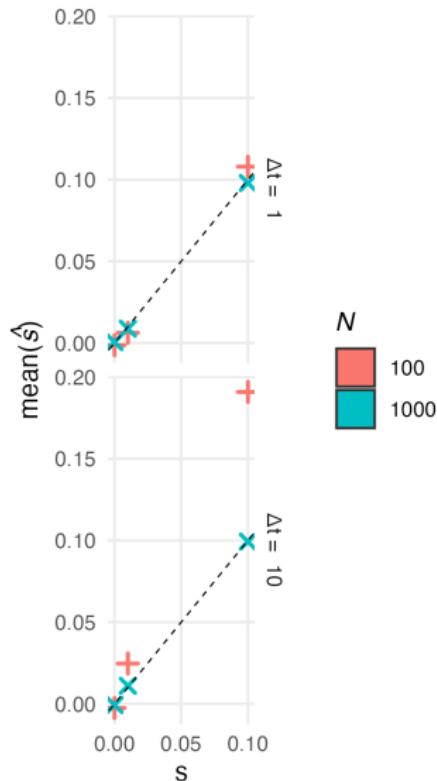
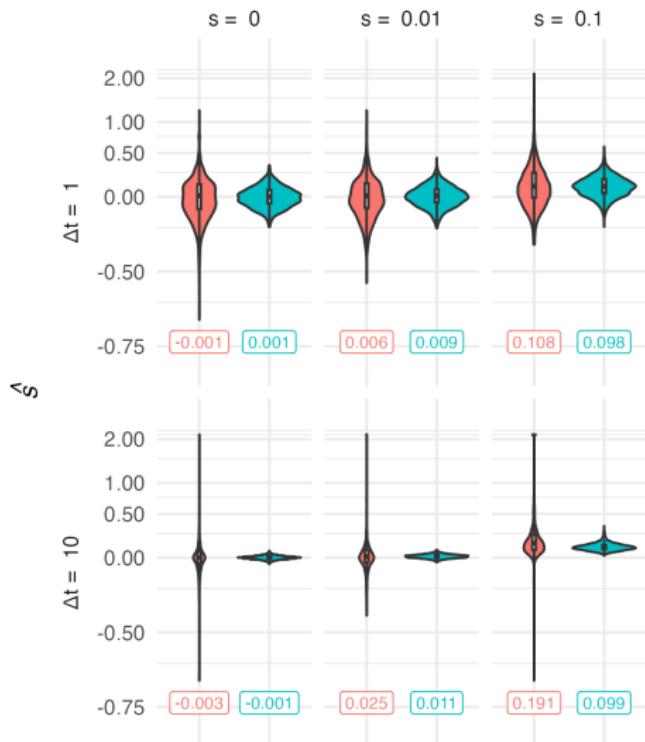
Estimation de s sachant le vrai N



Estimation de N en présence de sélection



Estimation de s sachant \hat{N} (possiblement biaisé)



Conclusions et perspectives

- Performance actuelle
 - Bonne estimation de N dans des scénarios avec peu de sélection.
 - Bonne estimation de s .
- Perspectives
 - Tester des stratégies d'échantillonage plus variées.
 - **Estimation jointe** de N et s pour limiter les biais sur N .
 - Tailles de population variables.