

Gryllus bimaculatus

Most of the available insect genomes are from **Holometabolan** insects, which are insects that undergo complete metamorphosis like beetles and flies. In contrast, the **Hemimetabolans**, which undergo gradual metamorphosis, and represent the basal branches of the insect tree, have very few genomic resources available (**Figure 1**).

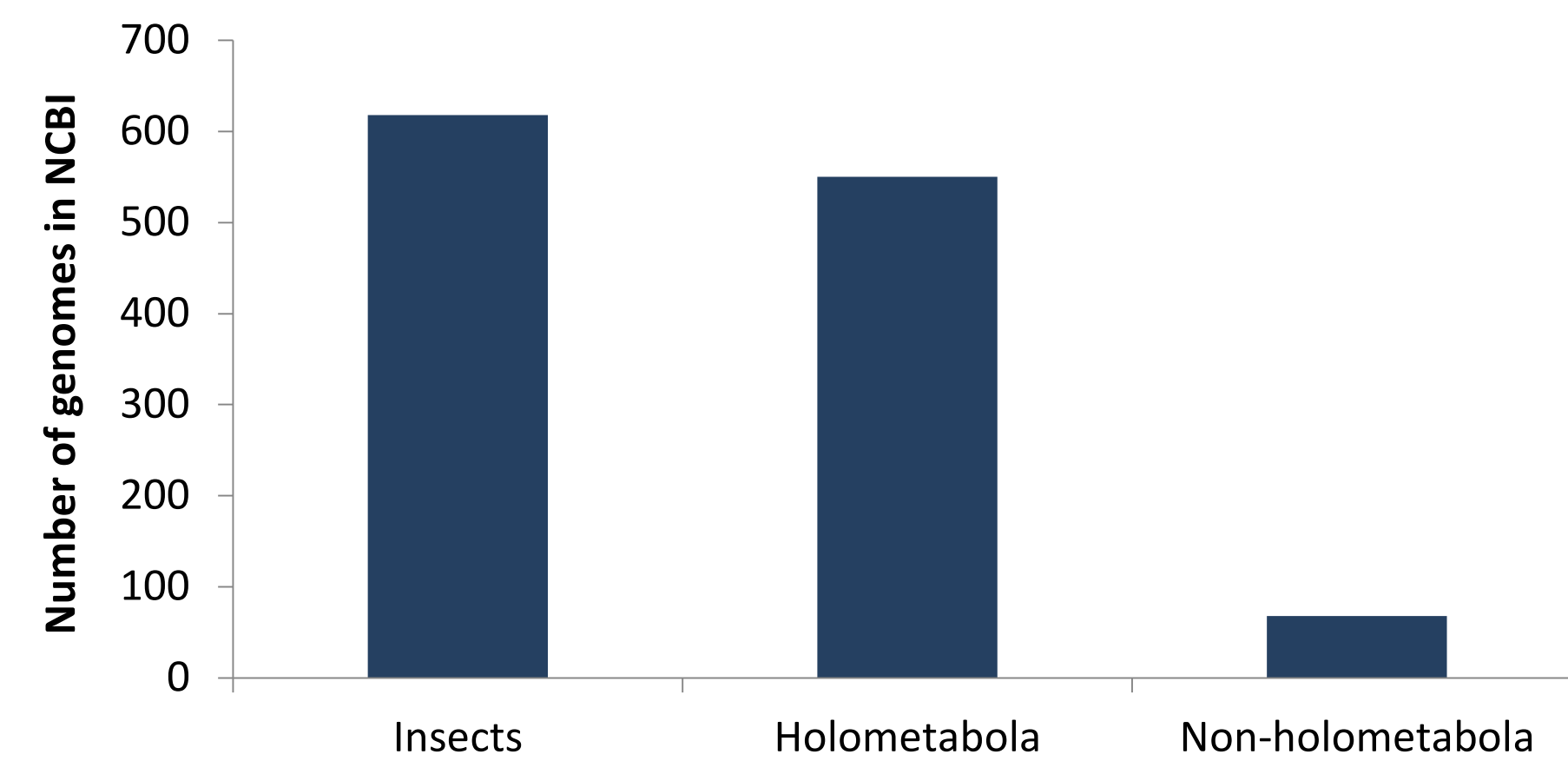


Figure 1: Number of genomes listed in NCBI for all insects, holometabolans (“Endopterygota”) and non-holometabolans (insect genomes - holometabolans).

There is only one available annotated genome for the Orthopteran order, that of *Locusta migratoria*, which is the largest insect genome sequenced until this date with ~7Gb.

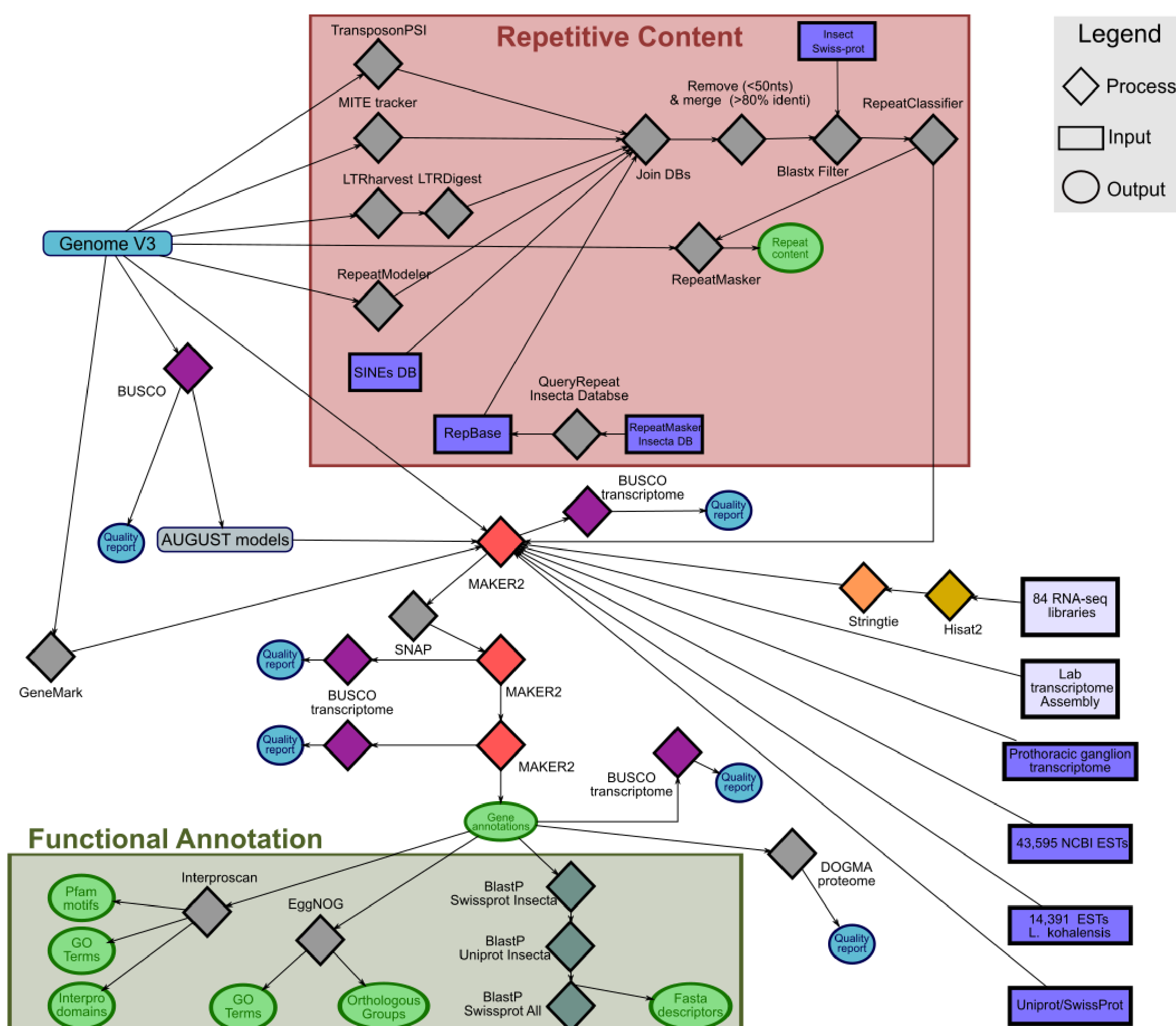


Figure 2: Photo of an adult female of *Gryllus bimaculatus*.

Gryllus bimaculatus

- Order: Orthoptera.
- Aka. two-spotted cricket.
- Hemimetabolan insect.
- Experimental model for functional genetics, development and evolution.

Genome annotation Pipeline



Results

Assembly Stats

	Gbi	Lmi	Lko	Bger	Zna
Year	2019	2014	2019	2018	2014
Scaffolds	47,877	564,328	148,784	24,820	31,662
Len. (Gb)	1.66	6.91	1.6	2.04	0.49
Avg.scaff. (Kb)	34.630	12.24	10.72	82.08	15.32
N50 (Kb)	6,287	325	583	1,056	751
N's (Kb)	56,571	1,094,108	31,618	326,821	20,751
L50	71	3,607	753	576	194

Table 1: Descriptive statistics of our newly assembled *Gryllus bimaculatus* genome (Gbi) in comparison with recently published hemimetabolan genomes of *Locusta migratoria* (Lmi), *Blattella germanica* (Bger) and *Zootermopsis nevadensis* (Zna).

Assembly Completeness

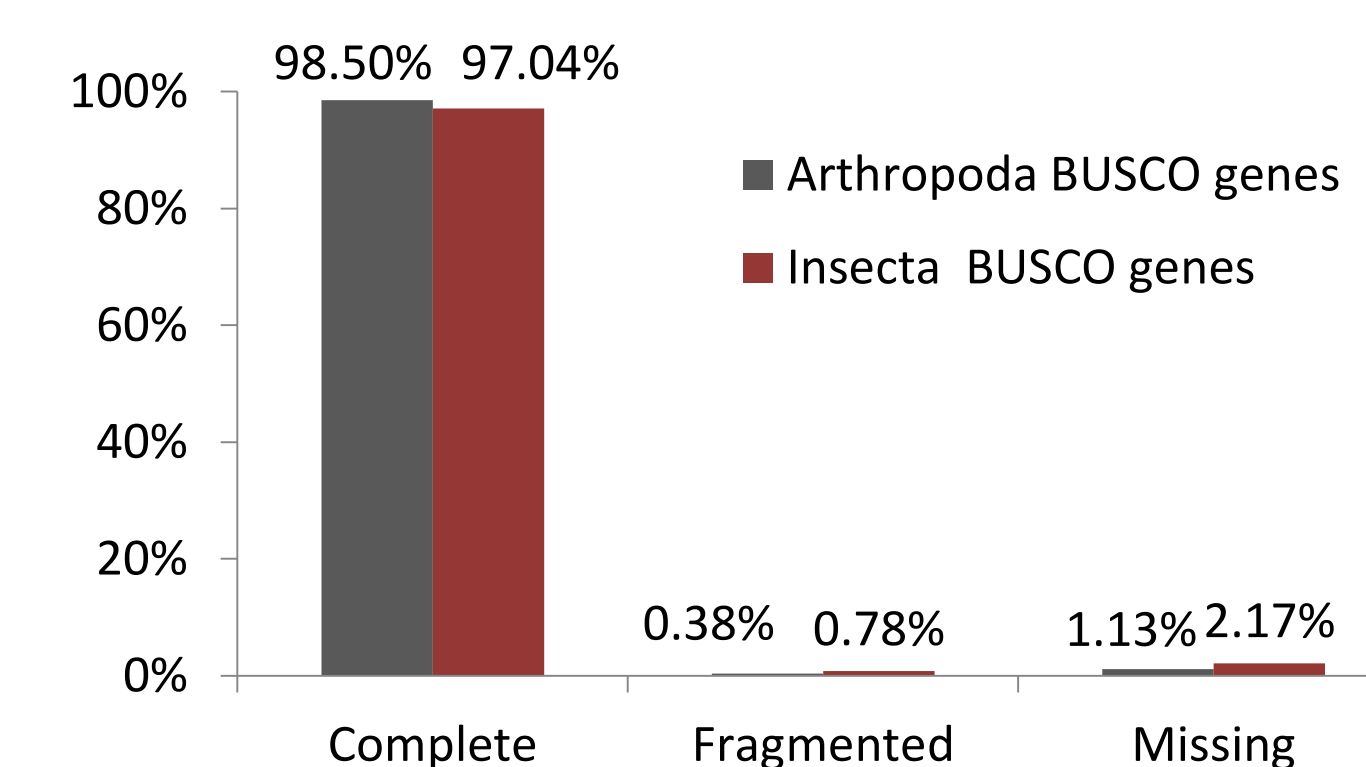


Figure 3: Percentage of complete, fragmented and missing BUSCO genes in the *G. bimaculatus* genome assembly. Analysis performed at arthropod level (1,066 genes) and insect level (1,658 genes).

Transposable Elements

Represent the 34% of the Genome

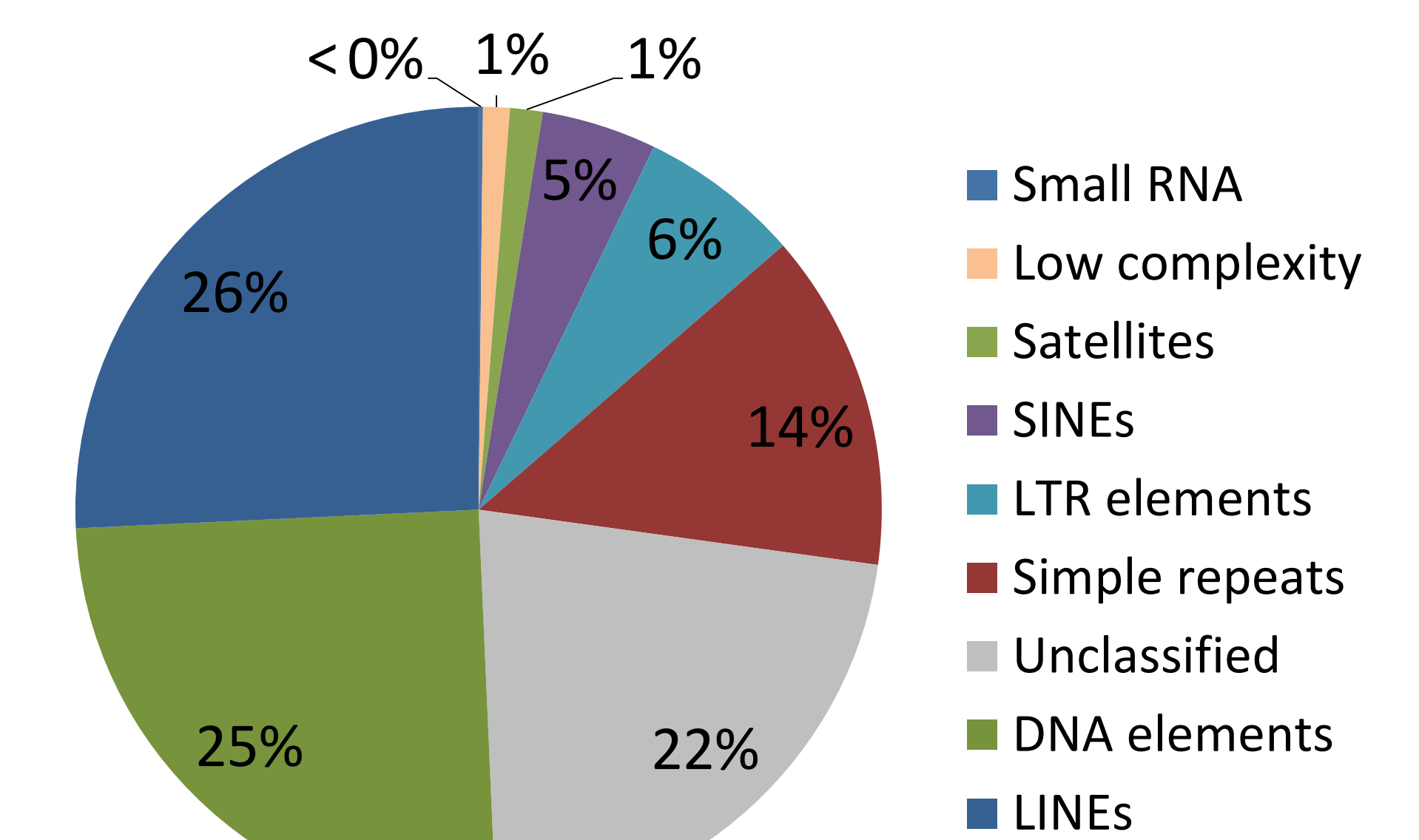
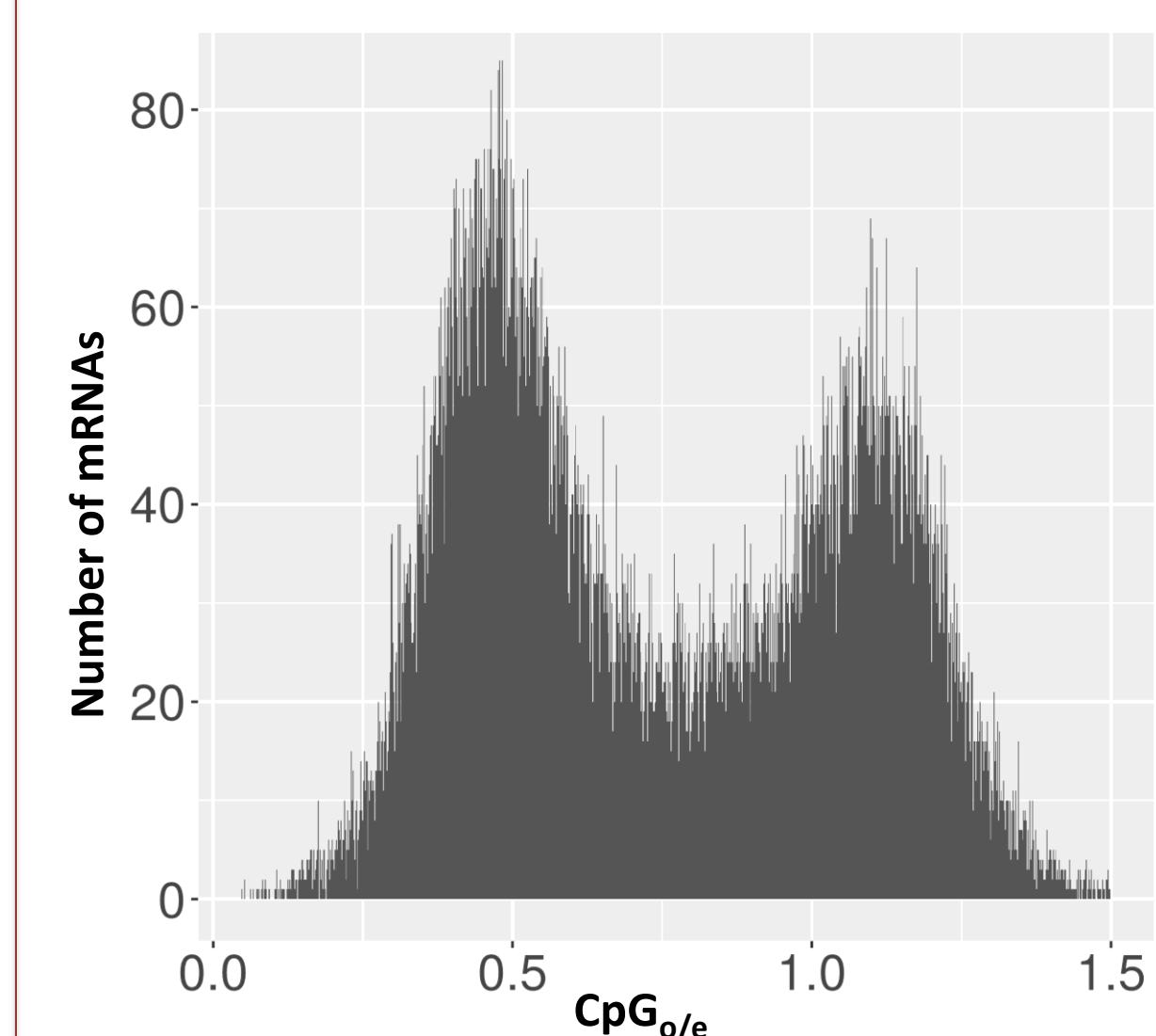


Figure 4: The ~34% of the *G. bimaculatus* genome is composed of transposable elements (TEs). The RepeatMasker software using a custom repeat library reported the percentage of genome accounted for each of the main TE families.

Genome Methylation



Identification of DNMT1 & DNMT3 genes

Figure 7: The ratio of observed CpGs to expected CpGs within mRNAs ($\text{CpG}_{\text{obs/e}}$) is an indicator of DNA methylation events. As in other hemimetabolous insect genomes, *G. bimaculatus* genome contains two groups of genes, one with low $\text{CpG}_{\text{obs/e}}$ indicating high methylation levels and another with high $\text{CpG}_{\text{obs/e}}$ indicating low methylation levels.

Transcription Factors

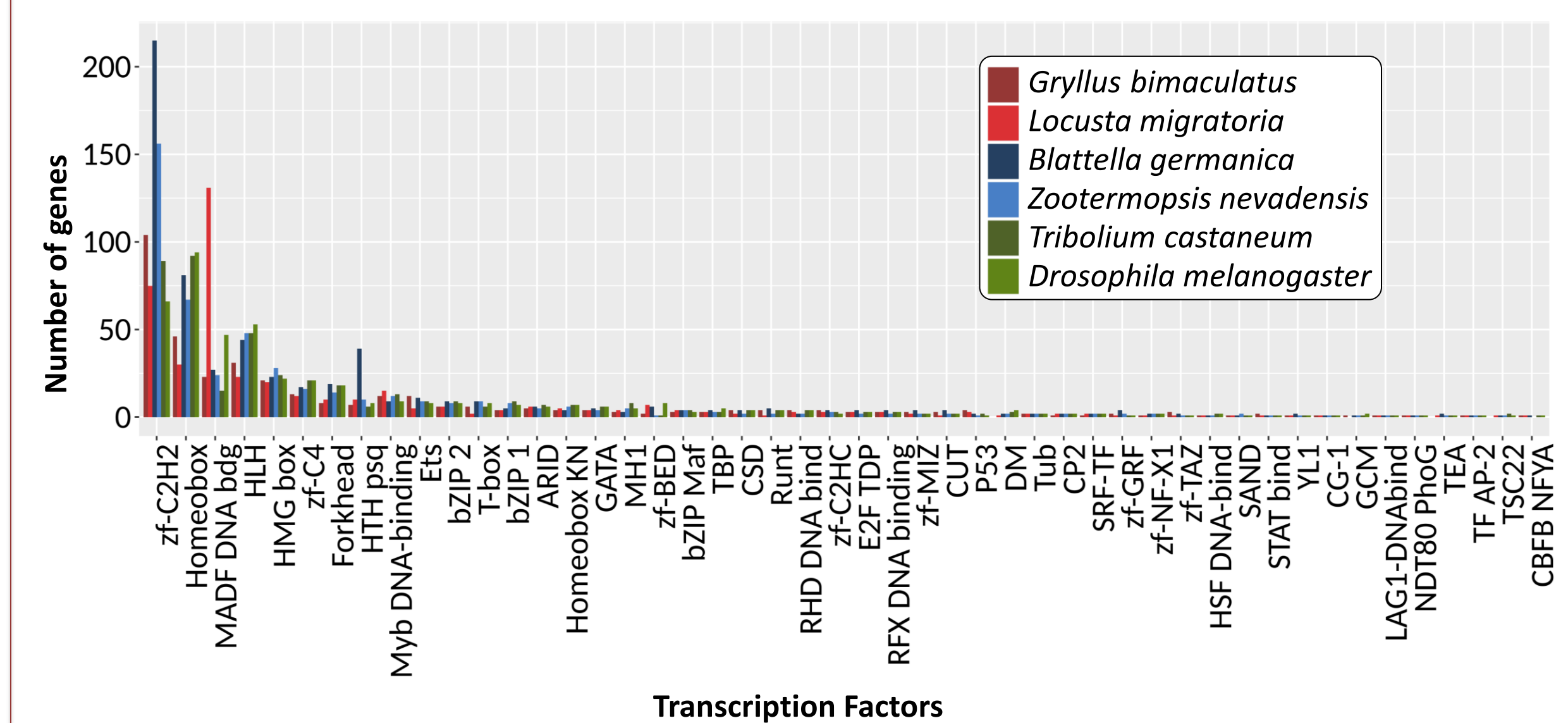


Figure 7: Number of genes coding for a protein with a PFAM motif corresponding to a transcription factor in the Orthopterans; *Gryllus bimaculatus* and *Locusta migratoria*, in the Blattodeans; *Blattella germanica* and *Zootermopsis nevadensis*, and the Endopterygota; *Tribolium castaneum* and *Drosophila melanogaster*.

