IFT3295 - Démo 9

Phylogénie

Généralité sur les arbres phylogénétiques

1. Expliciter la différence entre une monophylie, une paraphylie et une polyphylie.
2. Soit les deux séquences nucléotidiques suivantes : $S_1 = ATCGAGGAA$ et $S_2 = ATGAGTGA$, calculer la divergence génétique entre ces deux séquences. Pour deux séquences nucléotidiques aléatoire à quelle pourcentage d'identité s'attend t'on en moyenne ? À partir de votre réponse précédente, proposez un argument pour l'utilisation des séquences protéiques à la place des séquences d'ADN lors de la construction des arbres phylogénétique.
3. Montrer qu'un arbre binaire non-enraciné à n feuilles possède $n - 2$ noeuds internes et $2n - 3$ branches. En déduire le nombre d'enracinement possible qu'on peut obtenir à partir d'un tel arbre.
4. Démontrez que le nombre total d'arbres binaires non-enracinés à n feuilles est de $(2n - 5)!!$

Vrai ou Faux

Répondez par Vrai ou Faux aux affirmations suivantes :

- On peut utiliser tout ensemble de séquences homologues pour la construction d'un arbre phylogénétique d'espèces.
- La divergence génétique observée est un bon estimateur de la distance évolutive réelle entre séquences.
- Le modèle markovien d'évolution de Kimura considère le même taux de substitution entre nucléotide.

- La plupart des mutations au sein d'une population se font sous un processus de sélection positive.
- Le ratio d_n/d_s entre distance non-synonyme et distance synonyme permet d'identifier le type de sélection affectant les séquences
- La théorie de l'horloge moléculaire stipule que le taux d'évolution d'une protéine est constant à travers les différentes lignées.

Calcul de Dn/Ds

Considérez les deux séquences "alignées" suivantes :

S1 = ACT CCG AAC GGG GCG TTA GAG TTG AAA CCC GTT AGA

S2 = ACG CCG ATC GGC GCG ATA GGG TTC AAG CTC GTA CGA

Calculez le ratio d_n/d_s à partir de cet alignement. Pour ce faire, aidez vous du code génétique standard suivant :

NB : notez que la méthode de "comptage" qui sera utilisée au cours de la démo est une approximation très naïve et n'est pas celle utilisée en recherche

		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G

Third letter