SLiM Projects

Houston, we have a problem subproject

TYPE THIS UP ALONG WITH A ROADMAP AND SEND TO GIDEON ON MONDAY

- figures
- next steps for sims
  - two phases: debugging & automating
    - debugging: make sure structure of code "correct" somehow
    - automating: for a given set of parameters
      - make sure coeff of var of abund is low for both species
      - make sure spatial structure occurs in both species (which stat?)
      - make sure genome long enough wrt to recomb rate
      - record associated stats in metadata with each run
  - next steps for downstream analysis:
    - make sure ld thinning is not too strong and not too weak (check ld threshold)
    - get scale of absolute divergence to see if the species is reprod isolated at large enough dists

Simulate n replicated data sets (with $n \leq 1000$) for:

- four combinations of dispersal distances (short/long)
  - neutral scenario (no interaction)
  - four combinations of selection strengths (weak/strong)
    - four combinations of architecture (monogenic/oligogenic)

That's $4n$ neutral sims + $128n$ non-neutral sims. Assuming each sim takes ten minutes implies a total of $2.2n$ hours if run in serial.

With these data apply the nuismerian t-test approach and:

- compute distribution of type-1 error rates for each neutral scenario
- compute distributions of type-1 and type-2 error rates for each non-neutral scenario

Make four nested four panel figures for each error rate where the four plots correspond to the monogenic/oligogenic combinations, the exterior panels for each plot correspond to weak/strong coevolution and the plots in each panel display the estimated error rate as bivariate function of dispersal distances.

Conclude by describing scenarios where the t-test approach fails due to spurious interspecific correlations generated by spatially autocorrelated allele frequencies. Identify scenarios where the method works. Compare with Dybdal-Jenkins-Nuismer ppr.

Simulation development considerations:

- 1k rep limit
- Tests to figure out if sim doing what we want

  - want to be seeing spatial structure in both host and parasite
    * can check this with neutral sims (but it seems p obvious so i dont think this is necessary)
  - while making test for spatial hom abund densities, think about future app to study sp het abun densities
    * check to see if spatial var in trait value covaries with abund densitie in each species
    * what is the coef of var for abund density?

- is genome long enough wrt to recomb rate?
  - also consider if ld filtering threshold is too conservative
    * distance in bp's from selected site: how does that affect interspp sp corr?
    * want to make sure ld thinning is not too strong and not too weak

- might be helpful to get scale of absolute divergence to see if the species is reprod isolated at large enough dists

- consider different boundary conditions if we don't see a strong as pattern as we'd like
  - with absorbing boundaries

Genomic signal of coevolution in continuous space **subproject**

Find statistics with high power and low false-positive rates:

- Do community detection algorithms detect larger communities for the causal loci than for spuriously correlated neutral loci?
- Does analyzing the ild matrix at different resolutions of spatial discretization tell us anything?
- If we make parasites move to host locations for infection, can this spatial paring be used to obtain a "cleaner" signal than spatial discretization methods?
- Do the cca-spatial whitening stuff and see what pops out.

Potential third subproject ***Method to identify coev loci***

Other idea

- spatiotemporal correlations of host-parasite pop densities

Friendly reviewers

- louis zaman
- anurag agrawral (for intro at least)
- scott
- fields? (maybe better as non-friendly reviewer)
- debarre?
- send intro to coev gp and check if they want to coauthor (send slack message to gideon to review)