

ETUDE D'HISTOIRES ÉVOLUTIVES DU RIZ PUBLIÉES

V. Berry (juin 2018)



HISTOIRES EN COMPÉTITION

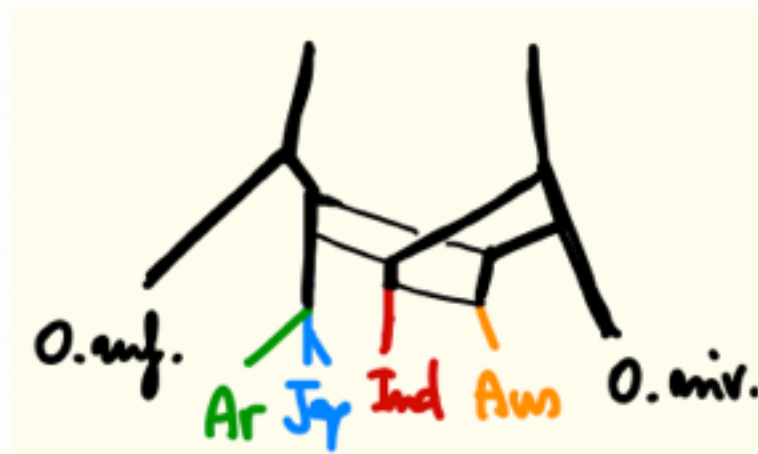
- Nombreuses publications ces 10 dernières années
- Pas d'accord sur l'histoire évolutive des sous-espèces de riz
- Sur 6 sous-espèces en autorisant 2 introgressions : 37 millions de topologies possibles !



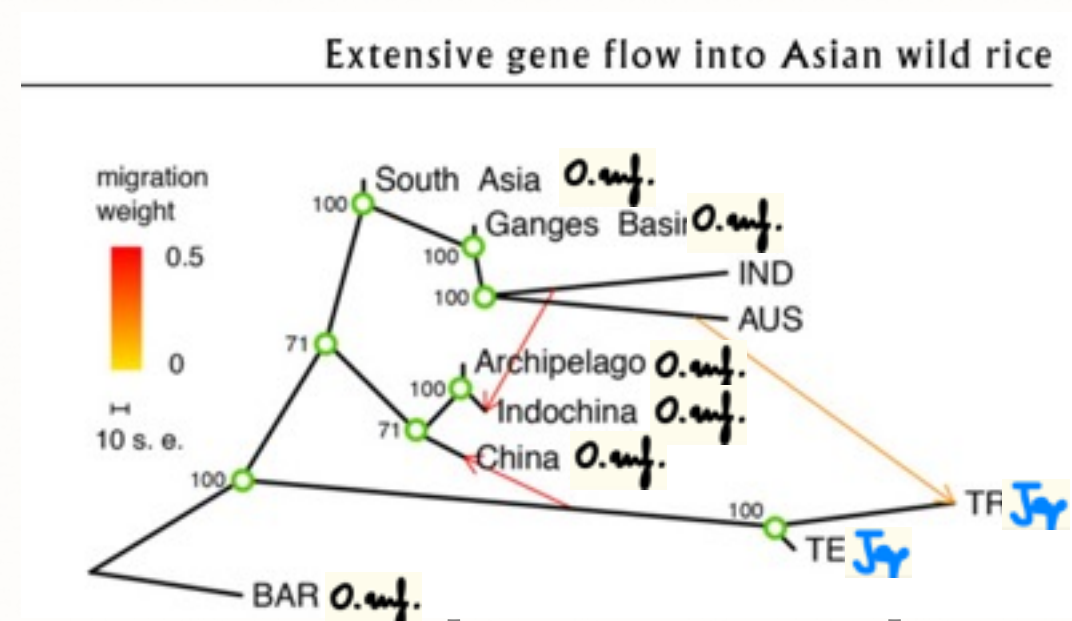
[Huang et al 2012]



[Civan et al 2015]



[Choi et al 2017]



[Wang et al 2017]

SÉLECTION DE 12 VARIÉTÉS DE RIZ

[Huang et al 2012]

[Wang et al 2017]

Core 20, Core 50, 115 variétés (Jean-Christophe G.)



Aus (CX227, IRIS_313-11058)

Aro (IRIS_313-11258, ...-11062)

Indica (IRIS_313-11643, CX225)

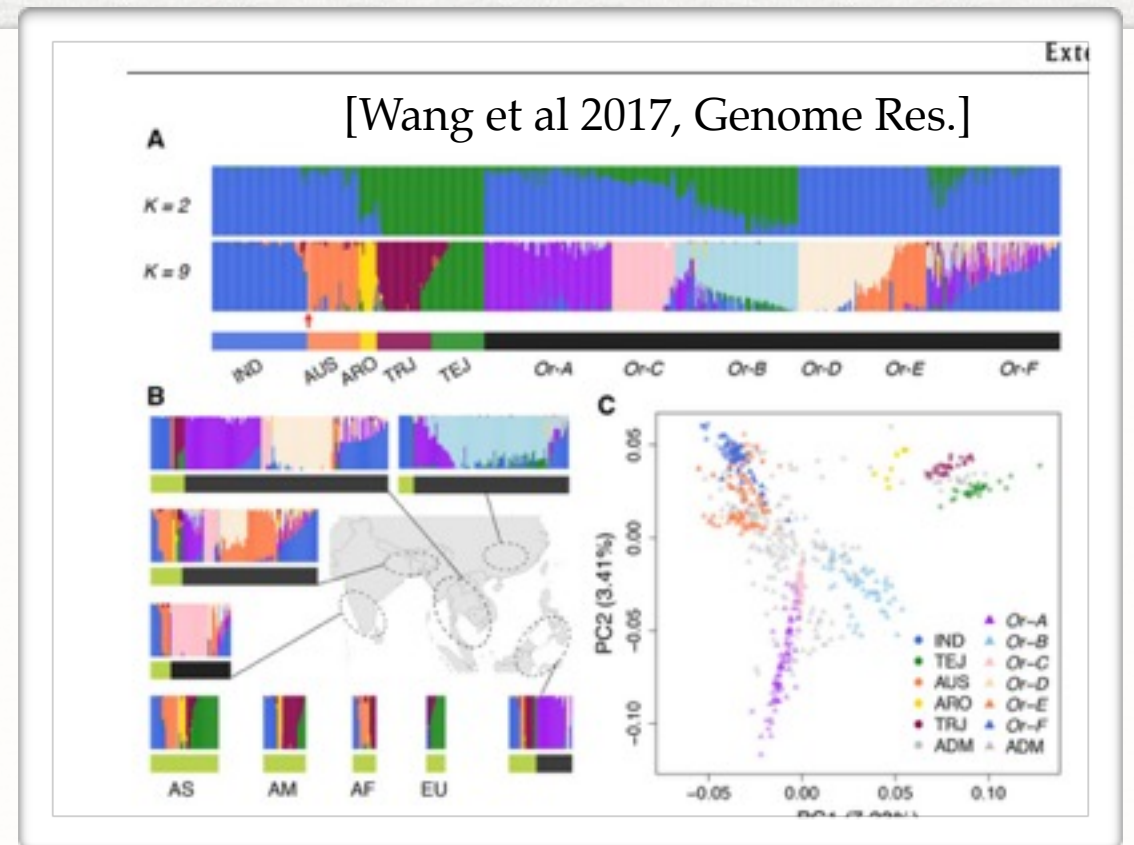
Japonica (B182, IRIS_313-10642)

O. Rufipogon I (W3105, W1559)

O. Rufipogon III (W1943, W2036)

MÉTHODES D'INFÉRENCE

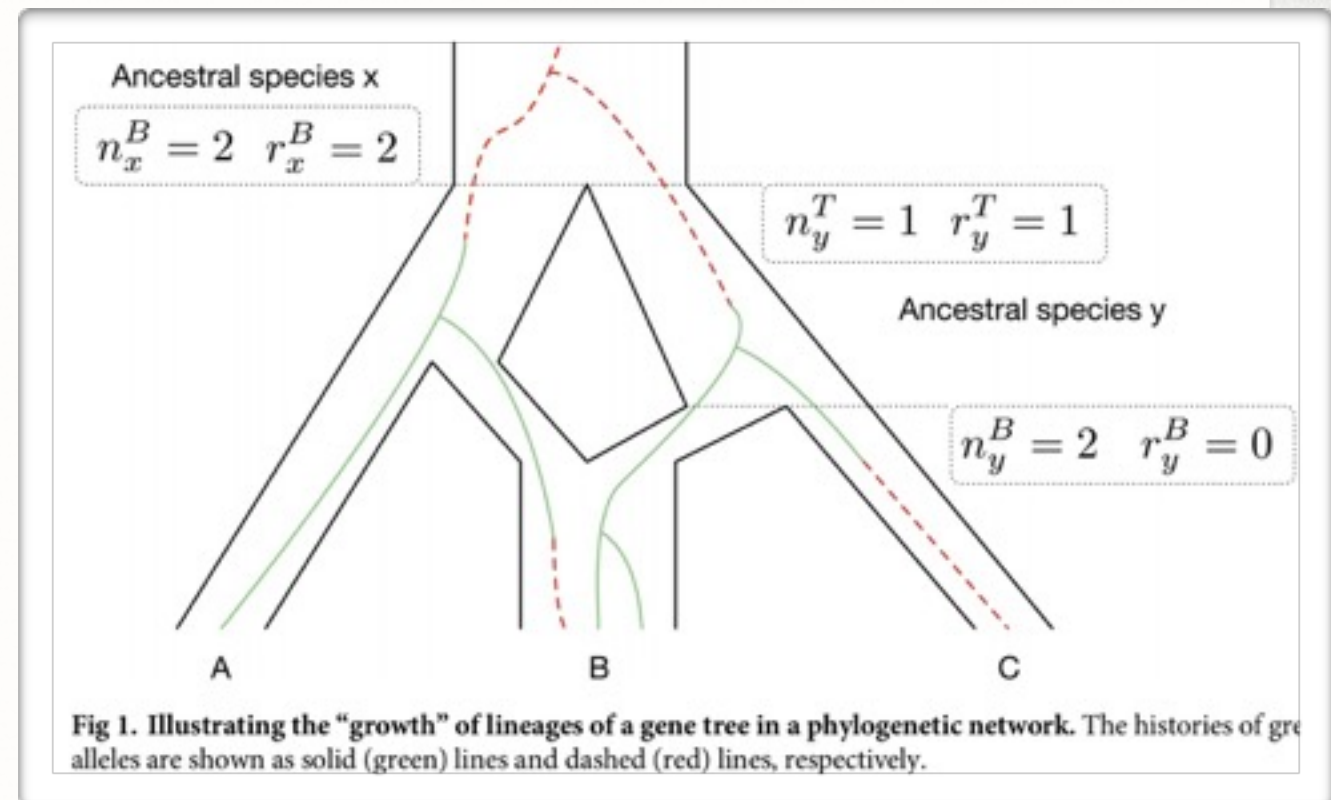
- Inférence de ces histoires par méthodes de **clustering** et **lissage statistique** (fineStructure, Admixture, NGSadmix, ...)



- Autre approche possible = méthode **probabiliste** basée sur un modèle évolutif explicite
- Au début du projet *Genome Harvest*, pas de méthodes d'inférence par modèle évolutif explicite
- Maintenant, c'est possible [Wen et al 2018, [Rabier et al](#)]
- Ces méthodes permettent-elles d'arbitrer entre les différentes histoires proposées par les études précédentes pour le riz ?

MÉTHODE DE TEST

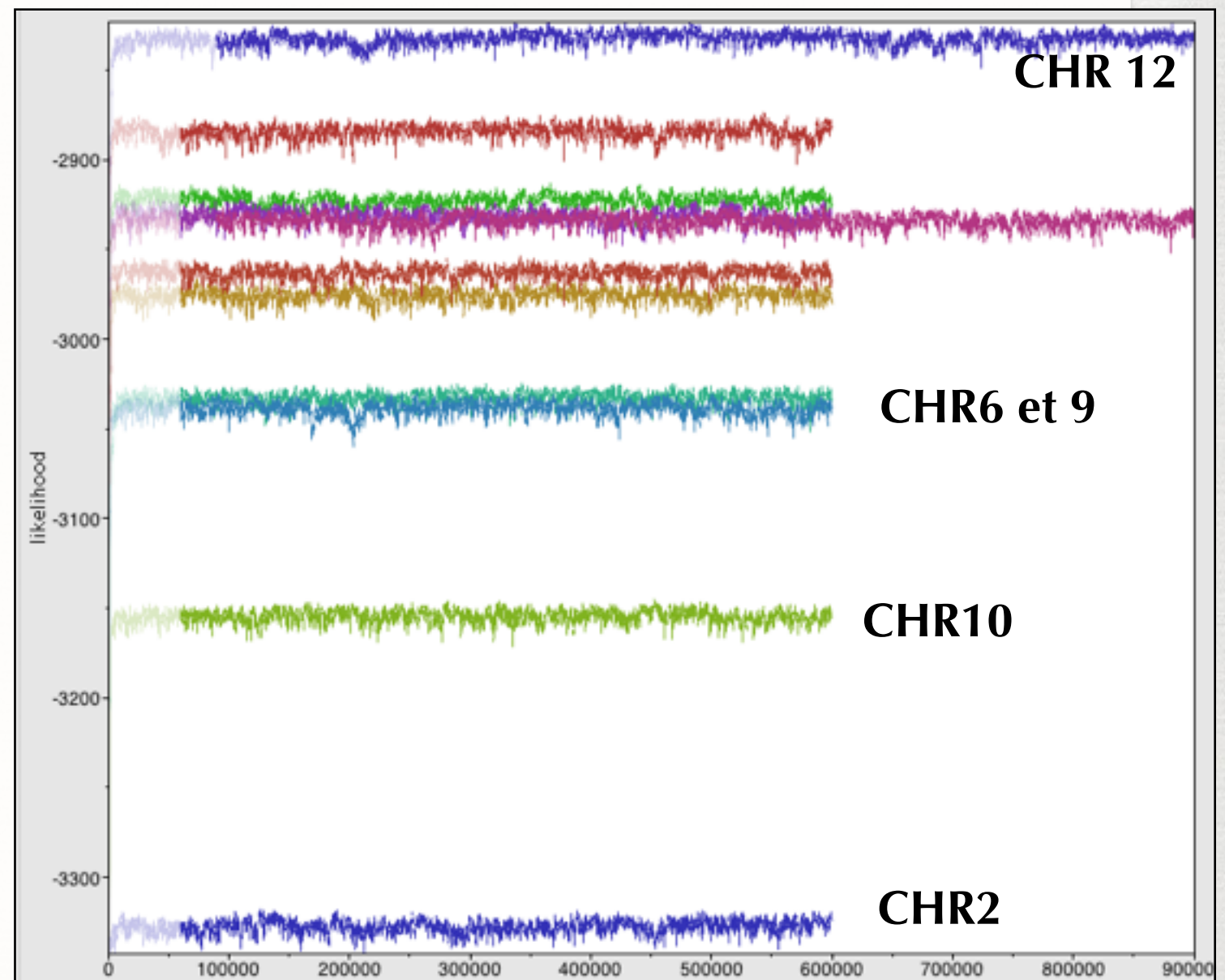
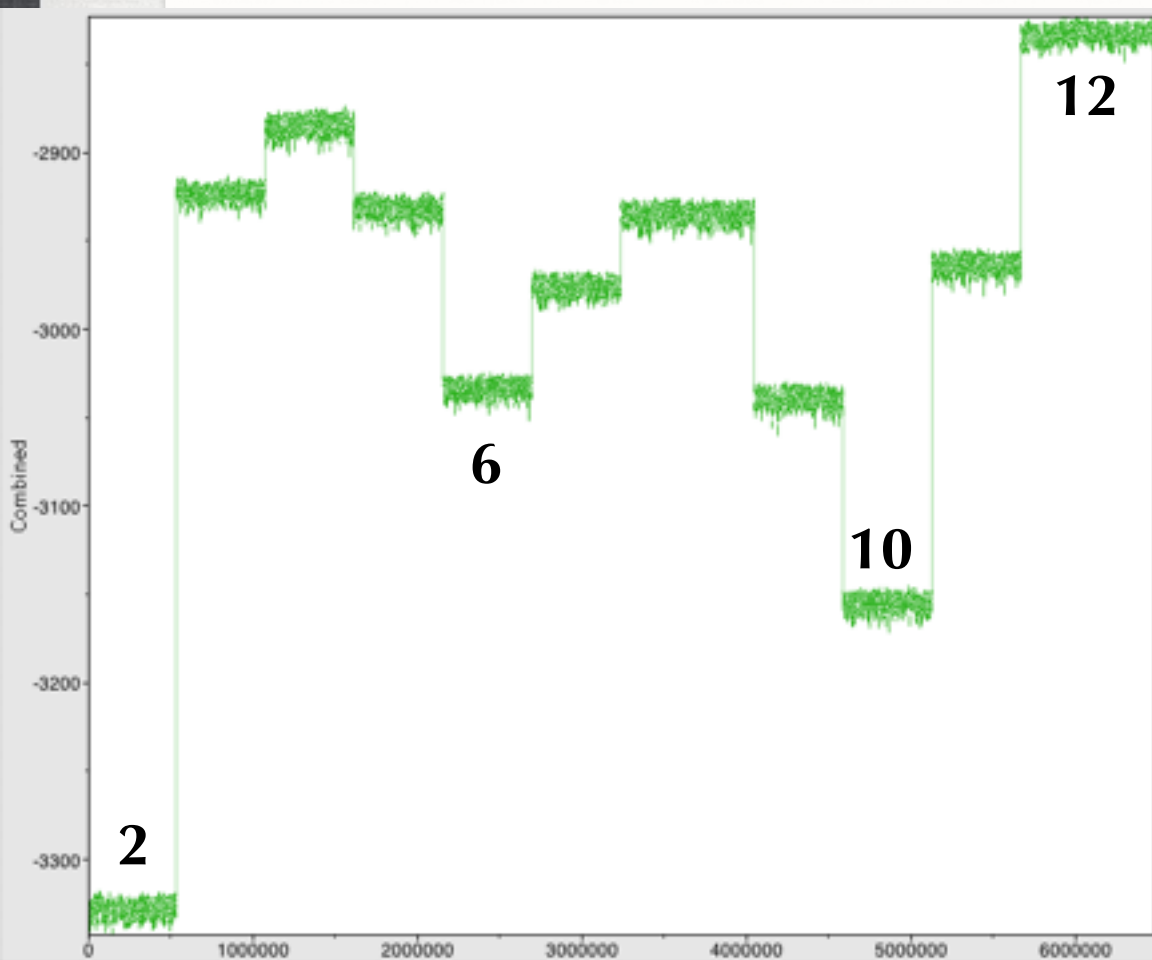
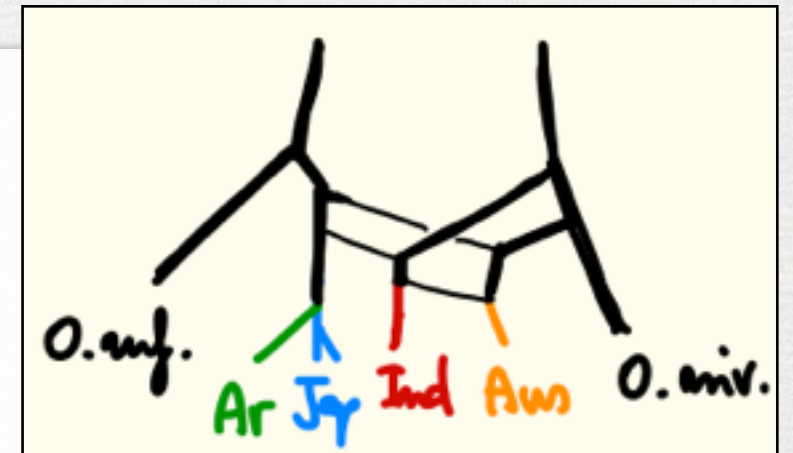
- **Modèle évolutif explicite** sous forme de réseau phylogénétique, permettant des introgressions, le tri de lignée (ILS) et des mutations [Rabier et al] et permettant plusieurs variétés par sous-espèce
- Inférence bayésienne
- Programme d'inférence intégré dans **BEAST**
- Sélection de variétés utilisées dans les **études précédentes**
- Sélection de sites homozygotes
- Scripts pour extraire des sous-ensembles de 1000 et 10,000 SNPs
- Inférence à **topologie fixe**



[Wen et al 2018, PLOS Comp. Biol.]

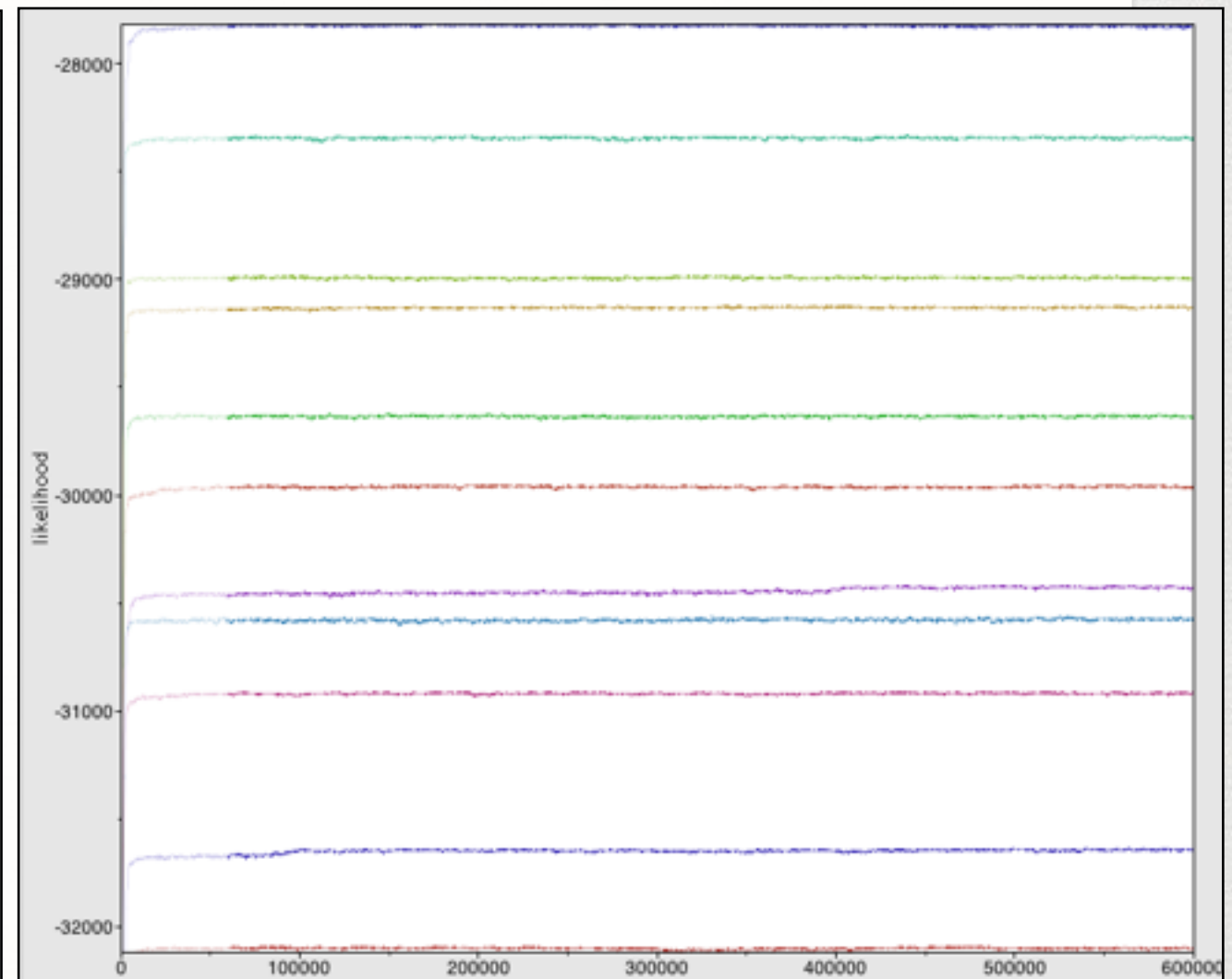
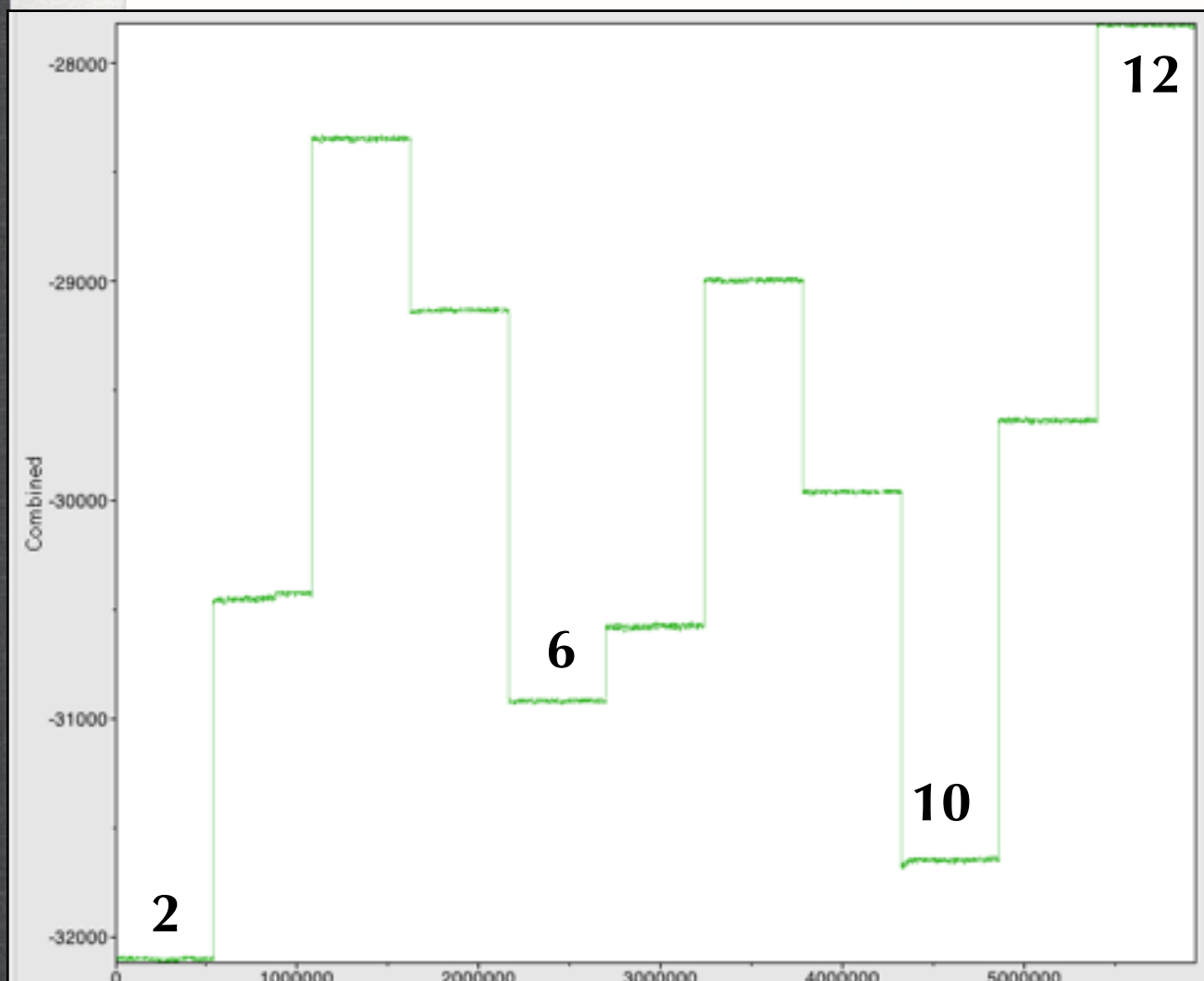
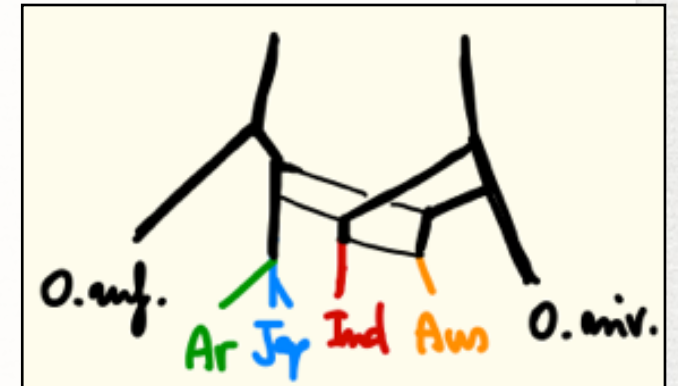
RÉSULTATS 1K SNPS - HISTOIRE [CHOI ET AL 2017]

- Convergence du MCMC
- Fortes disparités % CHR



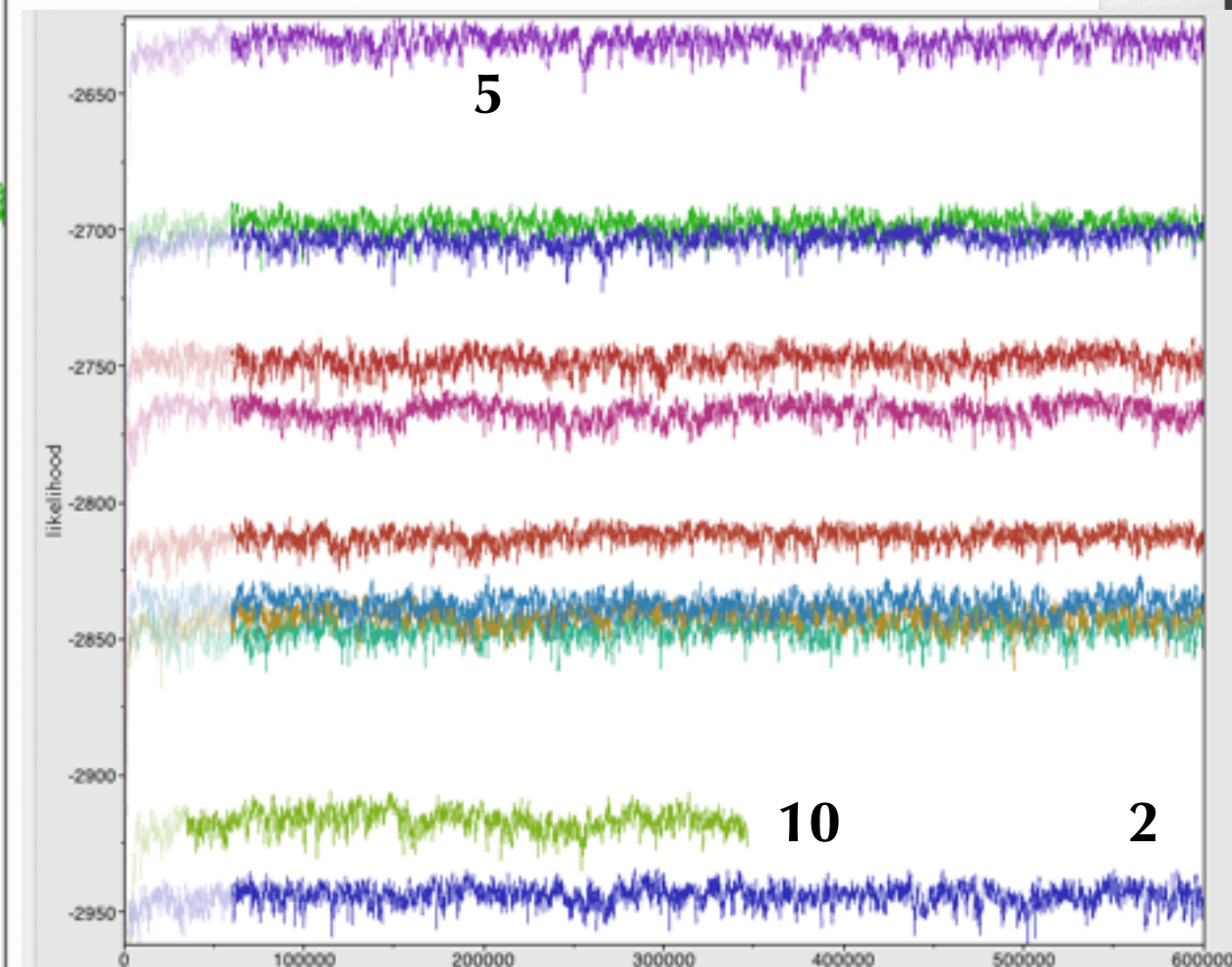
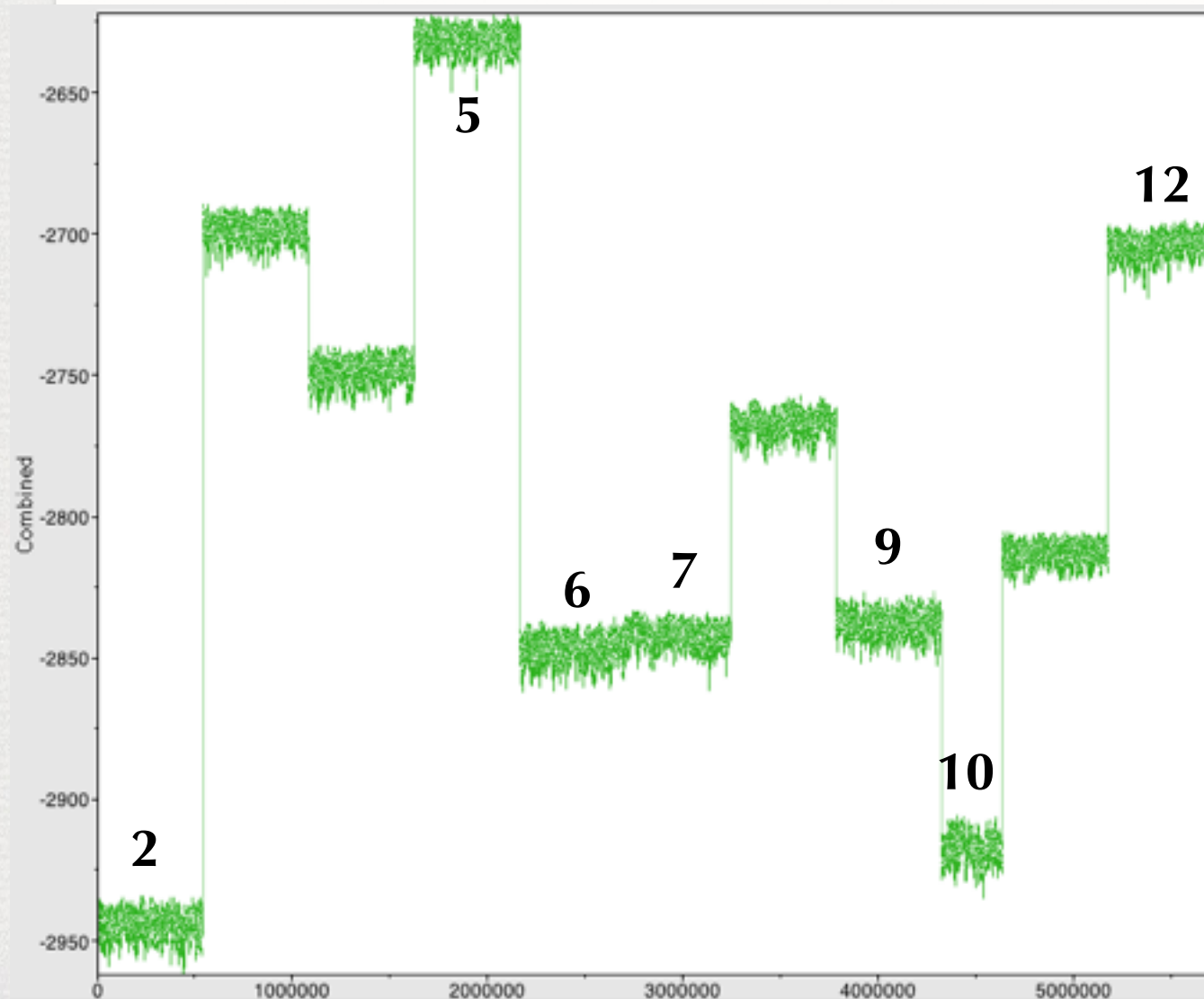
RÉSULTATS 10K SNPS - HISTOIRE [CHOI ET AL 2017]

- (non-convergence du MCMC : ça tourne encore)
- Profil relatif des CHR similaire à 1k SNP



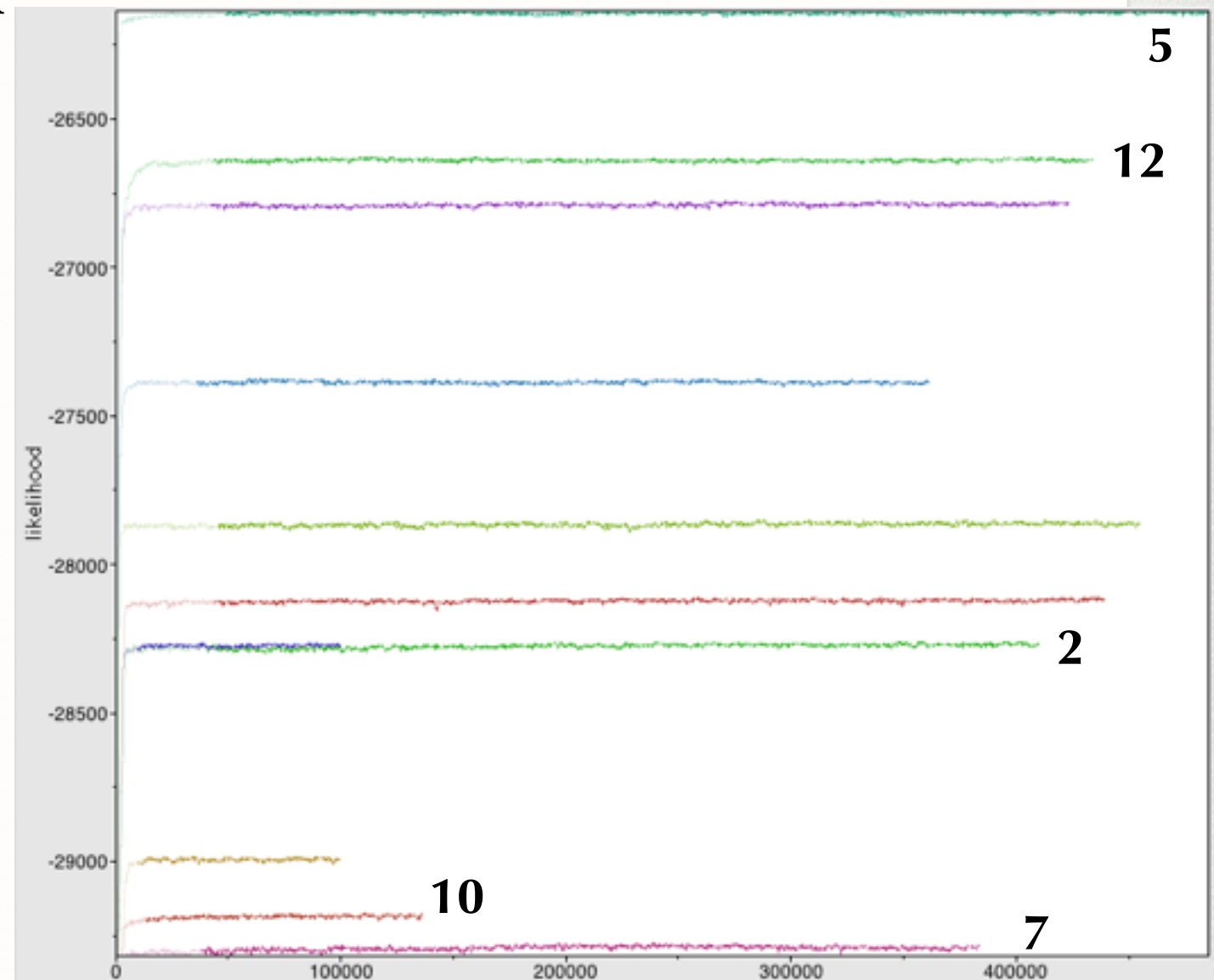
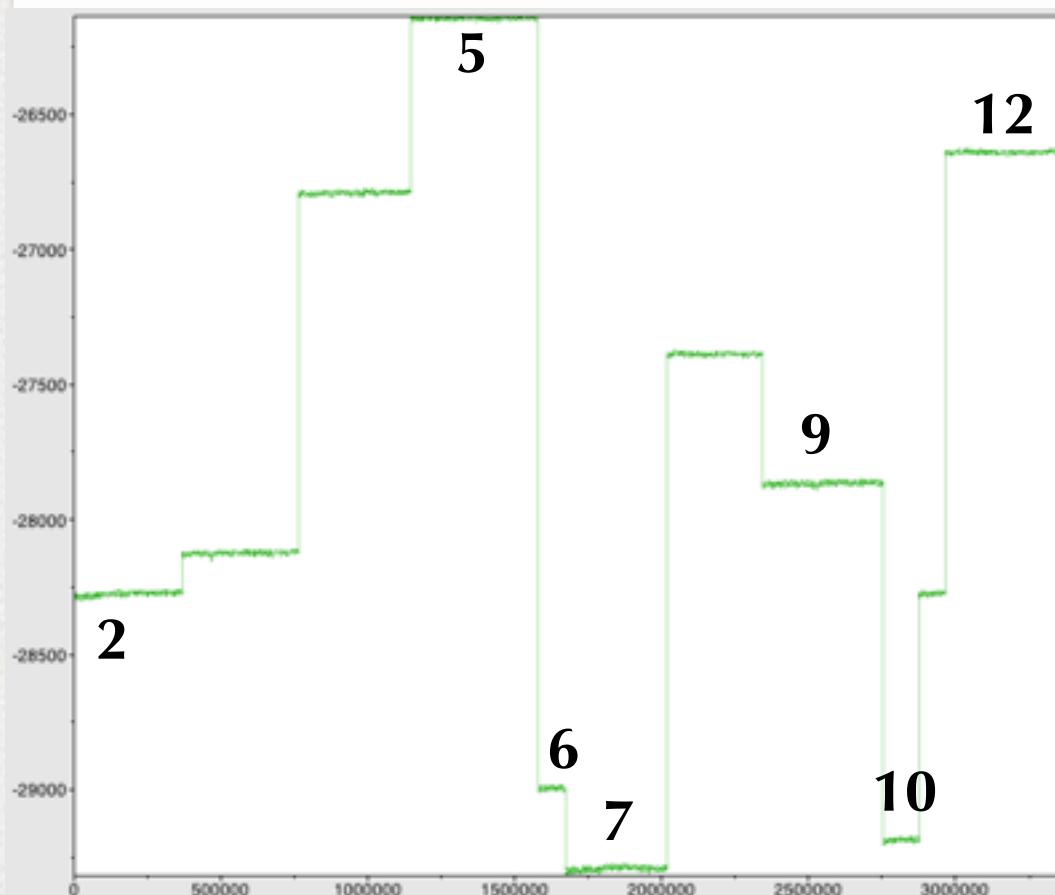
RÉSULTATS 1K SNPS - HISTOIRE [HUANG ET AL 2012]

- Convergence sur 2/3 des CHR : ça tourne
- Plus d'étapes MCMC nécessaires que [Choi et al]
- Fortes disparités % CHR
- Profil différent de ceux obtenus depuis la topologie [Choi et al]



RÉSULTATS 10K SNPS - HISTOIRE [HUANG ET AL 2012]

- Convergence pas du tout obtenue encore
- Besoin de plus d'étapes que pour 1000 SNPs
- Fortes disparités % CHR

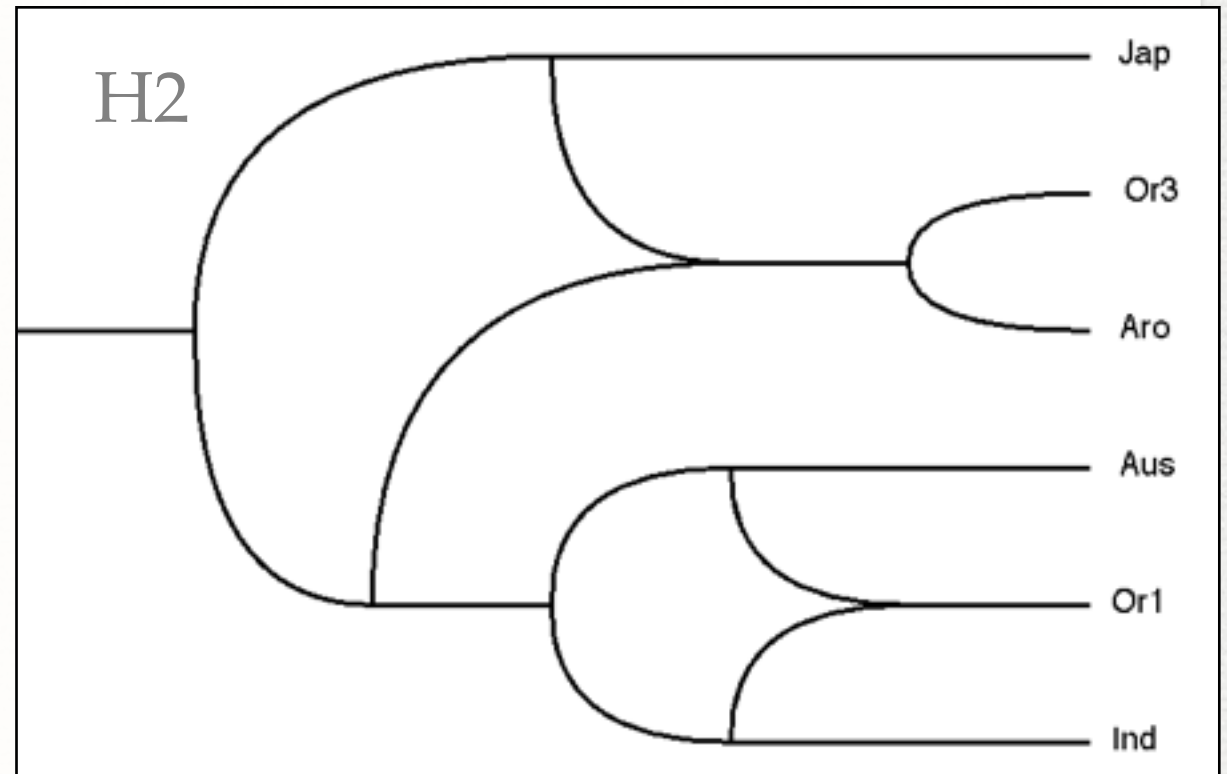
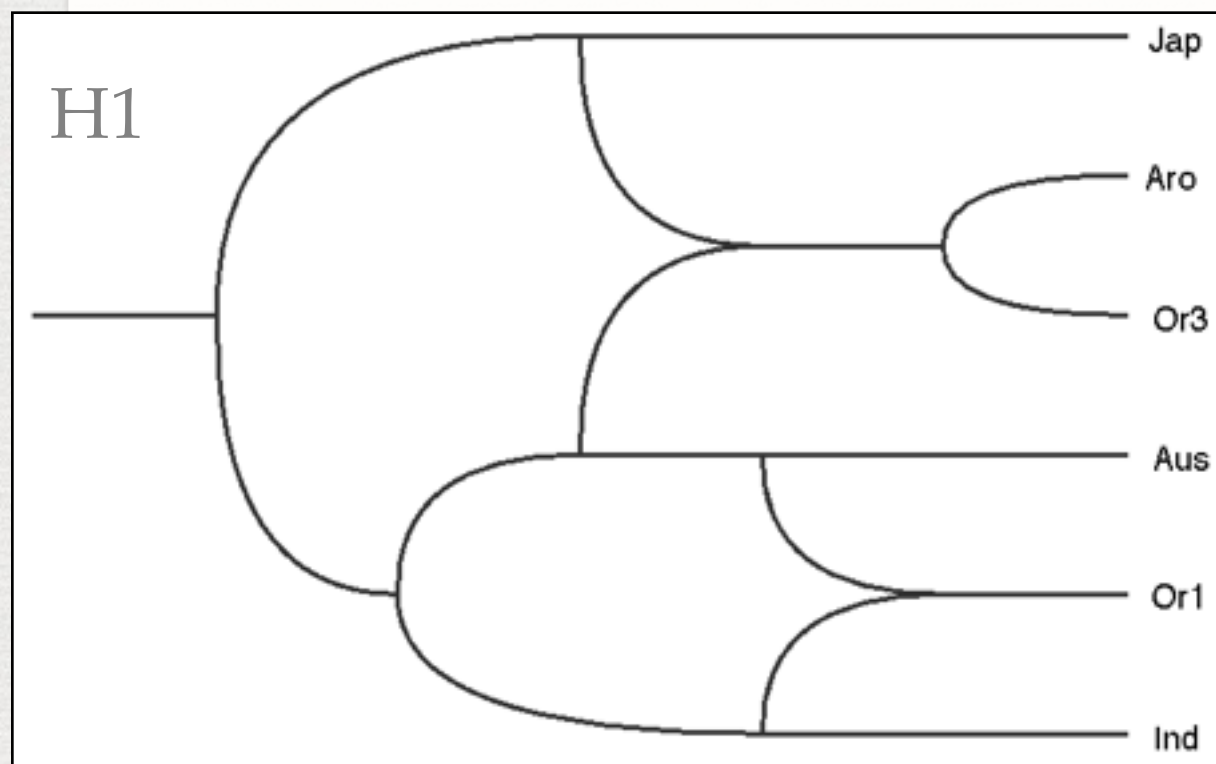


A FAIRE ENCORE

- Sur un même chromosome prendre des ensembles de SNPs de même taille mais situés à des endroits différents
- Etudier des fragments de CHR
- Conclusion pour **ces** variétés quand convergence (la préférence de variétés n'est pas forcément la préférence d'une sous-espèce complète)
- Etudier d'autres sélections de variétés pour mesurer le variabilité des résultats obtenus ; définir des variétés typiques de chaque sous-espèce : clustering
- . . .

TOPOLOGIES PROPOSÉES SI LIBRE

- depuis 20k SNPs du CHR2 (**convergence non encore atteinte** après 111k pas du MCMC)



- H2 : Aro et Japonica liées, d'un autre côté Indica et Aus [Huang et al 2012, Choi et al 2017]
- Or. Ruf. 3 plus lié à Japonica et Aro [Huang et al 2012]
- Or. Ruf. 1 plus lié à Aus et Indica [Huang et al 2012]
- H1 propose Aro contenant du matériel de Aus et de Jap (cf Joao et Jean-Christophe)
- Or. Ruf pas externes : choix de variantes sauvages proches des cultivées (Jean-Christophe) + introgressions de cultivées dans des sauvages [Wang et al 2017]