We're happy to help you run your basic variant analysis here at MiGS, but it is important that you know how your data was handled should you choose to publish using the data.

After your genomic DNA has been sequenced, your data consists of two sequencing read files: R1 and R2. These files are the forward and reverse reads for your sample and while the variant calling software used by MiGS does not take into consideration the paired nature of these reads, both are used in the variant calling process. These reads have already been trimmed to remove the adapters from the sequences and will not contain any index reads (these reads are separate files and are not typically provided to you as very few users have any need for them).

During the intake process for your order, you provided a file or a link to a file that you wished to be used as a reference. By default, we will always select .gbff or .gff formats if they are available as these will include annotations (fasta files do not include these). MiGS uses the variant caller breseq (currently version 0.36.1) to align and compare your sequencing data to this file using the following commands:

breseq -r [reference file] [sample name]_.fastq.gz [sample name]_fastq.gz -o [output file name]

Your data analysis is now complete and will be uploaded to your Box folder. Each sample is contained in a compressed file format. **We recommend that you download the entire file to your local computer and then extract the compressed contents. Begin looking at your analysis by opening the output folder and selecting the index.html file.** This file will allow you to navigate through not only the reads but also the evidence used to make these variant predictions. Additional information on the software can be found at:
https://barricklab.org/twiki/pub/Lab/ToolsBacterialGenomeResequencing/documentation/

Please use the following citation for breseq:

Deatherage, D.E., Barrick, J.E. (2014) Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol. Biol. 1151: 165–188.