mitoCaller Tutorial

Prepared by Yong Qian and Jun Ding
Updated on October 25, 2015

---

1. Unzip file “mitoCaller.zip”. In the current directory, there will be several folders: mitoCaller, libStatGen, genome_ref, tclap-1.2.0

2. Go to folder “mitoCaller” and compile the mitoCaller package by doing:
   make
   If it is successful, the last line of output would be similar to:
   make[1]: Leaving directory `.../.../mitoCaller/src'
   In the folder mitoCaller/bin, an executable file “mitoCaller” will be generated.

3. Run the following command with your own specifications to analyze one individual’s sequencing data (i.e., one bam file):

   bin/mitoCaller -m -b mtDNA.reads.only.bam -r reference.genome > mtDNA.calls.output.summary
   NOTE: -m is MANDATORY for the mtDNA variant calling!

   - **mtDNA.reads.only.bam** is the bam file with mtDNA sequencing reads only, it can be generated using the whole-genome bam file by the following command: samtools view -bh whole.genome.bam MT > mtDNA.reads.only.bam
   - **reference.genome** is the reference genome; a copy of the human reference genome (GRCh37 assembly with decoy sequences, as available in the 1000 Genomes Project ftp site, ftp://ftp.1000genomes.ebi.ac.uk) is provided in genome_ref/human.g1k.v37.fa
   - **mtDNA.calls.output.summary** is the output file. Refer to "mtDNA_variant_caller_output_annotation.xlsx" for detailed annotation of the output file and how to extract homoplasmies and heteroplasmies from the output file.