

What to do on the Cloud

April 18, 2012

This document describes what you should do to get started analyzing your data on the cloud.

1 Software to install

- Install the development version of R and needed packages.
 - http://webpages.uidaho.edu/msettles/courses/bcb504Sp12/lectures/Cloud_install/R-build-devel
- Install sequence mapping software (only one of is necessary)
 - BWA** <http://bio-bwa.sourceforge.net/>
 - Bowtie2** <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>
- Install samtools (vcftools)
 - <http://samtools.sourceforge.net/>

2 Move your data from Central Storage to you cloud server

- Copy your reads over
- Copy your assembled sequences over

3 Analyze your data

- Map your reads (produces SAM file)
- Call variants with samtools
 - use samtools view to convert sam to BAM
 - use samtools sort to sort
 - use samtools mpileup and bcftools to call variants [<http://samtools.sourceforge.net/cns0.shtml>]
 - use samtools depth to produce coverage by quality information