**Data processing**
- raw 454 sequence data
  - sort by tag and primer (Ruby scripts)
  - screen for indel errors (emelineV2.5.bf, batchAlignMPI.bf)
  - merge by subject and annotate sample dates (merge_alignments.py)
  - sample with replacement (random_sampler.py)
  - multiple sequence alignment (MUSCLE, Se-Al)
  - convert to NEXUS, censor stop codons (convertToNexus.bf)

**Phylogenetic reconstruction**
- convert to BEAST XML
  - parallel execution of BEAST (mpibeast.py)
  - check MCMC convergence (R coda)
  - extract trees as Newick strings (parseTreeLog.py)
  - codon ancestral reconstruction (fit_codon_model.bf, fit_codon_model.ibf, _mpiFitCodon.bf)
  - identify indel polymorphisms (scanForIndels.py)
  - indel ancestral reconstruction (fit_indels_v2.bf)

**Analysis and visualization**
- draw 2D histograms (v3trends.R)
- draw trees as PostScript (viz_g2p_shifts.py)
- compute MCT and MARCS (map_g2p_shifts.py)
- HIV coreceptor usage prediction
  - map mutations to tree (map_in_time.bf, batch_MPI_map.bf)
- find MRCA of CXCR4 variants (findMRCAofX4.bf)
- extract times of MRCAs (getTimesX4MRCA.py)