I. Basic Tools

- A. VPN access sitelicense.arizona.edu/vpn
- B. Tutorials
  - 1. UA computer based training: Unix Shell, Perl, etc. uits.arizona.edu/services/uacbt
  - 2. UC Riverside Manuals: R/Bioconductor, Galaxy, Emboss manuals.bioinformatics.ucr.edu/home
- C. Literature searches
  - 1. PubMed www.ncbi.nlm.nih.gov/pubmed
  - 2. Web of Knowledge (need to be on UA network) apps.webofknowledge.com
- D. Text Editors
  - 1. TextWrangler (Mac) www.barebones.com/products/textwrangler
  - 2. Notepad++ (Win) http://notepad-plus-plus.org/
  - 3. Komodo (Win, Mac, Linux) www.activestate.com/komodo-edit
  - 4. VIM/nano/pico(/gedit) (server-side)

II. Databases

- A. 2012 Nucleic Acids Research Database Summary (1,380 examples) www.oxfordjournals.org/nar/database/a
- B. National Center for Biotechnology Information (NCBI) www.ncbi.nlm.nih.gov/guide/all/#databases
  - 1. Sequence databases
    - a) Nucleotide (nt) www.ncbi.nlm.nih.gov/nuccore
    - b) Non-redundant Protein (nr) www.ncbi.nlm.nih.gov/protein
    - c) Reference Sequence (refseq) www.ncbi.nlm.nih.gov/RefSeq
  - 2. Other database examples
    - a) Conserved Domains (CDD) www.ncbi.nlm.nih.gov/cdd
    - b) Cluster of Orthologous Groups (COG) www.ncbi.nlm.nih.gov/COG
    - c) dbSNP Short Genetic Variations www.ncbi.nlm.nih.gov/snp
    - d) Genome www.ncbi.nlm.nih.gov/sites/genome
    - e) Taxonomy www.ncbi.nlm.nih.gov/taxonomy
- C. Other general databases
  - 1. European Bioinformatics Institute (EBI) www.ebi.ac.uk/Databases
    - a) UniProt, protein sequences and functional information
    - b) InterPro, protein sequence analysis and classification
  - 2. Pfam, collection of protein families pfam.sanger.ac.uk
  - 3. Kyoto Encyclopedia of Genes and Genomes (KEGG) www.genome.jp/kegg

4. The Gene Ontology www.geneontology.org
5. BioCyc, pathway/genome database biocyc.org
6. Genomes OnLine Database (GOLD), genomics metatdata www.genomesonline.org

D. Examples of specialized databases
  1. JGI genome portal genome.jgi-psf.org
  2. UCSC genome browser genome.ucsc.edu
  3. Ensemble, vertebrate centric uswest.ensembl.org
  4. The Arabidopsis Information Resource (TAIR) www.arabidopsis.org
  5. Ribosomal Database Project rdp.cme.msu.edu

III. Annotation and Analysis

A. Bulk Annotation
  1. BLAST (Web and stand-alone) blast.ncbi.nlm.nih.gov
  2. Blast2GO (Win, Mac, Linux) www.blast2go.org
  3. InterProScan (Web and stand-alone) www.ebi.ac.uk/Tools/pfa/iprscan
  4. PfamScan (Web and stand-alone) pfam.sanger.ac.uk/search
  5. KAAS, KEGG automatic annotation server (Web) www.genome.jp/tools/kaas
  6. RAST, Rapid Annotation using Subsystem Technology, prokaryote-centric (Web) rast.nmpdr.org
  7. MG-RAST, metagenomics analysis server (Web) metagenomics.anl.gov

B. Other Resources
  1. Galaxy, a web-based platform for computational biomedical research main.g2.bx.psu.edu
  2. Emboss, software for molecular biology (Mac, Linux) emboss.sourceforge.net
  3. ExPASy Bioinformatics Resource Portal (Web) expasy.org
  4. CBS Prediction Servers (Web and stand-alone) www.cbs.dtu.dk/services
  5. List of Population Genetics/Genomics tools softlinks.amnh.org/microsatellites
  6. List of Molecular Evolution tools softlinks.amnh.org/molevol

C. Annotation Visualization
  1. Cytoscape (Win, Mac, Linux) Network analysis and visualization www.cytoscape.org
  2. Artemis annotation browser (Win, Mac, Linux) www.sanger.ac.uk/resources/software/artemis
  3. Argo annotation browser (Win, Mac, Linux) www.broadinstitute.org/annotation/argo

4. GBrowse, Generic Genome Browser gmod.org/wiki/GBrowse#Logo
5. JBrowse (Java) Genome Browser jbrowse.org

IV. **Phylogenetics**

A. Alignments
  1. Comprehensive list softlinks.amnh.org/alignment
  2. Mafft (Web, Win, Mac, Linux) mafft.cbrc.jp/alignment/server
  3. Others interesting alignment programs
     a) Opal, IGERT fellow Travis Wheeler opal.cs.arizona.edu
     b) SATe, iterative sequence alignment and tree building phylo.bio.ku.edu/software/sate/sate.html
     c) RevTrans, uses virtual translation to align DNA sequences www.cbs.dtu.dk/services/RevTrans
  4. Genome alignment
     a) Mauve gel.ahabs.wisc.edu/mauve
     b) MUMmer mummer.sourceforge.net
  5. Alignment Editors/Visualizers
     a) SeaView (Win, Mac, Linux) pbil.univ-lyon1.fr/software/seaview
     b) BioEdit (Win) www.mbio.ncsu.edu/bioedit/bioedit
     c) GeneDoc (Win) www.nrbsc.org/gfx/genedoc

B. Tree Building
  1. Comprehensive list softlinks.amnh.org/phylogeny
  2. Felsenstein’s List evolution.genetics.washington.edu/phylip/software.html
  3. Maximum likelihood
     a) RAxML(Cipres, ICE, stand-alone) www.exelixis-lab.org
     b) GARLI (Cipres, ICE, stand-alone) www.bio.utexas.edu/faculty/antisense/garli/Garli.html
  4. Bayesian
     a) MrBayes (Cipres, stand-alone) mrbayes.sourceforge.net
     b) BEAST (Win, Mac, Linux) beast.bio.ed.ac.uk
  5. Tree viewers
     a) FigTree (Win, Mac, Linux) tree.bio.ed.ac.uk/software/figtree
     b) Dendroscope (Win, Mac, Linux) ab.inf.uni-tuebingen.de/software/dendroscope
     c) PhyloWidget (Web) www.phylowidget.org/full

C. Other helpful phylogenetic programs

        1. PAML, used to test hypotheses and estimate parameters (Win, Mac, Linux) abacus.gene.ucl.ac.uk/software/paml
        2. Gblocks, automated trimming of sequence alignments (Web and stand-alone) molevol.cmima.csic.es/castresana/Gblocks.html
        3. ARB, for ribosomal sequences www.arb-home.de
    D. Online servers
        1. Cipres Portal www.phylo.org/sub_sections/portal
        2. Phylogeny.fr, Phylogenetic analysis for the non-specialist www.phylogeny.fr

V. High Throughput

A. SEQanswers, the next generation sequencing community seqanswers.com

B. Software Suites
   1. SOAP, Short Oligonucleotide Analysis Package soap.genomics.org.cn/about.html
   2. 454 GS Data Analysis Software, a.k.a. Newbler (ICE) 454.com/products/analysis-software/index.asp
   3. CLC Genomics Workbench ($$$ but trial available) www.clcbio.com/index.php?id=1240

C. Quality Control (filter and trim genomic and metagenomic sequence data)
   1. PRINSEQ (Web and stand-alone) prinseq.sourceforge.net/index.html
   2. FastQC (Java) www.bioinformatics.bbsrc.ac.uk/projects/fastqc

D. *De novo* Assembly
   1. ABySS (ICE) www.bcgsc.ca/platform/bioinfo/software/abyss plus Trans-
   2. ABySS for transcriptomes
   3. Velvet (ICE) www.ebi.ac.uk/~zerbino/velvet plus Oases for transciptomes
   4. Mira (ICE) chevreux.org/projects_mira.html
   5. Trinity, transciptomes trinityrnaseq.sourceforge.net

E. Sequence Clustering
   1. SEED (Win, Linux, Mac) manuals.bioinformatics.ucr.edu/home/seed
   2. CD-HIT (Web and stand-alone) weizhong-lab.ucsd.edu/cd-hit

F. Mapping (resequencing, SNP and expression analysis)
   1. Bowtie (ICE) bowtie-bio.sourceforge.net/index.shtml
      a) TopHat, spliced read mapper for RNA-Seq tophat.cbcb.umd.edu
      b) Cufflinks, Transcript assembly and expression analysis for RNA-Seq cufflinks.cbcb.umd.edu
   2. Sanger Institute Software www.sanger.ac.uk/resources/software
      a) SSAHAsnp, a polymorphism detection tool

   b) SMALT, DNA sequence mapper

3. SeqMonk, visualize and analyze mapped sequence data
   www.bioinformatics.bbsrc.ac.uk/projects/seqmonk
4. BWA, Burrows-Wheeler Aligner bio-bwa.sourceforge.net