Application Notes Genomics & Informatics Vol. 2(3) 147-148, September 2004

MitGEN: Single Nucleotide Polymorphism DB Browser for Human Mitochondrial Genome

Hyun Seok Park[1,2,*] and Seung Uk Lee[1]

[1]Institute of Bioinformatics, Macrogen Inc., Seoul 153-023, Korea, [2]Department of Computer Science, Ewha Womans University, Seoul 120-75, Korea

Abstract

Recently completed mitochondrial genome databses from public resources provide us with a better understanding of individual mitochondrial genomes for population genomics. By determining the substitution rate of the genomic sequences, it is plausible to derive dates on the phylogenetic tree and build a chronology of events in the evolution of human species. MitGEN is specially designed as a mitochondrial genome browser for analyzing, comparing and visualizing single nucleotide polymorphism for human mitochondrial genomes between human races for comparative genomics. It is a standalone application and is available free for non-commercial work.

Availability : http://snp.macrogen.co.kr

Keywords: visualization, mitochondrial genome, bioinformatics

Summary

The human mitochondrial genome is a closed circular double-strand DNA molecule of 16,569 bp in length. As sequences of the public mitochondrial genome increase, researchers infer diverse origins of human species to build evolutionary trees through data mining of large scale SNP data. Molecular anthropologists have been comparing the DNA of living humans of diverse origins to build evolutionary trees. This work is useful for researchers not only to reveal the secret of human revolutions, but also to identify associations between SNPs. However, the main difficulty for an efficient exploitation lies in the fact that there are not many efficient visualization tools developed so far. Consequently, there is a strong need for a user-friendly system allowing the user to handle and combine large and diverse SNP data sets from the different mitochondrial databases.

*Corresponding author: E-mail hspark@macrogen.co, Tel +82-2-2113-7007, Fax +82-2-2113-701
Accepted 13 August 2004

General Features of MitGEN: MitGEN is a mitochondrial genome browser for comparing genomic variations of individual sequences on a pair-wise format. Version 1.1, which was released in September 2004, is the latest version of MitGEN for Microsoft Windows platforms. MitGEN provides zooming capabilities, extensive sequence-alignment capabilities, and the option of visualizing multiple mitochondrial genomes in the main window.

The large middle frame in figure 1 is the main Genome Viewer. The 'zoom' icon on the narrow bottom frame provides zoomable and scrollable feature display down to sequence level optimized for display of large regions of genome for moving within and between scaffolds (or fragments) as in the bottom left corner of

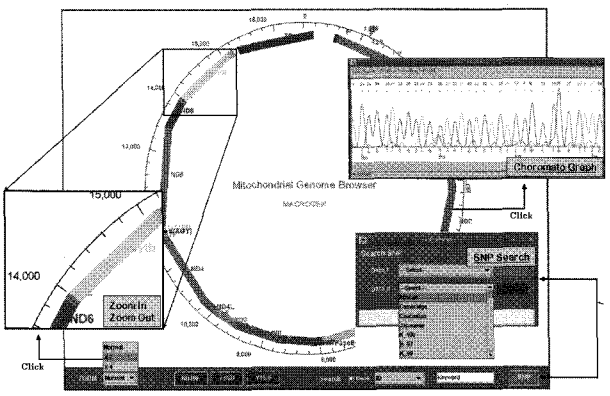

Fig. 1. General Features of MitGEN

figure 1. Users also want to check the evidence for individual DNA sequences by inspecting the chromatograms from which the base calls were deduced. On clicking the specific genome area, a pop-up window will be shown to display the chromatography along with base-calling quality values as in the top right corner of figure 1. To click on 'View Sequence' will display up to 1000 bps around the specified spot as in the top right corner, where the sequences can be inspected by scrolling the bar in the bottom frame. To click on 'SNP' invokes the genome comparison and SNP search mode between individuals or among races as in figure 2.

Comparing Individual Genomes: It is frequent to compare the standard Cambridge data with other mitochondrial DNA. MitGEN makes it possible to compare two mitochondrial genomes for comparative studies. As was previously mentioned, clicking on 'SNP' will invoke the SNP search

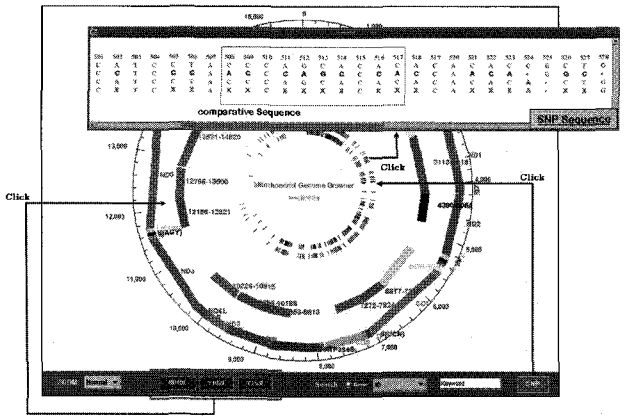

Fig. 2. Comparison Mode of MitGEN

mode between individuals or among races. For example, the main window in figure 2 displays two circular genomes rather than one. The outer circle represents the standard Cambridge data, and the inner circle represents the data in comparison. The Genome Browser in figure 2 shows that MitGEN can display not only the completed g enomes but also the fragments of partially sequenced genomes. However, the individual genomes should be pre-aligned according to the Cambridge Reference Sequence before being displayed. MitGEN provides the option of displaying multiple sequence alignments among the same races or similar groups. The different SNP sites will be displayed in different colors as shown in the top corner of figure 2. Clicking on certain features of MitGEN also searches our database for listings of genes or SNP sites. Users can connect directly to the mitochondrial web sites or databases for the latest human mitochondrial annotation.