

Abstract

The proportion of A vs. T and G vs. C is rarely equal in bacterial genomes. This phenomenon is called «AT skew» and «GC skew», respectively.

Questions

What is the magnitude of these skews across Firmicutes?

Does this skew change suddenly or gradually during evolution?

Which branches of bacterial phylogeny carry changes in skews?

Is there any correlation between AT skew and any other characteristics of bacteria?

Materials and methods

NCBI RefSeq (for genomes); DoriC databases (for origin of replication);

Python v. 3.6.1 for calculation of skews; phytools package (R v. 3.3.0) for ancestral mapping

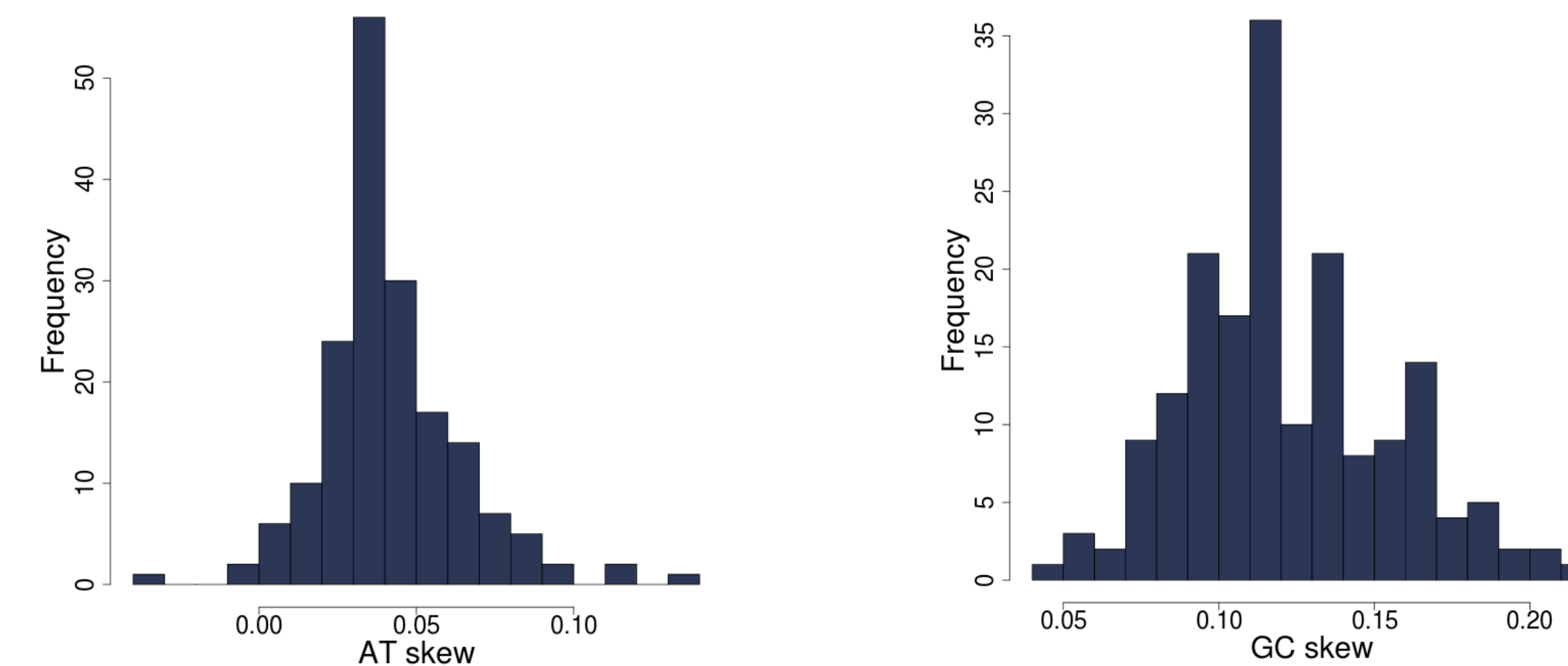
Conclusions

- Related bacteria have similar skew.
- In the common ancestor of all Firmicutes A prevailed over T; and G prevailed over C on the leading strand.
- In the history of Firmicutes some independent changes in AT skew and GC skew took place.
- The major skew change happened in the evolution history of *Clostridium*.
- In the Firmicutes history both sudden and gradual skew changes are present.
- There is a strong correlation between AT and GC skew (0.5), as well as between GC skew and GC content (-0.63). These correlations suggest that there is an unknown factor which affects both AT and GC skew in the same way.

Distribution of AT skew and GC skew in Firmicutes

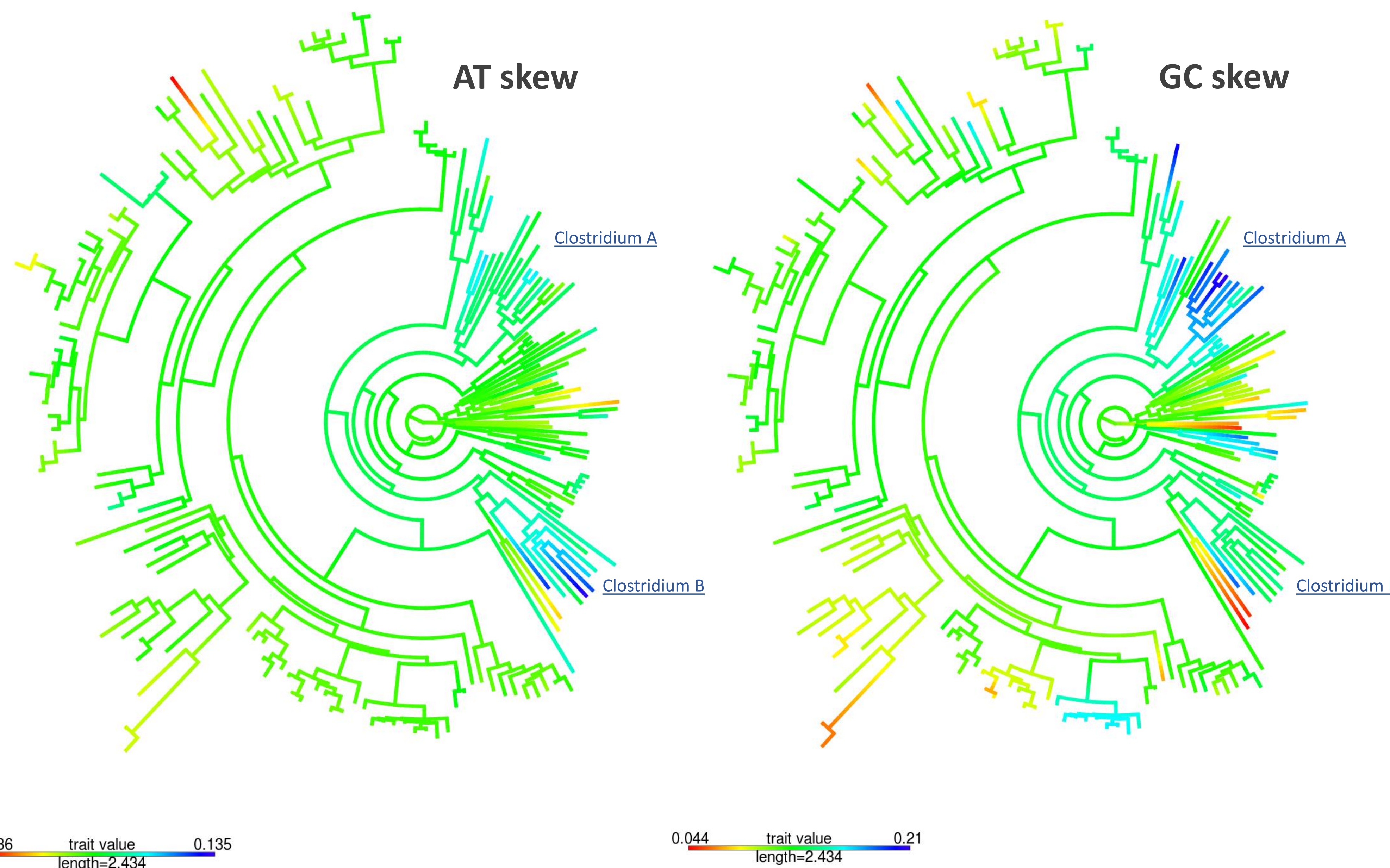
$$\text{AT skew} = \frac{A_{\text{lead}} - T_{\text{lead}}}{A_{\text{lead}} + T_{\text{lead}}}$$

$$\text{GC skew} = \frac{G_{\text{lead}} - C_{\text{lead}}}{G_{\text{lead}} + C_{\text{lead}}}$$

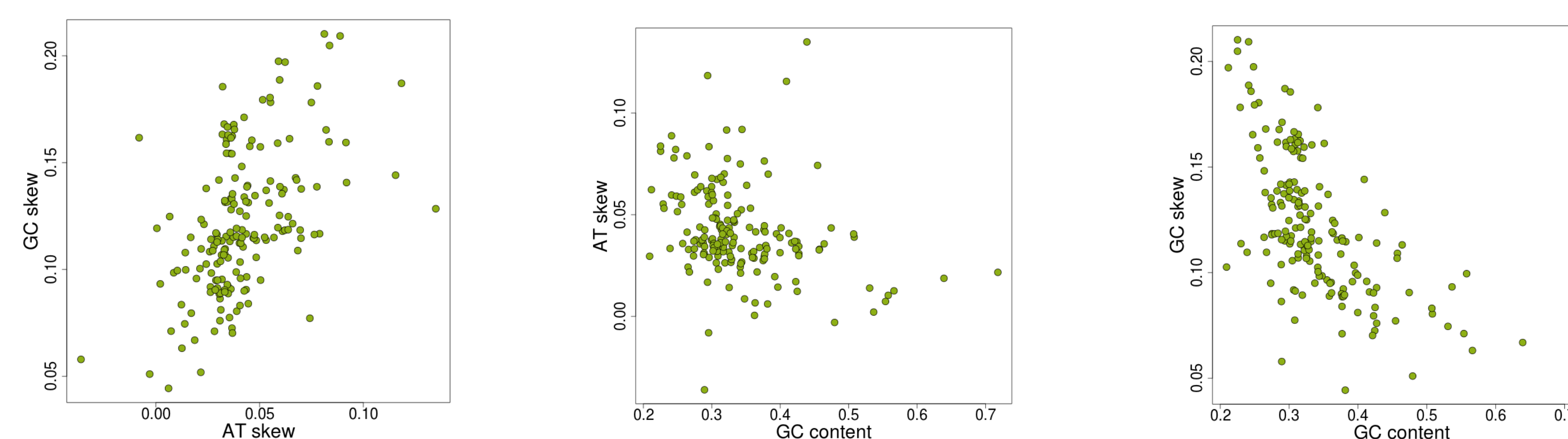


У большинства бактерий *Firmicuta* А превалирует над Т и G превалирует над С

Ancestral reconstruction of AT and GC skew



Correlation between At skew, GC skew and GC content



Введение

Представленность пар соответствующих друг другу нуклеотидов (А-Т или G-C) в любом бактериальном геноме редко одинакова. Это наблюдение получило название АТ-перекоса и GC-перекоса.

Проблемные вопросы

Какова амплитуда этих «перекосов»?

Был ли «перекос» внезапным?

На каких ветках бактериальной филогении произошел этот «перекос»?

Коррелирует ли «перекос» в геноме с другими характеристиками бактерий?

Методы и материалы

Базы данных NCBI, RefSeq, DoriC, Язык Python v. 3.6.1, пакет phytools (R v. 3.3.0)

Заключение

- Родственные бактерии имеют сходный «перекос».
- У общего предка всех *Firmicuta* А превалировало над Т и G превалировало над С на лидирующей цепи.
- В ходе эволюции произошло несколько независимых изменений и в АТ-«перекосе», и в GC-«перекосе».
- Можно заметить наибольшее изменение «перекоса» в эволюционной истории *Clostridium*.
- В истории *Firmicuta* присутствуют как внезапные, так и постепенные изменения в АТ-«перекосе», и в GC-«перекосе».
- Существует сильная корреляция между АТ-сдвигом и ПС-сдвигом (0.5), а также между GC-сдвигом и GC-контентом (-0.63). Между АТ-сдвигом и GC-контентом корреляция слабее (-0.3).