SOURCE OF GENOMES AND ANNOTATIONS USED FOR ALIGNMENT OF READS

Refer to 09\_MakeReference.sh for code used to pull files from the ensembl database and concatenate the parasite and human genome.

If a new version has been released, these files will not be under current. They can be retrieved by using the version number if the links below are broken. Consult with a bioinformatician on retrieving large files on the computing cluster using the ensembl database FTP functionality.

***Plasmodium falciparum* genome (ASM276v2):**

[ftp://ftp.ensemblgenomes.org/pub/protists/current/fasta/plasmodium\_falciparum/dna/**Plasmodium\_falciparum.ASM276v2.dna.toplevel.fa.gz**](ftp://ftp.ensemblgenomes.org/pub/protists/current/fasta/plasmodium_falciparum/dna/Plasmodium_falciparum.ASM276v2.dna.toplevel.fa.gz)

***Plasmodium falciparum* annotations (ASM276v2.52):**

[ftp://ftp.ensemblgenomes.org/pub/protists/current/gtf/plasmodium\_falciparum/**Plasmodium\_falciparum.ASM276v2.52.gtf.gz**](ftp://ftp.ensemblgenomes.org/pub/protists/current/gtf/plasmodium_falciparum/Plasmodium_falciparum.ASM276v2.52.gtf.gz)

***Homo sapiens* genome (GRCh38):**

[ftp://ftp.ensembl.org/pub/release-100/fasta/homo\_sapiens/dna/**Homo\_sapiens.GRCh38.dna.primary\_assembly.fa.gz**](ftp://ftp.ensembl.org/pub/release-100/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz)

***Homo sapiens* annotations (GRCh38.105):**

[ftp://ftp.ensembl.org/pub/current\_gtf/homo\_sapiens/**Homo\_sapiens.GRCh38.105.gtf.gz**](ftp://ftp.ensembl.org/pub/current_gtf/homo_sapiens/Homo_sapiens.GRCh38.105.gtf.gz)