

Project Proposal: Detecting tracts of introgression

Problem: Locating segments of the genome that have recently moved from one species into a close relative.

Systems that would be ideal: An ideal system would have an area of recent secondary contact after a period of spatial isolation (e.g., a species that recently invaded an area where a close relative is established). Introgression should then be recent, meaning tracts of introgressed DNA will be large. Inference would be easiest if other isolated populations persist elsewhere. E.g.:



There are species of *Drosophila*, *Anopheles*, and *Arabidopsis* that almost certainly satisfy these criteria.

Project steps:

- 1) Identify fixed differences/allele frequency differences between species using isolated populations.
 - i. Align sp1 (isolated) w/ sp2 (isolated)
 - ii. Pare down to non-identical regions
 - iii. Compare to other w/in species samples to see if diffs are fixed.
 - iv. Create dictionary of alleles & positions
- 2) Simulate introgression.
 - i. Cut chunks from sp1 and put them in sp2 background
- 3) Scan genomes of sp1 & sp2 in same location & look for alleles that belong to the 'wrong' species.
- 4) Looking for introgression using allele frequency differences would require more work. Probably set up a probability threshold for e/ SNP