README for the *Callicarpa americana* Genome Assembly, Annotation and Data Files

Data Files:

Genome Assembly:

car_asm.fa - pseudomolecules and unanchored scaffolds

Genome Annotation – High confidence gene models:

car.hc_gene_models.cdna.fa
Transcript sequences (cDNA) of high confidence gene models

car.hc_gene_models.cds.fa
Coding sequences (CDS) of high confidence gene models

car.hc_gene_models.pep.fa
Protein sequences of high confidence gene models

car.hc_gene_models.gff3
High confidence gene models annotation in GFF3 format

Genome Annotation – High confidence representative gene models:

High confidence representative gene models are a subset of the high confidence gene model set. Each representative gene model is the isoform with the longest CDS at each locus.

car.hc_gene_models.repr.gene_model.list.txt
List of representative high confidence gene model ids

car.hc_gene_models.repr.gff3
Representative high confidence gene models annotation in GFF3 format

car.hc_gene_models.repr.pep.fa
Protein sequences of the representative high confidence gene models

Genome Annotation – Working gene models:

The set of working gene models contains all loci and isoforms from the annotation pipeline and may include artifacts such as partial gene models.

car.working_models.cdna.fa
Transcript sequences (cDNA) of working gene models

car.working_models.cds.fa
Coding sequences (CDS) of working gene models

car.working_models.pep.fa
Protein sequences of working gene models

car.working_models.gff3
Working gene models annotation in GFF3 format

car.working_models.func_anno.txt
Functional annotation for the working gene models (tab delimited text file)

Gene Expression Matrix

A matrix containing the high confidence gene model transcript abundance values (TPM) for each library of the tissue RNA-seq atlas.

Callicarpa_RNA-Seq_TPM_expression_matrix.txt (tab delimited text file)

Callicarpa_RNA-Seq_TPM_expression_matrix.xlsx (Excel file)

BUSCO results

Output files from the BUSCO run on the final assembly:

full_table_car_final_asm_busco.tsv
missing_busco_list_car_final_asm_busco.tsv
short_summary_car_final_asm_busco.txt