Bioinformatics software listing

The following software packages are installed on the RCAC compute clusters.

To access them, use the `module` command as described in the [RCAC user manual](#).  

First, you need to configure it by entering the following commands:

```
module load bioinfo
```

After that, you can use the `module load` command to access the software you want to use. For instance, `module load blast` will enable the NCBI BLAST software.  

Contact the [Bioinformatics Core Director](#) if you want additional software installed.

1. **ABySS**
   
   module name: `abyss`
   version 2.0.2-maxk64
   
   Other versions: 1.3.2, 1.3.6-conte-maxk64, 1.3.6-gcc-maxk64, 1.3.6-gcc-maxk96, 1.3.6-gcc-maxk128, 1.3.6-gcc-maxk256, 1.3.6-intel-maxk64, 1.3.6-intel-maxk96, 1.3.6-intel-maxk128, 1.3.6-intel-maxk256, 1.3.7-gcc-maxk64, 1.3.7-gcc-maxk96, 1.3.7-gcc-maxk128, 1.3.7-gcc-maxk256, 1.3.7-gcc-maxk64-nosparseshash, 1.3.7-gcc-maxk96-nosparseshash, 1.3.7-gcc-maxk128-nosparseshash, 1.3.7-gcc-maxk256-nosparseshash, 1.3.7-gcc-maxk64-nosparseshash, 1.3.7-gcc-maxk96-nosparseshash, 1.3.7-gcc-maxk128-nosparseshash, 1.3.7-gcc-maxk256-nosparseshash, 1.3.7-intel-maxk64, 1.3.7-intel-maxk96, 1.3.7-intel-maxk128, 1.3.7-intel-maxk256, 1.3.8-maxk64, 1.3.8-maxk96, 1.3.8-maxk128, 1.3.8-maxk256, 1.5.0-gcc-maxk64, 1.5.0-gcc-maxk64-nosparseshash, 1.5.0-gcc-maxk96, 1.5.0-gcc-maxk96-nosparseshash, 1.5.0-gcc-maxk128, 1.5.0-gcc-maxk128-nosparseshash, 1.5.0-gcc-maxk256, 1.5.0-gcc-maxk256-nosparseshash, 1.5.0-intel-maxk64, 1.5.0-intel-maxk96, 1.5.0-intel-maxk128, 1.5.0-intel-maxk256, 1.5.1-gcc-maxk64, 1.5.1-gcc-maxk96, 1.5.1-gcc-maxk128, 1.5.1-gcc-maxk256, 1.5.1-intel-maxk64, 1.5.1-intel-maxk96, 1.5.1-intel-maxk128, 1.5.1-intel-maxk256, 1.5.2-gcc-maxk64, 1.5.2-gcc-maxk96, 1.5.2-gcc-maxk128, 1.5.2-gcc-maxk256, 1.5.2-gcc-maxk64-nosparseshash, 1.5.2-gcc-maxk96-nosparseshash, 1.5.2-gcc-maxk128-nosparseshash, 1.5.2-gcc-maxk256-nosparseshash, 1.5.2-intel-maxk64, 1.5.2-intel-maxk96, 1.5.2-intel-maxk128, 1.5.2-intel-maxk256, 1.5.2-nompi-maxk64, 1.5.2-nompi-maxk96, 1.5.2-nompi-maxk128, 1.5.2-nompi-maxk256, 1.5.2-nompi-maxk64-nosparseshash, 1.5.2-nompi-maxk96-nosparseshash, 1.5.2-nompi-maxk128-nosparseshash, 1.5.2-nompi-maxk256-nosparseshash, 1.5.3-gcc-maxk64, 1.5.3-gcc-maxk64-nosparseshash, 1.5.3-gcc-maxk96, 1.5.3-gcc-maxk96-nosparseshash, 1.5.3-gcc-maxk128, 1.5.3-gcc-maxk128-nosparseshash, 1.5.3-gcc-maxk256, 1.5.3-gcc-maxk256-nosparseshash, 1.5.3-intel-maxk64, 1.5.3-intel-maxk96, 1.5.3-intel-maxk128, 1.5.3-intel-maxk256, 1.5.4-gcc-maxk64, 1.5.4-gcc-maxk64-nosparseshash, 1.5.4-gcc-maxk96, 1.5.4-gcc-maxk96-nosparseshash, 1.5.4-gcc-maxk128, 1.5.4-gcc-maxk128-nosparseshash, 1.5.4-gcc-maxk256, 1.5.4-gcc-maxk256-nosparseshash, 1.5.4-intel-maxk64, 1.5.4-intel-maxk96, 1.5.4-intel-maxk128, 1.5.4-intel-maxk256

**Assembly By Short Sequences**

http://www.bcgsc.ca/platform/bioinfo/software/abyss

Info in /group/bioinfo/apps/apps/abyss-2.0.2-maxk64/share/doc/abyss

Use "man ABYSS" (also man pages for abyss-pe and abyss-tofasq)

Info: http://www.bcgsc.ca/downloads/abyss/doc/

This version is compiled with support for maximum k-mer size of 64

Uses gcc compiler suite

Depends on gcc
Depends on Boost
Depends on openmpi

1. **ABySS samtobreak**

   module name: `abyss-samtobreak`
   version 1.3.7

   Calculate contig and scaffold contiguity and correctness metrics

   Part of ABySS: http://www.bcgsc.ca/platform/bioinfo/software/abyss

   Info probably in /group/bioinfo/apps/apps/abyss-samtobreak-1.3.7/share/doc/abyss

   Info: http://www.bcgsc.ca/downloads/abyss/doc/

1. **agrep**

   module name: `agrep`
   version 2.04

   Fast text searching allowing errors

   ftp://ftp.cs.arizona.edu/agrep

   Info: /group/bioinfo/apps/apps/agrep-2.04/README

   Info: man agrep

   Depends on gcc

1. **ALLPATHS-LG**

   module name: `allpathslg`
   version 52488

   Other versions: 40969, 41634, 42667, 48686, 52415
Whole genome shotgun assembler
http://www.broadinstitute.org/software/allpaths-lg/blog/

Depends on gcc
Depends on gmp
Depends on picard-tools

1. **Amalgamated likelihood estimation (ALE)**
   module name: ALE
   version 2017-09-21
   Probabilistically explore reconciled gene trees
   https://github.com/ssolo/ALE
   Depends on Boost/1.65.1
   Depends on BppSuite

2. **AMOS**
   module name: amos
   version 3.1.0
   Other versions: 3.0.0
   AMOS (A Modular, Open-Source whole genome assembler)
   Info in /group/bioinfo/apps/apps/amos-3.1.0/share/doc/
   Depends on gcc
   Depends on perl
   Depends on blat
   Depends on MUMmer
   Depends on cross_match
   Depends on lucy
   Depends on rhel5libs

3. **AmpliconNoise**
   module name: rdp_classifier
   version 2.12
   Other versions: 2.2, 2.4
   Bayesian classifier
   http://sourceforge.net/projects/rdp-classifier/
   http://rdp.cme.msu.edu/
   Info: /group/bioinfo/apps/apps/rdp_classifier_2.12/README
   Docs: /group/bioinfo/apps/apps/rdp_classifier_2.12/docs/
   Usage: java -Xmx1g -jar $RDP_JAR_PATH
   Depends on java

4. **ANNOVAR**
   module name: annovar
   version 2015-03-22
   Functionally annotate genetic variants detected from diverse genomes
   Forum: https://groups.google.com/forum/#!forum/annovar
   Please cite ANNOVAR paper in your publication:
   Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from
   next-generation sequencing data, Nucleic Acids Research, 38:e164, 2010
   Depends on perl

5. **antiSMASH (antibiotics and Secondary Metabolite Analysis SHell)**
   module name: antiSMASH
version 4.0.2

Genome-wide analysis of gene clusters in bacterial and fungal genomes
https://bitbucket.org/antismash/antismash/overview
Web service: http://antismash.secondarymetabolites.org
Manual: http://docs.antismash.secondarymetabolites.org/

Includes PFAM and ClusterBlast 20170105_v8_31 databases

1. Apache Ant
   module name: apache-ant
   version 1.9.3
   Other versions: 1.8.4

   Build of Java applications
   http://ant.apache.org/

   Depends on java

1. ARAGORN
   module name: ARAGORN
   version 1.2.38

   tRNA (and tmRNA) detection
   http://mbio-serv2.mbioekol.lu.se/ARAGORN/

   Depends on gcc

1. ARC (Assembly by Reduced Complexity)
   module name: ARC
   version 1.1.3

   Pipeline for iterative, reference guided de novo genome assemblies
   http://ibest.github.io/ARC/
   Forum: https://groups.google.com/forum/#!forum/arc-assembly

   Depends on blat/35-fastq
   Depends on bowtie2
   Depends on spades
   Depends on biopython/2.7.12

1. Argtable
   module name: argtable
   version 2-13

   ANSI C command line parser.
   http://argtable.sourceforge.net/
   man pages available

   Depends on gcc

1. Artemis
   module name: Artemis
   version 15.0.0

   Genome Browser and Annotation Tool
   http://www.sanger.ac.uk/resources/software/artemis/
   Includes act, art and dnaplotter
   Needs X11 support to run GUI

   Depends on java
1. **Aspera Connect Client including ascp**
   module name: *aspera-connect*
   version 3.1.1.70545
   Other versions: 3.0.1.64802

   High speed data transfer client for NCBI SRA and other databases
   http://asperasoft.com/software/transfer-clients/connect-web-browser-plug-in/
   http://downloads.asperasoft.com/download_connect/
   Firefox plug-in in /group/bioinfo/apps/apps/aspera-connect-3.1.1.70545/firefox-plugin
   Copy it to ~/.mozilla/plugins/ to use it
   OpenSSH public key: /group/bioinfo/apps/apps/aspera-connect-3.1.1.70545/etc/asperaweb_id_dsa.openssh

   Depends on gcc

1. **AUGUSTUS**
   module name: *AUGUSTUS*
   version 3.2.1
   Other versions: 2.7, 3.0.3, 3.1

   Ab initio gene prediction in eukaryotic genomic sequences
   http://bioinf.uni-greifswald.de/augustus/
   Info: /group/bioinfo/apps/apps/augustus-3.2.1/README.TXT
   Info: /group/bioinfo/apps/apps/augustus-3.2.1/docs
   Web interface: http://bioinf.uni-greifswald.de/augustus/submission

   The program may need to alter its configuration files, so you might need put them somewhere else. Follow the procedure below to put the config files in your scratch space:

   $ mkdir $RCAC_SCRATCH/augustus
   $ cp -rp $AUGUSTUS_CONFIG_PATH $RCAC_SCRATCH/augustus
   $ export AUGUSTUS_CONFIG_PATH=$RCAC_SCRATCH/augustus/config

   Depends on gcc
   Depends on bamtools/2.4.0

1. **Automatically Tuned Linear Algebra Software (ATLAS)**
   module name: *ATLAS*
   version 3.10.1

   http://math-atlas.sourceforge.net/
   compiled with gcc 4.7.2

   Depends on gcc

1. **bambam**
   module name: *bambam*
   version 1.4

   Tools for DNA next-gen sequence analysis
   http://udall-lab.byu.edu/Research/Software/BamBam
   Info: /group/bioinfo/apps/apps/bambam-1.4/README

   Depends on gcc
   Depends on perl
   Depends on bamtools/2.4.0
   Depends on htslib/1.3.2
   Depends on zlib/1.2.8

1. **BamM**
   module name: *BamM*
version 1.7.3

Metagenomics-focused BAM file manipulation
https://github.com/ecogenomics/BamM
http://ecogenomics.github.io/BamM/

Depends on biopython/2.7.12
Depends on bwa/0.7.12
Depends on samtools/1.2

1. **BAMStats**
   module name: **BAMStats**
   version 1.25

   Interactive desktop GUI tool for summarising Next Generation Sequencing alignments
   http://bamstats.sourceforge.net/

   Command line wrapper provided for BAMStats-GUI (e.g. use "BAMStats-GUI" to start it with 2 GB of memory)

   Specify the amount of RAM for the programs like this:

   java -Xmx6g -jar /group/bioinfo/apps/apps/BAMStats-1.25/BAMStats-GUI-1.25.jar

   Running the equivalent command line tool:

   java -Xmx4g -jar /group/bioinfo/apps/apps/BAMStats-1.25/BAMStats-1.25.jar -i

   Between 2 and 8 GB of memory, depending on reference size, is typically required (specified with the -Xmx flag).

   Depends on java

1. **bamtools**
   module name: **bamtools**
   version 2.4.0
   Other versions: 2.1.0-466bc50, 2.2.3, 2.3.0

   Programmer API and an end-user toolkit for handling BAM files
   https://github.com/pezmaster31/bamtools

   Depends on gcc

1. **bamUtil**
   module name: **bamUtil**
   version 1.0.10
   Other versions: 1.0.3

   Perform operations on SAM/BAM files
   http://genome.sph.umich.edu/wiki/BamUtil
   Info: http://genome.sph.umich.edu/wiki/BamUtil#Programs

   Depends on gcc

1. **Barrnap**
   module name: **barrnap**
   version 0.8

   Predicts the location of ribosomal RNA genes in genomes
   http://www.vicbioinformatics.com/software.barrnap.shtml
   https://github.com/tseemann/barrnap/

   Depends on gcc
   Depends on HMMER
1. **Batman**
   module name: `batman`
   version alpha-0.2.3
   Other versions: 20090617
   
   A Bayesian Tool for Methylation Analysis
   http://td-blade.gurdon.cam.ac.uk/software/batman/
   Manual: http://td-blade.gurdon.cam.ac.uk/software/batman/batmanual-alpha-0.2.3.pdf
   You need to be able to write to a MySQL database to use this tool.
   See: http://www.itap.purdue.edu/tlt/careeraccount/mysql.cfm
   
   Depends on java
   Depends on mysql-connector-java/5.0.5

1. **BayeScan**
   module name: `BayeScan`
   version 2.1
   
   Detecting natural selection from population-based genetic data
   http://cmpg.unibe.ch/software/bayescan/index.html
   Info: http://vcftools.sourceforge.net/docs.html
   
   Depends on R-bioconductor

1. **BBMap/BBTools**
   module name: `BBMap`
   version 37.38
   
   Other versions: 31.32, 31.40, 31.56, 32.06, 32.07, 32.14, 32.15, 32.23, 32.27, 32.32, 33.04, 33.08, 33.11, 33.12, 33.13, 33.21, 33.34, 33.40b, 33.41, 33.43, 33.44, 33.46, 33.51, 33.54, 33.57, 33.73b, 34.23, 34.24, 34.25, 34.26, 34.30, 34.33, 34.38, 34.41, 34.46, 34.48, 34.49, 34.56, 34.64, 34.65, 34.72, 34.79, 34.83, 34.86, 34.90, 34.92, 34.94, 34.97, 35.02, 35.07, 35.10, 35.14, 35.24, 35.27, 35.29, 35.32, 35.34, 35.37, 35.38, 35.40, 35.43, 35.48, 35.50, 35.51, 35.59, 35.66, 35.68, 35.69, 35.74, 35.76, 35.80, 35.82, 35.84, 35.85, 35.90, 35.91, 35.92, 36.02, 36.11, 36.14, 36.20, 36.27, 36.28, 36.30, 36.32, 36.38, 36.47, 36.49, 36.59, 36.62, 36.64, 36.66, 36.67, 36.71, 36.73, 36.76, 36.77, 36.81, 36.84, 36.85, 36.86, 36.92, 36.99, 37.02, 37.09, 37.10, 37.17, 37.22, 37.23, 37.24, 37.25, 37.28, 37.31, 37.32, 37.33, 37.34, 37.36
   
   Short read aligner, and other bioinformatic tools
   http://sourceforge.net/projects/bbmap/
   Info at /group/bioinfo/apps/apps/BBMap-37.38/docs
   Guides at /group/bioinfo/apps/apps/BBMap-37.38/docs/guides
   In particular, see /group/bioinfo/apps/apps/BBMap-37.38/docs/guides/BBMapGuide.txt
   
   Depends on java
   Depends on samtools

1. **bcftools**
   module name: `bcftools`
   version 1.6
   
   Other versions: 1.3, 1.3.1, 1.4, 1.5
   
   Reading/writing BCF2/VCF/gVCF files and calling/filtering/summarising SNP and short indel sequence variants
   http://www.htslib.org/
   Includes support for "bcftools polysomy ..."
   Info: man bcftools
   
   Depends on gcc
   Depends on gsl/2.3
   Depends on htslib/1.6

1. **bcl2fastq**
   module name: `bcl2fastq`
   version 2.19.1.403
   
   Other versions: 1.8.3, 1.8.4, 2.16.0.10, 2.18.0.12
   
   illumina BCL2FASTQ Conversion Software
   BCL conversion and demultiplexing of both unzipped and zipped bcl files
1. **BEAGLE**
   module name: beagle
   version 2.1.2
   Other versions: 1260
   General purpose library for evaluating the likelihood of sequence evolution on trees
   https://github.com/beagle-dev/beagle-lib
   http://code.google.com/p/beagle-lib/
   Depends on gcc
   Depends on java

1. **Beagle**
   module name: Beagle
   version 4.1
   Genotype calling, genotype phasing, imputation of ungenotyped markers, and identity-by-descent segment detection
   https://faculty.washington.edu/browning/beagle/beagle.html
   Utilities: https://faculty.washington.edu/browning/beagle_utilities/utilities.html
   Start it with: java [options] -jar $BEAGLE_JAR [arguments]
   Also includes bref; use java [options] -jar $BREF_JAR [arguments]
   Depends on java/8

1. **bedops**
   module name: bedops
   version 2.0
   Other versions: 1.2.5
   Highly scalable genomic feature operations
   http://code.google.com/p/bedops/
   Docs: http://code.google.com/p/bedops/wiki/summaryTable
   Examples: http://code.google.com/p/bedops/wiki/UsageExamples
   Forum: http://bedops.uwencode.org/forum
   Depends on biopython/2.7.3

1. **BEDTools**
   module name: BEDTools
   version 2.26.0-35-g6114307
   Other versions: 2.16.2, 2.17.0, 2.20.1, 2.21.0, 2.25.0
   The swiss army knife for genome arithmetic
   https://github.com/ark5x/bedtools2/
   Bugs: https://github.com/ark5x/bedtools2/issues
   Forum at http://groups.google.com/group/bedtools-discuss
   Depends on gcc
   Depends on samtools

1. **Bellman's GAP compiler**
   module name: gapc
   version 2014.10.07
   GAP-L to C++ translator/compiler
http://gapc.eu/compiler.html
Depends on gcc

1. **BESST**
   module name: *BESST*
   version 1.3.8
   Other versions: 1.1.2
   Scaffold for genomic assemblies
   https://github.com/ksahlin/BESST
   Depends on biopython/2.7.8

1. **BFAST**
   module name: *bfast*
   version 0.7.0a
   BLAT-like Fast Accurate Search Tool
   Depends on gcc

1. **bioawk**
   module name: *bioawk*
   version 1.0
   BWK awk modified for biological data
   https://github.com/lh3/bioawk
   Depends on gcc

1. **Bioinformatics utilities**
   module name: *ucsc-util*
   version Nov-10-2014
   Miscellaneous UCSC Genome Bioinformatics utilities
   http://genome.ucsc.edu/index.html
   http://hgdownload.software.ustc.edu/downloads.html
   Manual: http://hgdownload.cse.ustc.edu/admin/exe/linux.x86_64/FOOTER
   For help on the bigBed and bigWig applications see:
   http://genome.ucsc.edu/goldenPath/help/bigBed.html
   http://genome.ucsc.edu/goldenPath/help/bigWig.html
   Depends on gcc

1. **biom-format**
   module name: *biom-format*
   version 1.1.2
   Python module; format for representing counts of observations (e.g.,
   OTUs, KO categories, lipid types) in one or more biological samples
   (e.g., microbiome samples, genomes, metagenomes)
   http://biom-format.org
   Depends on biopython/2.7.3

1. **biotoolbox**
   module name: *biotoolbox*
   version 1.19
   Other versions: 1.8.4, 1.8.6
   Tools for querying and analysis of genomic data
http://code.google.com/p/biotoolbox/
Wiki: http://code.google.com/p/biotoolbox/w/list

Depends on perl/5.20.1

1. **Bismark**
   
   module name: bismark
   
   version 0.18.2
   
   Other versions: 0.10.1, 0.14.2, 0.14.5, 0.17.0
   
   Map bisulfite converted sequence reads and determine cytosine methylation states
   
   http://www.bioinformatics.babraham.ac.uk/projects/bismark/
   
   https://github.com/FelixKrueger/Bismark
   
   Manual: /group/bioinfo/apps/apps/Bismark-0.18.2/Docs/README.md
   
   
   Depends on perl
   
   Depends on bowtie
   
   Depends on bowtie2
   
   Depends on samtools

1. **BLASR**
   
   module name: blasr
   
   version 1.3.1
   
   https://github.com/PacificBiosciences/blasr
   
   The PacBio long read aligner
   
   Depends on HDF5/1.8.14

1. **blast2html**
   
   module name: blast2html
   
   version 82b8c9722996
   
   Blast XML to HTML conversion tool
   
   https://toolshed.g2.bx.psu.edu/repository/view_repository?changeset_revision=82b8c9722996&id=f32220daa6f46e04
   
   Info: /group/bioinfo/apps/apps/blast2html-82b8c9722996/README.rst
   
   Depends on biopython/2.7.8

1. **BLAT**
   
   module name: blat
   
   version 35
   
   Other versions: 34, 35-fastq
   
   Alignment tool like BLAST
   
   http://genome.ucsc.edu/FAQ/FAQblat.html
   
   
   Depends on gcc

1. **BLESS**
   
   module name: BLESS
   
   version 0.15
   
   Bloom-filter-based Error Correction Tool for NGS reads
   
   http://sourceforge.net/projects/bless-ec/
   
   Info: /group/bioinfo/apps/apps/BLESS-0.15/README
   
   Depends on gcc

1. **boost**
   
   module name: Boost
   
   version 1.54.0
   
   Other versions: 1.50.0, 1.54.0, 1.63.0, 1.65.1
Portable C++ source libraries
http://www.boost.org/

Depends on gcc

1. **Bowtie**
   module name: bowtie
   version 1.2.0
   Other versions: 0.11.3, 0.12.7, 0.12.8, 0.12.9, 1.0.0, 1.0.1, 1.1.0, 1.1.1, 1.1.2

   An ultrafast memory-efficient short read aligner
   Manual: /group/bioinfo/apps/apps/bowtie-1.2-try2/MANUAL
   Tutorial: /group/bioinfo/apps/apps/bowtie-1.2-try2/TUTORIAL

   Depends on tbb

2. **Bowtie2**
   module name: bowtie2
   version 2.3.2
   Other versions: 2.0.0-beta5, 2.0.0-beta6, 2.0.0-beta7, 2.0.1, 2.0.2, 2.0.3, 2.0.4, 2.0.5, 2.0.6, 2.1.0, 2.2.0, 2.2.1, 2.2.2, 2.2.3, 2.2.4, 2.2.5, 2.2.6, 2.2.7, 2.2.8, 2.2.9, 2.3.0, 2.3.1

   Fast and sensitive read alignment
   Manual: /group/bioinfo/apps/apps/bowtie2-2.3.2/MANUAL

   Depends on gcc
   Depends on tbb
   Depends on zlib
   Depends on perl

1. **BppSuite & Bio++ libraries**
   module name: BppSuite
   version 2.2.0

   Bioinformatics libraries & programs (sequence analysis, phylogenetics, molecular evolution and population genetics)
   http://biopp.univ-montp2.fr/wiki/index.php/Main_Page
   https://github.com/BioPP

   Depends on gcc
   Depends on texinfo

1. **BRAKER1**
   module name: BRAKER
   version 1.9
   Other versions: 1.8

   Pipeline for unsupervised RNA-Seq-based genome annotation
   http://bioinf.uni-greifswald.de/augustus/downloads/
   Info: /group/bioinfo/apps/apps/BRAKER-1.9/README.braker
   Example data: $BRAKER_PATH/example

   Run the "init_augustus" script from AUGUSTUS/3.1 before using BRAKER
   That will set $AUGUSTUS_CONFIG_PATH to point to directory where program can write files.

   Depends on AUGUSTUS/3.1
   Depends on GeneMark/4.32
   Depends on samtools/1.2
   Depends on perl
1. **BreakDancer**
   module name: *BreakDancer*
   version 1.1.2
   Genome-wide detection of structural variants from next generation paired-end sequencing reads
   Depends on *perl*

1. **Bridger**
   module name: *Bridger*
   version 2014-12-01
   Framework for de novo transcriptome assembly using RNA-seq data
   https://github.com/fmaguire/Bridger_Assembler
   Depends on *Boost*

1. **Brownie**
   module name: *brownie*
   version 2017-09-07
   Corrects sequence reads
   [https://github.com/biointec/brownie](https://github.com/biointec/brownie)
   Depends on *gcc*

1. **BUSCO (Benchmarking Universal Single-Copy Orthologs)**
   module name: *BUSCO*
   version 2.0-2017-02-15
   Other versions: 1.1b1
   Assess genome assembly and annotation completeness
   Invoke script with “BUSCO”
   Depends on *biopython/2.7.12*
   Depends on *AUGUSTUS/3.2.1*
   Depends on *blast*
   Depends on *HMMER/3.1b2*
   Depends on *EMBOSS*

1. **bwa**
   module name: *bwa*
   version 0.7.15
   Other versions: 0.5.9, 0.6.1, 0.6.2, 0.7.5a, 0.7.10, 0.7.12
   Burrows-Wheeler Aligner (BWA)
   or using “man bwa”
   Info: /group/bioinfo/apps/apps/bwa-0.7.15/README.md
   Info: /group/bioinfo/apps/apps/bwa-0.7.15/doc/README.md
   Info: /group/bioinfo/apps/apps/bwa-0.7.15/doc/NEWS.md
   Depends on *gcc*
   Depends on *perl*
   Depends on *zlib*

1. **bzip2**
1. **bzip2**
   - Module name: bzip2
   - Version: 1.0.6
   - Description: bzip2 is a freely available, patent free, high-quality data compressor
     Includes shared library
     http://www.bzip.org/
     Man pages available

   Depends on gcc

1. **CAMSA**
   - Module name: CAMSA
   - Version: 1.0.0
   - Other versions: 1.1.0b14
   - Comparative Analysis and Merging of Scaffold Assemblies
     https://github.com/compbiol/CAMSA

   Depends on biopython/2.7.12

1. **Canu**
   - Module name: canu
   - Version: 1.4
   - Description: Fork of the Celera Assembler, designed for high-noise single-molecule sequencing (such as the PacBio RS II or Oxford Nanopore MinION)
     https://github.com/marbl/canu
     Info: http://bionxiv.org/content/early/2016/08/24/071282
     Info: https://www.youtube.com/watch?v=iiLj1vpuCRA

   Depends on perl
   Depends on java/8

1. **CAP3**
   - Module name: CAP3
   - Version: 12.21.07
   - Genome assembly
     http://seq.cs.iastate.edu/

   Depends on gcc

1. **CASAVA**
   - Module name: CASAVA
   - Version: 1.8.2
   - Other versions: 1.8.0
   - Consensus Assessment of Sequence And VARIation
     Processes sequencing reads provided by RTA or OLB
     http://www.illumina.com/support/sequencing/sequencing_software/casava.ilmn
     Info at /group/bioinfo/apps/apps/CASAVA-1.8.2/share/CASAVA-1.8.2/docs/casava

   Depends on gcc

1. **CASPER** (Context-Aware Scheme for Paired-End Read)
   - Module name: CASPER
   - Version: 0.8.2
   - Pair-end read merger
     http://best.snu.ac.kr/casper/

   Depends on gcc
   Depends on jellyfish
1. **CD-HIT**
   module name: cd-hit
   version 4.6.5-2016-0304
   Other versions: 3.1.1, 4.6-2012-04-25, 4.6.1-2012-08-27, 4.6.1-2012-08-27-openmp

   Clustering and comparing protein or nucleotide sequences
   Includes "AuxTools"
   http://weizhong-lab.ucsd.edu/cd-hit/
   https://github.com/weizhongli/cdhit
   Docs: http://weizhong-lab.ucsd.edu/cd-hit/doc.php
   Info: /group/bioinfo/apps/apps/cd-hit-v4.6.5-2016-0304/doc

   Depends on gcc

2. **cdbtools**
   module name: cdbtools
   version 0.99

   FASTA file indexing and entry retrieval
cdbfasta and cdbyank programs
   Info: /group/bioinfo/apps/apps/cdbtools-0.99/README

3. **CEAS**
   module name: CEAS
   version 1.0.2
   Other versions: 1.0.2a, 1.0.2b

   Cis-regulatory Element Annotation System
   http://liulab.dfci.harvard.edu/CEAS/

   Depends on biopython/2.7.3

4. **CEGMA (Core Eukaryotic Genes Mapping Approach)**
   module name: cegma
   version 2.5
   Other versions: 2.4.010312

   Pipeline for building a set of gene annotations in virtually any eukaryotic genome
   http://korflab.ucdavis.edu/Datasets/cegma/
   Info: /group/bioinfo/apps/apps/cegma-2.5/README.md
   Info: /group/bioinfo/apps/apps/cegma-2.5/release_notes.md

   Depends on gcc
   Depends on blast/2.2.29+
   Depends on HMMER
   Depends on genewise
   Depends on geneid

5. **Celera Assembler**
   module name: wgs-assembler
   version 8.3rc1

   de novo whole-genome shotgun (WGS) DNA sequence assembler
   License: /group/bioinfo/apps/apps/wgs-8.3rc1/LICENSE.txt

   Depends on gcc

6. **CellProfiler**
   module name: CellProfiler
   version 2.2.0
http://cellprofiler.org/
Quantitatively measure phenotypes from images automatically

Depends on gcc
Depends on tcl
Depends on igraph
Depends on HDF5/1.8.14
Depends on java
Depends on R-bioconductor/3.1.0

1. **Censor**
   module name: censor
   version 4.2.29
   Compares and masks protein or nucleotide sequences
   http://www.girinst.org/downloads/software/censor/
   Includes copy of Repbase
   No support for WU-BLAST; only censor.ncbi available

   Depends on perl
   Depends on blast/2.2.28+

1. **CentroidFold**
   module name: centroid_fold
   version 0.0.9
   Predict RNA secondary structures based on a generalized centroid estimator
   MCL algorithm finds cluster structure in graphs
   http://www.ncrna.org/software/centroidfold
   http://www.ncrna.org/software/centroidfold/download/
   Info: /group/bioinfo/apps/apps/centroid_fold-0.0.9/README.html

1. **CFITSIO**
   module name: CFITSIO
   version 3.370
   FITS File Subroutine Library
   http://heasarc.gsfc.nasa.gov/fitsio/
   man pages available

   Depends on gcc

1. **CIRCexplorer**
   module name: CIRCexplorer
   version 1.1.10
   Other versions: 1.1.3, 2.2.6
   Circular RNA analysis toolkits
   http://yanglab.github.io/CIRCexplorer/
   FAQ: https://github.com/YangLab/CIRCexplorer/wiki/FAQ

   Depends on biopython/2.7.12
   Depends on tophat
   Depends on STAR
   Depends on BEDTools
   Depends on samtools

1. **Circos**
   module name: Circos
   version 0.69-2
   Other versions: 0.67-1

   Visualize data in a circular layout
   http://www.circos.ca/
   Manual: http://www.circos.ca/documentation/
<table>
<thead>
<tr>
<th>Package</th>
<th>Module Name</th>
<th>Version</th>
<th>Description</th>
<th>Dependencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>clearcut</td>
<td>clearcut</td>
<td>1.0.9</td>
<td>Reference implementation for the Relaxed Neighbor Joining (RNJ) algorithm</td>
<td>perl/5.20.1</td>
</tr>
<tr>
<td>CleaveLand</td>
<td>CleaveLand</td>
<td>4.3</td>
<td>Pipeline for finding cleaved small RNA targets using degradome data</td>
<td>perl, bowtie, ViennaRNA, samtools/1.2, R-bioconductor/3.1.2</td>
</tr>
<tr>
<td>CLIPper</td>
<td>CLIPper</td>
<td>0.2.0</td>
<td>A tool to identify CLIP-seq peaks?</td>
<td>biopython/2.7.11</td>
</tr>
<tr>
<td>CLUMPP</td>
<td>CLUMPP</td>
<td>1.1.2</td>
<td>Deals with label switching and multimodality problems in population-genetic cluster analyses</td>
<td>gcc</td>
</tr>
<tr>
<td>Clustal Omega</td>
<td>clustalo</td>
<td>1.2.1</td>
<td>Multiple sequence alignment</td>
<td>argtable</td>
</tr>
<tr>
<td>ClustalW2</td>
<td>clustalw</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**clearcut**

- Module name: clearcut
- Version: 1.0.9

Reference implementation for the Relaxed Neighbor Joining (RNJ) algorithm

http://www.mothur.org/wiki/Clearcut

Info: /group/bioinfo/apps/apps/clearcut-1.0.9/README

Info: /group/bioinfo/apps/apps/clearcut-1.0.9/examples

**CleaveLand**

- Module name: CleaveLand
- Version: 4.3
- Other versions: 3.0.1

Pipeline for finding cleaved small RNA targets using degradome data

a.k.a PARE (Parallel Analysis of RNA Ends) and GMUCT (Genome-wide Mapping of Uncapped Transcripts).

http://axtell-lab-psu.weebly.com/cleaveland.html

Info: /group/bioinfo/apps/apps/CleaveLand-4.3/CleaveLand4_TUTORIAL.pdf

Tutorial: http://axtelldata.bio.psu.edu/data/CleaveLand4_Tutorial/

Depends on perl

Depends on bowtie

Depends on ViennaRNA

Depends on samtools/1.2

Depends on R-bioconductor/3.1.2

**CLIPper**

- Module name: CLIPper
- Version: 0.2.0

A tool to identify CLIP-seq peaks?

https://github.com/YeoLab/clipper

https://github.com/YeoLab/clipper/wiki/CLIPper-Home

Forum: https://groups.google.com/forum/?fromgroups#!forum/clipper-discuss

Depends on biopython/2.7.11

**CLUMPP**

- Module name: CLUMPP
- Version: 1.1.2

Deals with label switching and multimodality problems in population-genetic cluster analyses

http://www.stanford.edu/group/rosenberglab/clumpp.html


Depends on gcc

**Clustal Omega**

- Module name: clustalo
- Version: 1.2.1
- Other versions: 1.1.0

Multiple sequence alignment

http://www.clustal.org/omega/

Depends on argtable

**ClustalW2**

- Module name: clustalw
- Version: 2.1
Multiple sequence alignment
http://www.ebi.ac.uk/Tools/msa/clustalw2/

Depends on gcc

1. **CNVer**
   module name: `CNVer`
   version 0.8.1

   Method for Detecting Copy Number Variation
   http://compbio.cs.toronto.edu/CNVer/
   Forum: http://groups.google.com/group/cnver-users
   Manual: /group/bioinfo/apps/apps/cnver-0.8.1/README-0.8.1.txt
   Manual: http://compbio.cs.toronto.edu/cnver/README-0.8.1.txt
   hg18 data: /group/bioinfo/apps/apps/cnver-0.8.1/hg18comp.tar.gz
   hg19 data: /group/bioinfo/apps/apps/cnver-0.8.1/hg19comp.tar.gz

   Depends on gcc

1. **CNVnator**
   module name: `CNVnator`
   version 0.3

   http://avrilonomics.blogspot.com/2013/01/using-cnvnator-to-find-copy-number.html
   Identify copy number variations (CNVs) in a genome
   Citation: http://www.ncbi.nlm.nih.gov/pubmed/21324876

   Depends on ROOT

1. **consed**
   module name: `consed`
   version 29.0
   Other versions: 23.0, 24.0, 25.0, 26.0, 27.0, 28.0

   A sequence finishing package
   http://bozeman.mbt.washington.edu/consed/consed.html
   Info in /group/bioinfo/apps/apps/consed-29.0/README.txt
   Needs X11 support for GUI

   Depends on cross_match/1.090518
   Depends on phred

1. **CONTIGuator**
   module name: `CONTIGuator`
   version 2.7.5

   https://github.com/combogenomics/CONTIGuator
   Bacterial genomes finishing tool for structural insights on draft genomes

   Depends on biopython/2.7.12
   Depends on blast
   Depends on perl
   Depends on MUMmer
   Depends on primer3

1. **Control-FREEC**
   module name: `freec`
   version 7.0
   Other versions: 5.5, 5.9

   A tool for detection of copy-number changes and allelic imbalances (including LOH) using deep-sequencing data
   http://bioinfo-out.curie.fr/projects/freec
Depends on gcc

1. **corset**
   module name: corset
   version 1.03

   Clustering de novo assembled transcripts and counting overlapping reads
   [https://code.google.com/p/corset-project/](https://code.google.com/p/corset-project/)
   Example: [https://code.google.com/p/corset-project/wiki/Example](https://code.google.com/p/corset-project/wiki/Example)
   Forum: [https://groups.google.com/forum/#!forum/corset-project](https://groups.google.com/forum/#!forum/corset-project)

   Depends on gcc

1. **cortex_var**
   module name: cortex_var
   version 1.0.5.21
   Other versions: 1.0.5.20

   Tool for genome assembly and variation analysis
   [http://cortexassembler.sourceforge.net/index_cortex_var.html](http://cortexassembler.sourceforge.net/index_cortex_var.html)

   Depends on gsl/1.15
   Depends on stampy/1.0.26
   Depends on vcf tools/0.1.12b

1. **cppunit**
   module name: cppunit
   version 1.13.2
   Other versions: 1.12.0, 1.12.1

   C++ port of the JUnit framework for unit testing
   [https://www.freedesktop.org/wiki/Software/cppunit/](https://www.freedesktop.org/wiki/Software/cppunit/)

   Depends on gcc

1. **CRI-MAP Improved**
   module name: crimap
   version 2.507
   Other versions: 2.4, 2.504a

   Genetic linkage analysis of diploid species
   Forum: [http://www.animalgenome.org/community/crimap-users](http://www.animalgenome.org/community/crimap-users)
   Info in /group/bioinfo/apps/apps/crimap-2.507/doc directory

   Depends on gcc

1. **CRISP**
   module name: CRISP
   version 122713

   Multi-sample variant caller for high-throughput pooled sequence data
   [https://sites.google.com/site/vibansal/software/crisp](https://sites.google.com/site/vibansal/software/crisp)
   Info: /group/bioinfo/apps/apps/CRISP-122713/README

1. **cross_match**
   module name: cross_match
   version 1.080812
   Other versions: 1.090518

   Compare DNA sequences
   [http://www.phrap.org/phredphrapconsed.html](http://www.phrap.org/phredphrapconsed.html)
1. **CrossMap**  
   module name: CrossMap  
   version 0.2.4  
   Conversion of genome coordinates (or annotation files) between different assemblies  
   Depends on biopython/2.7.11

1. **Cufflinks**  
   module name: cufflinks  
   version 2.2.1  
   Other versions: 0.7.0, 1.0.3, 1.1.0, 1.2.0, 1.2.1, 1.3.0, 2.0.0, 2.0.2, 2.1.0, 2.1.1, 2.2.0, 2.2.1-t1  
   Transcript assembly, differential expression, and differential regulation for RNA-Seq  
   [http://cufflinks.cbcb.umd.edu/](http://cufflinks.cbcb.umd.edu/)  
   FAQ: [http://cufflinks.cbcb.umd.edu/faq.html](http://cufflinks.cbcb.umd.edu/faq.html)  
   Depends on zlib

1. **cURL and libcurl**  
   module name: curl  
   version 7.46.0  
   Command line tool and library for transferring data with URL syntax  
   [http://curl.haxx.se/](http://curl.haxx.se/)  
   man curl  
   Depends on gcc

1. **cutadapt**  
   module name: cutadapt  
   version 1.13  
   Other versions: 1.1a, 1.1b, 1.2.1, 1.9.1, 1.12  
   Removes adapter sequences from DNA sequencing reads  
   [https://github.com/marcelm/cutadapt/](https://github.com/marcelm/cutadapt/)  
   [https://pypi.python.org/pypi/cutadapt/](https://pypi.python.org/pypi/cutadapt/)  
   Depends on biopython/2.7.12

1. **Cytoscape**  
   module name: Cytoscape  
   version 2.8.3  
   Network Data Integration, Analysis, and Visualization in a Box  
   [http://www.cytoscape.org/what_is_cytoscape.html](http://www.cytoscape.org/what_is_cytoscape.html)  
   Depends on java

1. **DANPOS**  
   module name: DANPOS  
   version 2.2.2  
   A toolkit for Dynamic Analysis of Nucleosome and Protein Occupancy by Sequencing
1. **DBG2OLC**
   - module name: **DBG2OLC**
   - version 12-08-2015
   - Other versions: 12-08-2105

   Assembly of Large Genomes Using the Compressed Overlap Graph
   [https://sourceforge.net/projects/dbg2olc/](https://sourceforge.net/projects/dbg2olc/)

2. **dbHT-Trans**
   - module name: **dbHT-Trans**
   - version 14-Jul-2016

   Filtering the Protein-Encoding Transcripts
   [https://github.com/chengroup/dbHT-Trans](https://github.com/chengroup/dbHT-Trans)

   This software doesn't seem to work
   Need to configure MySQL database prior to using this software
   Need to copy script into your home directory and run from there:
   ```
   $ cp -rp $dThome ~/.
   $ cd ~/dbHT-Trans-14-Jul-2016
   $ # edit config.txt with MySQL server info
   $ python config.py
   $ ./dbHT-Trans-Operator ...
   ```
   
   Depends on **biopython/2.7.11**

3. **DeconSeq**
   - module name: **deconseq**
   - version 0.4.3
   - Other versions: 0.4.1, 0.4.2

   DECONamination of SEQuence data using a modified version of BWA
   Info at /group/bioinfo/apps/apps/deconseq-standalone-0.4.3/README
   FAQ: [http://deconseq.sourceforge.net/faq.html](http://deconseq.sourceforge.net/faq.html)

   Based on BWA (Burrows-Wheeler Aligner) Version: 0.5.9-r16
   Probably need to make copy of deconseq.pl and DeconSeqConfig.pm files
   from /group/bioinfo/apps/apps/deconseq-standalone-0.4.3
   Modify the DeconSeqConfig.pm file as necessary.

   Depends on **perl**

4. **deepTools**
   - module name: **deeptools**
   - version 2.2.4
   - Other versions: 1.5.12

   Tools to process and analyze deep sequencing data.
   [https://github.com/fidelram/deepTools](https://github.com/fidelram/deepTools)
   Wiki: [https://github.com/fidelram/deepTools/wiki](https://github.com/fidelram/deepTools/wiki)

   Depends on **biopython/2.7.8**
   Depends on **samtools**
1. **DejaGnu**
   module name: dejagnu
   version 1.6
   Framework for testing other programs
   https://www.gnu.org/software/dejagnu/

1. **DELLY**
   module name: DELLY
   version 0.7.2
   Other versions: 0.6.1
   https://github.com/tobiasrausch/delly
   Structural variant discovery by integrated paired-end and split-read analysis
   Forum: http://groups.google.com/d/forum/delly-users
   Use "delly" to run the program.
   To use multithreaded, OpenMP version, use "delly-parallel"
   Set the environment variable OMP_NUM_THREADS to the number of threads you want to use before running.
   Depends on PyVCF

1. **DETONATE (DE novo TranscriptOme rNa-seq Assembly with or without the Truth Evaluation)**
   module name: detonate
   version 1.10
   Other versions: 1.8.1
   Evaluate de novo transcriptome assemblies
   http://deweylab.biostat.wisc.edu/detonate/
   Tutorial: /group/bioinfo/apps/apps/detonate-1.10/VIGNETTE
   Tutorial example data at /group/bioinfo/apps/apps/detonate-1.10/rsem-eval/true_transcript_length_distribution
   Depends on gcc
   Depends on bowtie
   Depends on blat
   Depends on rsem

1. **DIALIGN-TX**
   module name: DIALIGN-TX
   version 1.0.2
   Segment-based multiple sequence alignment
   http://dialign-tx.gobics.de/
   Depends on perl
   Depends on gcc

1. **DIAMOND**
   module name: diamond
   version 0.8.36
   Other versions: 0.6.12, 0.7.1, 0.7.8, 0.7.9, 0.9.6
   Align short DNA sequencing reads to a protein reference database such as NCBI-NR
   http://ab.inf.uni-tuebingen.de/software/diamond/
   Includes Mar 3 2017 NCBI nr database data at $DIAMONDBDB/nr.dmnd
   Includes Feb 23 2017 NCBI refseq_protein data at $DIAMONDBDB/refseq_protein.dmnd

1. **diffReps**
   module name: diffReps
version 1.55.6
Other versions: 1.55.4

Differential analysis for ChIP-seq with biological replicates
https://github.com/shenlab-sinai/diffreps
Forum: https://groups.google.com/forum/?fromgroups%5C#forum/diffreps-discuss

Depends on perl/5.20.1

1. **DiNuP**  
   module name: DiNuP  
   version 1.3  
   A Systematic Approach to Identify Regions of Differential Nucleosome Positioning
   
   Depends on biopython/2.7.8

1. **DISCOVAR**  
   module name: discovar  
   version 52488  
   Other versions: 48686, 52415  
   Genome assembler and variant caller  
   http://www.broadinstitute.org/software/discovar/blog/  
   FAQ: http://www.broadinstitute.org/software/discovar/blog/?page_id=23  
   License: /group/bioinfo/apps/apps/discovar-52488/share/doc/discovar/LICENSE
   Forum: https://groups.google.com/a/broadinstitute.org/forum/?hl=en&fromgroups#!forum/discovar-user-forum  
   Manual: https://docs.google.com/document/d/1U_o-Z0dJ0QKJn86AV20_YHiFzUfW9c57eh3YjkINc/edit?usp=sharing  
   Example data in /group/bioinfo/apps/apps/discovar-52488/examples
   
   Depends on gcc  
   Depends on picard-tools
   Depends on samtools
   Depends on biopython/2.7.3

1. **DISCOVAR de novo**  
   module name: discovardenovo  
   version 52488  
   Other versions: 52415  
   Large genome assembler  
   http://www.broadinstitute.org/software/discovar/blog/  
   FAQ: http://www.broadinstitute.org/software/discovar/blog/?page_id=23  
   Forum: https://groups.google.com/a/broadinstitute.org/forum/?hl=en&fromgroups#!forum/discovar-user-forum  
   Manual: https://docs.google.com/document/d/1U_o-Z0dJ0QKJn86AV20_YHiFzUfW9c57eh3YjkINc/edit?usp=sharing  
   Example data in /group/bioinfo/apps/apps/discovardenovo-52488/examples
   
   Depends on gcc  
   Depends on samtools
   Depends on jemalloc

1. **DISLIN**  
   module name: DISLIN  
   version 11.0  
   Other versions: 10.5

   Plotting library for displaying data  
   http://www.mps.mpg.de/disin/  
   
   Probably need X11 support  
   
   Depends on gcc

1. **distruct**  
   module name: distruct
version 1.1

Graphically display results produced by the genetic clustering program "Structure"
http://www.stanford.edu/group/rosenberglab/distruct.html

Depends on gcc

1. DLCpar
   module name: DLCpar
   version 1.0

Reconciliation method for inferring gene duplications, losses, and coalescence
Python module
https://www.cs.hmc.edu/~yjw/software/dlcpar/

Depends on biopython/2.7.12

1. dna-bison
   module name: dna-bison
   version 0.3.3b
   Other versions: 0.2, 0.2.4, 0.3.0, 0.3.1, 0.3.2, 0.3.2b, 0.3.3

Bisulfite alignment On Nodes of a cluster
http://sourceforge.net/projects/dna-bison/
wiki: http://sourceforge.net/p/dna-bison/wiki/Home/
Tutorial: /group/bioinfo/apps/apps/dna-bison-0.3.3b/bison_tutorial

Depends on gcc
Depends on openmpi
Depends on biopython/2.7.8
Depends on bowtie2
Depends on R-bioconductor/3.1.0

1. DOTS-Finder
   module name: DOTS-Finder
   version 1.0

Driver Oncogenes and Tumor Suppressors Finder
http://cgsb.genomics.iit.it/wiki/projects/DOTS-Finder
http://sourceforge.net/projects/dotsfinder/
Sample files; /group/bioinfo/apps/apps/DOTS-Finder-1.0/sample-files

Depends on biopython/2.7.8

1. ea-utils
   module name: ea-utils
   version 1.1.2-537
   Other versions: 1.1.2-408

FASTQ processing utilities
http://code.google.com/p/ea-utils/
wiki: http://code.google.com/p/ea-utils/w/list

Depends on gcc
Depends on bamtools
Depends on gsl

1. EDGE-pro
   module name: EDGE-pro
   version 1.3.1

Prokaryotic transcriptome analysis system
http://www.ccb.jhu.edu/software/EDGE-pro/index.shtml
Manual: http://www.ccb.jhu.edu/software/EDGE-pro/MANUAL
Manual: /group/bioinfo/apps/apps/EDGE_pro_v1.3.1/MANUAL
Example data: /group/bioinfo/apps/apps/EDGE_pro_v1.3.1/example

Modified edge.pl to remove -s switch

The edge.pl script does not need to be in working directory nor should you invoke it as PATH/edge.pl as described in manual.

Depends on gcc
Depends on bowtie2

1. **edittag**
   module name: *edittag*
   version 1.1

   Edit distance sequence tags and helpers
   https://github.com/faircloth-lab/edittag/
   http://faircloth-lab.github.io/edittag/

   Depends on biopython/2.7.12
   Depends on mod-primer3

1. **elinks**
   module name: *elinks*
   version 0.12pre5

   text mode web browser
   http://elinks.or.cz/
   man: elinks, elinkskeys, elinks.conf
   copied from elinks-0.12-0.21.pre5.el6_3.x86_64

   Depends on gcc

1. **EMBOSS**
   module name: *EMBOSS*
   version 6.6.0-t2
   Other versions: 6.4.0, 6.5.7, 6.6.0

   The European Molecular Biology Open Software Suite
   http://emboss.sourceforge.net/
   Info: http://emboss.sourceforge.net/docs/
   Info: http://emboss.sourceforge.net/news/
   Info: /group/bioinfo/apps/apps/EMBOSS-6.6.0-t2/share/EMBOSS/doc/

   Includes REBASE 702 data

   Get Jemboss GUI with "runJemboss.sh"; need X11 support

   Depends on gcc
   Depends on java

1. **EMIRGE (Expectation-Maximization Iterative Reconstruction of Genes from the Environment)**
   module name: *EMIRGE*
   version 0.61.0

   Reconstructs full length ribosomal genes from short read sequencing data.
   https://github.com/csmiller/EMIRGE
   Issues: https://github.com/csmiller/EMIRGE/issues
   Forum: https://groups.google.com/forum/#!forum/emirge-users

   Depends on biopython/2.7.12
   Depends on bowtie/1.1.2
   Depends on vsearch
   Depends on usearch/8

1. **EMMAX**
module name: **EMMAX**
version 20120210

Statistical test for large scale human or model organism association mapping accounting for the sample structure.
http://genome.sph.umich.edu/wiki/EMMAX

Depends on gcc
Depends on ATLAS

1. **epic**
   module name: **bioepic**
   version 0.1.25
   Diffuse domain ChIP-Seq caller based on SICER
   https://github.com/biocore-ntnu/epic
   Forum: https://groups.google.com/forum/#!forum/epic-chip-seq

   Depends on biopython/2.7.12
   Depends on jellyfish

1. **EvidentialGene (tr2aacds.pl)**
   module name: **EvidentialGene**
   version 2017-02-13
   mRNA Transcript Assembly
   http://arthropods.eugenes.org/EvidentialGene/
   http://arthropods.eugenes.org/EvidentialGene/about/EvidentialGene_trassembly

   Just set up for tr2aacds.pl script
   Probably doesn't work

   Depends on perl
   Depends on exonerate
   Depends on cd-hit
   Depends on blast

1. **Exonerate**
   module name: **exonerate**
   version 2.2.0
   Generic tool for sequence alignment
   http://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate
   Use "man exonerate" for more info
   Users guide: http://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate-user-guide
   Manual: http://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate-manual

   Depends on gcc

1. **expat**
   module name: **expat**
   version 2.1.0
   XML Parser
   http://sourceforge.net/projects/expat/
   man pages available

   Depends on gcc

1. **eXpress**
   module name: **express**
   version 1.5.1
   Streaming quantification for high-throughput sequencing
   http://bio.math.berkeley.edu/eXpress/
   Info: /group/bioinfo/apps/apps/express-1.5.1/README
1. **FALCON**
   module name: *FALCON*
   version 2017-06-28

   Experimental PacBio diploid assembler
   https://github.com/PacificBiosciences/FALCON
   Info: https://github.com/PacificBiosciences/FALCON/wiki

   This probably does not work

   Depends on biopython/2.7.12

1. **FASconCAT**
   module name: *FASconCAT*
   version 1.0

   Concatenate sequence fragments into a supermatrix file for phylogenetic purposes
   http://zfmk.de/web/Forschung/Abteilungen/AG_Wgele/Software/FASconCAT/index.en.html

   Depends on perl

1. **FAST-iCLIP**
   module name: *FAST-iCLIP*
   version 0.9.1
   Other versions: 0.9.3

   Fully Automated and Standardized (FAST) iCLIP analysis pipeline.
   https://github.com/ChangLab/FAST-iCLIP

   Depends on biopython/2.7.11
   Depends on bowtie2
   Depends on STAR
   Depends on BEDTools
   Depends on fastx

1. **FASTA**
   module name: *fasta*
   version 36.3.5e
   Other versions: 35, 35.4.12

   Identify statistically significant sequence similarity that can be used to infer homology
   Similar to NCBI BLAST
   http://fasta.bioch.virginia.edu/fasta_www2/fasta_down.shtml
   Manual: man fasta36

   Depends on gcc

1. **Fasta Utilities**
   module name: *fasta_utilities*
   version 30-Nov-2015
   Other versions: 9-May-2014

   A collection of scripts developed to interact with fasta, fastq and sam/bam files.
   https://github.com/jimhester/fasta_utilities
   Use man command to read provided man pages

   Depends on perl
   Depends on samtools
1. **FastaGrep**
   module name: fastagrep
   version 2.0

   Search oligonucleotide binding sites from genomic sequences
   http://bioinfo.ut.ee/?page_id=167
   Manual at /group/bioinfo/apps/apps/fastagrep-2.0/README_fg.txt

   Depends on gcc

1. **fastq_screen**
   module name: fastq_screen
   version 0.5.2
   Other versions: 0.4.2

   Contamination Screening for large data sets
   http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/
   Info: /group/bioinfo/apps/apps/fastq_screen_v0.5.2/README.txt
   Info: /group/bioinfo/apps/apps/fastq_screen_v0.5.2/RELEASE_NOTES.txt

   Need to create configuration file and add databases:
   - cp /group/bioinfo/apps/apps/fastq_screen_v0.5.2/fastq_screen.conf.example ~/fastq_screen.conf
   - edit fastq_screen.conf in your home directory
   - fastq_screen --conf ~/fastq_screen.conf

   Depends on perl
   Depends on bowtie2

1. **FastQC**
   module name: fastqc
   version 0.11.5
   Other versions: 0.10.0, 0.10.1, 0.11.2

   Quality control application for high throughput sequence data
   http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

   Depends on java

1. **fastStructure**
   module name: fastStructure
   version 1.0.x

   Framework for inferring population structure from SNP genotype data.
   http://rajanil.github.io/fastStructure/
   Use "structure.py" to start program

   Depends on biopython/2.7.11
   Depends on gsl

1. **FastTree**
   module name: FastTree
   version 2.1.7
   Other versions: 2.1.3, 2.1.4

   Infers approximately-maximum-likelihood phylogenetic trees from
   alignments of nucleotide or protein sequences
   http://www.microbesonline.org/fasttree/

1. **FASTX-Toolkit**
   module name: fastx
   version 0.0.14
   Other versions: 0.0.13, 0.0.13.2
FASTQ/A short-reads pre-processing tools
http://hannonlab.cshl.edu/fastx_toolkit/index.html
Info: http://hannonlab.cshl.edu/fastx_toolkit/commandline.html

Depends on gcc
Depends on perl
Depends on libgtextutils

1. FinisherSC
   module name: FinisherSC
   version 2.1

   FinisherSC: a repeat-aware tool for upgrading de novo assembly using long reads
   https://github.com/kakitone/finishingTool

   Automatically loads MUMmer software; use "ml show MUMmer" to determine path which is needed to run finisherSC.py

   Depends on MUMmer
   Depends on biopython/2.7.12
   Depends on java/7

1. FLASH (Fast Length Adjustment of SHort reads)
   module name: FLASH
   version 1.2.11
   Other versions: 1.2.7, 1.2.8, 1.2.10

   Merge mates from fragments that are shorter than twice the read length
   http://ccb.jhu.edu/software/FLASH/

   Depends on gcc

1. flexbar
   module name: flexbar
   version 2.4

   Flexible barcode and adapter removal
   https://github.com/seqan/flexbar
   Info: /group/bioinfo/apps/apps/flexbar-2.4/README.txt

1. Fold-Grammars
   module name: fold-grammars
   version 2014-03-17

   Components for dynamic programming problems covering RNA secondary structure predictions
   http://bibiwsserv.cebitec.uni-bielefeld.de/fold-grammars/
   Includes: pKiss and KnotInFrame
   http://bibiserv2.cebitec.uni-bielefeld.de/pkiss?id=pkiss_pkiss_manual_manual
   http://bibiserv2.cebitec.uni-bielefeld.de/knotinframe?id=knotinframe_manual_manual

   Depends on gapc

1. FragGeneScan
   module name: FragGeneScan
   version 1.30
   Other versions: 1.16, 1.19

   Finds (fragmented) genes in short reads
   http://omics.informatics.indiana.edu/FragGeneScan/
   Example data in /group/bioinfo/apps/apps/FragGeneScan-1.30/example
   Training data in /group/bioinfo/apps/apps/FragGeneScan-1.30/train

   Depends on gcc
1. **FusionCatcher**
   module name: *FusionCatcher*
   version 0.99.6a
   Other versions: 0.99.4c, 0.99.5a, test

   Finder of somatic fusion-genes in RNA-seq data.
   https://github.com/ndaniel/fusioncatcher

   For Ensembl data release 84
   You will need to use “fusioncatcher-build” command to create index files as described in the manual

   Depends on biopython/2.7.8
   Depends on bowtie/1.1.2
   Depends on STAR/2.5.1b
   Depends on bowtie2/2.2.9
   Depends on bwa/0.7.12
   Depends on seqtk/1.0-r82b
   Depends on velvet/1.2.10
   Depends on sra-toolkit/2.5.1
   Depends on samtools/0.1.19
   Depends on picard-tools/1.129
   Depends on parallel/20150522

1. **GAEMR (Genome Assembly Evaluation, Metrics and Reporting)**
   module name: *GAEMR*
   version 1.0.1

   Report on a genome assembly's completeness, correctness, and contiguity
   http://www.broadinstitute.org/software/gaemr/

   Depends on biopython/2.7.8
   Depends on bwa/0.7.12
   Depends on bowtie2/2.2.5
   Depends on blast/2.2.30+
   Depends on MUMmer/3.23
   Depends on mammer/1.2
   Depends on samtools/0.1.18
   Depends on picard-tools/1.123
   Depends on rdp_classifier/2.4

1. **GAM-NGS**
   module name: *GAM-NGS*
   version 1.1

   Genomic Assemblies Merger for NGS
   https://github.com/vice87/gam-NGS
   Example: /group/bioinfo/apps/apps/gam-NGS-1.1/example
   Ref: http://www.biomedcentral.com/1471-2105/14/S7/S6

   Depends on Boost
   Depends on bwa
   Depends on samtools/1.0

1. **GapCloser**
   module name: *GapCloser*
   version 1.12-r6

   Close gaps which emerge from scaffolding process by SOAPdenovo or other assembler using the abundant pair relationships of short reads.
   http://soap.genomics.org.cn/soapdenovo.html

   Depends on gcc
1. **GapFiller**
   module name: `GapFiller`
   version 1.11
   Closing gaps within pre-assembled scaffold
   Tutorial at /group/bioinfo/apps/apps/GapFiller-1.11/tutorial.pdf
   License at /group/bioinfo/apps/apps/GapFiller-1.11/license.pdf
   Depends on gcc

1. **GATB-Core**
   module name: `gatb-core`
   version 1.2.2
   Genome Analysis Toolbox with de-Bruijn graph
   https://gatb.inria.fr/software/gatb-core/
   http://gatb-core.gforge.inria.fr/doc/api/
   Depends on gcc
   Depends on cppunit/1.12.1

1. **GATK (Genome Analysis Toolkit; Full version)**
   module name: `GATK`
   version 3.8.0
   Other versions: 2.2.15, 2.2.16, 2.3.0, 2.3.1, 2.3.3, 2.3.4, 2.3.5, 2.3.9, 2.4.3, 2.4.9, 2.5.2, 2.6.4, 2.6.5, 2.7.1, 2.7.2, 2.7.4, 2.8.1, 3.0.0, 3.1.1, 3.2.0, 3.2.2, 3.3.0, 3.4.0, 3.4.46, 3.5, 3.6.0, 3.7.0, nightly-2015-08-23
   Analyse next-generation resequencing data
   http://www.broadinstitute.org/gatk/
   Forum: http://gatkforums.broadinstitute.org/
   Info: https://software.broadinstitute.org/gatk/documentation/quickstart
   GenomeAnalysisTK.jar file at /group/bioinfo/apps/apps/GATK-3.8.0
   Run the script "GenomeAnalysisTK" to use it.
   Or, use "java -jar $GATK -h", for example.
   License terms at /group/bioinfo/apps/apps/GATK-3.8.0/LICENSE
   Info: http://gatkforums.broadinstitute.org/discussion/6495/version-highlights-for-gatk-version-3-5
   Depends on java/8

1. **GATK Queue**
   module name: `GATK-Queue`
   version 3.8.0
   Other versions: 3.2.2, 3.3.0, 3.4.0, 3.4.46, 3.5, 3.6
   Scripting framework for defining multi-stage genomic analysis pipelines
   http://gatkforums.broadinstitute.org/discussion/1306/overview-of-queue
   Queue.jar file at /group/bioinfo/apps/apps/GATK-Queue-3.8.0
   Run the script "Queue" to use it.
   License terms at /group/bioinfo/apps/apps/GATK-Queue-3.8.0/LICENSE
   Depends on GATK/3.8.0

1. **GATK-Lite (Genome Analysis Toolkit; Public version)**
   module name: `GATK-lite`
   version lite-2.3.9
   Other versions: lite-2.2.13, lite-2.2.16, lite-2.3.0, lite-2.3.1, lite-2.3.4, lite-2.3.5
   Analyse next-generation resequencing data
   http://www.broadinstitute.org/gatk/
   Forum: http://gatkforums.broadinstitute.org/
   GenomeAnalysisTKLite.jar file at /group/bioinfo/apps/apps/GATK-lite-2.3.9
   Use "GenomeAnalysisTKLite" to use it.
   Depends on java
1. **GBS-SNP-CROP**
   module name: **GBS-SNP-CROP**
   version 2.0

   GBS SNP Calling Reference Optional Pipeline
   [https://github.com/halelab/GBS-SNP-CROP](https://github.com/halelab/GBS-SNP-CROP)
   Forum: [https://groups.google.com/forum/#!forum/gbs-snp-crop](https://groups.google.com/forum/#!forum/gbs-snp-crop)
   Tutorial data available in /group/bioinfo/apps/apps/GBS-SNP-CROP-2.0/tutorial

   This probably does not work.

   Depends on **pear/0.9.6**
   Depends on **java/8**
   Depends on **usearch/8.0.1623**
   Depends on **trimmomatic/0.33**
   Depends on **bwa/0.7.12**
   Depends on **samtools/1.2**

1. **GDAL (Geospatial Data Abstraction Library)**
   module name: **GDAL**
   version 2.2.1
   Other versions: 2.1.3

   Translator library for raster and vector geospatial data formats

   Depends on **gcc**
   Depends on **biopython/2.7.12**

1. **GEM (GEnome Multi-tool) Library**
   module name: **GEM**
   version 20130406-045632
   Other versions: 1, 20130406-045632

   Next-generation sequence data handling tools
   [http://big.crg.cat/services/gem_genome_multi_tool_library](http://big.crg.cat/services/gem_genome_multi_tool_library)
   [http://algorithms.cnag.cat/wiki/The_GEM_library](http://algorithms.cnag.cat/wiki/The_GEM_library)
   Info in /group/bioinfo/apps/apps/GEM-20130406-045632-Intel/man

   Depends on **gcc**

1. **geneid**
   module name: **geneid**
   version 1.4.4

   Predict genes in anonymous genomic sequences designed with a hierarchical structure
   [http://genome.crg.es/software/geneid/](http://genome.crg.es/software/geneid/)
   Help: geneid -h

   Depends on **gcc**

1. **GeneMark-ES/ET**
   module name: **GeneMark**
   version 4.32
   Other versions: 2.5m

   Gene Prediction in Eukaryotes
   [http://exon.gatech.edu/GeneMark/](http://exon.gatech.edu/GeneMark/)

   Depends on **perl**

1. **Genepop**
   module name: **Genepop**
   version 4.3
Other versions: 4.1.4

Population-genetics software for exact tests and ecumenism
http://kimura.univ-montp2.fr/~rousset/Genepop.htm

Depends on gcc

1. GenePRIMP (Gene PRediction IMprovement Pipeline)
   module name: GenePRIMP
   version 0.3

   Identify & correct erroneous gene calls and missed genes
   http://geneprimp.jgi-psf.org/about.html
   Software license at /group/bioinfo/apps/apps/GenePRIMP-0.3/license.txt
   Info at /group/bioinfo/apps/apps/GenePRIMP-0.3/README

   Depends on blast/2.2.26

1. GenomeMapper
   module name: GenomeMapper
   version 0.4.4

   Short read mapping tool designed for accurate read alignments
   http://1001genomes.org/software/genomemapper_singleref.html
   Manual:

1. GenomeThreader
   module name: GenomeThreader
   version 1.6.5

   Compute gene structure predictions
   http://genomethreader.org/

   Depends on Vmatch/2.2.4

1. GenomeTools
   module name: GenomeTools
   version 1.5.9
   Other versions: 1.3.9, 1.4.1, 1.5.8

   Bioinformatics tools collection including LTRharvest
   http://genometools.org/
   Manuals: /group/bioinfo/apps/apps/gt-1.5.9/doc

   LTRharvest: http://www.zbh.uni-hamburg.de/forschung/genominformatik/software/Ltrharvest.html

   Also includes gtToGenePred to convert a GTF file to a genePred

   Depends on gcc

1. Genovo
   module name: genovo
   version 0.4

   Metagenomic de novo Sequencing
   http://cs.stanford.edu/group/genovo/
   Manual: /group/bioinfo/apps/apps/genovo-0.4/MANUAL

   Depends on gsl
1. **GFF3_to_GTF**
   module name: *GFF3_to_GTF*
   version 1.0.0

   Converts data from GFF3 format to GTF format
   https://toolshed.g2.bx.psu.edu/repository/display_tool?repository_id=afcb6456d8e300ed&tool_config=%2Fsr%2Ftoolshed%2Fmain%2Fvar%2Fdata%2Frepos%2F000%2Frepo_21%2Ffml_gff_converter_programs%2Fgalaxy%2Fgff3_to_gtf.xml&changeset_revision=c6cf723c2ed4
   Info: /group/bioinfo/apps/apps/GFF3_to_GTF-1.0.0/README.md

   Depends on biopython/2.7.8

1. **gff3togtf**
   module name: *gff3togtf*
   version 5c6f33e20fccc

   Tools for converting genome annotation between GTF, BED, GenBank and GFF
   https://github.com/vipints/GFFtools-GX/
   https://toolshed.g2.bx.psu.edu/repository?repository_id=afcb6456d8e300ed
   Info: /group/bioinfo/apps/apps/gff3togtf-5c6f33e20fccc/README.md

1. **GFOLD**
   module name: *gfold*
   version 1.1.4

   Generalized fold change for ranking differentially expressed genes from RNA-seq data
   https://bitbucket.org/feeldead/gfold/
   Info: man gfold
   Info: http://compbio.tongji.edu.cn/~fengjx/GFOLD/gfold.html

   Depends on gsl/1.15

1. **Glasgow Haskell Compiler**
   module name: *ghc*
   version 6.12.3
   Other versions: 7.6.3

   Compiler and interactive environment for the functional language Haskell
   http://www.haskell.org/ghc/
   http://www.haskell.org/ghc/download_ghc_6_12_3
   man ghc

   Depends on gcc

1. **GLIMMER**
   module name: *Glimmer*
   version 3.0.2

   GLIMMER (Gene Locator and Interpolated Markov ModelER) is a system for finding genes in microbial DNA
   http://www.cbcb.umd.edu/software/glimmer/

   Depends on gcc

1. **GlimmerHMM**
   module name: *GlimmerHMM*
   version 3.0.1

   Gene finder based on a Generalized Hidden Markov Model (GHMM)
   http://www.cbcb.umd.edu/software/GlimmerHMM/
1. **GMAP and GSNAP**
   module name: gmap
   version 2016-05-01

   A Genomic Mapping and Alignment Program for mRNA and EST Sequences
   http://research-pub.gene.com/gmap/
   Manual: /group/bioinfo/apps/apps/gmap-2016-05-01/README

   No genome databases for GMAP and GSNAP are installed. You can build your own or download prebuilt ones from the web site.

   Depends on gcc
   Depends on perl/5.20.1
   Depends on zlib
   Depends on bzip2

2. **GMcloser**
   module name: GMcloser
   version 1.6
   Other versions: 1.5.1

   Closing the gaps in scaffolds with preassembled contigs
   https://sourceforge.net/projects/gmcloser/

   Includes gmvalue version 1.2

   Depends on MUMmer/3.23
   Depends on blast
   Depends on perl
   Depends on bowtie2
   Depends on yass

3. **GMP**
   module name: gmp
   version 5.0.4
   Other versions: 5.0.3

   GNU MP Bignum Library
   http://gmplib.org/

   Depends on gcc

4. **GMvalue**
   module name: GMvalue
   version 1.3

   Create genomic sequence assembly and evaluate its accuracy
   https://sourceforge.net/projects/gmcloser/

   Depends on MUMmer/3.23
   Depends on perl

5. **GNU Autoconf**
   module name: autoconf
   version 2.69

   M4 macros to make shell scripts to configure software
   http://www.gnu.org/software/autoconf/

   Depends on m4
1. GNU Automake
   module name: automake
   version 1.14

   Automatically generate Makefile.in
   http://www.gnu.org/software/automake/

   Depends on autoconf

1. GNU Libtool
   module name: libtool
   version 2.4.2

   Generic library support script
   https://www.gnu.org/software/libtool/

1. GNU M4
   module name: m4
   version 1.4.17

   Unix macro processor
   http://www.gnu.org/software/m4/

1. GNU MPFR
   module name: mpfr
   version 3.1.2

   Library for multiple-precision floating-point computations with correct rounding
   http://www.mpfr.org/

   Depends on gcc
   Depends on gmp

1. GNU parallel
   module name: parallel
   version 20170722
   Other versions: 20130122, 20130722, 20150522, 20150622

   Shell tool for executing jobs in parallel
   http://www.gnu.org/software/parallel/
   Manual: use "man parallel"
   Also man pages for niceload, sem and sql

1. GNU Scientific Library (GSL)
   module name: gsl
   version 2.3
   Other versions: 1.15

   Numerical library for C and C++ programmers
   http://www.gnu.org/software/gsl/
   man -s3 gsl

   Depends on gcc

1. GNU Texinfo
   module name: texinfo
   version 6.5

   The GNU Documentation System
1. **Gnuplot**  
   module name: `gnuplot`  
   version 4.6.6  
   Other versions: 4.6.1, 4.6.2, 5.0.5  

   Portable, command-line driven graphing utility  
   http://www.gnuplot.info/  
   Manual: man gnuplot  
   FAQ: /group/bioinfo/apps/apps/gnuplot-4.6.6/FAQ.pdf  
   FAQ: http://www.gnuplot.info/faq/index.html

   Depends on gcc

1. **gpg-error**  
   module name: `libgpg-error`  
   version 1.12

   Library with error codes and descriptions shared by most GnuPG related software  
   http://www.gnupg.org/download/

   Depends on gcc

1. **GraPhlAn**  
   module name: `graphlan`  
   version 2017-01-21  

   Make circular representations of taxonomic and phylogenetic trees  
   https://bitbucket.org/nsegata/graphlan/wiki/Home  
   Example data: /group/bioinfo/apps/apps/GraPhlAn-2017-01-21/examples

   Depends on biopython/2.7.12

1. **Graphviz**  
   module name: `graphviz`  
   version 2.38.0

   Graph Visualization Software  
   http://www.graphviz.org/  

   Depends on gcc

1. **GREAT (Genomic Regions Enrichment of Annotations Tool)**  
   module name: `GREAT`  
   version 1.2

   Predicts functions of cis-regulatory regions  
   http://bejerano.stanford.edu/help/display/GREAT/Home  
   Forum: http://bejerano.stanford.edu/forum  
   Manual: /group/bioinfo/apps/apps/greatTools-1.2/README.txt

   Includes the core calculation engine tools:  
   - createRegulatoryDomains to calculate regulatory domains  
   - calculateBinomialP to calculate genomic region-based binomial p-values

   Depends on gcc
1. **GroopM**
   - module name: *GroopM*
   - version 0.3.4
   - Metagenomic binning suite
   - https://github.com/ecogenomics/GroopM
   - http://ecogenomics.github.io/GroopM/

   Depends on biopython/2.7.12
   Depends on BamM
   Depends on numexpr

1. **GS FLX**
   - module name: 454
   - version 2.9
   - Other versions: 2.6, 2.7, 2.8
   - Roche 454 Sequencing System Software
   - http://www.454.com/
   - Manuals in /group/bioinfo/apps/apps/454-2.9/doc

   Depends on gcc
   Depends on java/6

1. **gstreamer**
   - module name: gstreamer
   - https://gstreamer.freedesktop.org/
   - Open source multimedia framework
   - Includes plugins-base

   Depends on gcc

1. **gtextutils**
   - module name: libgtextutils
   - version 0.7
   - Other versions: 0.6
   - GNU Text Utilities

   Depends on gcc

1. **gVCFtools**
   - module name: gvcftools
   - version 0.15
   - Tools to create and analyze gVCF files
   - https://sites.google.com/site/gvcftools/home

   Depends on gcc

1. **HALC**
   - module name: HALC
   - version 2017-09-06
   - High Throughput Algorithm for Long Read Error Correction
   - https://github.com/lanl001/halc
   - Info: /group/bioinfo/apps/apps/HALC-2017-09-06/Readme.md

   Depends on blasr
   Depends on LoRDEC

1. **HAllA: Hierarchical All-against-All significance testing**
module name: **halla**
version 0.6.3

Statistical method for discovery of significant relationships among data features with high power.
Python module
http://huttenhower.sph.harvard.edu/halla
Forum: https://groups.google.com/forum/#!forum/halla-users

Depends on biopython/2.7.11

1. **hapsembler**
   module name: **hapsembler**
   version 2.21

   Haplotype-specific genome assembly toolkit
   Designed for genomes that are rich in SNPs and other types of polymorphism
   http://compbio.cs.toronto.edu/hapsembler/

   Depends on gcc

1. **Haskell Platform**
   module name: **haskell-platform**
   version 2013.2.0.0-ghc-7.6.3

   Environment for the functional language Haskell
   http://www.haskell.org/platform/

   Depends on ghc/7.6.3

1. **HDF5**
   module name: **HDF5**
   version 1.8.19
   Other versions: 1.8.14

   Data model, library, and file format for storing and managing data
   Clustering de novo assembled transcripts and counting overlapping reads
   http://www.hdfgroup.org/HDF5/

1. **HGTector**
   module name: **HGTector**
   version 0.2.1

   Genome-wide prediction of horizontal gene transfer events based on Blast hit distribution statistics
   https://github.com/DittmarLab/HGTector

   Feb 4 2017 release

   Depends on perl
   Depends on R-bioconductor/3.3.0

1. **HISAT2**
   module name: **hisat2**
   version 2.0.5
   Other versions: 2.0.4

   Graph-based alignment of next generation sequencing reads to a population of genomes
   http://ccb.jhu.edu/software/hisat2/index.shtml
   FAQ: http://ccb.jhu.edu/software/hisat2/faq.shtml

   Depends on gcc

1. **HMMER**
   module name: **HMMER**
version 3.1b2
Other versions: 2.3.2, 3.0, 3.1, 3.1b1

Biosequence analysis using profile hidden Markov models
http://hmmer.janelia.org/
Help: http://hmmer.janelia.org/help
Info: /group/bioinfo/apps/apps/hmmer-3.1b2/release-notes

Depends on gcc

1. HOMER (Hypergeometric Optimization of Motif EnRichment)
   module name: HOMER
   version 4.9
   Other versions: 4.1, 4.3, 4.7, 4.7b

   Suite of tools for Motif Discovery and next-gen sequencing analysis
   http://homer.salk.edu/homer/

   Installed data:
   ORGANISMS
   + fly-o v5.10 Drosophila melanogaster (fly) accession and ontology information
   + mouse-o v5.10 Mus musculus (mouse) accession and ontology information
   + human-o v5.10 Homo sapiens (human) accession and ontology information
   + arabidopsis-o v5.10 Arabidopsis thaliana (arabidopsis) accession and ontology information
   PROMOTERS
   + human-p v5.5 human promoters (human)
   + fly-p v5.5 fly promoters (fly)
   GENOMES
   + tair10 v5.10 arabidopsis genome and annotation (tair10)
   + mm10 v5.10 mouse genome and annotation for UCSC mm10
   + hg38 v5.10 human genome and annotation for UCSC hg38
   + hg19 v5.10 human genome and annotation for UCSC hg19

   Depends on perl
   Depends on samtools
   Depends on R

1. HTSeq
   module name: htseq
   version 0.7.0
   Other versions: 0.5.3p3, 0.6.1

   Analysing high-throughput sequencing data with Python
   https://github.com/simon-anders/htseq

   Depends on biopython/2.7.12

1. htslib
   module name: htslib
   version 1.6
   Other versions: 1.3, 1.3.1, 1.3.2, 1.4, 1.5

   A C library for reading/writing high-throughput sequencing data
   http://www.htslib.org/
   Includes these programs: bgzip, htsfile, tabix
   man pages for: htsfile, tabix, sam, faidx, vcf

   Depends on perl
   Depends on curl
   Depends on xz

1. HUMAnN2: The HMP Unified Metabolic Analysis Network 2
   module name: humann2
version 0.10.0

Pipeline for profiling microbial pathways from sequencing data
http://huttenhower.sph.harvard.edu/humann2

Includes full databases from:
http://huttenhower.sph.harvard.edu/humann2_data/chocophlan/full_chocophlan_plus_viral.v0.1.1.tar.gz
http://huttenhower.sph.harvard.edu/humann2_data/uniprot/uniref_annotated/uniref90_annotated_1_1.tar.gz

Forum: https://groups.google.com/forum/#!forum/humann-users
Tutorial: https://bitbucket.org/biobakery/biobakery/wiki/humann2

Depends on metaphlan2
Depends on diamond
Depends on RAPsearch
Depends on usearch/8.0

1. HUMAnN: The HMP Unified Metabolic Analysis Network
module name: humann
version 0.99

Pipeline for measuring microbial pathways in metagenomic data
https://huttenhower.sph.harvard.edu/humann

Need to copy files into your own space before running:
$ cd $RCAC_SCRATCH
$ tar zxf $HUMANN_PKG
$ cd humann-0.99
$ less README.text # start reading at "INPUTS" section

Depends on biopython/2.7.12
Depends on blast
Depends on R-bioconductor/3.3.0
Depends on MinPath
Depends on maq

1. hyb
module name: hyb
version Nov2014

Pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data
https://github.com/gkudla/hyb
Info: man hyb

Depends on blat
Depends on flexbar
Depends on blast
Depends on bowtie2
Depends on ViennaRNA
Depends on unafold

1. i-ADHoRe
module name: i-adhore
version 3.0.01-mpi-b
Other versions: 3.0.01, 3.0.01-mpi

Detect degenerated homology relations within and between different genomes
Manual: /group/bioinfo/apps/apps/i-adhore-3.0.01-mpi-b/documentation/iADHoRe_manual_3_0.pdf
Includes support for MPI

Depends on gcc
Depends on openmpi

1. IDBA
module name: idba
version 1.1.2
Other versions: 0.19, 1.1.0, 1.1.1, 1.1.3

Iterative De Bruijn Graph De Novo short read assembler specially designed with highly uneven sequencing depth
https://github.com/loneknightpy/idba
IDBA, IDBA-UD, IDBA-Hybrid and IDBA-Tran all in one package

Depends on gcc
Depends on blat

1. IDBA-UD
module name: idba_ud
version 1.0.9

Iterative De Bruijn Graph De Novo Assembler for Short Reads Sequencing data with Highly Uneven Sequencing Depth.

Depends on biopython/2.7.2

1. IDR (Irreproducible Discovery Rate)
module name: idr
version 2.0.3

Measures consistency between replicates in high-throughput experiments
https://www.encodeproject.org/software/idr/
https://sites.google.com/site/anshulkundaje/projects/idr
Version info: https://groups.google.com/forum/#!topic/idr-discuss/A7PaMnzoFwg

Depends on biopython/3.5.1

1. igraph
module name: igraph
version 0.7.1

The network analysis package
http://igraph.org/c/

Depends on gcc

1. IGV (Integrative Genomics Viewer)
module name: igv
version 2.3.60
Other versions: 2.1.13, 2.1.30, 2.3.12, 2.3.57

Visualization tool for interactive exploration of genomic datasets
http://www.broadinstitute.org/software/igv/home
FAQ: http://www.broadinstitute.org/software/igv/FAQ
Forum: https://groups.google.com/forum/#!forum/igv-help
Info: /group/bioinfo/apps/apps/IGV_2.3.60/readme.txt
Start with "igv.sh"
Need X11 display: More config info at:
https://www.rcac.purdue.edu/compute/carter/guide/#accounts_login_x11

Depends on java

1. IGV Tools
module name: igvtools
version 2.3.60
Other versions: 2.1.7, 2.1.24, 2.3.12, 2.3.57

Utilities for preprocessing data files
http://www.broadinstitute.org/software/igv/home
FAQ: http://www.broadinstitute.org/software/igv/FAQ
Forum: https://groups.google.com/forum/#!forum/igv-help
Info: /group/bioinfo/apps/apps/IGVTools-2.3.60/igvtools_readme.txt
Need X11 display for igvtools_gui

Depends on java

1. IMAGE
   module name: image
   version 2.33
   Other versions: 2.31
   Iterative Mapping and Assembly for Gap Elimination
   Close gaps in any draft assembly using Illumina paired end reads
   Note: This software doesn't work

Depends on gcc

1. ImageMagick
   module name: ImageMagick
   version 6.8.9-8
   Other versions: 6.8.6-8
   Software suite to create, edit, compose, or convert bitmap images
   http://www.imagemagick.org/script/index.php

Depends on gcc

1. indelMINER
   module name: indelMINER
   version 1.0
   Identify indels in whole-genome resequencing project
   http://www.biomedcentral.com/1471-2105/16/42

Depends on gcc

1. Infernal ("INFERence of RNA ALignment")
   module name: infernal
   version 1.1.2
   Other versions: 1.0.2, 1.1.1
   Builds probabilistic profiles of the sequence and secondary structure of an RNA family
   http://eddylab.org/infernal/
   Inf: /group/bioinfo/apps/apps/infernal-1.1.2-linux-intel-gcc/documentation/
   Ref: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3810854/

Depends on gcc

1. InParanoid
   module name: InParanoid
   version 4.1
   Identify orthologs
   http://inparanoid.sbc.su.se/cgi-bin/index.cgi
   Need to make copy of InParanoid files to your own directory and run from there.
   For example:
   cd $RCAC_SCRATCH
   tar zxf $INPARANOID
   cd inparanoid-4.1
   # consult README file on how to use software
1. iNPS
   module name: iNPS
   version 1.0
   Accurate nucleosome positioning from sequencing data
   http://www.picb.ac.cn/hanlab/iNPS.html
   Depends on perl
   Depends on java
   Depends on blast/2.2.26

1. InterProScan
   module name: InterProScan
   version 5.25-64.0
   Other versions: 5-RC7, 5.1-44.0, 5.2-45.0, 5.3-46.0, 5.4-47.0, 5.5-48.0, 5.6-48.0, 5.7-48.0, 5.8-49.0, 5.9-50.0, 5.10-50.0, 5.11-51.0, 5.12-52.0, 5.13-52.0, 5.14-53.0, 5.15-54.0, 5.16-55.0, 5.17-56.0, 5.18-57.0, 5.19-58.0, 5.20-59.0, 5.21-60.0, 5.22-61.0, 5.23-62.0, 5.24-63.0
   Scans a range of protein signatures against your sequence
   https://github.com/ebi-pf-team/interproscan/wiki
   Manual: https://github.com/ebi-pf-team/interproscan/wiki/HowToRun
   Notes: https://github.com/ebi-pf-team/interproscan/wiki/Interproscan5_24_63_ReleaseNotes
   InterPro version 63.0 data (includes Panther 11.1 data)
   Uses pre-calculated match lookup service at http://www.ebi.ac.uk
   Includes support for TMHMM, Phobius and SignalP
   To adjust parameters in interproscan.properties file, you need to copy
   the software and run it in your own directory. See instructions at
   /group/bioinfo/apps/apps/interproscan-5.25-64.0/README.local
   Depends on perl
   Depends on biopython/2.7.8
   Depends on java/8
   Depends on signalp
   Depends on tmhmm
   Depends on phobius

1. IPknot
   module name: ipknot
   version 0.0.2
   RNA pseudoknot prediction based on maximizing expected accuracy
   http://www.ncrna.org/software/ipknot/
   Info: /group/bioinfo/apps/apps/ipknot-0.0.2/README.rd

1. iPlant iRODS iCommands
   module name: icommands
   version 3.0
   Storing and accessing iPlant data with iRODS
   https://pods.iplantcollaborative.org/wiki/display/start/Storing+Your+Data+with+iPlant+and+Accessing+that+Data
   https://pods.iplantcollaborative.org/wiki/display/start/Using+icommands
   https://pods.iplantcollaborative.org/wiki/display/start/Sharing+Your+Data
   https://www.irods.org/index.php/What_is_iRODS%3F

1. IQ-TREE
   module name: IQ-TREE
   version 1.5.5
1. **Infer phylogenetic trees by maximum likelihood**
   
   http://www.iqtree.org/
   
   https://github.com/Cibiv/IQ-TREE
   
   
   Includes both iqtree and iqtree-omp

---

1. **Isaac Aligner**
   - module name: isaac_aligner
   - version 01.13.06.27
   
   Illumina/HiSeq Human WGS analysis workflow
   
   https://github.com/sequencing/isaac_aligner
   
   
   License: /group/bioinfo/apps/apps/issac_aligner-01.13.06.27/illumina_Open_Source_License_1.pdf
   
   Depends on gcc
   
   Depends on gnuplot

---

1. **Isaac Variant Caller**
   - module name: isaac_variant_caller
   - version 1.0.4
   
   Detect SNVs and small indels from the aligned sequencing reads of a single diploid sample.
   
   https://github.com/sequencing/isaac_variant_caller
   
   License: /group/bioinfo/apps/apps/isaac_variant_caller-1.0.4/illumina_Open_Source_License_1.pdf
   
   Depends on gcc
   
   Depends on gvcftools

---

1. **ITSx**
   - module name: ITSx
   - version 1.0.11
   
   Extract ITS1 and ITS2 (fungal molecular barcode) subregions from ITS sequences.
   
   http://microbiology.se/software/itsx/
   
   
   
   Depends on HMMER

---

1. **jabba**
   - module name: jabba
   - version 1.0.0
   
   Read error correction tool for PacBio data
   
   https://github.com/biointec/jabba
   
   Info: https://github.com/biointec/jabba/wiki
   
   Depends on gcc

---

1. **Java**
   - module name: java
   - version 1.8.0_144
   
   Other versions: 1.8.0_31, 1.6.0_38, 1.6.0_39, 1.6.0_43, 1.6.0_45, 1.7.0_04, 1.7.0_05, 1.7.0_07, 1.7.0_09, 1.7.0_10, 1.7.0_11, 1.7.0_13, 1.7.0_17, 1.7.0_21, 1.7.0_25, 1.7.0_40, 1.7.0_45, 1.7.0_51, 1.7.0_55, 1.7.0_67, 1.7.0_71, 1.7.0_75, 1.7.0_79, 1.8.0_05, 1.8.0_11, 1.8.0_20, 1.8.0_25, 1.8.0_31, 1.8.0_45, 1.8.0_60, 1.8.0_65, 1.8.0_73, 1.8.0_77, 1.8.0_91, 1.8.0_101, 1.8.0_111, 1.8.0_121, 1.8.0_131, 6, 7, 8
   
   Java Standard Edition Development Kit (JDK) for Linux on Intel 64 machines
   
   
   Critical Patch Update (CPU) release--fixes vulnerabilities and critical bugs
1. **Jellyfish**
   module name: *jellyfish*
   version 2.2.6
   Other versions: 1.1.3, 1.1.4, 1.1.11, 2.1.1, 2.2.4

   Fast, Parallel k-mer Counting for DNA
   https://github.com/gmarcais/Jellyfish
   http://www.genome.umd.edu/jellyfish.html
   Manual: "man jellyfish"

   Depends on gcc

1. **jemalloc**
   module name: *jemalloc*
   version 3.4.1

   Memory allocator with fragmentation avoidance and scalable concurrency support
   man -s3 jemalloc

   Depends on gcc

1. **JIGSAW:**
   module name: *jigsaw*
   version 3.2.10

   Gene prediction using multiple sources of evidence
   http://cbcb.umd.edu/software/jigsaw/
   Info in /group/bioinfo/apps/apps/jigsaw-3.2.10/README.txt

   Depends on gcc
   Depends on perl

1. **jModelTest**
   module name: *jModelTest*
   version 2.1.4

   Statistical selection of best-fit models of nucleotide substitution
   http://code.google.com/p/jmodeltest2/

   Run as: java -jar $JMODELTEST_HOME/jModelTest.jar ...
   May need to copy to your home directory in order to run:
   cp -r $JMODELTEST_HOME ~/.
cd jModelTest-2.1.4
   [modify conf/jmodeltest.conf as needed]
   java -jar jModelTest.jar ...

   Need X11 support to run GUI

   Depends on java
   Depends on phyml

1. **KAAS - KEGG Automatic Annotation Server**
   module name: *KAAS*
   version 2.0

   Functional annotation of genes by BLAST comparisons against the manually curated KEGG GENES database
   http://www.genome.jp/tools/kaas/
   To use: review instructions at /group/bioinfo/apps/apps/KAAS-2.0/readme.txt
   Then, copy /group/bioinfo/apps/apps/KAAS-2.0/kaas.config to your working directory
   Issue all KAAS commands from that directory

   Depends on ruby
Depends on perl
Depends on HMMER
Depends on blast/2.2.26

1. **kallisto**
   - module name: kallisto
   - version 0.43.1
   - Other versions: 0.43.0
   - Quantify abundances of transcripts from RNA-Seq data
     - https://pachterlab.github.io/kallisto/
     - Ref: http://www.nature.com/nbt/journal/v34/n5/full/nbt.3519.html

   Depends on gcc

1. **KARECT (KAUST Assembly Read Error Correction Tool)**
   - module name: karect
   - version 1.0
   - Correction of Substitution, Insertion and Deletion Errors for Next-generation Sequencing Data
     - https://github.com/aminallam/karect
     - http://aminallam.github.io/karect/

   Depends on gcc

1. **kentUtils**
   - module name: kentUtils
   - version 302.1.0
   - UCSC command line bioinformatic utilities
     - http://genome.cse.ucsc.edu/index.html
     - https://github.com/ENCODE-DCC/kentUtils

   Depends on gcc

1. **khmer**
   - module name: khmer
   - version 1.4.1
   - Other versions: 2012-07-06
   - In-memory nucleotide sequence k-mer counting, filtering, graph traversal and more
     - Forum: http://lists.idyll.org/listinfo/khmer
     - Info in /group/bioinfo/apps/apps/khmer-1.4.1/doc
     - Citation: /group/bioinfo/apps/apps/khmer-1.4.1/CITATION

   Depends on gcc
   Depends on tcl
   Depends on igraph
   Depends on HDF5/1.8.14
   Depends on R-bioconductor/3.1.0

1. **KmerGenie**
   - module name: kmergenie
   - version 1.6982
   - Other versions: 1.5692
   - Estimates the best k-mer length for de novo assembly
     - http://kmergenie.bx.psu.edu/

   Built with maximum kmer size = 200

   Depends on biopython/2.7.8

1. **KOBAS (KEGG Orthology Based Annotation System)**
module name: KOBAS
version 2.1.0

Identify statistically significantly enriched pathways, human
diseases, and functional terms for an input set of genes using
biological knowledge from well-known pathway databases, disease
databases, and Gene Ontology.

http://kobas.cbi.pku.edu.cn/home.do
Tutorial: http://kobas.cbi.pku.edu.cn/help.do
Tutorial: /group/bioinfo/apps/apps/kobas-2.1.0/doc/tutorial.txt
Release notes: /group/bioinfo/apps/apps/kobas-2.1.0/doc/Release_Note.txt

Depends on biopython/2.7.8
Depends on blast

1. kSNP
   module name: meraculous
   version 2.0.5

   Distributed and scalable assembler for eukaryotic genomes
   http://jgi.doe.gov/data-and-tools/meraculous/

   Depends on gcc
   Depends on Boost
   Depends on perl

1. L_RNA_scaffolder
   module name: L_RNA_scaffolder
   version 14Feb2014

   Long transcriptome reads to scaffold genome

   Depends on gcc
   Depends on perl/5.18.1

1. LAPACK Linear Algebra PACKage and BLAS
   module name: lapack
   version 3.4.2

   http://www.netlib.org/lapack/
   compiled with gcc 4.7.2

   Depends on gcc

1. LAST
   module name: LAST
   version 801

   Finds similar regions between sequences
   http://last.cbrc.jp/

   Depends on gcc

1. LASTZ
   module name: LASTZ
   version 1.04.00

   Pairwise DNA sequence aligner
   https://github.com/lastz/lastz
   http://www.bx.psu.edu/~rsharris/lastz/
   Manual: http://www.bx.psu.edu/~rsharris/lastz/README.lastz-1.04.00.html
Includes lastz_32

Depends on biopython/2.7.12

1. **LEfSe:** (Linear discriminant analysis effect size)
   
   module name: lefse
   version 2017-02-03

   Find biomarkers between two or more groups using relative abundances
   Info in step #4 in https://bitbucket.org/nsegata/metaphlan/wiki/MetaPhlAn_Pipelines_Tutorial
   Tutorial: https://bitbucket.org/biobakery/biobakery/wiki/lefse
   Run it online: https://huttenhower.sph.harvard.edu/galaxy/

   Depends on biopython/2.7.12
   Depends on R-bioconductor/3.3.0

1. **libcfu**
   
   module name: libcfu
   version 0.03

   Multithreaded software library
   https://sourceforge.net/projects/libcfu/

   Depends on gcc

1. **libgmp**
   
   module name: libgmp
   version 1.5.3

   General purpose cryptographic library
   http://www.gnu.org/software/libgmp/

   Depends on libgmp-error

1. **libsvm**
   
   module name: libsvm
   version 322

   Support vector classification, regression and distribution estimation.
   http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html
   https://github.com/cjlin1/libsvm

   Depends on gcc

1. **libuuid**
   
   module name: libuuid
   version 2.17.0

   Universally Unique ID library
   Libs in /group/bioinfo/apps/apps/libuuid-2.17.0/lib64
   Header/include files in /group/bioinfo/apps/apps/libuuid-2.17.0/include

   Depends on gcc

1. **libxml2**
   
   module name: libxml2
   version 2.9.1

   XML C parser and toolkit developed for the Gnome project
   http://www.xmlsoft.org/

   Depends on gcc
1. **libxslt**  
   module name: **libxslt**  
   version 1.1.28  
   
   XSLT C library developed for the GNOME project  
   XSLT is a an XML language to define transformation for XML  
   [http://xmlsoft.org/XSLT/](http://xmlsoft.org/XSLT/)  
   
   Depends on gcc  
   Depends on libxml2

1. **IncRNA-screen**  
   module name: **IncRNA-screen**  
   version 0.1  
   
   Pipeline for computationally screening putative IncRNA transcripts over multimodal datasets  
   [https://github.com/NYU-BFX/IncRNA-screen](https://github.com/NYU-BFX/IncRNA-screen)  
   
   Warning: this probably does not work.  
   
   First, put a copy of the program files somewhere you can manage:  
   
   $ cd $RCAC_SCRATCH  
   $ cp -r /group/bioinfo/apps/apps/IncRNA-screen-0.1 ./  
   $ cd IncRNA-screen-0.1/IncRNA-screen  
   
   Next, follow the directions in the “Setting up” section of the web site for where to put your data and hints on how to run.  
   
   Depends on biopython/2.7.11  
   Depends on samtools/0.1.19  
   Depends on R-bioconductor/3.3.0

1. **IncRScan-SVM**  
   module name: **IncRScan-SVM**  
   version 1.0.1  
   
   Predict long non-coding RNAs (IncRNAs) or protein coding transcripts using a support vector machine (SVM)  
   [https://sourceforge.net/projects/incrscansvm/files/](https://sourceforge.net/projects/incrscansvm/files/)  
   
   Need to run from your own directory.  
   Use "setup_IncRScan-SVM $RCAC_SCRATCH" to run from your scratch directory  
   
   Depends on biopython/2.7.11  
   Depends on libsvm

1. **LoRDEC**  
   module name: **LoRDEC**  
   version 0.5.3  
   
   A hybrid error correction program for long, PacBio reads  
   [https://gatb.inria.fr/software/lordec/](https://gatb.inria.fr/software/lordec/)  
   Warning: This software might not work  
   
   Depends on gcc

1. **LoRMA**  
   module name: **LoRMA**  
   version 0.4  
   
   Correct sequencing errors in long reads  
   [https://www.cs.helsinki.fi/u/lmsalmel/LoRMA/](https://www.cs.helsinki.fi/u/lmsalmel/LoRMA/)  
   Warning: This software might not work
1. **LTR_FINDER**
   module name: *LTR_FINDER*
   version 1.0.6
   Other versions: 1.0.5
   
   Find full-length LTR retrotransposons in genome sequences
   http://tlife.fudan.edu.cn/ltr_finder/
   Info: /group/bioinfo/apps/apps/LTR_FINDER.x86_64-1.0.6/help.pdf
   
   Depends on *gcc*

1. **Lucy**
   module name: *lucy*
   version 1.20p
   
   DNA sequence quality and vector trimming tool
   http://lucy.sourceforge.net/
   Manual: man lucy
   
   Depends on *gcc*

1. **lz4**
   module name: *lz4*
   version 131
   Other versions: 122
   
   Extremely Fast Compression algorithm
   https://github.com/Cyan4973/lz4
   http://fastcompression.blogspot.com/p/lz4.html
   man lz4
   
   Depends on *gcc*

1. **MACS**
   module name: *macs*
   version 2.1.0.20150420
   Other versions: 1.4.2a, 1.4.2b, 2.0.10, 2.0.10-2.7.2, 2.1.0
   
   Model-based Analysis for ChIP-Seq
   https://github.com/taoliu/MACS/
   Info: /group/bioinfo/apps/apps/MACS-2.1.0.20150420/README.rst
   
   Depends on *biopython/2.7.12*

1. **MAFFT**
   module name: *mafft*
   version 6.903
   
   Multiple alignment program for amino acid or nucleotide sequences
   http://mafft.cbrc.jp/alignment/server/index.html
   man pages for mafft and mafft-homologs

1. **MAKER**
   module name: *MAKER*
   version 2.31.9-nompi
   Other versions: 2.31.9-nompi
   
   Genome annotation pipeline
   http://www.yandell-lab.org/software/maker.html
   Info: http://www.yandell-lab.org/software/maker_docs.html
   Info: /group/bioinfo/apps/apps/MAKER-2.31.9-nompi/README
This version does not support MPI.

Depends on **SNAP**
Depends on **RepeatMasker**
Depends on **exonerate**
Depends on **blast**
Depends on **AUGUSTUS**
Depends on **GeneMark**

1. **MapSplice**
   - module name: *MapSplice*
   - version 2.2.1

Mapping RNA-seq data to reference genome for splice junction discovery that depends only on reference genome
http://www.netlab.uky.edu/p/bioinfo/MapSplice2

Depends on **gcc**

1. **MAQ**
   - module name: *maq*
   - version 0.7.1

MAQ: Mapping and Assembly with Qualities
http://maq.sourceforge.net/
http://sourceforge.net/projects/maq/files/maq
FAQ: http://maq.sourceforge.net/faq.shtml

Depends on **gcc**

1. **MaSuRCA (Maryland Super Read Celera Assembler)**
   - module name: *MaSuRCA*
   - version 3.2.3
   - Other versions: 2.1.0, 3.1.3, 3.2.2, 3.2.2_RC1

Genome assembler
http://www.genome.umd.edu/masurca.html
Info: man jellyfish

Create sample config file with: masurca -g config.txt
Put in assembly directory and modify as needed

Depends on **gcc**
Depends on **Boost**
Depends on **perl**

1. **MATLAB Compiler Runtime (MCR)**
   - module name: *MCR*
   - version 7.14

Shared libraries that enables the execution of compiled MATLAB applications
http://www.mathworks.com/products/compiler/mcr/
Corresponds to MATLAB R2010b

1. **matrix2png**
   - module name: *matrix2png*
   - version 1.2.2

Visualizations of microarray data and many other data types
http://www.chibi.ubc.ca/matrix2png/
Examples at http://www.chibi.ubc.ca/matrix2png/matrix2png-samples.html
1. **Mauve**
   - module name: *mauve*
   - version 2015-02-13
   - Other versions: 2.3.1

   [http://darlinglab.org/mauve/mauve.html](http://darlinglab.org/mauve/mauve.html)

   Construct multiple genome alignments in the presence of
   large-scale evolutionary events such as rearrangement and inversion


   Note: Needs X11 support for GUI (not needed with command line
   utilities mauveAligner or progressiveMauve)

   Depends on java

1. **MaxBin**
   - module name: *MaxBin*
   - version 2.2.3

   Binning assembled metagenomic sequences based on an Expectation-Maximization algorithm

   [https://downloads.jbei.org/data/microbial_communities/MaxBin/MaxBin.html](https://downloads.jbei.org/data/microbial_communities/MaxBin/MaxBin.html)

   FAQ: [https://downloads.jbei.org/data/microbial_communities/MaxBin/MaxBin.html#faq](https://downloads.jbei.org/data/microbial_communities/MaxBin/MaxBin.html#faq)

   Depends on gcc

   Depends on bowtie2/2.2.3

   Depends on HMMER/3.1b2

   Depends on idba/1.1.1

   Depends on FragGeneScan/1.30

   Depends on R-bioconductor/3.3.3

1. **mbuffer**
   - module name: *mbuffer*
   - version 20150412

   Tool for buffering data streams

   [http://www.maier-komor.de/mbuffer.html](http://www.maier-komor.de/mbuffer.html)

   Info: /group/bioinfo/apps/apps/mbuffer-20150412/README

   Info: man mbuffer

   Depends on gcc

1. **mcl**
   - module name: *mcl*
   - version 14-137
   - Other versions: 12-068

   The Markov Cluster Algorithm (MCL)

   MCL algorithm finds cluster structure in graphs


   man pages available

   Depends on gcc

1. **MCScanX**
   - module name: *MCScanX*
   - version 2366066597

   Multiple Collinearity Scan toolkit

   [http://chibba.pgmml.uga.edu/mcsan2/](http://chibba.pgmml.uga.edu/mcsan2/)


   Depends on gcc

   Depends on java

   Depends on perl
1. **mdust**
   module name: *mdust*
   version 2006-10-17

   Mask Low-Complexity DNA Sequences
   [https://github.com/lh3/mdust](https://github.com/lh3/mdust)
   [ftp://occams.dfci.harvard.edu/pub/tgi/software/seqclean/](ftp://occams.dfci.harvard.edu/pub/tgi/software/seqclean/)

1. **MeDUSA**
   module name: *medusa*
   version 1.0

   Pipeline for MeDIP-seq data analysis
   [http://www2.cancer.ucl.ac.uk/medicalgenomics/medusa/](http://www2.cancer.ucl.ac.uk/medicalgenomics/medusa/)
   Info at /group/bioinfo/apps/apps/medusa-1.0/README
   Copy and change the config file in
   /group/bioinfo/apps/apps/medusa-1.0/config_files/medusa_template.cfg

   Depends on perl
   Depends on R-bioconductor
   Depends on bwa
   Depends on samtools
   Depends on fastqc
   Depends on USeq
   Depends on BEDTools

1. **Medusa**
   module name: *Medusa*
   version 1.6

   [https://github.com/combogenomics/medusa](https://github.com/combogenomics/medusa)
   [http://combo.dbe.unifi.it/medusa](http://combo.dbe.unifi.it/medusa)

   A draft genome scaffoldor that uses multiple reference genomes in a graph-based approach
   Run using:
   ```
   java -jar $MEDUSA_JAR -f test/reference_genomes/ -i test/Rhodobacter_target.fna -v
   ```

   Depends on biopython/2.7.12
   Depends on java
   Depends on MUMmer

1. **MeFiT (Merging and Filtering Tool for Paired-End Reads)**
   module name: *MeFiT*
   version 1.0

   Pipeline to merge overlapping paired-end reads, calculate merge statistics, and filter reads for quality
   [https://github.com/nisheth/MeFiT](https://github.com/nisheth/MeFiT)
   Example data in /group/bioinfo/apps/apps/MeFiT-1.0/sample
   Probably need to copy /group/bioinfo/apps/apps/MeFiT-1.0/casper.params

   Depends on biopython/2.7.11
   Depends on CASPER

1. **MEGAN (MEtaGenome ANalyzer) Community Edition**
   module name: *MEGAN*
   version 6.8.0
   Other versions: 4.7.0.4, 5.1.4, 5.2.3, 6.4.8, 6.5.7

   Software for analyzing metagenomes
   [http://ab.inf.uni-tuebingen.de/software/megan5/](http://ab.inf.uni-tuebingen.de/software/megan5/)
   Need X11 support for program to run (ssh -X)
   Configured to use 8 GB of memory
Depends on java/8

1. **MEME suite (Multiple Em for Motif Elicitation)**
   module name: meme
   version 4.10.1

   Discovers novel, ungapped motifs in nucleotide or protein sequences
   http://meme-suite.org/index.html
   This includes the command-line programs but not the web server

   Public web server: http://meme-suite.org/tools/meme
   Info in /group/bioinfo/apps/apps/meme-4.10.1/doc
   Forum: https://groups.google.com/forum/#!forum/meme-suite

   Depends on openmpi/1.8.1_gcc-4.7.2
   Depends on perl/5.20.1
   Depends on biopython/2.7.8

1. **MeQA**
   module name: MeQA
   version 1.0.0

   A pipeline for MeDIP-seq data quality assessment and analysis
   http://life.tongji.edu.cn/meqa/
   Need to copy /group/bioinfo/apps/apps/MeQA-1.0.0/example-config.cfg
   to your project dir and modify to suit your needs

   Depends on R-bioconductor
   Depends on biopython/2.7.2
   Depends on bwa
   Depends on samtools
   Depends on BEDTools
   Depends on samstat
   Depends on CEAS

1. **MetaCV**
   module name: metacv
   version 2.2.9

   Composition based classification for short metagenomic sequences
   http://metacv.sourceforge.net/
   Need to install prokaryotes database from website (>16GB compressed)

   Depends on gcc

1. **MetaPhlAn: Metagenomic Phylogenetic Analysis**
   module name: metaphlan2
   version 2.6.0

   Profiling the composition of microbial communities from metagenomic shotgun sequencing data.
   http://huttenhower.sph.harvard.edu/metaphlan2

   Precomputed BLAST and bowtie2 mpa databases at
   - /group/bioinfo/apps/apps/metaphlan-2.6.0/db_v20

   Forum: https://groups.google.com/forum/#!forum/metaphlan-users
   FAQ: https://bitbucket.org/nsegata/metaphlan/wiki/FAQ
   Tutorial: http://bitbucket.org/biobakery/biobakery/wiki/metaphlan2

   Depends on biopython/2.7.12
   Depends on blast
   Depends on bowtie2
1. **MetaTrans**
   module name: *MetaTrans*
   version 1.0
   Metatranscriptomic sequence analysis pipeline
   http://www.metatrans.org/
   You need to make a copy of the program and data into a directory owned by you before you can use it.
   Depends on R-bioconductor/3.3.0
   Depends on biopython/2.7.11
   Depends on usearch

1. **MetaVelvet**
   module name: *MetaVelvet*
   version 1.2.01
   A short read assembler for metagenomics
   http://metavelvet.dna.bio.keio.ac.jp/
   Built with MAX KMER LENGTH=99 and CATEGORIES=4
   Depends on velvet

1. **methylQA**
   module name: *methylQA*
   version 0.1.5
   Methylation sequencing data quality assessment tool
   http://methylqa.sourceforge.net/index.php
   Tutorial: http://methylqa.sourceforge.net/tutorial.php
   Depends on gcc
   Depends on R-bioconductor/3.1.2
   Depends on texlive

1. **METIS**
   module name: *metis*
   version 5.0.2
   Serial Graph Partitioning and Fill-reducing Matrix Ordering
   http://glaros.dtc.umn.edu/gkhome/metis/metis/overview
   Depends on gcc

1. **MFEprimer**
   module name: *MFEprimer*
   version 2.0-246757
   A fast thermodynamics-based program for checking PCR primer specificity
   https://github.com/quwubin/MFEprimer
   On-line: http://biocompute.bmi.ac.cn/CZlab/MFEprimer-2.0/
   Info: /group/bioinfo/apps/apps/MFEprimer-2.0-246757/README.md
   Probably need to install in home directory if you want to use your own custom database
   Depends on biopython/2.7.8

1. **Microbiome Utilities Portal of the Broad Institute**
   module name: *microbiomeutil*
   version r20110519
Provides microbiome utilities:
- Sequence alignment (NAST-iEr)
- Chimera detection (ChimeraSlayer, WigeoN)
- Operational taxonomic unit OTU binning (TreeChopper)
- Sequence assembly (AmosCmp16Spipeline)
- database of reference 16S sequences

http://microbiomeutil.sourceforge.net/
Info: /group/bioinfo/apps/apps/microbiomeutil-r20110519/docs

Depends on blast/2.2.26
Depends on cdbtools

1. **MicroRazerS**
   module name: MicroRazerS
   version 1.0
   Fast mapping short RNAs onto a reference genome
   http://www.seqan.de/projects/MicroRazerS/
   Mailing list: https://lists.fu-berlin.de/pipermail/seqan-dev/

   Depends on gcc

1. **MinCED (Mining CRISPRs in Environmental Datasets)**
   module name: minced
   version 0.2.0
   Find CRISPRs in genomes or metagenomes
   https://github.com/ctSkennerton/minced/

   Depends on java

1. **Minia**
   module name: minia
   version 1.5418
   Other versions: 1.5418-maxk128, maxk32, maxk128
   Short-read assembler based on a de Bruijn graph
   http://minia.genouest.org/
   This version is compiled with support for maximum k-mer size of 128

   Depends on gcc

1. **Miniconda**
   module name: conda
   version 4.3.21
   Other versions: 4.3.11
   Package, dependency and environment management for any language
   https://github.com/conda/conda
   https://conda.io/docs/index.html

1. **MinPath (Minimal set of Pathways)**
   module name: MinPath
   version 1.2
   Biological pathway reconstructions using protein family predictions
   http://omics.informatics.indiana.edu/MinPath/

   Depends on biopython/2.7.12

1. **MIRA**
   module name: mira
version 4.0.2
Other versions: 3.4.1.1, 3.9.1, 3.9.3, 3.9.4, 3.9.5, 3.9.7, 3.9.8, 3.9.9, 3.9.10, 3.9.11, 3.9.12, 3.9.14, 3.9.15, 3.9.16, 3.9.18, 4.0, 4.0.1, 4.0rc1, 4.0rc2, 4.0rc3, 4.0rc4, 4.0rc5, 4.9.1, 4.9.3, 4.9.4, 4.9.5

Whole genome shotgun and EST sequence assembler
http://sourceforge.net/p/mira-assembler/wiki/Home/
Docs at /group/bioinfo/apps/apps/mira-4.0.2/docs
Mailing lists: http://www.chevreux.org/mira_mailinglists.html

Depends on gcc
Depends on tcl

1. **miRanda**
   module name: miRanda
   version 3.3a
   microRNA Target Detection Software
   http://www.microrna.org/microrna/getDownloads.do
   Manual: man miranda

   Depends on gcc

1. **miRDeep**
   module name: mirdeep
   version 2.0.0.8
   Other versions: 2.0.0.4, 2.0.0.7
   Discovering known and novel miRNAs from deep sequencing data
   Tutorial: /group/bioinfo/apps/apps/mirdeep-2.0.0.8/TUTORIAL

   Depends on perl
   Depends on bowtie
   Depends on ViennaRNA
   Depends on randfold
   Depends on squid

1. **miRDeep**
   module name: miRDeep-star
   version 37
   miRNA identification from RNA sequencing data
   http://www.australianprostatecentre.org/research/software/mirdeep-star

   Configured to use 8GB of memory
   Need X11 support for program to run (ssh -X)

   You need to make a copy of the program and data into a directory owned by you before you can use it. The steps below will copy everything into a directory in your home directory:

   ```
   $ cd $HOME
   $ cp -r $MDS_HOME $HOME # 6 GB of data will be copied
   ```

   Now you have your own copy of the program and data. You only need to do the above steps one time. Every time you want to run the program, do the following steps:

   ```
   $ cd miRDeep_star_v37
   $ ./miRDeep_star.sh # starts the program
   ```

   Depends on java/8
1. **miRDeep-P**  
   module name: *miRDeep-P*  
   version 1.3  
   A computational tool for analyzing the microRNA transcriptome in plants  
   http://faculty.virginia.edu/lilab/miRDP/  
   Manual at /group/bioinfo/apps/apps/miRDP-1.3/miRDP_manual1.3.pdf  
   Depends on perl  
   Depends on bowtie  
   Depends on ViennaRNA  
   Depends on randfold

2. **MIREAP**  
   module name: *mireap*  
   version 0.2  
   Reap miRNAs from deeply sequenced smRNA library  
   http://sourceforge.net/projects/mireap/  
   Depends on ViennaRNA  
   Depends on perl/5.16.1

3. **miRExpress**  
   module name: *miRExpress*  
   version 2.1.4  
   Other versions: 2.1.3  
   Analyzing high-throughput sequencing data for profiling microRNA expression  
   http://mirexpress.mbc.nctu.edu.tw/  
   Usage: /group/bioinfo/apps/apps/miRExpress-2.1.4/README  
   Data avail for download from http://mirexpress.mbc.nctu.edu.tw/  
   Depends on gcc

4. **miRNAkey**  
   module name: *miRNAkey*  
   version 1.2  
   A pipeline microRNA Deep Sequencing data analysis  
   http://ibis.tau.ac.il/miRNAkey/  
   Need to copy /group/bioinfo/apps/apps/miRNAkey-1.2 to your project directory and run from there  
   Depends on java  
   Depends on bwa  
   Depends on fastx  
   Depends on perl

5. **miRquant**  
   module name: *miRquant*  
   version 2.0  
   Detect miRNAs in smRNA-seq data  
   https://github.com/Sethupathy-Lab/miRquant  
   Tutorial: https://github.com/Sethupathy-Lab/miRquant/blob/master/tutorial/TUTORIAL.md  
   Need to make copy of miRquant files to your own directory and run from there.  
   For example:  
   cd $RCAC_SCRATCH  
   cp -rp $MIQUANT_HOME .  
   cd miRquant-2.0  
   # download genomes, modify files in "configuration" directory, etc
1. **MITE-Hunter**  
   module name: *MITE-Hunter*  
   version 11-2011  
   Find miniature inverted-repeat transposable elements from genomic sequences  
   [http://target.iplantcollaborative.org/mite_hunter.html](http://target.iplantcollaborative.org/mite_hunter.html)  
   [https://github.com/jburnette/MITE-Hunter](https://github.com/jburnette/MITE-Hunter)  
   Depends on blast/2.2.26  
   Depends on mdust  
   Depends on muscle  

1. **MITObim**  
   module name: *MITObim*  
   version 1.8  
   Other versions: 1.6, 1.7  
   Mitochondrial baiting and iterative mapping  
   [https://github.com/chrishah/MITObim](https://github.com/chrishah/MITObim)  
   Info: /group/bioinfo/apps/apps/MITObim-1.8/README.md  
   Test data: /group/bioinfo/apps/apps/MITObim-1.8/testdata1.tgz  
   Depends on perl  
   Depends on mira/4.0.2  

1. **MOCAT (Metagenomics Analysis Toolkit)**  
   module name: *MOCAT*  
   version 2.0.0  
   Pipeline for processing Illumina metagenomic data  
   [http://mocat.embl.de/index.html](http://mocat.embl.de/index.html)  
   Tutorial: [http://mocat.embl.de/tutorial.html](http://mocat.embl.de/tutorial.html)  
   Depends on perl  
   Depends on GeneMark  
   Depends on usearch/8  

1. **mod-primer3**  
   module name: *mod-primer3*  
   version 2.2.3  
   Modified primer3 software used by edittag  
   [https://github.com/faircloth-lab/mod-primer3](https://github.com/faircloth-lab/mod-primer3)  
   Thermodynamic parameter files located in:  
   /group/bioinfo/apps/apps/mod-primer3-2.2.3/src/primer3_config/  
   Depends on gcc  

1. **Mono**  
   module name: *mono*  
   version 4.0.2  
   Cross platform, open source .NET framework  
   [http://www.mono-project.com/](http://www.mono-project.com/)  
   Depends on gcc  

---  

**Depends on cutadapt/1.12**  
**Depends on BEDTools/2.25.0**  
**Depends on bowtie/1.1.0**  
**Depends on SHRiMP/2.2.2**  
**Depends on R-bioconductor/3.3.0**
1. **MOTHUR**
   module name: *mothur*
   version 1.39.3
   Other versions: 1.25.0, 1.25.1, 1.32.1, 1.34.4, 1.37.2, 1.37.6

   Microbial ecology tools
   mothur and uchime (not GUI version)
   http://www.mothur.org/
   wiki: http://www.mothur.org/wiki
   forum: http://www.mothur.org/forum

   Depends on gcc

1. **MrBayes**
   module name: *mrbayes*
   version 3.2.6
   Other versions: 3.2.1, 3.2.5

   Bayesian Inference of Phylogeny
   http://mrbayes.sourceforge.net/download.php

   Depends on beagle/2.1.2
   Depends on gcc
   Depends on openmpi

1. **MSPC (Multiple Sample Peak Calling)**
   module name: *MSPC*
   version 1.1.0.0
   Other versions: 1.0.0.0

   Using combined evidence from replicates to evaluate ChIP-seq peaks
   http://www.bioinformatics.deib.polimi.it/genomic_computing/MSPC/index.html
   Info: http://www.bioinformatics.deib.polimi.it/genomic_computing/MSPC/documentation.html

   Wrapper script provided; use "MSPC" command

   Depends on mono

1. **Mugsy**
   module name: *Mugsy*
   version 1r2.3

   Multiple whole genome alignment tool
   http://mugsy.sourceforge.net/

   Depends on gcc
   Depends on muscle

1. **MultiQC**
   module name: *MultiQC*
   version 0.8

   Aggregate results from bioinformatics analyses across many samples into a single report
   http://multiqc.info/

   Depends on biopython/2.7.11

1. **MUMmer**
   module name: *MUMmer*
   version 3.23

   Ultra-fast alignment of large-scale DNA and protein sequences
   http://mummer.sourceforge.net/
   Tutorial: http://mummer.sourceforge.net/examples
1. **MUSCLE**
   module name: *muscle*
   version 3.8.31
   Other versions: 3.8.425
   Multiple sequence alignment (faster than clustalw)
   http://www.drive5.com/muscle/
   Docs: http://www.drive5.com/muscle/docs.htm

1. **MySQL Connector/J**
   module name: *mysql-connector-java*
   version 5.1.21
   Other versions: 5.0.5, 5.0.8
   MySQL Connector/J is the official JDBC driver for MySQL
   http://www.mysql.com/products/connector/
   Docs: /group/bioinfo/apps/apps/mysql-connector-java-5.1.21/docs

Depends on java

1. **NCBI BLAST**
   module name: *blast*
   version 2.6.0+
   Other versions: 2.2.22, 2.2.25+, 2.2.26, 2.2.26+, 2.2.27+, 2.2.28+, 2.2.30+, 2.2.31+, 2.3.0+, 2.4.0+, 2.5.0+
   Basic Local Alignment Search Tool
   Manuals: http://www.ncbi.nlm.nih.gov/books/NBK1762
   Includes RMBlast (for use with RepeatMasker)

Depends on gcc

1. **NCBI BLAST CGI programs**
   module name: *wwwblast*
   version 2.2.26
   BLAST CGI programs (legacy version)
   Docs: /group/bioinfo/apps/apps/wwwblast-2.2.26/docs
   This "legacy" version of NCBI BLAST is deprecated; see
   for reasons to use BLAST+ instead
   These are the wwwblast CGI programs. They are designed to be run by a web server.
   Depends on gcc

1. **NCBI Entrez Direct**
   module name: *edirect*
   version 1.40
   Entrez utilities on the UNIX command line
   Docs: http://www.ncbi.nlm.nih.gov/books/NBK179288/
1. **NCBI Magic-BLAST**  
   module name: *magic-blast*  
   version 1.3.0  
   Other versions: 1.1.0, 1.2.0  
   Map next-gen RNA/DNA sequencing runs against a whole genome or transcriptome  
   [https://ncbi.github.io/magicblast/](https://ncbi.github.io/magicblast/)  
   Info: /group/bioinfo/apps/apps/ncbi-magicblast-1.3.0/README  
   Depends on gcc

1. **NCBI nseg**  
   module name: *nseg*  
   version 1995-09-28  
   Might have something to do with repeat library filtering?  
   No official instructions or documentation?  
   Mentioned in RepeatScout README:  
   [http://bix.ucsd.edu/repeatscout/readme.1.0.5.txt](http://bix.ucsd.edu/repeatscout/readme.1.0.5.txt)  
   Depends on gcc

1. **NcFTP client**  
   module name: *ncftp*  
   version 3.2.6  
   FTP client  
   Depends on gcc

1. **nc coils**  
   module name: *nc coils*  
   version 2.2  
   Secondary structure predictions from protein sequences  
   The algorithm was published in Lupas, van Dyke & Stock, Predicting coiled coils from protein sequences Science, 252, 1162-1164, 1991  
   Depends on gcc

1. **NetBeans IDE**  
   module name: *netbeans*  
   version 7.4  
   Develop applications with Java  
   [https://netbeans.org/](https://netbeans.org/)  
   [https://netbeans.org/community/releases/74/](https://netbeans.org/community/releases/74/)  
   Info: [https://netbeans.org/kb/index.html](https://netbeans.org/kb/index.html)  
   Needs X11 support  
   Depends on java

1. **NetBox**  
   module name: *netbox*  
   version 22-May-2014  
   Perform network analysis on human interaction networks  
Must make private copy of database files before using:
$ module load netbox
$ mkdir netbox_db
$ cd netbox_db
$ cp -r $NETBOX_HOME/db ./.
$ cp -r $NETBOX_HOME/config ./.
$ export NETBOX_HOME=`pwd`

If using csh or tcsh, change last line to:
% setenv NETBOX_HOME `pwd`

Depends on java
Depends on biopython/2.7.3

1. **NetCDF (Network Common Data Format)**
   module name: NetCDF
   version 4.4.1.1
   https://www.unidata.ucar.edu/software/netcdf/
   Software libraries and data formats for the creation, access, and sharing of array-oriented scientific data.

   Depends on gcc
   Depends on curl/7.46.0
   Depends on zlib/1.2.11
   Depends on HDF5/1.8.19

1. **Netpbm**
   module name: netpbm-progs
   version 10.35.58-10.el5.x86_64

   Toolkit for manipulation of graphic images
   http://netpbm.sourceforge.net/
   http://pkgs.org/centos-5-rhel-5/centos-rhel-x86_64/netpbm-progs-10.35.58-10.el5.x86_64.rpm/download/
   Manual pages available

1. **NextClip**
   module name: nextclip
   version 0.7

   Quality analysis and read preparation for Nextera long mate pair (LMP) libraries
   http://www.tgac.ac.uk/nextclip/
   Needs TeX Live 2012 or newer software to generate PDF report.

   Depends on gcc
   Depends on bwa
   Depends on R-bioconductor
   Depends on texlive

1. **ngs-bits**
   module name: ngs-bits
   version May-27-2016

   Short-read sequencing tools
   https://github.com/imgag/ngs-bits
   Includes SeqPurge

   Depends on biopython/2.7.8

1. **ngsplot**
   module name: ngsplot
   version 2.61
   Other versions: 2.41.3, 2.47

   Visualize next-generation sequencing (NGS) samples at functional genomic regions
   https://code.google.com/p/ngsplot/
Example data: /group/bioinfo/apps/apps/ngsplot-2.61/example
Depends on R-bioconductor/3.1.0
Depends on biopython/2.7.8

1. **Novocraft**
   module name: novocraft
   version 3.02.12
   Other versions: 2.07.05, 2.07.13, 2.08.01, 2.08.02
   Next-Generation Sequencing analysis suite
   http://www.novocraft.com/products/novoalign/
   Manual: some PDF files in /group/bioinfo/apps/apps/novocraft-3.02.12

1. **NPS (Nucleosome Positioning from Sequencing)**
   module name: NPS
   version 1.3.2
   Identify nucleosome positions given histone-modification ChIP-seq or
   nucleosome sequencing at the nucleosome level.
   http://liulab.dfci.harvard.edu/NPS/
   Might need to copy program files to home directory and run there:
   mkdir ~/NPS
   cd ~/NPS
   cp $NPS_HOME/* ~/NPS
   Depends on biopython/2.7.3

1. **NSeq**
   module name: NSeq
   version 1.0
   Nucleosome positioning
   https://github.com/songlab/NSeq
   Depends on java

1. **numexpr**
   module name: numexpr
   version 2.6.2
   Other versions: 1.7.3
   Fast numerical array expression evaluator for Python, NumPy, PyTables, pandas, bcolz and more
   https://github.com/pydata/numexpr
   Depends on biopython/2.7.12

1. **Oases**
   module name: oases
   version 0.2.08
   Other versions: 0.02.05, 0.02.06, 0.02.08
   De novo transcriptome assembler for very short reads
   http://www.ebi.ac.uk/~zerbino/oases/
   Manual at /group/bioinfo/apps/apps/oases_0.2.08/doc/OasesManual.pdf
   Includes colorspace version oases_de as well as oases
   Build parameters:
   - MAXKMERLENGTH=99
   - CATEGORIES=4
   - OPENMP=1
   Depends on gcc
1. **OpenMPI**
   module name: *OpenMPI*
   version 1.8.8_gcc-4.7.2
   Other versions: 1.6.5_gcc-4.7.2

   A High Performance Message Passing Library
   http://www.open-mpi.org/

   Depends on *gcc/4.7.2*

1. **OrthoFiller**
   module name: *OrthoFiller*
   version 1.1.4

   Identify missing annotations for evolutionarily conserved genes
   https://github.com/mpdunne/orthofiller

   Depends on OrthoFinder
   Depends on BEDTools/2.25.0
   Depends on R-bioconductor/3.4.0
   Depends on AUGUSTUS
   Depends on HMMER
   Depends on gff3toggf

1. **OrthoFilter**
   module name: *OrthoFinder*
   version 1.1.8

   Accurate inference of orthologous gene groups made easy
   https://github.com/davidemms/OrthoFinder
   Example data: /group/bioinfo/apps/apps/OrthoFinder-1.1.8/ExampleData

   Depends on blast
   Depends on mcl
   Depends on FastME
   Depends on DLCpar
   Depends on maff
   Depends on FastTree

1. **OrthoMCL DB**
   module name: *orthomcl*
   Ortholog Groups of Protein Sequences from Multiple Genomes
   version 2.0.9
   Other versions: 2.0.3, 2.0.8

   http://orthomcl.org/orthomcl/
   Info: /group/bioinfo/apps/apps/orthomcl-2.0.9/doc/OrthoMCLEngine/Main/UserGuide.txt
   Need to use relational database (like MySQL)

   Depends on perl
   Depends on mcl
   Depends on blast

1. **PAGIT**
   module name: *PAGIT*
   version 1.0

   Post Assembly Genome Improvement Toolkit
   Generate sequence by ordering contigs, closing gaps, correcting errors and transferring annotation.
   http://www.sanger.ac.uk/resources/software/pagit/
   Note: This program doesn't work
1. **PALS (Pairwise Aligner for Long Sequences)**
   module name: PALS
   version 1.0

   Finds local alignments between DNA sequences
   http://drive5.com/pals/
   Warning: PALS is obsolete and no longer supported

   Depends on gcc

1. **PAML (Phylogenetic Analysis by Maximum Likelihood)**
   module name: PAML
   version 4.8

   Phylogenetic analyses of DNA or protein sequences using maximum likelihood.
   http://abacus.gene.ucl.ac.uk/software/paml.html
   Manuals in /group/bioinfo/apps/apps/PAML-4.8/doc/
   Examples in /group/bioinfo/apps/apps/PAML-4.8/examples/

   You may need to copy data and control files into your working directory:
   $ cp $paml_data/* paml/work/dir
   $ cd paml/work/dir
   $ baseml ...

   Depends on gcc

1. **PANDAseq**
   module name: pandaseq
   version 2.8.1
   Other versions: 0.6t, 2.0, 2.3, 2.3-RDP-1.0.3, 2.5, 2.6

   PAired-eND Assembler for DNA sequences
   https://github.com/neufeld/pandaseq
   Paper: http://www.biomedcentral.com/1471-2105/13/31/
   Wiki at https://github.com/neufeld/pandaseq/wiki
   man pages for pandaseq, pandaxs, pandaseq-checkid, pandaseq-diff and pandaseq-hang

   Depends on gcc

1. **Parallel-meta**
   module name: parallel-meta
   version 2.4

   Metagenomic analysis pipeline
   https://github.com/Comp-Bio-Group/Parallel-META
   Need to run with X11 support to generate plots:
   https://www.rcac.purdue.edu/compute/conte/guide/#accounts_login_x11

   Depends on gcc
   Depends on R-bioconductor/3.1.0
   Depends on HMMER
   Depends on blast/2.2.26
   Depends on velvet/1.2.10
   Depends on FragGeneScan/1.19

1. **ParsInsert**
   module name: ParsInsert
   version 1.04

   Produces both a phylogenetic tree and taxonomic classification for
   sequences for microbial community sequence analysis
   http://parsinsert.sourceforge.net/
1. **PartitionFinder**
   module name: `partitionfinder`
   version 2.1.1
   Discovers optimal partitioning schemes for DNA sequences
   Depends on biopython/2.7.12

1. **PASA_Lite (PASA without MySQL dependency)**
   module name: `PASA_Lite`
   version 0.1.0
   Eukaryotic genome annotation tool
   [https://github.com/PASApipeline/PASA_Lite](https://github.com/PASApipeline/PASA_Lite)
   Depends on perl
   Depends on gmap
   Depends on fasta/35
   Depends on seqclean
   Depends on blat

1. **PASHA: Parallelized Short Read Assembly**
   module name: `PASHA`
   version 1.0.10
   Parallel short read assemblers for large genomes
   [https://sourceforge.net/projects/pasha/](https://sourceforge.net/projects/pasha/)
   [http://pasha.sourceforge.net/homepage.htm](http://pasha.sourceforge.net/homepage.htm)
   Depends on gcc
   Depends on openmpi
   Depends on tbb

1. **patchelf**
   module name: `patchelf`
   [http://nixos.org/patchelf.html](http://nixos.org/patchelf.html)
   Modify the dynamic linker and RPATH of ELF executables
   Depends on gcc

1. **PBSuite**
   module name: `PBSuite`
   version 15.8.24
   Other versions: 14.9.9
   Software for Long-Read Sequencing Data from PacBio
   Info: /group/bioinfo/apps/apps/PBSuite_15.8.24/docs
   Depends on blasr
   Depends on biopython/2.7.8

1. **PCAP (Parallel Contig Assembly Program)**
   module name: `PCAP`
   version 06-07-05
   Assembles large genomes
   [http://seq.cs.iastate.edu/](http://seq.cs.iastate.edu/)
   [https://banana-slug.soe.ucsc.edu/bioinformatic_tools:cap3 PCAP](https://banana-slug.soe.ucsc.edu/bioinformatic_tools:cap3 PCAP)
   Manual: See directory /group/bioinfo/apps/apps/PCAP-06-07-05
read files: README, Doc, Doc.rep, Distributed.doc

To use autopcap, read the file autopcap.doc
You probably need to copy all the files in /group/bioinfo/apps/apps/PCAP-06-07-05
into your home directory and put your data into a subdirectory
(like was done in the "example" subdirectory.)

Depends on gcc

1. **PCRE**
   module name: `pcre`
   version 8.38

   Perl Compatible Regular Expressions
   http://www.pcre.org/
   man pages available
   Info: /group/bioinfo/apps/apps/pcre-8.38/share/doc

   Build options were:
   - `--enable-utf`
   - `--enable-unicode-properties`
   - `--enable-jit --disable-cpp`
   - `--enable-pcretest-libreadline`
   - `--enable-pcregrep-libz`
   - `--enable-pcregrep-libbz2`

   Depends on gcc
   Depends on zlib
   Depends on bzip2

1. **PeakSplitter**
   module name: `PeakSplitter`
   version 1.0

   Subdivision of ChIP-seq/ChIP-chip regions into discrete signal peaks
   http://www.ebi.ac.uk/bertone/software.html
   Usage: /group/bioinfo/apps/apps/PeakSplitter-1.0/README

   Depends on java

1. **PEAR - Paired-End reAd mergeR**
   module name: `pear`
   version 0.9.10
   Other versions: 0.9.6

   Pair-end read merger
   http://sco.h-its.org/exelixis/web/software/pear/

1. **PEP scaffold**
   module name: `PEP-scaffold`
   version Mar-28-2016

   Scaffolding tool using (homologous) proteins to scaffold genome fragments

1. **Perl**
   module name: `perl`
   version 5.20.1
   Other versions: 5.16.0, 5.16.1, 5.18.1

   Programming language
   http://www.perl.org/
   Manual: http://www.perl.org/docs.html
   Includes support for perl threads
   Built as shared perl library
Bioperl 1.6.924
Bio-SamTools-1.39 using samtools 0.1.19 library
Bio-BigFile-1.07 using Sep 2 2014 Jim Kent library

Depends on gcc
Depends on gnuplot
Depends on tidyp
Depends on gsl/2.3
Depends on ImageMagick/6.8.9-8
Depends on expat

1. Phobius
   module name: phobius
   version 1.01
   A combined transmembrane topology and signal peptide predictor
   http://phobius.sbc.su.se/
   Depends on gcc

1. Phobos
   module name: phobos
   version 3.3.12
   Tandem repeat search tool for complete genomes
   http://www.ruhr-uni-bochum.de/ecoenv/cm/cm_phobos.htm
   Info: /group/bioinfo/apps/apps/phobos-v3.3.12-linux/Quickstart.txt
   Depends on gcc

1. phrap
   module name: phrap
   version 1.080812
   Other versions: 1.090518
   Assemble shotgun DNA sequence data
   http://www.phrap.org/phredphrapconsed.html
   Includes cross_match
   Depends on cross_match/1.080812

1. phred
   module name: phred
   version 0.020425.c
   Assigns quality values to DNA sequences
   http://www.phrap.org/phredphrapconsed.html
   Manual: /group/bioinfo/apps/apps/phred-0.020425.c/PHRED.DOC
   Includes daev

1. Phusion Assembler
   module name: phusion-pipeline
   version 2.1c
   Assemble genome sequences from whole genome shotgun (WGS) reads
   http://www.sanger.ac.uk/resources/software/phusion/
   Info: /group/bioinfo/apps/apps/phusion_pipeline_v2.1c/README.1st
   Info: /group/bioinfo/apps/apps/phusion_pipeline_v2.1c/releaseNote_v2.1c
   Info: /group/bioinfo/apps/apps/phusion_pipeline_v2.1c/doc
   Depends on gcc
   Depends on phrap/1.090518
1. **Phusion2**
   module name: `phusion2`
   version 3.0
   Genome assembly pipeline based on read clustering
   Info: /group/bioinfo/apps/apps/phusion2-3.0/how_to_make_mates
   Depends on gcc

1. **PHYLIP** (PHYlogeny Inference Package)**
   module name: `PHYLIP`
   version 3.696
   Package of programs for inferring phylogenies (evolutionary trees)
   Depends on gcc

1. **PhyloBayes**
   module name: `PhyloBayes`
   version 4.1c
   Phylogenetic reconstruction using infinite mixtures
   [http://megasun.bch.umontreal.ca/People/lartillot/www/index.htm](http://megasun.bch.umontreal.ca/People/lartillot/www/index.htm)
   Depends on gsl/2.3
   Depends on gcc

1. **PhyML**
   module name: `phyml`
   version 20120412
   Other versions: 20131016
   Estimate maximum-likelihood phylogenies
   Forum: [https://groups.google.com/forum/#!forum/phyml-forum](https://groups.google.com/forum/#!forum/phyml-forum)
   Includes 20131031.patch
   No support for MPI
   Depends on gcc

1. **Picard**
   module name: `picard-tools`
   version 2.9.0
   Other versions: 1.63, 1.67, 1.72, 1.74, 1.80, 1.87, 1.88, 1.89, 1.90, 1.91, 1.92, 1.93, 1.94, 1.95, 1.96, 1.107, 1.108, 1.109, 1.110, 1.111, 1.112, 1.113, 1.114, 1.115, 1.117, 1.118, 1.119, 1.122, 1.123, 1.124, 1.125, 1.126, 1.127, 1.128, 1.129, 1.134, 1.135, 2.0.1, 2.3.0, 2.8.2
   Java based command-line utilities for SAM file manipulation
   Issues: [https://github.com/broadinstitute/picard/issues](https://github.com/broadinstitute/picard/issues)
   Picard-tools JAR file is at /group/bioinfo/apps/apps/picard-tools-2.9.0
   Use "PicardCommandLine" command for summary
   Or, use "java [jvm-options] -jar $PICARD/picard.jar [options]" for more control
   Depends on gcc
1. **PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States)**
   module name: **PICRUSt**
   version 1.0.0
   Predict metagenome functional content from marker gene surveys and full genomes.
   http://picrust.github.io/picrust/
   Forum: https://groups.google.com/group/picrust-users/subscribe?note=1&hl=en&noredirect=true&pli=1
   Tests: /group/bioinfo/apps/apps/PICRUSt-1.0.0/tests
   Depends on gcc
   Depends on tcl
   Depends on igraph
   Depends on HDFS/1.8.14
   Depends on R-bioconductor/3.1.0

1. **PILER (Parsimonious Inference of a Library of Elementary Repeats)**
   module name: **PILER**
   version 1.0
   Searches a genome sequence for repetitive elements
   http://www.drive5.com/piler/
   Depends on gcc
   Depends on PALS
   Depends on muscle

1. **Pilon**
   module name: **Pilon**
   version 1.22
   Automated Assembly Improvement and Variant Calling
   http://software.broadinstitute.org/software/pilon/
   Start program with:
   java -Xmx16G -jar $PILON_JAR [...other options...]
   Depends on java

1. **Pindel**
   module name: **Pindel**
   version 0.2.5b9, 20160729 (github branch downloaded May 9 2017)
   Other versions: May-9-2017, Oct-6-2014
   Detect breakpoints of large deletions, medium sized insertions, inversions, tandem duplications and other structural variants at single-based resolution from next-gen sequence data
   https://github.com/genome/pindel
   Depends on htslib/1.4

1. **PIPTS**
   module name: **PIPTS**
   version 1.3.4
   Pipeline for fungal internal transcribed spacer (ITS) analysis from Illumina sequencers
   https://github.com/hsgweon/pipits
   Depends on biopython/2.7.11
   Depends on fastx
   Depends on vsearch
   Depends on ITSx
1. **platanus**  
   module name: platanus  
   version 1.2.4  
   A de novo sequence assembler for highly heterozygous diploids from massively parallel shotgun sequencing data  
   http://platanus.bio.titech.ac.jp/?page_id=180  
   Info: http://platanus.bio.titech.ac.jp/?page_id=2  
   Depends on gcc

1. **PLEK**  
   module name: PLEK  
   version 1.2  
   Predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme  
   http://www.ibiomedical.net/plek/installation.html  
   Depends on biopython/2.7.8

1. **plink**  
   module name: plink  
   version 1.07  
   Whole genome association analysis toolset  
   http://pngu.mgh.harvard.edu/~purcell/plink  
   Info: /group/bioinfo/apps/apps/plink-1.07/README.txt  
   Depends on gcc

1. **PolyPhen-2 (Polymorphism Phenotyping v2)**  
   module name: PolyPhen  
   version 2.2.2r405a  
   Predicts possible impact of amino acid substitutions on the structure and function of human proteins using straightforward physical and evolutionary comparative considerations  
   http://genetics.bwh.harvard.edu/pph2/dokuwiki/start  
   Docs: http://genetics.bwh.harvard.edu/pph2/dokuwiki/docs  
   Will write temp files to $RCAC_SCRATCH/PPH_TEMP  
   Temp file directories will be created by loading this module.  
   To customize, copy config files with:  
   cp -r $PPH/config ~/.pph  
   then make changes  
   See /group/bioinfo/apps/apps/polyphen-2.2.2r405a/README for more info  
   Modified to use zebrafish instead of human data.  
   Depends on perl/5.18.1  
   Depends on blast/2.2.29+  
   Depends on java/6

1. **PoPoolation2**  
   module name: popoolation2  
   version 1201  
   Compare allele frequencies for SNPs between two or more populations  
   https://sourceforge.net/projects/popoolation2/  
   Info: https://sourceforge.net/p/popoolation2/wiki/Main/  
   The Java JAR is at: $POPOOLATION2_DIR/mpileup2sync.jar
Depends on perl
Depends on java
Depends on samtools
Depends on bwa

1. **pplacer**
   module name: **pplacer**
   version v1.1.alpha13

   Places query sequences on a fixed reference phylogenetic tree to
   maximize phylogenetic likelihood or posterior probability according to
   a reference alignment

   Depends on gcc

2. **Primer Prospector**
   module name: **pprospector**
   version 1.0.1

   Pipeline of programs to design and analyze PCR primers
   Info: /group/bioinfo/apps/apps/pprospector-1.0.1/ChangeLog

   Depends on biopython/2.7.8
   Depends on ViennaRNA
   Depends on rdp_classifier

3. **Primer3**
   module name: **primer3**
   version 2.3.7
   Other versions: 1.1.4, 2.3.4, 2.3.6

   Program for designing PCR primers
   Manual: /group/bioinfo/apps/apps/primer3-2.3.7/primer3_manual.htm
   Release notes: /group/bioinfo/apps/apps/primer3-2.3.7/release_notes.txt
   Thermodynamic parameter files located in:
   /group/bioinfo/apps/apps/primer3-2.3.7/primer3_config/

   Depends on gcc

4. **PROBer**
   module name: **PROBer**
   version 0.1.3-May-27b-2016
   Other versions: 0.1.2, 0.1.3-May-12-2016, 0.1.3-May-27-2016

   Quantitative modeling of transcriptome-wide RNA structure-probing experiments
   [https://github.com/pachterlab/PROBer](https://github.com/pachterlab/PROBer)
   Ref: /group/bioinfo/apps/apps/PROBer-0.1.3-May-27b-2016/README.md
   Ref: /group/bioinfo/apps/apps/PROBer-0.1.3-May-27b-2016/README.tools

   Depends on gcc

5. **Prodigal l (Prokaryotic Dynamic Programming Genefinding Algorithm)**
   module name: **prodigal**
   version 2.6.3
   Other versions: 2.60

   Microbial (bacterial and archaeal) gene finding program
   [https://github.com/hyattpd/prodigal](https://github.com/hyattpd/prodigal)
   Info: /group/bioinfo/apps/apps/prodigal-2.6.3/README.md
   Changelog: /group/bioinfo/apps/apps/prodigal-2.6.3/CHANGES
   Forum: [http://groups.google.com/group/prodigal-discuss](http://groups.google.com/group/prodigal-discuss)
Depends on gcc

1. **Program to Assemble Spliced Alignments (PASA)**
   module name: PASA
   version r20140417
   Other versions: 2.0.0, r2012-06-25, r20130605p1, r20130907

   Eukaryotic genome annotation tool
   http://pasa.sourceforge.net
   Info: /group/bioinfo/apps/apps/PASA_r20140417/docs
   Need MySQL database to store results

   Create environment variable $PASACONF with full path to your PASA configuration file; e.g.: export PASACONF=~/pasa.conf

   Depends on perl
   Depends on gmap
   Depends on fasta/35
   Depends on seqclean
   Depends on blat

1. **Prokka ("prokaryotic annotation")**
   module name: Prokka
   version 1.12

   Rapid annotation of prokaryotic genomes
   http://www.vicbioinformatics.com/software.prokka.shtml
   https://github.com/tseemann/prokka

   Depends on perl
   Depends on blast
   Depends on parallel
   Depends on tbl2asn
   Depends on HMMER
   Depends on infernal
   Depends on mummer
   Depends on signalp
   Depends on prodigal
   Depends on ARAGORN
   Depends on barrnap
   Depends on minced

1. **proovread**
   module name: proovread
   version 2.14.0

   PacBio hybrid error correction through iterative short read consensus
   https://github.com/BioInf-Wuerzburg/proovread

   Depends on gcc
   Depends on perl
   Depends on blast
   Depends on samtools

1. **ProtTest**
   module name: prottest
   version 3.3-20130716

   Select best-fit models of protein evolution
   http://code.google.com/p/prottest3/
   Need to copy to your home directory in order to run:
   cp -r $PROTTEST_SRC ~/.
   Follow usage instructions at:

   Depends on java
1. **Pscan_ChIP**
   module name: *Pscan_ChIP*
   version 1.1

   Find transcription factor-binding site motifs in ChIP-Seq experiments.
   http://159.149.160.51/pscan_chip_dev/
   Help: http://159.149.160.51/pscan_chip_dev/help.html
   Help: /group/bioinfo/apps/apps/Pscan_ChIP-1.1/README.txt

   Depends on *gsl*

2. **pyfasta**
   module name: *pyfasta*
   version 0.5.2

   Fast, memory-efficient, pythonic (and command-line) access to fasta sequence files
   https://pypi.python.org/pypi/pyfasta/

   Depends on *biopython/2.7.8*

3. **Pyicoteo**
   module name: *Pyicoteo*
   version 2.0.7

   Tools for analysing high-throughput sequencing (HTS) data.
   It was mainly developed using Solexa/Illumina mapped reads.
   http://regulatorygenomics.upf.edu/Software/Pyicoteo/index.html

   Depends on *biopython/2.7.11*

4. **PyQt**
   module name: *PyQt*
   version 4.9.4

   Python bindings for Qt application framework
   http://www.riverbankcomputing.co.uk/software/pyqt/intro

   Depends on *biopython*

5. **pyRAD**
   module name: *pyRAD*
   version 3.0.61
   Other versions: 2.15

   Analyze RAD, ddRAD, GBS, paired-end ddRAD and paired-end GBS data sets
   http://dereneaton.com/software/pyrad/

   Depends on *biopython/2.7.8*
   Depends on *muscle*
   Depends on *vsearch*

6. **pysam**
   module name: *pysam*
   version 0.7.7
   Other versions: 0.6, 0.7

   Python interface for the SAM/BAM sequence alignment and mapping format
   http://code.google.com/p/pysam/

   Depends on *biopython/2.7.3*

7. **pyScaf**
   module name: *pyScaf*
   version 0.12a
Genome assembly scaffolding
https://github.com/lpryszcz/pyScaf

Depends on biopython/2.7.11
Depends on LAST

1. **Python with bioinformatics modules**
   
   module name: biopython
   
   version 2.7.2
   
   Other versions: 2.7.3, 2.7.8, 2.7.11, 2.7.12, 3.5.1
   
   http://python.org/
   
   Install info: http://docs.python.org/install/index.html
   
   Depends on python/2.7.2

1. **PyVCF**
   
   module name: PyVCF
   
   version 0.6.7
   
   Variant Call Format (VCF) parser for Python
   
   https://pypi.python.org/pypi/PyVCF

   Depends on biopython/2.7.8

1. **QDD**
   
   module name: QDD
   
   version 3.1.2
   
   Other versions: 3.1.2b
   
   Select microsatellite markers and design primers
   
   http://net.imbe.fr/~emeglecz/qdd.html
   
   You might need to copy and modify the configuration file before running the programs:
   
   `cp -p $QDD_PATH/set_qdd_default.ini $RCAC_SCRATCH`
   
   `export QDD_CONF=$RCAC_SCRATCH/set_qdd_default.ini`
   
   `edit $RCAC_SCRATCH/set_qdd_default.ini`
   
   Depends on perl
   
   Depends on RepeatMasker/4.0.5
   
   Depends on clustalw
   
   Depends on primer3/2.3.6

1. **QIIME (Quantitative Insights Into Microbial Ecology)**
   
   module name: Qiime
   
   version 1.9.1
   
   Other versions: 1.6.0, 1.7.0, 1.8.0, 1.9.0, 2, 2-2017.2, 2-2017.8
   
   Comparison and analysis of microbial communities
   
   http://qiime.org/
   
   http://qiime.org/documentation_index.html
   
   Forum: http://forum.qiime.org/
   
   Info: http://qiime.wordpress.com/
   
   Config file info: http://qiime.org/install/qiime_config.html
   
   Parallel QIIME: http://qiime.org/tutorials/parallel_qiime.html
   
   Includes PhyloToAST (qiime-tools) version 1.1.2
   
   https://github.com/smdabdoub/phylotoast
   
   Info: http://qiime-tools.readthedocs.org/
   
   Includes prepare_taxa_charts.py
   
   Ref: https://github.com/Xcelris-Labs-Ltd/Publication-ready-taxonomic-charts-from-QIIME

   Depends on gcc
   
   Depends on tcl
   
   Depends on igraph
Depends on R-bioconductor/3.1.2
Depends on HDF5/1.8.14
Depends on rdp_classifier/2.2
Depends on microbiomeutil/r20110519
Depends on cd-hit/3.1.1
Depends on mothur/1.25.0
Depends on clearcut/1.0.9
Depends on RAxML/standard-5_7_2012
Depends on infernal/1.0.2
Depends on muscle/3.8.31
Depends on usearch/5.2.236
Depends on usearch/6.1.544
Depends on rtax/0.984
Depends on AmpliconNoise/1.27b
Depends on gsl
Depends on ghc
Depends on blast/2.2.22
Depends on Cytoscape
Depends on sourcetracker/0.9.5
Depends on ea-utils

1. QoRTs
   module name: QoRTs
   version 1.1.8
   Other versions: 1.1.2, 1.2.42

   Analysis, quality control, and data management of RNA-Seq datasets
   https://github.com/hartleys/QoRTs
   Vignette: http://hartleys.github.io/QoRTs/doc/QoRTs-vignette.pdf
   Info: http://hartleys.github.io/QoRTs/index.html
   FAQ: http://hartleys.github.io/QoRTs/FAQ.html

   Start java utility using "QoRTs" wrapper script.
   Or, using predefined JAR path:
   java -jar $QoRTs_JAR...

   Depends on R-bioconductor/3.3.0

1. Qt
   module name: Qt
   version 5.6.2
   Other versions: 5.4.0

   Cross-platform application framework
   https://www.qt.io/
   Includes qmake & QtCreator

   Depends on gcc

1. QTL Cartographer
   module name: QTLCart
   version 1.17j

   Suite of programs to map quantitative traits using a map of molecular markers.
   http://statgen.ncsu.edu/qtlcart/
   Example data at /group/bioinfo/apps/apps/QTLCart-1.17j/example

   Depends on gcc
   Depends on gnuplot

1. Quake
   module name: quake
   version 0.3.5

   Fix substitution sequencing errors in Illumina reads with deep coverage
http://www.cbcb.umd.edu/software/quake/index.html

Depends on gcc
Depends on Boost
Depends on jellyfish
Depends on R-bioconductor/3.1.0

1. QualiMap
   module name: qualimap
   version 2.2.1
   Other versions: 0.6, 1.0, 2.1.1, 2.1.3

   Graphical and command-line interface for quality control
   of alignment sequencing data
   http://qualimap.bioinfo.cipf.es/

   Documentation and mailing list on web site
   Need X11 support if you want to use GUI (ssh -X ...)

   Depends on java
   Depends on R-bioconductor/3.1.2

1. QUAST
   module name: quast
   version 3.2
   Other versions: 2.3

   Quality Assessment Tool for Genome Assemblies
   http://bioinf.spbau.ru/QUAST

   You need to install QUAST in your home directory as described in the manual.

   Depends on perl/5.20.1
   Depends on biopython/2.7.8
   Depends on java
   Depends on blast

1. QuEST
   module name: quest
   version 2.422
   Other versions: 2.4

   ChIP-Seq data analysis
   http://www.stanford.edu/~valouev/QuEST/QuEST.html
   Forum: http://groups.google.com/group/chipseq

   Depends on gcc
   Depends on perl

1. Quickmerge
   module name: quickmerge
   version 0.2

   A simple and fast metassembler and assembly gap filler designed for long molecule based assemblies.
   https://github.com/mahulchak/quickmerge

   Depends on MUMmer

1. Quip
   module name: quip
   version 1.1.1-3-g02d3228
   Other versions: 1.1.1

   Compresses next-generation sequencing data in the FASTQ and SAM/BAM formats
http://www.cs.washington.edu/homes/dcjones/quip/
Info: https://github.com/dcjones/quip#readme
or using "man quip"

Depends on gcc

1. **R with Bioconductor libraries**
   module name: R-bioconductor
   version 2.15.1
   Other versions: 2.14.1, 2.15.3, 3.0.2, 3.1.0, 3.1.2, 3.2.3, 3.3.0, 3.3.3, 3.4.0

   Statistics processing language with bioinformatics libraries
   http://www.r-project.org/
   http://www.bioconductor.org/

   Depends on gcc
   Depends on rhel5libs

1. **RADtools**
   module name: RADtools
   version 1.2.2

   RAD (Restriction site Associated DNA) Sequencing
   https://www.wiki.ed.ac.uk/display/RADSequencing/Home

   Depends on perl

1. **RandA**
   module name: RandA
   version 1.1.2

   Non-coding RNA profiling based on deep sequencing
   http://ibis.tau.ac.il/RandA/

   Depends on java
   Depends on bwa
   Depends on perl
   Depends on R-bioconductor
   Depends on ea-utils

1. **Randfold**
   module name: randfold
   version 2.0

   Minimum free energy of folding randomization test software
   randfold version 2 (C version)

   Usage: randfold

   methods:
   -s simple mononucleotide shuffling
   -d dinucleotide shuffling
   -m markov chain 1 shuffling

   Example: randfold -d let7.tfa 999

   Depends on gcc

1. **RAPsearch (Reduced Alphabet based Protein similarity Search)**
   module name: RAPsearch
   version 2.23
   Other versions: 2.12, 2.16
Fast protein similarity search
http://omics.informatics.indiana.edu/mg/RAPSearch2/
Info: /group/bioinfo/apps/apps/RAPsearch-2.23/readme

Depends on gcc

1. Rascaf
   module name: Rascaf
   https://github.com/mourisl/Rascaf
   Scaffolding with RNA-seq read alignment

Depends on gcc

1. RAxML (standard edition)
   module name: RAxML
   version standard-5_7_2012
   Other versions: 8.2.9
   RAxML (Randomized Accelerated Maximum Likelihood)
   Maximum Likelihood based inference of large phylogenetic trees
   http://sco.h-its.org/exelixis/software.html
   Info: /group/bioinfo/apps/apps/RAxML-standard-5_7_2012/README
   Forum: https://groups.google.com/forum/?fromgroups#!forum/raxml

Depends on gcc

1. Ray
   module name: ray
   version 2.3.1b
   Other versions: 2.1.0, 2.2.0, 2.3.0, 2.3.0-zlib, 2.3.1
   Parallel genome assemblies for parallel DNA sequencing
   http://denovoassembler.sourceforge.net/index.html
   Manual: /group/bioinfo/apps/apps/ray-2.3.1b/Documentation
   Docs: /group/bioinfo/apps/apps/ray-2.3.1b/MANUAL_PAGE.txt
   Forum: http://denovoassembler.sourceforge.net/faq.html
   Max k-mer length = 256
   Includes libz support for uncompressing fastq.gz files

Depends on gcc
Depends on openmpi

1. REAGO
   module name: REAGO
   version 1.1
   An assembly tool for 16S ribosomal RNA recovery from metagenomic data
   https://github.com/chengyuan/reago-1.1

   Might need to copy contents of "cm" directory from /group/bioinfo/apps/apps/reago-1.1 to your working directory.

Depends on biopython/2.7.11
Depends on infernal/1.1.1
Depends on GenomeTools

1. REAPR (Recognising Errors in Assemblies using Paired Reads)
   module name: REAPR
   version 1.0.16
   Evaluates the accuracy of a genome assembly using mapped paired end reads
   http://www.sanger.ac.uk/resources/software/reapr/

Depends on perl
Depends on R-bioconductor
Depends on Artemis
1. **RECON**
   module name: **RECON**
   version 1.08
   Other versions: 1.07

   Identifies repeat families from biological sequences
   http://selab.janelia.org/recon.html
   RepeatModeler version mentioned on:
   http://www.repeatmasker.org/RepeatModeler.html
   Docs: /group/bioinfo/apps/apps/RECON-1.08/00README
   Demos: /group/bioinfo/apps/apps/RECON-1.08/Demos/

   Depends on perl

1. **Redundans**
   module name: **Redundans**
   version 0.13a-try3
   Other versions: 0.13a, 0.13a-try2, 2016-11-18

   Assembly pipeline for heterozygous/polymorphic genomes
   https://github.com/Gabaldonlab/redundans

   Feb 1 2017 release

   Depends on biopython/2.7.11
   Depends on perl

1. **RepARK**
   module name: **RepARK**
   version 1.2

   Repetitive motif detection by Assembly of Repetitive K-mers
   https://github.com/PhKoch/RepARK
   Docs: /group/bioinfo/apps/apps/RepARK-1.2/README

   Depends on jellyfish/1.1.11
   Depends on velvet/1.2.07

1. **RepeatMasker**
   module name: **RepeatMasker**
   version 4.0.7
   Other versions: 3.3.0, 4.0.0, 4.0.1, 4.0.2, 4.0.3, 4.0.5, 4.0.6

   Screens DNA sequences for interspersed repeats and low complexity
   http://www.repeatmasker.org/
   Manual: /group/bioinfo/apps/apps/RepeatMasker-4.0.7/repeatmasker.help
   FAQ: http://www.repeatmasker.org/faq.html

   Dfam database version Dfam_2.0
   RepeatMasker Combined Database: Dfam_Consensus-20170127, RepBase-20170127

   Installed Search Engines:
   1. Default = RMBlast (NCBI BLAST with RepeatMasker extensions)
   2. CrossMatch
   3. HMMER & DFAM

   Genomic Datasets:
   http://www.repeatmasker.org/genomicDatasets/RMGenomicDatasets.html

   Depends on trf
   Depends on perl
   Depends on cross_match
   Depends on blast/2.2.28+
   Depends on HMMER
1. **RepeatModeler**
   module name: *RepeatModeler*
   version 1.0.9
   Other versions: 1.0.7, 1.0.8
   De-novo repeat family identification and modeling package
   http://www.repeatmasker.org/RepeatModeler.html
   Manual: /group/bioinfo/apps/apps/RepeatModeler-1.0.9/README
   Depends on **RepeatMasker**
   Depends on **RepeatScout**
   Depends on **RECON**

2. **RepeatScout**
   module name: *RepeatScout2*
   version 1.0.5
   This is probably a fork of RepeatScout
   https://github.com/DobzhanskyCenter/RepeatScout2
   Info: /group/bioinfo/apps/apps/RepeatScout2-1.0.5/README
   The -freq file can only have two fields per line
   Depends on **perl**
   Depends on **trf**
   Depends on **nseg**

3. **RGT (Regulatory Genomics Toolbox)**
   module name: *RGT*
   version 0.10.0
   http://www.regulatory-genomics.org/
   Integrative analysis of high throughput regulatory genomics data
   Includes THOR to detect and analyze differential peaks in ChIP-seq data
   http://www.regulatory-genomics.org/thor-2/basic-instruction/
   Depends on **gcc**

4. **riboFrame**
   module name: *riboFrame*
   version 1.0
   Improved method for microbial taxonomy profiling from non targeted metagenomics
   http://bioserver2.sbsc.unifi.it/bioinfo/riboframe.html
   Example data in /group/bioinfo/apps/apps/riboFrame-1.0/sample
   Depends on **HMMER**
   Depends on **rdp_classifier**
   Depends on **R-bioconductor/3.3.0**

5. **RIGERJ**
   module name: *rigerj*
   version 2.0.1
   Java implementation of RIGER and other gene enrichment methods, useful for RNAi and CRISPR screen analysis
   https://github.com/broadinstitute/rigerj
   Run using:
   java -jar $RIGERJ_JAR ...
   Depends on **java**

6. **rMATS**
   module name: *rMATS*
   version 3.0.9
   Multivariate Analysis of Transcript Splicing (MATS)
1. **RMBlas**
   module name: `rmblast`
   version 2.2.28
   Other versions: 1.2, 2.2.27

   RepeatMasker compatible version of the standard NCBI BLAST
   [http://www.repeatmasker.org/RMBlast.html](http://www.repeatmasker.org/RMBlast.html)

   Depends on blast/2.2.28+

1. **RNA-SeQC**
   module name: `RNA-SeQC`
   version 1.1.8

   Computes a series of quality control metrics for RNA-seq data

   Use "RNA-SeQC-cmd" command for summary
   Or, use "java [jvm-options] -jar $RNA_SeQC/RNA-SeQC.jar [options]" for more control

   Depends on java
   Depends on bwa

1. **RNAhybrid**
   module name: `RNAhybrid`
   version 2.1.2

   Find minimum free energy hybridization of a long and a short RNA
   [http://bibiserv.techfak.uni-bielefeld.de/rnahybrid](http://bibiserv.techfak.uni-bielefeld.de/rnahybrid)
   Manual: man RNAhybrid

   Depends on gcc

1. **RNAmmer**
   module name: `mammer`
   version 1.2

   Predicts ribosomal RNA genes in prokaryotic genome sequences
   [http://www.cbs.dtu.dk/services/RNAmmer/](http://www.cbs.dtu.dk/services/RNAmmer/)
   Info: man mammer

1. **maQUAST**
   module name: `maQUAST`
   version 1.4.0

   Evaluate quality of transcriptome assemblies using reference genome and gene database

   Depends on gcc
   Depends on biopython/2.7.12
   Depends on blat
   Depends on gmap
   Depends on blast
1. Rockhopper
   module name: Rockhopper
   version 2.0.3

   Analysis of bacterial RNA-seq data
   FAQ: http://cs.wellesley.edu/~btjaden/Rockhopper/FAQ.html
   Need X11 support for program to run (ssh -X)
   Use "Rockhopper" command to start GUI with default settings

   Depends on java/7

1. ROOT
   module name: ROOT
   version 6.02.03-rhel6

   Data analysis framework
   https://root.cern.ch/drupal/
   Binary-download; production version
   Requires gcc >= 4.8

   Depends on gcc/4.8.2

1. RSEG
   module name: rseg
   version 0.4.9

   Analyze ChIP-Seq data
   http://smithlabresearch.org/software/rseg/

   Depends on gcc

1. RSEM
   module name: rsem
   version 1.3.0
   Other versions: 1.1.19, 1.1.20, 1.1.21, 1.2.0, 1.2.7, 1.2.11, 1.2.12, 1.2.15, 1.2.18, 1.2.25, 1.2.30

   RSEM (RNA-Seq by Expectation-Maximization)
   http://deweylab.github.io/RSEM/
   Forum: https://groups.google.com/forum/#!forum/rsem-users

   Depends on gcc
   Depends on R-bioconductor/3.3.3
   Depends on perl
   Depends on bowtie
   Depends on bowtie2
   Depends on STAR

1. RSeQC
   module name: RSeQC
   version 2.6.3

   Evaluate high throughput sequence data especially RNA-seq data
   http://rseqc.sourceforge.net/

   Depends on biopython/2.7.8
   Depends on R-bioconductor/3.1.2

1. RStudio Desktop Open Source Edition
   module name: RStudio
   version 1.0.143
   Other versions: 1.1.240, 1.1.244
1. **rtax**
   - module name: rtax
   - version 0.984
   - Other versions: 0.983
   - Taxonomic classification of short paired-end sequence reads from the 16S ribosomal RNA gene.
   - http://dev.davidsoergel.com/trac/rtax/
   - Docs: http://dev.davidsoergel.com/trac/rtax/wiki/QuickStart
   - Forum: http://dev.davidsoergel.com/trac/rtax/discussion/1
   - Used by QIIME; ref: http://qiime.org/tutorials/rtax.html

2. **Ruby**
   - module name: ruby
   - version 2.3.0
   - Other versions: 2.0.0-p195
   - Programming language

3. **Sailfish**
   - module name: sailfish
   - version 0.10.0
   - Other versions: 0.9.2
   - Rapid Mapping-based Isoform Quantification from RNA-Seq Reads
   - http://www.cs.cmu.edu/~ckingsf/software/sailfish/
   - https://github.com/kingsfordgroup/sailfish

4. **Salmon**
   - module name: salmon
   - version 0.8.2
   - Other versions: 0.8.0
   - Highly-accurate & wicked fast transcript-level quantification from RNA-seq reads using lightweight alignments
   - https://github.com/COMBINE-lab/salmon

5. **samstat**
   - module name: samstat
   - version 1.08
   - Displaying sequence statistics for next generation sequencing
   - http://samstat.sourceforge.net/
1. **SAMtools**
   module name: `samtools`
   version 1.6
   Other versions: 0.1.16, 0.1.18, 0.1.19, 0.2.0-devel-29-July-2014, 1.0, 1.1, 1.2, 1.3, 1.3.1, 1.4, 1.5

   Utilities for manipulating alignments in the SAM format
   http://www.htslib.org/
   Man pages for: samtools, wgsim
   Support: http://www.htslib.org/support/#lists

   Depends on `htslib/1.6`
   Depends on `bcftools/1.6`

1. **SCaFoS (Selection, Concatenation and Fusion of Sequences)**
   module name: `scafos`
   version 1.2.5

   Phylogenetic analysis tool
   http://megasun.bch.umontreal.ca/Software/scafos/scafos.html
   Needs X11 support to run

   Depends on `perl`
   Depends on `tree-puzzle`

1. **segemehl**
   module name: `segemehl`
   version 0.2.0

   Short read mapping with gaps
   http://www.bioinf.uni-leipzig.de/Software/segemehl/

   Depends on `gcc`

1. **SeqAlto**
   module name: `seqalto`
   version 0.5-r123
   Other versions: 0.5-r117

   Fast and accurate read aligner for moderately long Illumina reads
   http://www.stanford.edu/group/wonglab/seqalto/
   Info: /group/bioinfo/apps/apps/seqalto-0.5-r123/README.txt

1. **seqbuster**
   module name: `seqbuster`
   version 44

   Small RNA analysis of deep sequencing data
   http://code.google.com/p/seqbuster/
   Documentation on web site
   Copy /group/bioinfo/apps/apps/seqbuster-44 to a private directory and work from it

   Depends on `BEDTools`
   Depends on `java`
   Depends on `R-bioconductor`
   Depends on `blast`
   Depends on `ViennaRNA`

1. **SeqPrep**
   module name: `SeqPrep`
Merger overlapping, paired end Illumina reads into a single read and/or trim adapters.
https://github.com/jstjohn/SeqPrep
Depends on gcc

1. seqtk
   module name: seqtk
   version 1.0-r82b
   Other versions: 1.0-r68e
   Parsed FASTA and FASTQ files
   https://github.com/ndaniel/seqtk
   Info: /group/bioinfo/apps/apps/seqtk-1.0-r82b/README.md
   Depends on gcc

1. Sequence Cleaner
   module name: seqclean
   version 2011-02-22
   Automated trimming and validation of ESTs or other DNA sequences
   http://compbio.dfci.harvard.edu/tgi/software/
   http://sourceforge.net/projects/seqclean/
   Depends on perl
   Depends on blast/2.2.26

1. SequenceServer
   module name: SequenceServer
   version 1.0.7
   Server with web front end for BLAST+
   http://www.sequenceserver.com/
   Documentation: http://www.sequenceserver.com/doc
   Forum: https://groups.google.com/forum/#!forum/sequenceserver
   Probably doesn't work
   Depends on ruby/2.3.0
   Depends on blast

1. SeqyClean
   module name: seqyclean
   version 1.3.13
   Clean Roche 454 and Illumina NGS data
   http://cores.ibest.uidaho.edu/software/seqyclean
   Depends on gcc

1. SGA String Graph Assembler
   module name: SGA
   version 0.10.12
   De novo sequence assembler using string graphs
   http://genome.cshlp.org/content/22/3/549
   https://github.com/jts/sga
   FAQ: https://github.com/jts/sga/wiki/FAQ
   Depends on pysam

1. ShellCheck
   module name: shellcheck
version 0.3.8
A shell script static analysis tool
https://github.com/koalaman/shellcheck
Depends on haskell-platform

1. SHORE (Short Read)
   module name: SHORE
   version 0.9.3
   Mapping and analysis pipeline for short read data produced on the Illumina platform.
   http://1001genomes.org/software/shore.html
   Info: /group/bioinfo/apps/apps/shore-0.9.3/share
   Depends on Boost
   Depends on gsl
   Depends on xz
   Depends on GenomeMapper

1. SHOREmap
   module name: SHOREmap
   version 3.4
   Other versions: 3.0
   Fast and accurate identification of causal mutations in plants
   http://bioinfo.mpipz.mpg.de/shoremap/index.html
   Depends on DISLIN
   Depends on xz
   Depends on SHORE

1. SHRiMP
   module name: SHRiMP
   version 2.2.3
   Other versions: 2.2.2
   SHort Read Mapping: aligning genomic reads against a target genome
   http://compbio.cs.toronto.edu/shrimp/
   Info in /group/bioinfo/apps/apps/SHRiMP-2.2.3/README
   Depends on gcc

1. SICER
   module name: SICER
   version 1.1
   A clustering approach for identification of enriched domains from histone modification ChIP-Seq data
   http://home.gwu.edu/~wpeng/Software.htm
   Forum: http://groups.google.com/group/sicer-users
   Depends on biopython/2.7.8

1. SignalP
   module name: signalp
   version 4.1c
   Other versions: 4.1b
   Predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms
   http://www.cbs.dtu.dk/services/SignalP/
   http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?signalp
   License: /group/bioinfo/apps/apps/signalp-4.1c/signalp-4.1.licence.txt
   Manual page: "man signalp"
Default MAX_ALLOWED_ENTRIES changed from 10,000 to 2,000,000

1. SiLiX
   module name: silix
   version 1.2.8
   ultra fast SIngle Linkage Clustering of Sequences
   http://lbbe.univ-lyon1.fr/Overview.html
   Info: man pages for silix and silixx

   Depends on gcc

1. SInC
   module name: SInC
   version 2014-01-20
   An accurate & fast error-model based simulator for SNPs, Indel & CNVs
   https://sourceforge.net/projects/sincsimulator/

   Depends on gcc

1. SMRT Analysis
   module name: smrtanalysis
   version 2.3.0.140936
   https://github.com/PacificBiosciences/SMRT-Analysis
   Genomic analysis from PacBio
   Info: https://github.com/PacificBiosciences/SMRT-Analysis/wiki


   To use interactively, first run "smrtshell" command.

   To use in batch bash job, first run "source $SMRT_ROOT/current/etc/setup.sh"

   This is not compatible with other software modules!

1. SNAP
   module name: snap
   version 0.15
   Other versions: 0.13.4, 0.13.11
   Scalable Nucleotide Alignment Program
   http://snap.cs.berkeley.edu/
   Info: http://snap.cs.berkeley.edu/#documentation
   Info: /group/bioinfo/apps/apps/snap-0.15/snap-0.15-manual.pdf
   Info: /group/bioinfo/apps/apps/snap-0.15/snap-0.15-quickstart.pdf

1. SNAP (Semi-HMM-based Nucleic Acid Parser)
   module name: SNAP
   version 2013-02-16
   Gene prediction tool.
   http://korflab.ucdavis.edu/software.html
   Info: /group/bioinfo/apps/apps/SNAP-2013-02-16/00README

   Depends on gcc

1. snpEff and snpSift
   module name: snpEff
   version 4.3p
   Other versions: 3.3F, 3.6, 4.0e, 4.1, 4.1d, 4.2, 4.3
Genetic variant annotation and effect prediction toolbox
http://snpeff.sourceforge.net/index.html

Start program with:
java -jar $snpEff_root/snpEff.jar ...  

Might need to change configuration options. Make copy of config file from /group/bioinfo/apps/apps/snpEff-4.3p/snpEff.config and modify as needed. Then, reference your modified config file as:
java -jar $snpEff_root/snpEff.jar -c path/to/modified/snpEff.config ...

See manual for more information

Also includes SnpSift
See: http://snpeff.sourceforge.net/SnpSift.html
Start program with:
java -jar $snpEff_root/SnpSift.jar ...

Also includes ClinEff version 1.0g:
See: http://www.dnaminer.com/clineff.html
Start program with:
java -jar $clinEff_root/ClinEff.jar ...

Depends on java

1. **SNPhylo**

   module name: **SNPhylo**
   version 20140701

   Pipeline to generate a phylogenetic tree from SNP data
http://chibba.pgml.uga.edu/snphylo/

   Depends on **PHYLIP**
   Depends on biopython/2.7.8
   Depends on R-bioconductor/3.1.2
   Depends on muscle

1. **SOAPaligner**

   module name: **SOAPaligner**
   version 2.21

   SOAP (Short Oligonucleotide Analysis Package)
   Short oligonucleotide alignment
http://soap.genomics.org.cn/soapaligner.html
   Info: http://soap.genomics.org.cn/soapaligner.html#commopt2

1. **SOAPdenovo**

   module name: **SOAPdenovo**
   version 240
   Other versions: 1.05

   SOAP (Short Oligonucleotide Analysis Package)
   Assemble Illumina GA short reads
http://soap.genomics.org.cn/soapdenovo.html
   Info: http://soap.genomics.org.cn/soapdenovo.html#comm2

1. **SOAPdenovo-Trans**

   module name: **SOAPdenovo-Trans**
   version 1.03

   SOAP (Short Oligonucleotide Analysis Package)
   De novo transcriptome assembler
http://soap.genomics.org.cn/SOAPdenovo-Trans.html
http://sourceforge.net/projects/soapdenovotrans/
Info: /group/bioinfo/apps/apps/SOAPdenovo-Trans-1.03/MANUAL
Example data: /group/bioinfo/apps/apps/SOAPdenovo-Trans-1.03/example

Depends on gcc

1. SOAPec
   module name: SOAPec
   version 2.03
   Other versions: 2.02
   Corrects sequencing errors based on kmer frequency spectrum (KFS)
   Info: /group/bioinfo/apps/apps/SOAPec_bin_v2.03/ReadMe
   Info: /group/bioinfo/apps/apps/SOAPec_bin_v2.03/NEWS

1. SOAPfuse
   module name: SOAPfuse
   version 1.26
   Genome-wide detection of fusion transcripts from paired-end RNA-Seq data
   http://soap.genomics.org.cn/soapfuse.html
   To run SOAPfuse; you need to make your own copy of the program and data files.
   Those files are located in /group/bioinfo/apps/apps/SOAPfuse-v1.26
   Read /group/bioinfo/apps/apps/SOAPfuse-v1.26/README.Purdue for more information
   You will need to download and install index files
   Depends on perl

1. SOAPfusion
   module name: SOAPfusion
   version 0.3
   Other versions: 0.2
   Genome-wide discovery of gene fusion events from RNA-seq data.
   http://soap.genomics.org.cn/SOAPfusion.html
   To run SOAPfusion; you probably need to make your own copy of the program and data files.
   Those files are located in /group/bioinfo/apps/apps/SOAPfusion-v0.3
   Read /group/bioinfo/apps/apps/SOAPfusion-v0.3/README.Purdue for more information
   You will need to download and install index files

1. SOAPsplice
   module name: SOAPsplice
   version 1.8
   SOAP (Short Oligonucleotide Analysis Package)
   Detection of splice junction sites from RNA-Seq
   http://soap.genomics.org.cn/soapsplice.html
   Info: http://soap.genomics.org.cn/soapsplice.html#commopt2

1. SOAPsv
   module name: SOAPsv
   version 1.02
   SOAP (Short Oligonucleotide Analysis Package)
   Detecting structural variation by whole genome de novo assembly
   http://soap.genomics.org.cn/SOAPsv.html
   Manual at /group/bioinfo/apps/apps/SOAPsv-1.02/SOAPSV-Pipeline-Doc.pdf

1. SolexaQA
module name: *SolexaQA*
version 1.1.13

Calculate quality data from Illumnia/Solexa FASTQ files
http://solexaqa.sourceforge.net/

Depends on *matrix2png*
Depends on *R-bioconductor*

1. **Some RHEL 5 shared libraries**
   module name: *rhel5libs*
   version 1.0

   This is a copy of the following shared libraries from RHEL5 RPMs that are bundled with RCAC RHEL 5 community clusters.
   - blas-3.0-38.el5
   - libgfortran-4.1.2-52.el5
   - mysql-5.0.95-1.el5_7.1
   - openssl-0.9.8e-22.el5
   - tk-8.4.13-5.el5_1.1
   - db4-4.3.29-10.el5_5.2
   - expat-1.95.8-11.el5_8

   Depends on *gcc*

1. **SortMeRNA**
   module name: *SortMeRNA*
   version 2.1b
   Other versions: 2.1

   Biological sequence analysis tool for filtering, mapping and OTU-picking NGS reads
   http://bioinfo.lifl.fr/RNA/sortmerna/

   Depends on *gcc*

1. **SourceTracker**
   module name: *sourcetracker*
   version 0.9.5

   Bayesian approach to estimating the proportion of a novel community that comes from a set of source environments
   http://sourceforge.net/projects/sourcetracker/

1. **SPAdes**
   module name: *spades*
   version 3.11.0
   Other versions: 2.2.1, 2.3.0, 3.0.0, 3.1.0, 3.1.1, 3.5.0, 3.5.0-p1, 3.5.0-p2, 3.6.0, 3.6.1, 3.6.2, 3.7.0, 3.8.0, 3.9.0, 3.9.1, 3.10.1

   Genome assembler for multicell and single-cell MDA bacteria
   http://cab.spbu.ru/software/spades/

   Depends on *gcc*
   Depends on *biopython/2.7.12*

1. **Sparc**
   module name: *Sparc*
   version 2016-04-22

   Sparsity-based consensus algorithm for long erroneous sequencing reads
   https://github.com/yechengxi/Sparc
This probably doesn't work

Depends on blasr
Depends on biopython/2.7.8

1. SparseAssembler
   module name: SparseAssembler
   version 2015-12-08
   Sparse k-mer Graph for Memory Efficient Genome Assembly
   https://sites.google.com/site/sparseassembler/
   https://sourceforge.net/projects/sparseassembler/

1. sparsehash
   module name: sparsehash
   version 2.0.2
   Other versions: 2.0.1, 2.0.2-intel
   An extremely memory-efficient hash_map implementation
   http://code.google.com/p/sparsehash/?redir=1
   Info: /group/bioinfo/apps/apps/sparsehash-2.0.2/share/doc/sparsehash-2.0.2
   Depends on gcc

1. SplAdder
   module name: spladder
   version Apr-09-2015
   Splice Adder: augments genome annotation using RNA-Seq alignment data.
   http://raetschlab.org/suppl/spladder
   This is the Python language version, not the MATLAB one.
   Depends on biopython/2.7.8

1. SpliceGrapher
   module name: splicegrapher
   version 0.2.4
   Other versions: 0.1.0, 0.2.0
   Create splice graphs from RNA-Seq data, guided by gene models and EST data
   http://splicegrapher.sourceforge.net/index.html
   Tutorial at http://splicegrapher.sourceforge.net/tutorial.html
   Depends on biopython/2.7.8

1. SQUID
   module name: squid
   version 1.9g
   C function library for sequence analysis
   http://eddylab.org/software.html
   Documentation on web site and man pages
   Depends on gcc

1. SRA Toolkit
   module name: sra-toolkit
   version 2.8.0
   Other versions: 2.1.6, 2.3.5, 2.4.2, 2.5.1, 2.5.2, 2.5.7
   Tools for using data in the INSDC Sequence Read Archives
Depends on perl

1. **SSPACE**
   module name: *SSPACE*
   version 3.0
   Other versions: 2.0

Scaffolding pre-assembled contigs using paired-read data
http://www.baseclear.com/bioinformatics-tools

Depends on gcc

1. **SSPACE-LongRead**
   module name: *SSPACE-LongRead*
   version 1-1

Scaffolding pre-assembled contigs using long reads

1. **Stacks**
   module name: *stacks*
   version 1.35
   Other versions: 0.998, 0.999, 0.9991, 0.9995, 0.9996, 0.9999, 0.99991, 0.99993, 0.99996, 0.99997, 1.01, 1.02, 1.04, 1.05, 1.06, 1.08, 1.09, 1.10, 1.12, 1.13, 1.17, 1.18, 1.19, 1.20, 1.21, 1.29, 1.34

Pipeline for building loci out of a set of short-read sequenced samples
http://catchenlab.life.illinois.edu/stacks/
FAQ: http://catchenlab.life.illinois.edu/stacks/faq.php

You will need to configure a MySQL database to view results.

Depends on gcc
Depends on perl
Depends on samtools/0.1.19

1. **Stampy**
   module name: *stampy*
   version 1.0.26
   Other versions: 1.0.22, 1.0.23

Mapping of short reads from Illumina sequencing machines onto a reference genome
http://www.well.ox.ac.uk/project-stampy
Info: /group/bioinfo/apps/apps/stampy-1.0.26/README.txt

Depends on biopython/2.7.8
Depends on bwa

1. **Stanford Network Analysis Platform (SNAP)**
   module name: *SNAP_library*
   version 1.10

C++ library for analysis and manipulation of large networks
http://snap.stanford.edu/index.html
Docs: http://snap.stanford.edu/snap/doc.html

Includes taxid2tsa.pl and taxid2wgs.pl for TSA & WGS data searches
Info in /group/bioinfo/apps/apps/Snap-1.10/README.txt
Examples in /group/bioinfo/apps/apps/Snap-1.10/examples
Tutorial in /group/bioinfo/apps/apps/Snap-1.10/tutorials
Hint: Make a copy of everything in /group/bioinfo/apps/apps/Snap-1.10
then work in your copy of the tutorials directory

Depends on gcc
Depends on gnuplot

1. **STAR**
   
   module name: **STAR**
   version 2.5.2b
   Other versions: 2.4.0j, 2.4.0k, 2.4.1d, 2.4.2a, 2.5.0c, 2.5.1b

   RNA-seq aligner
   
   https://github.com/alexdobin/STAR
   Info: /group/bioinfo/apps/apps/STAR-2.5.2b/README.md
   Info: /group/bioinfo/apps/apps/STAR-2.5.2b/RELEASEnotes.md
   Info: /group/bioinfo/apps/apps/STAR-2.5.2b/CHANGES.md

   Depends on gcc

1. **storm**
   
   module name: **storm**
   version 0.20

   Object Relational Mapper for the Python programming language.
   Python module
   
   https://launchpad.net/storm

   Depends on biopython/2.7.3

1. **StringTie**
   
   module name: **stringtie**
   version 1.3.3
   Other versions: 1.0.1, 1.0.2

   Transcript assembly and quantification for RNA-Seq
   
   http://ccb.jhu.edu/software/stringtie/
   Info: /group/bioinfo/apps/apps/stringtie-1.3.3/README

   Depends on gcc

1. **Structure**
   
   module name: **Structure**
   version 2.3.4

   Use multi-locus genotype data to investigate population structure
   
   http://pritchardlab.stanford.edu/structure.html
   Need X11 support to see graphics
   Use "structure-gui" to start GUI/X11 version
   Use "structure" for command line version

   Depends on java
   Depends on CLUMPP
   Depends on distruct

1. **SUMATRA and SUMACLUST**
   
   module name: **sumatra**
   version 1.0.00

   Fast and exact comparison and clustering of sequences
   
   http://metabarcoding.org/sumatra
   Manual: /group/bioinfo/apps/apps/sumatra-1.0.00/sumatra_sumaclust_user_manual.pdf
1. **SUPPA**
   module name: SUPPA
   version 2.2.0
   Study splicing across multiple conditions
   https://bitbucket.org/regulatorygenomicsupf/suppa
   Depends on biopython/3.5.1

1. **SWIG**
   module name: swig
   version 3.0.8
   SWIG is a compiler that integrates C and C++ with other languages
   http://www.swig.org
   Info: /group/bioinfo/apps/apps/swig-3.0.8/README
   Depends on gcc

1. **T-COFFEE**
   module name: T-COFFEE
   version 9.03r1318
   Multiple Sequence Alignment Tools
   http://www.tcoffee.org/Projects/tcoffee/
   Just includes the t_coffee program
   Depends on gcc
   Depends on perl

1. **T-lex**
   module name: tlex
   version 2.0
   Computational pipeline that detects presence and/or absence of annotated individual transposable elements
   License and download: http://petrov.stanford.edu/Tlex_scripts/
   Depends on fastagrep
   Depends on cross_match
   Depends on blat
   Depends on SHRIMP
   Depends on maq
   Depends on RepeatMasker

1. **tabix**
   module name: tabix
   version 0.2.6
   Generic indexer for TAB-delimited genome position files
   http://samtools.sourceforge.net/
   Info: http://samtools.sourceforge.net/tabix.shtml
   or using "man tabix"
   Depends on gcc

1. **Target Finder**
   module name: TargetFinder
   version 1.7
   Other versions: 1.6
Plant small RNA target prediction tool
https://github.com/carringtonlab/TargetFinder
Info: /group/bioinfo/apps/apps/TargetFinder-1.7/README
Depends on perl
Depends on fasta/35

1. **TASR pipeline**
   module name: **TASR**
   version 1.1

   TASR (Transposon Annotation using Small RNAs)
   http://sourceforge.net/projects/tasr-pipeline/
   Info: /group/bioinfo/apps/apps/TASR-1.1/README
   Info: /group/bioinfo/apps/apps/TASR-1.1/TASR_tutorial
   License: /group/bioinfo/apps/apps/TASR-1.1/LICENSE
   Citation: http://nar.oxfordjournals.org/content/early/2015/03/26/nar.gkv257.full

   Start program with command "TASR"

   Need to specify location of usearch7 program which is:
   -usearchv /group/bioinfo/apps/apps/usearch-7.0.1090/bin/usearch7

   Depends on perl
   Depends on blast
   Depends on bowtie2/2.2.3
   Depends on trf
   Depends on usearch7
   Depends on silix

1. **TASSEL (Trait Analysis by Association, Evolution and Linkage)**
   module name: **TASSEL**
   version 5.2.36
   Other versions: 5.2.9

   Association mapping of complex traits in diverse samples.
   http://www.maizegenetics.net/#tassel/c17q9
   Forum: http://groups.google.com/group/tassel
   Needs X11 support for GUI:
   https://www.rcac.purdue.edu/compute/carter/guide/#accounts_login_x11

   Depends on java/8

1. **TBB (Intel Threading Building Blocks)**
   module name: **tbb**
   version 2017_20161128oss

   C++ template library for task parallelism
   https://www.threadingbuildingblocks.org/
   CPPFLAGS=-I/group/bioinfo/apps/apps/tbb2017_20161128oss/include
   LDFLAGS=-L/group/bioinfo/apps/apps/tbb2017_20161128oss/lib/intel64/gcc4.7

   Depends on gcc

1. **tbl2asn**
   module name: **tbl2asn**
   version 25.3
   Other versions: 22.9, 24.9

   Automates the submission of sequence records to GenBank
   Info: /group/bioinfo/apps/apps/tbl2asn-25.3(tbl2asn.txt

   Also includes table2asn_GFF version 1.19.259
1. **Tcl/Tk (Tool Command Language)**
   module name: *tcl*
   version 8.6.1

   Dynamic programming language
   Includes Tk GUI widget platform
   http://www.tcl.tk/
   Manuals: http://www.tcl.tk/doc/

   Depends on gcc

1. **Tedna**
   module name: *tedna*
   version 1.2.2-mid-k91
   Other versions: 1.2.2, 1.2.2-fast-k91, 1.2.2-slow-k91

   Lightweight de novo transposable element assembler
   It assembles the transposable elements directly from the raw reads.
   https://urgi.versailles.inra.fr/Tools/Tedna
   Documentation: https://urgi.versailles.inra.fr/content/download/2962/25436/file/doc.pdf
   Documentation: /group/bioinfo/apps/apps/tedna-1.2.2-mid-k91/doc.pdf

   Built using C++/mid hash implementation and
   max k-mer size of 91

   Depends on gcc

1. **TeX Live**
   module name: *texlive*
   version 2013

   TeX document production system
   http://www.tug.org/texlive/

   Depends on gcc

1. **The GNU C Library (glibc)**
   module name: *glibc*
   https://www.gnu.org/software/libc/
   Core libraries for GNU/Linux systems

1. **tidyp**
   module name: *tidyp*
   version 1.04

   Validate HTML
   http://tidyp.com/

   Depends on gcc

1. **TMHMM**
   module name: *tmhmm*
   version 2.0c

   Prediction of transmembrane helices in proteins
   http://www.cbs.dtu.dk/services/TMHMM/
   http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?tmhmm
   License: /group/bioinfo/apps/apps/tmhmm-2.0c/tmhmm-2.0c.licence.txt
   Manual: /group/bioinfo/apps/apps/tmhmm-2.0c/TMHMM2.0.html

   Depends on perl
1. **TopHat**
   module name: *tophat*
   version 2.1.1
   Other versions: 1.3.2, 1.3.3, 1.4.0, 1.4.1, 2.0.0, 2.0.4, 2.0.5, 2.0.6, 2.0.7, 2.0.8, 2.0.9, 2.0.10, 2.0.11, 2.0.13, 2.0.14, 2.1.0

   Spliced read mapper for RNA-Seq
   http://ccb.jhu.edu/software/tophat/index.shtml
   Forum: https://groups.google.com/forum/#forum/tuxedo-tools-users

   According to TopHat release notes: "TopHat has entered a low maintenance, low support stage as it is now largely superseded by HISAT2 which provides the same core functionality (i.e. spliced alignment of RNA-Seq reads), in a more accurate and much more efficient way."

   Use "module load hisat2" to use that software

   Depends on *gcc*
   Depends on *zlib*
   Depends on *bowtie2/2.2.9*
   Depends on *bowtie*
   Depends on *blast*

1. **Tracer**
   module name: *Tracer*
   version 1.6

   Analyze trace files generated by Bayesian MCMC runs of BEAST, MrBayes, LAMARC, etc.
   http://tree.bio.ed.ac.uk/software/tracer/

   Need X11 support to run. Consult:
   https://www.rcac.purdue.edu/knowledge/halstead/accounts/all#accounts_login_x11

   Depends on *java*

1. **trans-ABySS**
   module name: *trans-abyss*
   version 1.5.5
   Other versions: 1.3.2, 1.5.1, 1.5.3

   de novo assembly of RNA-Seq data using ABysS
   http://www.bcgsc.ca/platform/bioinfo/software/trans-abyss
   https://github.com/bcgsc/transabyss

   Tutorial: /group/bioinfo/apps/apps/transabys-1.5.5/TUTORIAL.txt
   Support: http://groups.google.com/group/trans-abys

   Depends on *abyss/1.5.2-maxk64*
   Depends on *biopython/2.7.8*
   Depends on *blat/35*
   Depends on *bowtie2*
   Depends on *gmap*
   Depends on *samtools/1.3*

1. **TransDecoder**
   module name: *TransDecoder*
   version 3.0.1
   Other versions: 2.0.1, 2.1.0, 3.0.0-beta

   Find Coding Regions Within Transcripts
   http://transdecoder.github.io/

   Depends on *perl*
1. **TransposonPSI**  
   module name: *TransposonPSI*  
   version 08222010  
   Use PSI-Blast to Mine (Retro-)Transposon ORF Homologies  
   Depends on perl  
   Depends on blast/2.2.26

1. **Transrate**  
   module name: *transrate*  
   version 1.0.1  
   de-novo transcriptome assembly quality analysis  
   Mailing list: [https://groups.google.com/forum/#!forum/transrate-users](https://groups.google.com/forum/#!forum/transrate-users)  
   Bug tracker: [https://github.com/blahah/transrate/issues](https://github.com/blahah/transrate/issues)  
   Live chat: [https://gitter.im/blahah/transrate](https://gitter.im/blahah/transrate)  
   Depends on gcc

1. **Tree Puzzle**  
   module name: *tree-puzzle*  
   version 5.2  
   Maximum likelihood analysis for nucleotide, amino acid, and two-state data  
   Depends on gcc

1. **TreeFix**  
   module name: *TreeFix*  
   version 1.1.10  
   Statistically Informed Gene Tree Error Correction Using Species Trees  
   [https://www.cs.hmc.edu/~yjw/software/treefix/](https://www.cs.hmc.edu/~yjw/software/treefix/)  
   Depends on biopython/2.7.12

1. **TRF (Tandem Repeats Finder)**  
   module name: *trf*  
   version 4.09  
   Other versions: 4.04  
   Locate and display tandem repeats in DNA sequences  

1. **TrimAl**  
   module name: *trimal*  
   version 1.2r59  
   Other versions: 1.4.1  
   Removes spurious or poorly aligned regions from a multiple sequence alignment  
   Depends on gcc

1. **TrimGalore**  
   module name: *TrimGalore*  
   version 0.4.4
Adapter and quality trimming for FastQ files, with extra functionality for RRBS data
https://github.com/FelixKrueger/TrimGalore
Manuals in /group/bioinfo/apps/apps/TrimGalore-0.4.4/Docs

Depends on fastqc
Depends on cutadapt

1. **Trimmomatic**
   module name: trimmomatic
   version 0.36
   Other versions: 0.17, 0.20, 0.22, 0.30, 0.32, 0.33
   A flexible read trimming tool for Illumina NGS data
   Use "trimmomatic" to start the program
   Depends on java

1. **Trinity**
   module name: trinity
   version 2.4.0
   Other versions: 2.0.2, 2.0.3, 2.0.6, 2.1.0, 2.1.1, 2.2.0, 2.3.2, b20140127, b20140128, b20140316, b20140323, beta, r2012-01-25, r2012-03-17, r2012-04-27, r2012-05-18, r2012-06-08, r2012-10-05, r2013-02-15, r2013-02-16, r2013-02-25, r2013_08_14, r20131110, r20140413, r20140413p1, r20140717, r20150110beta
   RNA-Seq De novo Assembly
   https://github.com/trinityrnaseq/trinityrnaseq/wiki
   Forum: https://groups.google.com/forum/#!forum/trinityrnaseq-users
   Ref: /group/bioinfo/apps/apps/trinityrnaseq-2.4.0/Changelog.txt
   For parallel processing with PBS support; see:
   /group/bioinfo/apps/apps/trinityrnaseq-2.4.0/util/PBS/README
   If storing output in Lustre filesystem, configure output directory
   with "lfs setstripe -c 1" and submit job with "#PBS -l software=trinity"
   as described in http://www.rcac.purdue.edu/news/detail.cfm?NewsID=555
   Depends on gcc
   Depends on java
   Depends on perl
   Depends on samtools
   Depends on bowtie2
   Depends on R-bioconductor/3.1.0
   Depends on PASA
   Depends on rsem
   Depends on express

1. **Trinity (development version)**
   module name: trinity-dev
   version 2165
   Other versions: 1697
   RNA-Seq De novo Assembly
   http://trinityrnaseq.sourceforge.net/
   FAQ: http://trinityrnaseq.sourceforge.net/trinity_faq.html
   Ref: /group/bioinfo/apps/apps/trinityrnaseq_dev-2165/Release.Notes
   Ref: /group/bioinfo/apps/apps/trinityrnaseq_dev-2165/docs
   Depends on gcc
   Depends on java
   Depends on perl
   Depends on samtools
   Depends on bowtie
   Depends on R-bioconductor
1. **Trinotate**  
   module name: *trinotate*  
   version 3.0.1  
   Other versions: 2.0.2, r20131110  
   Transcriptome functional annotation and analysis  
   [http://trinotate.github.io/](http://trinotate.github.io/)  
   Prebuilt SQLite database with Uniprot (swissprot & uniref90) annotation information  
   at /group/bioinfo/apps/apps/Trinotate-3.0.1/data/Trinotate.sqlite  
   Copy that database to a directory owned by you and use it from there.  
   Depends on *trinity*  
   Depends on *blast*  
   Depends on *signalp*  
   Depends on *tmhmm*  
   Depends on *mammer*  

1. **tRNAscan-SE**  
   module name: *tRNAscan-SE*  
   version 1.3.1  
   Search for tRNA genes in genomic sequence  
   [http://lowelab.ucsc.edu/tRNAscan-SE/](http://lowelab.ucsc.edu/tRNAscan-SE/)  
   Info: man *tRNAscan-SE*  

1. **uclust**  
   module name: *uclust*  
   version 1.2.22q  
   Extreme high-speed sequence clustering, alignment and database search  
   [http://www.drive5.com/uclust/uclust_userguide_1_1_579.html](http://www.drive5.com/uclust/uclust_userguide_1_1_579.html)  

1. **UNAFold**  
   module name: *unafold*  
   version 3.8  
   RNA and DNA secondary structure prediction mainly by using thermodynamic methods  
   [http://mfold.rna.albany.edu/?q=node/60](http://mfold.rna.albany.edu/?q=node/60)  
   man pages available  
   docs: [http://mfold.rna.albany.edu/?q=unafold-man-pages](http://mfold.rna.albany.edu/?q=unafold-man-pages)  
   Depends on *gcc*  

1. **usearch**  
   module name: *usearch*  
   version 5.2.236  
   Other versions: 5.2.32, 6, 6.1.544, 7, 7.0.959, 7.0.1090, 8, 8.0.1517, 8.0.1623, 8.1.1861, 9, 9.0.2132, 9.2.64  
   High-throughput biological sequence analysis  
   License: [http://www.drive5.com/usearch/license32.html](http://www.drive5.com/usearch/license32.html)  
   32-bit version only; used by QIIME  

1. **USeq**  
   module name: *USeq*
version 8.4.0

Analysis of next generation signature sequencing data
http://useq.sourceforge.net/
Usage: http://useq.sourceforge.net/usage.html
Forum: https://lists.sourceforge.net/lists/listinfo/useq-users

Depends on java
Depends on R-bioconductor

1. util-linux  
   module name: util-linux  
   version 2.25.2

   Miscellaneous system utilities
   https://www.kernel.org/pub/linux/utils/util-linux/

   Depends on gcc

1. vcflib  
   module name: vcflib  
   version 03-Dec-2015

   Utilities and C++ library for parsing and manipulating VCF files
   https://github.com/ekg/vcflib
   Info: /group/bioinfo/apps/apps/vcflib-03-Dec-2015/README.md

   Depends on gcc
   Depends on perl
   Depends on R-bioconductor/3.1.0

1. VCFtools  
   module name: vcf tools  
   version 0.1.14
   Other versions: 0.1.9, 0.1.11, 0.1.12, 0.1.12a, 0.1.12b, 0.1.13

   Validate, merge, compare and calculate basic population genetic statistics from VCF files
   https://vcftools.github.io/index.html
   Info: https://vcftools.github.io/examples.html
   Info: man vcftools

   Depends on htslib

1. Velvet  
   module name: velvet  
   version 1.2.10
   Other versions: 1.2.03, 1.2.07, 1.2.10-kmer151

   Sequence assembler for very short reads
   http://www.ebi.ac.uk/~zerbino/velvet/
   Includes colorspace support
   Build parameters:
   - MAXKMERLENGTH=99
   - CATEGORIES=4
   - OPENMP=1

   Depends on gcc

1. VelvetOptimizer  
   module name: VelvetOptimiser  
   version 2.2.4

   Automatically optimize parameters for the Velvet assembler
   http://bioinformatics.net.au/software.velvetoptimiser.shtml
Manual: /group/bioinfo/apps/apps/VelvetOptimiser-2.2.4/README

Depends on velvet
Depends on perl

1. ViennaRNA
   module name: ViennaRNA
   version 2.3.5
   Other versions: 1.8.5, 2.0.7, 2.1.1, 2.2.4
   RNA Secondary Structure Prediction and Comparison
   http://www.tbi.univie.ac.at/~ronny/RNA/
   Documentation on web site and man pages
   Depends on perl/5.20.1
   Depends on graphviz

1. Vmatch
   module name: Vmatch
   version 2.3.0
   Other versions: 2.2.4
   Large scale sequencer analysis software
   http://www.vmatch.de/
   Depends on gcc
   Depends on perl

1. vsearch
   module name: vsearch
   version 2.4.3
   Other versions: 1.4.7, 2.0.2
   Metagenomics tool; free replacement for usearch
   https://github.com/torognes/vsearch
   or use "man vsearch"

1. WebLogo
   module name: weblogo
   version 2.8.2
   Generates sequence logos, which illustrate
   amino acid or nucleic acid multiple sequence alignment
   http://weblogo.berkeley.edu/
   Depends on perl

1. WIGGLER (a.k.a. align2rawsignal)
   module name: WIGGLER
   version 2.0
   Creates genome-wide raw or normalized signal tracks from aligned sequencing reads (BAM/tagAlign)
   https://code.google.com/p/align2rawsignal/wiki/README
   March 2013 update
   Includes scripts for generating binary unique mappability tracks for
   each contig/chromosome. Need to run them from somewhere else:
   cp -rp /group/bioinfo/apps/apps/align2rawsignal-2.0/map /some/where/else
   Depends on MCR
1. **Wise2 (includes genewise)**
   
   module name: wise
   
   version 2.4.1
   
   Other versions: 2.2.3-rc7
   
   Align proteins or protein HMMs to DNA
   
   http://www.ebi.ac.uk/~birney/wise2/
   
   Docs: /group/bioinfo/apps/apps/wise2.4.1/docs
   
   Run it online at http://www.ebi.ac.uk/Tools/psa/genewise/
   
   Depends on gcc

1. **XZ Utils**

   module name: xz

   version 5.2.3

   Other versions: 5.0.5

   Data compression (xz, lzma) software with high compression ratio

   http://tukaani.org/xz/

   Info: /group/bioinfo/apps/apps/xz-5.2.3/share/doc/xz/

   man pages available

   Depends on gcc

1. **YASS**

   module name: yass

   version 1.14

   Genomic similarity search tool for nucleic (DNA/RNA) sequences

   http://bioinfo.lifl.fr/yass/

   Use "man yass" for usage info

   Depends on gcc

1. **zlib**

   module name: zlib

   version 1.2.11

   Other versions: 1.2.7, 1.2.8

   Free compression library

   http://www.zlib.net/

   man -s3 zlib

   Depends on gcc

Updated: Mon Oct 2 16:21:24 2017