Bioinformatics Applied to the Y-DNA Haplogroup Q-M242: Male Navajo Genetic Marker

Begay, Jordyn (School: Navajo Preparatory School)

My study applies the emerging technology of Bioinformatics, which makes use of biology, computer science, mathematics, and statistics to analyze and interpret biological data, including of DNA. I used the nucleotide search tool called BLAST (Basic Local Alignment Search Tool) to search one of the main repositories of genetic sequence information, GenBank. GenBank is administered by NCBI (National Center for Biotechnology Information). The application of this search tool was focused on the study of the Y (Male) DNA Haplogroup Q-M242. Q-M242 was chosen because it is a genetic marker in male Navajos, with an incidence rate between 92.3% and 96%. Via the use of Primer-BLAST, the forward and reverse primers of Q-M242 were used to recover the entire genetic sequence of 366 nucleotides, in FASTA format. I then applied three-letter (A, C, G, T) “triplet” nucleotide sequences, called DNA Codons, to identify the amino acid sequence in Q-M242. My primer results were checked via the EMBOSS revseq and transeq tools, which were provided by EMBL’s European Bioinformatics Institute (EMBL-EBI). The amino acids in Q-M242 were then tallied in histogram form via a PERL computer program. Indications are that Q-M242 originated some 30,000 years ago, in the Altai region of Central Asia, sandwiched between Russia, Kazakhstan, Mongolia, and China. From there, my people migrated eastward, across the Bering Strait, to North America.