**Databases and Results.**

**Databases:**

**Genomes available:**

*/scratch/projects/illorens/genome\_annotation*

**Blast databases for blasting against particular genomes:**

*/scratch/projects/illorens/data/blast\_db*

**Proteome databases:**

**Diamond proteins used in braker2:**

*/scratch/projects/illorens/data/proteomes/diamond\_proteins/proteins.fasta*

**BUSCO proteins:**

*/scratch/projects/illorens/data/proteomes/plants/\**

**Proteomes used for orthofinder:**

*/scratch/projects/illorens/data/proteomes/proteomes\_orthofinder/\**

**RNAseq results:**

*“/projects/ensa/plants/RNAseq/”* Within this folder there are:

4 species folders:

Alnglu, Begfuc, Casgla and Datglo containing RNAseq reads from those species.

*“*ncbi/” With raw reads downloaded from ncbi.

“raw\_data”: The RNAseq data generated by Jordan for the new species

“rnaseq\_course”:Raw data used for the RNAseq course.

**Results:**

**BUSCO output:**

*/scratch/projects/illorens/phylogenomics/[Species\_key]*

**iqtree\_out (note: These results were preliminary, and I recommend to run iqtree on a new list of proteins. I am leaving this folder as reference):**

*/scratch/projects/illorens/phylogenomics/out/iqtree\_out/\**

**Genome annotation results:**

*/scratch/projects/illorens/genome\_annotation/*

The folder contains two subfolders “data” and “out”.

“data/” contains the following folders and files:

“genome\_size/”: Each file contains the genome size for each scaffold/chromosome for each genome. These genome sizes can be generated using the scripts provided (“rbb.irl.p1.sh or rbb.irl.p2.sh”).

“proteins.fasta/”: Diamond protein sequences for braker2

“out/” contains the following folders and files:

“braker\_out/”: Outputs for BRAKER2 for each species”.

**Masked genomes:**

*“/projects/ensa/plants/final\_assemblies\_080920*/\*”

*“/projects/ensa/plants/ivan\_genomes*/\*”-These are genomes I downloaded from SRA and that I copied from Bruno’s folder.