Things I need to get:

Filter all data to some reasonable error threshold, e.g. 1e-3 or 1e-2 for highest of mapping/sequencing error rate (can alternatively use site-specific error rates…not sure how to make this useful since map Q is per read not per site)

Calculate ref and alt read counts (passing this threshold) in all 3 reference populations

Calculate ref and alt read counts for each diploid individual included in the sample

Calculate map distances between each marker – some kind of smoothing. NOTE: for each new chromosome, it doesn’t matter the 1st value (ignored)

Divide individuals by population (ID go into which pops) and make input files

Let it calculate number of gens since admixture, but can cap reasonable limits, up to 2 generations per year is maximum 130 generations, and a conservative minimum is just 10 generations. A good guess for 2014 samples would be 80 generations.

2 pulses: European mixture (could be very old – 500 generations) and then African admixture (see caps above)

500 colonies ~ Ne = 1000 There is an equation to calculate Ne from # colonies and avg # inseminations per queen (Wright 1933 inbreeding and homozygosis on m/f inequality effects I think; Kerr 1967 multiple alleles and genetic load in bees)