Dsim DNA- Ethanol Project-analysis plan

Project Leader Sarah Signor (written by AMM)

1. Dsim DNA collected
2. Dsim DNA samples transferred from USC to Rapid Genomics for library construction (100 bp reads, PE)
   1. Rapid Genomics QC- assess conc of DNA and QC gel of random samples to check quality
   2. Libraries quantified with Kappa kit
3. Completed libraries sent to USC for sequencing
4. Transfer DNA files from USC to UF HPC (rlts/McIntyre/\_\_\_\_\_\_\_\_\_\_\_\_\_)
5. For HISeq data, concatenate into single files
6. QC – copy from protocols discuss any potential issues
   1. Remove duplicates before aligning?
7. Align using GATK protocol- check BWA mem?
   1. This is sim- how is frankensim?? Adolfatto’s new w501 reference??? Do we need a new genome??? Talk to corbin?
   2. Focus on exons only or try to assemble whole genome?
8. Call snps/indels
9. Population genetic characteristics? David Lowrie??
10. Create vcf files- go to ASE pipeline