VERSION 5

OCT 24, 2017

OPEN ACCESS

DOI:
dx.doi.org/10.17504/protocols.io.ketcten

Protocol Citation: James E Thornton Jr 2017. Read Recruitment Using Bowtie2. **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.ketcten>

License: This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

Created: Oct 24, 2017

Last Modified: Feb 16, 2018

PROTOCOL integer ID:
8371

Read Recruitment Using Bowtie2 V.5

James E Thornton Jr¹

¹Hurwitz Lab

Metafunc Course 2017

James E Thornton Jr
Hurwitz Lab

ABSTRACT

This protocol details how to perform read recruitment to the contigs generated from the previous protocol using Bowtie2.

1 Log into the HPC.

Command

```
$ ssh hpc  
$ ocelote
```

PAIRED END CLEAN UP

2 For paired end only:

Move into your fasta directory. Ensure only 1 pair from each file is in the fasta directory. Keep the pair that is "_1.fasta". Move all "_2.fasta" files into the fastq directory.

Command

```
$ cd /rsgrps/bh_class/username/assembly/fastq  
$ mv *_2.fasta ../../fastq
```

Note

In the previous protocol we ended up only using 1 pair for assembly to simplify the process. In this protocol we only want to map the one pair that we used during assembly. Make sure the fasta directory only contains the _1.fasta of the pair.

3 Move into the assembly/megahit-out directory created from the previous protocol.

Command

```
$ cd /rsgrps/bh_class/username/assembly/megahit-out
```

4 Simplify the fasta headers of your final.contigs.fa file using fasta_renamer from the Fastx toolkit.

Command

```
$ module load fastx  
$ fastx_renamer -n COUNT -i final.contigs.fa -o fixed-contigs.fa
```

Note

Important: If you get 'This looks like a multi-line FASTA file', try this alternative method to rename sequence headers:

```
awk '/^>/{print ">" ++i; next}{print}' < contigs_in.fa > contigs_out.fa
```

- 5 Append your fixed-contigs.fa file with your partners fixed-contigs.fa file.

Command

```
$ cat fixed-contigs.fa /rsgrps/bh_class/partnerusername/assembly/megahit-out/fixed-contigs.fa  
> </ProtocolCommand>  
<ProtocolNote id=
```

- 6 Move into your project directory. Then create a 'read_recruit' directory. Move into that directory.

Command

```
$ cd /rsgrps/bh_class/username  
$ mkdir read_recruit  
$ cd !$
```

- 7 Create a bam and bowtie2 index directory.

Command

```
$ mkdir bam bt2_index
```

- 8 Move into the contig indexing directory. And create the contig index.

Command

```
$ cd bt2_index  
$ module load bowtie2  
$ bowtie2-build -f /rsgrps/bh_class/username/assembly/megahit-out/final-contigs.fa  
contig_index
```

Note

This step could take awhile depending on the size of your contig file.

- 9 Move into your bam directory found at /rsgrps/bh_class/username/read_recruit/bam

Command

```
$ cd ../bam
```

- 10 Make directories for standard error and standard out.

Command

```
$ mkdir std-err std-out
```

- 11 Copy the following script named bt2_align.sh. Edit the username variables found in the script.

Note

Many of you actually had your fasta files located in:
/rsgroups/bh_class/username/assembly/fasta

Which is fine, just make sure the script points to the correct location.

Command

```
#!/bin/bash
#PBS -W group_list=bh_class
#PBS -q windfall
#PBS -l select=1:ncpus=4:mem=15gb
#PBS -l pvmem=14gb
#PBS -l walltime=24:00:00
#PBS -l cput=24:00:00
#PBS -M netid@email.arizona.edu
#PBS -m bea
echo "my job_id is: ${PBS_JOBID}"
#####change here #####
FASTA_DIR=
```

Note

Important: CONTIGS variable is the contigs that you combined with your partner.

12 Submit the job.

Command

```
$ qsub -e std-err/ -o std-out/ bt2_align.sh
```

13 Upon job completion navigate to your bam std-err directory.

Command

```
$ cd /rsgrps/bh_class/username/read_recruit/bam/std-err
```

- 14** Cat the standard error file to view the alignment rate for each file.

Command

Your file name will differ

```
$ cat 881767.head1.cm.cluster.ER
```