

Домашняя 2

28.01.2025

1. Митохондриальная ДНК наследуется исключительно от матери. Если ее использовать для построения дерева, связывающего большое количество людей из разных этнических групп, то ее корень будет представлять "мать-прародительницу" (или митохондриальная Ева), от которой мы все унаследовали наши митохондрии. Кластеризация этнических групп на таком дереве может дать представление о физическом месторасположении Евы. В работе Канна, Стоункинга и Уилсона утверждалось, что Ева находится в Африке (что поддерживает теорию происхождения человека из Африки). Было построено корневое дерево, которое показывало связи между 147 людьми. Оцените, сколько топологически различных деревьев нужно было рассмотреть, если бы каждая возможность была действительно исследована. (Порядок величины количества различных деревьев на 147 листьях)
2. Скачайте 1240K панель с сайта группы David Reich (если у вас есть платный VPN) <https://reich.hms.harvard.edu>. Если нет, то с гугл-диска. Файлы .geno, .ind, .snr и .anno Установить AdmixTools с <https://github.com/DReichLab/AdmixTools> (необходимо openblas, gsl) Прочитайте RAEDME.3PopTest, чтобы сделать файл с параметрами (туда пропишутся пути до файлов 1240K+ путь до файлов с популяциями).
 - С помощью f_3 статистики с внешней группой Mbuti будем тестировать, кто ближе к русским Russia.НО в популяционном смысле. Для этого изучите файл .anno и выберите 5 ваших "любимых" СОВРЕМЕННЫХ неафриканских популяций. Посчитайте статистику

$$f_3(Mbuti.NO; Russia.NO, Pop).$$

Сравните их между собой. Согласуются ли полученные результаты с общеизвестными представлениями?

! Обратите внимание, что

$$f_3^{admixtools}(X1, X2; X3) = f_3(X3; X1, X2)$$

- Допустим, мы хотим построить укорененное дерево между популяциями Nganasan.HO, French.HO и Finnish.HO без перемешиваний. Возможно ли это? Аргументируйте свой ответ.
Hint: Посчитайте f3 статистики, наличие одной отрицательнозначной будет свидетельствовать о событии примешивания и невозможности такой реализации.
 - Постройте укорененное дерево с популяциями Mbuti.HO, French.HO, Lezgin.HO. Проверьте, возможно ли это (см предыдущую задачу). С помощью f3 статистик рассчитайте веса на ребрах графа.
3. (задача на +1 к итоговой оценке за курс) Разобрать https://github.com/simonhmartin/tutorials/blob/master/ABBA_BABA_whole_genome/README.md и рассказать на защите.