

Scott's biofuel genomes *de novo* assembly

Gang Fang

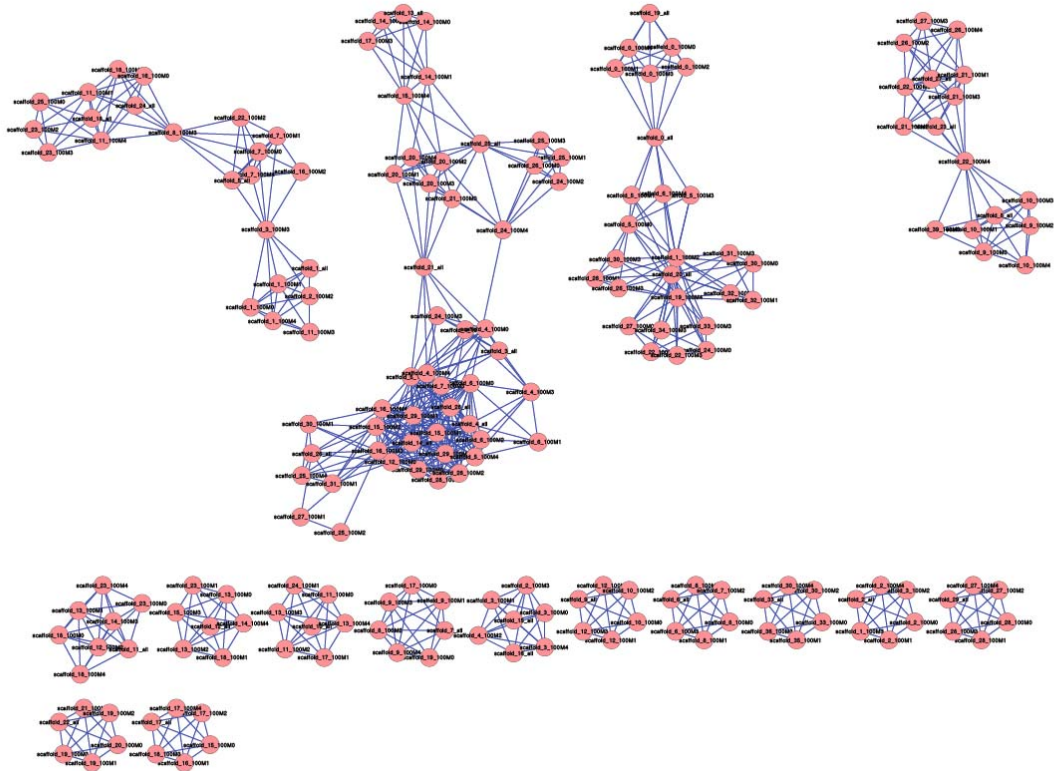
7 Fungal genomes and 2 *E. coli* genomes

- illumina fragment reads and jumping reads (3k)
- PacBio reads
- Fragment reads: Fungal genome 100bp; *E. coli* 75 bp
- Software: allpaths-lg, Velvet, SOAP and celera

allpaths-Ig assembly

Problem: Ribosomal RNAs and telomere sequences are not found in the assemblies, given ~200x genome coverage

- 1) sub-sampling of jumping reads yields different assemblies
- 2) Chimeric jumping read libraries?



Example:

