

# گزارش پروژه ی بیوانفورماتیک

هلیانه ضیائی ۹۴۱۰۴۷۷۸

۷ بهمن ۱۳۹۶

سوال ۱

•

سوال ۳

تفاوت میان دو الگوریتم UPGMA و NJ

با مشاهده ی نتایج متوجه میشویم که درخت های حاصل از الگوریتم UPGMA ریشه دار هستند و درخت های حاصل از الگوریتم NJ بدون ریشه. تفاوت دیگر در ارتفاع شاخه هاست که در الگوریتم UPGMA یکسان و در دیگری متفاوت است. اولین و مهم ترین تفاوت بین این دو الگوریتم این است که روش UPGMA یک درخت با ریشه و الگوریتم NJ یک درخت بدون ریشه خروجی میدهد. از طرفی الگوریتم UPGMA نرخ جهش را در تمام شاخه ها یکسان فرض میکند. به همین خاطر ارتفاع تمام شاخه ها در درخت یکسان است. الگوریتم UPGMA الگوریتم خیلی دقیقی نیست و تنها در صورتی که تحول با نرخ ثابت و دقیقی انجام شده باشد جواب میدهد و در مورد درخت های غیر ultrametric یعنی درخت هایی که فاصله ی تمام برگ ها تا ریشه یکسان نیست، خوب عمل نخواهد کرد. اما الگوریتم NJ این مشکل را نداشته اما در صورتی که درخت additive نباشد نتایج درستی به ما نخواهد داد. [۱، ۲]

روش پیشنهادی برای به دست آوردن یک درخت واحد

میتوان از هر ژنوم، مناطقی که هر کدام از ۷ ژن به آن منطبق شده اند را جدا کرد و در نهایت این ۷ تارا به هم چسباند تا یک ژنوم جدید به وجود آید. سپس میان این ژنوم های جدید که فقط شامل قسمت های منطبق شده به ژن ها هستند الگوریتم تطبیق سراسری را اجرا کرد. [۳]

روش دیگر برای این کار استفاده از تابع consensus از نرم افزار آر است که چندین درخت فاصله به عنوان ورودی دریافت کرده و یک درخت واحد را خروجی میدهد.

## سوال ۴

روش برای به دست آوردن فاصله ی زمانی دو به دو گونه ها در این قسمت ما مدل خود را مدل Jukes-cantor انتخاب میکنیم و از رابطه ی زیر برای به دست آوردن زمان بین هر دو گونه استفاده میکنیم:

$$\alpha t = -\frac{3}{4} \ln\left(1 - \frac{4}{3} \frac{N_d}{N}\right)$$

که در آن  $\alpha$  نرخ جهش،  $N_d$  تعداد جهش ها و  $N$  تعداد حروف رشته ی اولیه است. در این روش Indella در نظر گرفته نمی شوند اما با توجه به اینکه طول رشته ها در حدود همدیگر است خیلی مشکل ساز نیست. برای طول  $N$  میانگین طول دو رشته را در نظر گرفتیم و تعداد جهش ها را از ماتریس فاصله ی به دست آمده از همترازی سراسری قرار دادیم. [۴، ۵]

تخمین فاصله ی زمانی برای جهش بعدی

فاصله ی ویرایش بین دو گونه ی متفاوت از ویروس ابولا به طور متوسط ۶۰۰ تاست، اگر در نظر بگیریم که این مقدار لازم جهش برای ایجاد یک گونه ی جدید است و طول متوسط گونه ی این ویروس را حدودا ۱۸۹۰۰ بگیریم، آن گاه طبق مدل Jukes-cantor زمان لازم برای ایجاد یک گونه ی جدید حدودا ۲۱۷ سال است.

## مراجع

- [1] [https://en.wikipedia.org/wiki/Neighbor\\_joining#Advantages\\_and\\_disadvantages](https://en.wikipedia.org/wiki/Neighbor_joining#Advantages_and_disadvantages)
- [2] [https://www.researchgate.net/post/What\\_is\\_the\\_difference\\_between\\_UPGMA\\_and\\_NEJ\\_method\\_while\\_constructing\\_a\\_tree\\_using\\_a\\_MEGA\\_4\\_software](https://www.researchgate.net/post/What_is_the_difference_between_UPGMA_and_NEJ_method_while_constructing_a_tree_using_a_MEGA_4_software)
- [3] [https://www.researchgate.net/post/How\\_can\\_I\\_construct\\_one\\_phylogenetic\\_tree\\_using\\_two\\_genes\\_together](https://www.researchgate.net/post/How_can_I_construct_one_phylogenetic_tree_using_two_genes_together)
- [4] <https://mathcs.clarku.edu/~djoyce/java/Phyltree/mutations.html>
- [5] <http://www.cs.utoronto.ca/~brudno/csc2427/Lec4Notes.pdf>