

Galaxy

Rapport TP



CHENEL Hugo
NASRALLAH Oussama
PAPAIIS Sandro

M2 Bioinformatique
Année 2021-2022

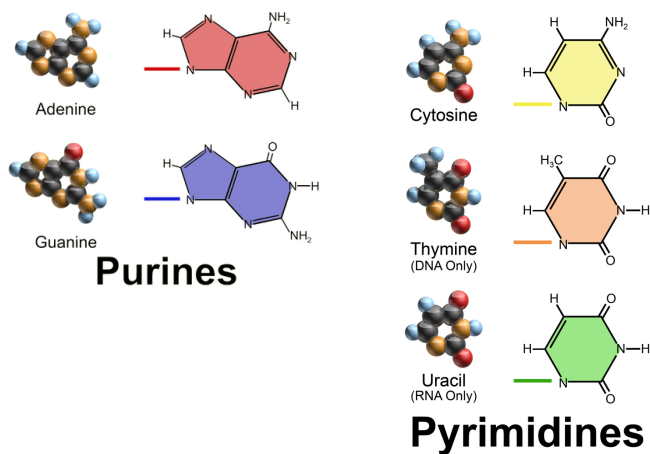
Responsable : MAMAN Sarah

Introduction

Des biais de composition dans les contextes des nucléotides, des codons et des acides aminés sont observés chez les bactéries, les champignons, les insectes, les plantes et les vertébrés, qui résultent probablement de forces de mutation et de sélection déséquilibrées et sont maintenus par les espèces dans leurs populations. Pour tout gène individuel, les biais de composition reflètent l'action de la mutation et de la sélection, qui est également liée à l'abondance des ARN de transfert iso-acceptants et aux efficacités catalytiques de leurs synthétases, et donc aux efficacités de traduction. Par exemple, les Actinobactéries sont caractérisées comme des "bactéries à fort contenu GC". Dans *Streptomyces coelicolor* A3, le contenu GC est de 72%. Par conséquent, l'analyse de la composition est d'une grande importance pour mieux comprendre la dynamique de la composition afin de fournir des preuves de l'évolution moléculaire.

La composition des nucléotides est très variable d'un génome à l'autre, et la teneur en guanine plus cytosine (G + C ou simplement GC) diffère considérablement d'une espèce à l'autre, notamment parmi les génomes bactériens, qui varie de 17 % (*Candidatus Carsonella ruddii* PV) à 75 % (*Anaeromyxobacter dehalogenans* 2CP-C). La dynamique des compositions nucléotidiques est également étroitement couplée aux compositions des codons et des acides aminés. Le facteur le plus probable qui détermine l'utilisation des codons/acides aminés est le biais mutationnel qui façonne constamment la composition des GC lorsque les génomes sont soit répliqués soit réparés par les ADN polymérases. En outre, des relations empiriques entre la teneur en GC et l'utilisation des codons/acides aminés ont été documentées chez de nombreuses espèces, fournissant ainsi de nombreuses données pour des études de modélisation et de simulation théoriques.

Dans une autre mesure, concernant le domaine de la biologie moléculaire, les expériences de réaction en chaîne par polymérase (PCR), la teneur en GC des oligonucléotides courts appelés amorces est souvent utilisée pour prédire leur température d'hybridation avec l'ADN matrice. Un niveau de contenu GC plus élevé indique une température de fusion relativement plus élevée.



ACIDES AMINÉS

ESSENTIELS

NON ESSENTIELS



Dans cette optique, nous nous sommes attelés à créer un outil nous permettant de calculer le taux de GC et AT, comptabiliser le nombre de bases puriques et pyrimidiques dans des séquences ADN et ARN d'intérêt ou bien les acides aminés au sein de séquences protéiques.