Два однотипных исследования пришли к противоположным выводам

Олег Балановский

Данная статья привлекает внимание прежде всего своим необычным выводом: что структура генофонда больше связана с лингвистическим родством популяций, чем с их географическим соседством. Ведь чаще всего в исследованиях генетиков получается наоборот. Обратная закономерность — большая связь генетики с лингвистикой, чем с географией (в таких, отчасти жаргонных, терминах принято обсуждать эту проблему) выявляется очень редко — например, в нашем исследовании Y-хромосомного генофонда и языков Кавказа. Напротив, большая связь генетики с географией, чем с лингвистикой, показана для многих других регионов и для самых различных генетических систем.

Особенно удивительно, что большая роль лингвистики, чем географии в структурировании генофонда выявлена авторами рассматриваемой статьи для Европы — региона, в котором важная роль географии была показана многократно, в том числе и по широкогеномным маркерам.

Любопытно, что название рассматриваемой статьи («Изменчивость геномов отражает лингвистическую изменчивость в Европе») очевидным образом полемизирует с первой знаменитой статьей по широкогеномным маркерам в Европе «Гены отражают географию в Европе» (Novembre et al., 2008).

Правда, в статье Novembre была просто показана удивительно большая связь генофонда с географией, и теоретически оставалась возможность, что связь с лингвистикой еще больше. Но все же вызывает удивление и даже настороженность что два исследования, опирающиеся на однотипные данные (изменчивость сотен тысяч аутосомных маркеров в европейских популяциях) приходят к противоположным выводам.

В то же время, формально вывод Лонгобарди с соавторами обоснован: частная корреляция генетики и лингвистики при исключения влиянии географии — 0,54, а частная корреляция генетики и географии при исключении влияния лингвистики намного меньше — 0.30. Это при анализе 12 индоевропейских популяций, а также басков, финнов и венгров. При исключении венгров роль лингвистики становится еще более рельефной, хотя должен сказать, что это может быть отчасти вызвано хорошо известным генетикам своеобразием финнов, которые и лингвистически относятся к другой семье. А учитывая, что лингвистические расстояния между семьями трудно выразить количественно, надежнее был бы анализ внутри одной семьи. Авторы его и провели — для индоевропейцев, но, что настораживает, не указали значения частных корреляций, а парные в таком анализе не слишком надежные индикаторы.

Также, на мой взгляд, сами генетические данные в рассматриваемой статье Лонгобарди с соавторами слабоваты, по крайней мере по современным меркам: взяты всего 15 популяций, хотя в Европе намного больше народов со своими языками. В качестве источника использована только одна публичная база данных (дополненная лишь данными из «1000 геномов»), хотя для специалистов доступны — пусть их надо потрудиться найти, скачать, совместить и проанализировать — данные на намного большему числу популяций.

Итак, мне представляется, что полученный авторами результат стоит принять во внимание, но не спешить ему безоговорочно доверять — не исключено, что последующие исследования того же вопроса дадут иные ответы, а учитывая быструю скорость накопления широкогеномных и полногеномных данных, ждать этого, скорее всего, осталось недолго.