***Idotea baltica* RNA-Seq Preprocessing**

Rcorrector 2015-11-06

**INSTALLING**

Change directory to USERAPPL and create new directory

cd $USERAPPL

mkdir rcorrector

Go to the new directory and download the installation package

cd rcorrector

git clone https:// github.com/mourisl/rcorrector.git

Compile

cd rcorrector

make

Usage

perl $USERAPPL/rcorrector/rcorrector/run\_rcorrector.pl

**RUNNING**

Using TRIMMOMATIC output (25-mers) (rcorrector.sh)

perl $USERAPPL/rcorrector/rcorrector/run\_rcorrector.pl \

-1 $WRKDIR/reptile\_data/trimmomatic\_unzipped/out\_fw\_paired.fastq \

-2 $WRKDIR/reptile\_data/trimmomatic\_unzipped/out\_rev\_paired.fastq \

-t 16 -k 25

\*\*\* ~3hrs (16 cores, 16GB RAM)

Using TRIMMOMATIC output (**20**-mers) (rcor\_20mers.sh)

perl $USERAPPL/rcorrector/rcorrector/run\_rcorrector.pl \

-1 $WRKDIR/reptile\_data/trimmomatic\_unzipped/out\_fw\_paired.fastq \

-2 $WRKDIR/reptile\_data/trimmomatic\_unzipped/out\_rev\_paired.fastq \

-t 16 -k 20

\*\*\* ~3hrs (16 cores, 12GB RAM)

**OUTPUT**

File err\_7584578.txt:

Stored 139073753 kmers

Weak kmer threshold rate: 0.006066

Bad quality threshold is ;

Processed 524041096 reads

Corrected 44876577 bases.

File names:

out\_fw\_paired.cor.fq

out\_rev\_paired.cor.fq

File err\_7590894.txt (20mers):

Stored 128297607 kmers

Weak kmer threshold rate: 0.009025

Bad quality threshold is ;

Processed 524041096 reads

Corrected 46535134 bases.

File names:

out\_fw\_paired.cor.fq

out\_rev\_paired.cor.fq

**RESULTS**

1,658,557 more bases were corrected with the smaller k-mer.

This is not a very significant change since there are 500M reads each with an average length of 125bp.