

TP6 : Préhension de sélection dans les séquences codantes

Benoit Nabholz (benoit.nabholz@umontpellier.fr)

Partie 1 : Divergence synonyme et non-synonyme

La du TP5, nous avons constaté qu'il existait une relation positive entre le nombre de gap par site et le nombre de substitution par site entre plusieurs gènes de mammifères (BRCA1, NFYA, PHF6, RAP2C, SEMA3F). Cette relation peut être causé par l'effet de la sélection naturelle ou de la mutation. Afin de discriminer entre ces deux hypothèses nous allons calculer les divergences synonymes (D_s) et non-synonymes (D_n).

- Reprenez le tableau 2 du TP précédent et ajouter deux colonnes D_n et D_s calculées entre l'homme et monodelphis. Le calcul s'effectue avec la fonction `kaks` du packages `seqinr`

```
library(seqinr)
BRCA1<-read.alignment("sequences/BRCA1.fasta", format="fasta")
kaks(BRCA1)
PHF6<-read.alignment("sequences/PHF6.fasta", format="fasta")
kaks(PHF6)
RAP2C<-read.alignment("sequences/RAP2C.fasta", format="fasta")
kaks(RAP2C)
SEMA3F<-read.alignment("sequences/SEMA3F.fasta", format="fasta")
kaks(SEMA3F)
```

-Est-ce que le ratio D_n/D_s est en accord avec l'hypothèse de mutation ou de sélection?

EXERCICE 2: Perte de fonction d'un gène

Objectif de l'exercice : Nous allons étudier l'évolution du gène de l'énaméline chez 6 espèces de mammifères placentaires. Ce gène code pour l'une des trois protéines impliquées dans la fabrication de l'émail (= la partie externe de la couronne des dents). Enameline (1104 amino-acides) est la plus grande des trois protéines qui constitue l'émail. Nous allons étudier l'effet de la perte des dents sur l'évolution de la séquence de l'Enameline.

Les données sont extraites de l'étude de Meredith et al. (2009).

Le tableau 1 présente les espèces étudiées.

Table 1: Présentation des espèces étudiées par Meredith et al. (2009)

Taxonomie	Espèces	Régime.alimentaire	Dents
Afrotheria	Orycteropus afer	Myrmécophage (Termivore)	Oui mais pas d'émail
Afrotheria	Loxodonta africana	Herbivore	Oui
Afrotheria	Procavia capensis	Herbivore	Oui
Laurasiatheria	Sus scrofa	Omnivore	Oui
Laurasiatheria	Manis pentadactyla	Myrmécophage (Termivore)	Non
Laurasiatheria	Lama glama	Herbivore	Oui

- 1) Ouvrir le fichier ENAM.fasta avec seaview. Il contient les 8 séquences de l'exon 10 du gène de l'énaméline. Visualiser les séquences au niveau des acides aminés à l'aide de l'onglet « Props » → « View as protein », puis lancer un alignement avec « Align » → « Align all ». **Que constatez-vous et proposez une interprétation?** « * » dans l'alignement correspondent à des codon stops)

- 2) A présent, nous allons utiliser un alignement qui respecte le cadre de lecture réalisé avec le programme **macse**. Cette alignement est stocké dans le fichier : “ENAM_align.fasta”
- 2.1) Calculez les divergences synonymes et non-synonymes pour une paire d’espèces sans émail chez les Afrotheria et avec émail. Remplissez le tableau 2 (présenté ci-dessous) avec vos résultats.

Table 2: D_n/D_s et présence d’émails

Présence.d.émails	Paire.d.espèces	Dn	Ds
NA	Lama / Loxodonta	NA	NA
NA	Lama / Orycterops	NA	NA
NA	Loxodonta / Manis	NA	NA
NA	Loxodonta / Sus	NA	NA

```
library(seqinr)
ENAM<-read.alignment("sequences/ENAM_align.fasta", format="fasta")
kaks(ENAM)
```

Reference

- Meredith RW, Gatesy J, Murphy WJ, Ryder OA, Springer MS. 2009. Molecular decay of the tooth gene Enamelin (ENAM) mirrors the loss of enamel in the fossil record of placental mammals. PLoS Genet. 5:e1000634.