Script to download the *Salmonella* Agona reference genome:

|  |
| --- |
| mkdir reference  cd reference  wget https://raw.githubusercontent.com/CFSAN-Biostatistics/snp-pipeline/master/snppipeline/data/agonaInputs/reference/NC\_011149.fasta |

Script to generate 1000 mutated genomes – assumes you have downloaded SNP mutator:

|  |
| --- |
| #!/bin/bash  mkdir -p replicates  snpmutator.py -r 1 -o summary.tsv -n 1000 reference/NC\_011149.fasta  mv \*.fasta replicates |

Script to generate simulated Illumina fastq files at 100x coverage – assumes software art installed:

|  |
| --- |
| #!/bin/bash  function runart  {  basefile=$1  inputFile=replicates/$basefile.fasta  outputDir=artsamples/$basefile  mkdir -p $outputDir  outputFile=$outputDir/${basefile}\_  art\_illumina --noALN -i $inputFile -p -l 250 -ss MS -f 100 -m 500 -s 10 -rs 1 -o $outputFile  }  export -f runart  seq 1 1000 | xargs -I @ -P 8 bash -c "runart NC\_011149\_mutated\_@" |

Script to generate simulated Illumina fastq files at 20x coverage – assumes the software art installed:

|  |
| --- |
| #!/bin/bash  function runart  {  basefile=$1  inputFile=replicates/$basefile.fasta  outputDir=artsamples/$basefile  mkdir -p $outputDir  outputFile=$outputDir/${basefile}\_  art\_illumina --noALN -i $inputFile -p -l 250 -ss MS -f 20 -m 500 -s 10 -rs 1 -o $outputFile  }  export -f runart  seq 1 1000 | xargs -I @ -P 8 bash -c "runart NC\_011149\_mutated\_@" |

Script to run the CFSAN SNP Pipeline – assumes CFSAN SNP Pipeline is installed along with Bowtie2 v2.2.2 , VarScan v2.3.9, and samtools v0.1.19:

|  |
| --- |
| mkdir snppipelineworkdir  copy\_snppipeline\_data.py configurationFile  echo 'Bowtie2Align\_ExtraParams="--reorder -q -X 547"' >> snppipeline.conf  run\_snp\_pipeline.sh –m soft -c snppipeline.conf -o snppipelineworkdir -s artsamples NC\_011149.fasta |