#!/bin/bash

### Antismash v6 for HPC ###

# Location: M3

# Inputs: fasta nt assemblies (folder)

# Requires antismash

# Version: antismash v6.0

# Newer version now available (v6.1)

# Array job

inp\_dir="../MAGs\_new\_names/"

op\_dir="./"

ext=".fa"

conda\_source="/monfs00/scratch/lwoo0007/WoodsL/miniconda/bin/activate"

conda\_env="/monfs00/scratch/lwoo0007/WoodsL/miniconda/conda/envs/antismash\_env"

script\_name="antismash\_array\_job.sh"

script\_name\_helper="antismash\_array\_helper.sh"

if [ ! -d ${op\_dir} ]

then

mkdir ${op\_dir}

fi

#Clearing old scripts

if [ -f ${script\_name} ]

then

rm ${script\_name}

fi

if [ -f ${script\_name\_helper} ]

then

rm ${script\_name\_helper}

fi

cat << EOF >> ${script\_name}

#!/bin/bash

#SBATCH --mail-type=END,FAIL

#SBATCH --mail-user=laura.woods1@monash.edu

#SBATCH -c 4

#SBATCH --time=06:00:00

#SBATCH --partition=short

#SBATCH --job-name ${script\_name}

source ${conda\_source} base

conda activate ${conda\_env}

inp\_dir="${inp\_dir}"

op\_dir="${op\_dir}"

ext="${ext}"

EOF

cat << 'EOF' >> ${script\_name}

echo "This job in the array has:"

echo "- SLURM\_JOB\_ID=${SLURM\_JOB\_ID}"

echo "- SLURM\_ARRAY\_TASK\_ID=${SLURM\_ARRAY\_TASK\_ID}"

# grab our filename from a directory listing

FILES=($(ls -1 ${inp\_dir}/\*${ext}))

FILENAME=${FILES[$SLURM\_ARRAY\_TASK\_ID]}

echo "My input file is ${FILENAME}"

sample=$(basename ${FILENAME} ${ext})

echo "COMMAND: antismash -c 8 --fullhmmer --cb-general --cb-knownclusters --cb-subclusters --asf --pfam2go --smcog-trees --tigrfam --rre --cc-mibig --html-start-compact --genefinding-tool prodigal-m --output-dir ${op\_dir}/${sample}\_antismash\_all\_options/ ${FILENAME}"

antismash -c 8 --fullhmmer --cb-general --cb-knownclusters --cb-subclusters --asf --pfam2go --smcog-trees --tigrfam --rre --cc-mibig --html-start-compact --genefinding-tool prodigal-m --output-dir ${op\_dir}/${sample}\_antismash\_all\_options/ ${FILENAME}

echo "Finished"

EOF

cat << EOF >> ${script\_name\_helper}

#!/bin/bash

inp\_dir="${inp\_dir}"

op\_dir="${op\_dir}"

ext="${ext}"

EOF

cat << 'EOF' >> ${script\_name\_helper}

# get count of files in this directory

NUMFILES=$(ls -1 ${inp\_dir}/\*${ext} | wc -l)

echo "There are a total of ${NUMFILES} input files"

# subtract 1 as we have to use zero-based indexing (first element is 0)

ZBNUMFILES=$(($NUMFILES - 1))

if [ $ZBNUMFILES != 0 ]; then

echo "COMMAND: sbatch --array=0-$ZBNUMFILES antismash\_array\_job.sh"

sbatch --array=0-$ZBNUMFILES antismash\_array\_job.sh

else

echo "No jobs to submit, since no input files in this directory."

fi

EOF

chmod +x ${script\_name}

chmod +x ${script\_name\_helper}

./antismash\_array\_helper.sh