BIOINFORMATICS SOFTWARE SOLUTIONS FOR NGS DATA ANALYSIS USING THE ION TORRENT PLATFORM

Narasimhan Rajagopalan, Sharon Wotton, Robert Lagace, Chien-Wei Chen, Joseph Chang, Reina Langit, Sharada Vijayachander, Carolina Dallet, Priti Sabadra, Gloria Lam, Human Identification Division, Thermos Fisher Scientific

Data analysis with Next Gen Sequencing is a very challenging task. Traditionally, analysts are used to working with STR data from CE where it was possible to interpret data by visual inspection. But with NGS, when analyzing a large set of markers, it becomes impossible to manually interpret the data. NGS systems produce orders of magnitude more data than traditional systems and it is a challenge to store and manage this data. Also, NGS provides greater data resolution than traditional methods.

Thermo Fisher Scientific has developed a software solution for analyzing and interpreting SNP, STR and mitochondrial data using the Ion Torrent™ platform. Users can obtain the final answers they look for when running these assays, interact with the data using a user-friendly UI, obtain an analysis report and export data for further analysis. These software run within the Torrent Suite™ software as plugins and can be launched automatically to provide a seamless data analysis workflow.

There are separate plugins for each of the three applications: HID_SNP_Genotyper for analyzing SNPs, HID_STR_Genotyper for analyzing STRs and variantCallerForMtDNA for mito and they are available for download and install from the Torrent Browser Plugin Store.

The HID_SNP_Genotyper** plugin performs genotype calling on a specified set of target and displays the most likely geo ancestry for a sample on a world map thereby presenting the final answer that the user is looking for in an easily interpretable manner.

The HID_STR_Genotyper* analyzes STR data using a novel algorithm to perform allele calling and reports both the allele length and sequence. Users can visualize the coverage data in a manner very similar to CE and also drill down to see the sequence variations. Alleles that are of the same length but are different in sequences are highlighted automatically.

variantCallerForMtDNA** performs variant calling and outputs the most likely haplogroup for each sample in the run.

Thermo Fisher is committed to providing an end-to-end NGS solution for forensic customers. The software solutions that were developed for this will be presented.

* For Forensic or Paternity Use Only.
** For Research, Forensic, Paternity and Cell Line Authentication. Not for use in Diagnostic and Therapeutic applications.

When used for purposes other than Human Identification the instruments cited are for Research Use Only.

Not for use in diagnostic procedures.