CNEE\_liftover\_beds/ includes bed files for each halictid species with the CNEE regions lifted over from the NMEL coordinates used as a reference when identifying CNEEs. CNEE\_gffs/ is the same information in gff format (which is what my analysis pipeline uses. CNEE\_fastas/ is fasta files of CNEE sequences for each species. cnees.index is an index file listing all of the sequences in each CNEE. cne\_proximity.txt indicates which gene, if any, the CNEE has been associated with. Note that this gene association is in NMEL space.

halictid\_cnee\_ortho.index is the index file produced and used by the analysis pipeline. Note that there is some filtering done by the first step in the pipeline so there are not quite as many loci present as in the raw cnees.index file.

halictid\_cnee\_fastas.tar.gz contains fasta files of each CNEE. There are over 170,000 of these. cnee\_halictid\_fsa.tar.gz is the same sequences aligned with FSA. And finally, cnee\_halictid\_fsa\_trimal.tar.gz contains the alignments filtered with TrimAl. These are the alignments used for analyses