Прочитана самая полная последовательность генома человека

Консорциум Telomere-to-Telomere (T2T) представил полную последовательность генома человека, которая содержит 3 055 миллиардов пар оснований, включая околоцентромерные участки, короткие плечи акроцентричных хромосом, длинные тандемные повторы и другие пробелы первой версии генома человека.

Секвенирована самая полная на сегодняшний день последовательность генома человека, статья с результатами этой работы размещена на сайте препринтов bioRxiv. Дело в том, что опубликованный в 2001 году геном человека по сути не является полным. Он содержит эухроматиновые, но не гетерохроматиновые участки ДНК, и в нем имеется множество пробелов, например не прочитанными остались регионы, примыкающие к центромерам и теломерам, а также короткие плечи акроцентрических хромосом (13,14,15,21 и 22); не прочитаны многие длинные тандемные повторы и сателлитные последовательности. Эти пробелы составляют примерно 151 мегабаз , или 8% от всего генома.

С тех пор появились новые технологии секвенирования, которые позволили прочитать недостающие участки человеческого генома. Эту работу осуществил международный консорциум Telomere-to-Telomere (T2T); для сборки *de novo* генома человека специалисты использовали гомозиготную клеточную линию CHM13.

Полный секвенированный геном человека (T2T-CHM13) насчитывает 3 055 миллиардов пар оснований; он содержит полностью прочитанные 22 аутосомы и X-хромосому. В новом сиквенсе исправлено множество ошибок и добавлено около 200 миллионов пар оснований. Последовательность T2T-CHM13 включает все околоцентромерные и околотеломерные участки, короткие плечи пяти акроцентрических хромосом, длинные тандемные повторы и другие непрочитанные ранее элементы.

В черновой последовательности T2T-CHM13 насчитывается 63 494 гена, из которых 19 969, предположительно, кодируют белки. Впервые секвенированные последовательности ДНК содержат 2 226 паралогичных копий генов, 115 из которых, предположительно, кодируют белки.

Основным ограничением генома T2T-CHM13 авторы считают отсутствие Y-хромосомы. Ее секвенирование осуществляется на следующем этапе работы с использованием клеточной линии HG002 с кариотипом 46 XY.

Источник:

Nurk S., et al. // The complete sequence of a human genome // bioRxiv, published May 27, 2021, DOI: 10.1101/2021.05.26.445798