README for the *Mitragyna speciosa* (Kratom) Genome Assembly, Annotation and Data

**Data Files:**

For the manuscript, gene and transcript IDs were formalized as follows:

Gene names:

“g” stands for gene

Transcript names:

Transcripts have the same format as gene names where isoforms are followed by ".t1", ".t2"........

**Genome Assembly:**

kratom\_final\_assembly.fa

Final genome assembly

**Genome Annotation – High confidence gene models:**

kratom.hc\_models.cdna.fa

Transcript sequences (cDNA) of high confidence gene models

kratom.hc\_models.cds.fa

Coding sequences (CDS) of high confidence gene models

kratom.hc\_models.pep.fa

Protein sequences of high confidence gene models

kratom.gene\_models.hc.gff3

High confidence gene models annotation in GFF3 format

**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.

kratom.working\_models.cdna.fa

Transcript sequences (cDNA) of working gene models

kratom.working\_models.cds.fa

Coding sequences (CDS) of working gene models

kratom.working\_models.pep.fa

Protein sequences of working gene models

kratom.working\_models.gff3

Working gene models annotation in GFF3 format