A Supermatrix Perspective on the Amphiesmenoptera Phylogeny

Bret M. Boyd[2,3] & Gordon Burleigh[1,3]

[1]Department of Biology, College of Liberal Arts and Sciences; [2]Genetics Institute; and [3]Florida Museum of Natural History: University of Florida, Gainesville, Florida

Abstract: The Amphiesmenoptera are a highly diverse clade of endopterygote insects that includes Lepidoptera (butterflies and moths) and Trichoptera (caddisflies). We built a comprehensive phylogenetic framework of Amphiesmenotpera that reflects both the rapidly growing sequence data and the computational advances that allow for large-scale phylogenetic analyses. We first assembled all potentially phylogenetic informative sequence data from Amphiesmenotpera and selected outgroups in GenBank. After pruning questionable or anomalous sequences and editing the alignments, homologous sequence clusters with more than 100 taxa were concatenated into a single phylogenetic "supermatrix" that contained 5609 taxa, representing by far the largest possible Amphiesmenoptera data set. Maximum likelihood phylogenetic analyses on the supermatrix revealed a credible phylogenetic hypothesis that raises intriguing evolutionary questions and suggests directions for future data collection. The supermatrix phylogenetic hypothesis provides not only a starting point for large scale evolutionary and ecological analyses of the hyper-diverse Amphiesmenoptera clade, but will hopefully be a resource to promote and guide future phylogenetic studies within Amphiesmenotpera.

Introduction: The superorder Amphiesmenotpera has been supported by fossil, morphological, rRNA sequence data. Lepidoptera and Trichoptera are supported as sister clades that shared an Amphiesmentopera ancestor during the Mesozoic (Grimaldi and Engle 2005, Henning 1981, Kristensen et al. 2007, Kristensen and Skalski 1999). Lepidoptera is a hyperdiverse order with 500,000 potential species and Trichoptera is smaller with 50,000 potential species (Grimaldi and Engle 2005, Kristensen et al. 2007). Previous phylogenetic studies within Amphiesmenotpera have inferred backbone relationships based on a small number of representative taxa and have focused within either the Lepidoptera or Trichoptera clades (Mutanen et al. 2010). We utilized all informative publically available data to assemble a more comprehensive phylogeny across all Amphiesmenotpera. Due to the patchy available data, we used a supermatrix approach to infer a comprehensive phylogenetic hypothesis. While the results derived from this study are not conclusive, our methods provide an unprecedented preliminary exploration of species-level relationships throughout the clade. This approach can help elucidate potential polyphyly of paraphyly of currently accepted clades that may not be evident from phylogenetic hypotheses that include only a small number of representative taxa. Members of the Amphiesmenotpera are of economic importance (crop pests and water quality indicator species), flagship species in conservation efforts, and the focus evolutionary and ecological studies. We hope that our supermatrix based phylogeny will promote and guide more comprehensive phylogenetic studies to better understand this highly diverse and important clade of insects.

Methods:

- Downloaded core nucleotide sequence data from www.ncbi.nlm.nih.gov
- Clustered sequences using an all-by-all blastn
- Aligned sequences within single linkage clusters using Muscle
- Manually edited alignments and removed problematic sequences
- Built phylogenetic trees for each cluster (loci) using RAxML
- Removed any questionable sequences from single locus trees
- Selected clusters that included 100 or more taxa
- Constructed supermatrix by concatenating alignments from selected clusters
- Built phylogenetic tree from supermatrix using RAxML
- Conducted bootstrap replicates using RAxML and Fast Tree

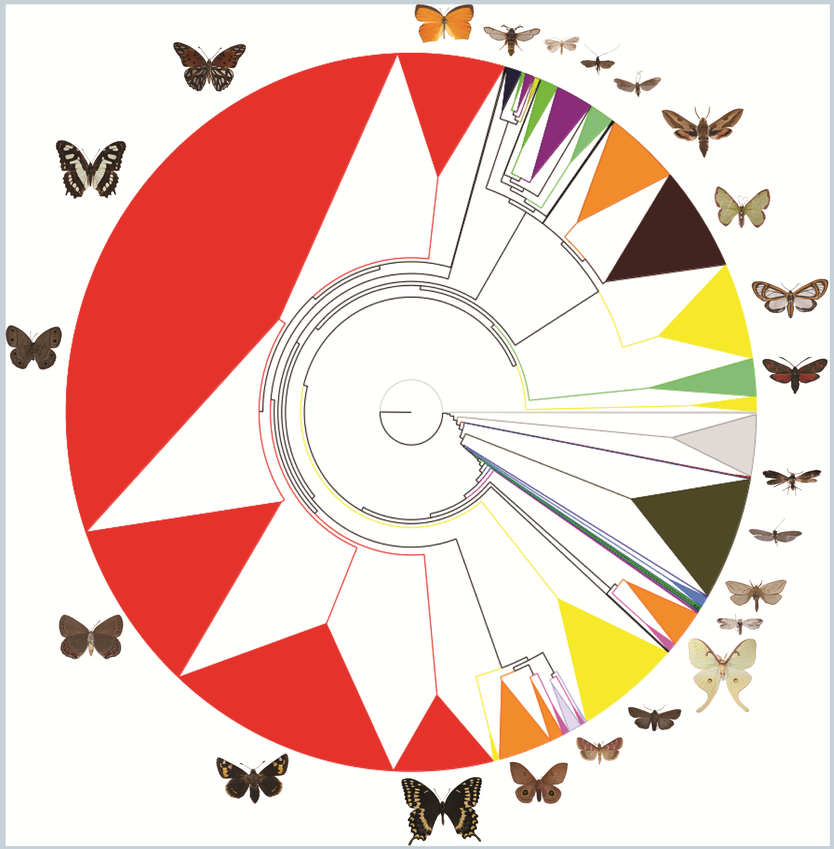

Figure 1: Phylogenetic reconstruction of Amphiesmenoptera using supermatrix generated in this study.

Results and Conclusions:

- Our data matrix and phylogenetic tree contains 5344 Amphiesmenoptera taxa (5020 from Lepidoptera and 324 from Trichoptera) and 265 outgroup taxa.
- Our data matrix contained thirteen clusters representing elongation factor 1 alpha, ITS, NDS, NADH subunit 1, ribosomal protein s5, ribosomal protein s2, 18SrRNA, GADPH (glyceraldehyde-3-phosphate dehydrogenase), rRNA, CAD (carbamoylphosphate synthetase), cytochrome b, MDH (Malate Dehydrogenase), 28SrRNA, COI, wingless, and 28SrRNA.
- Two primitive mandibulate moth families formed the base of the Amphiesmenopteran clade
- Trichoptera forms a clade within Lepidoptera
- Superfamilies Noctuoidea and Bombycoidea are supported as polyphyletic
- More work is needed to understand the relationships between Lepidoptera and Trichoptera
- More work is needed to understand the relationships among Ditrysian Lepidoptera

Acknowledgements: We would like to thank Marc Branham, Andy Warren, Delano Lewis, Jacqueline Miller, and Debbie Mathews for their research assistance and David Reed, Marta Wayne, and Wilfred Vermeris for their support and interest in our project.

Literature Cited:

Grimaldi, D., and Engle, M. S. 2005. *The Evolution of Insects.* Cambridge University Press, Cambridge. 755pp.

Hennig, W. 1981. *Insect Phylogeny.* John Wiley and Sons, Chinchester. 514pp.

Kristensen, N. P., Scoble, M. J., and Karsholt, O. 2007. Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. *Zootaxa* 1668:699-747.

Kristensen, N. P., Skalski, A. W. 1999. Phylogeny and palaeontology. In *Handbuch der Zoologie, a Natural History of the Phyla of the Animal Kingdom, Vol IV, Arthropoda: Insecta, Part 35, Lepidoptera, Moths and Butterflies, Vol. I: Evolution, Systematics, and Biogeography.* Edited Kristensen, N. P.. Walter de Gruyter, Berlin.

Mutanen, M., Wahlberg, N., Kaila, L. 2010. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. *Proceeding of the Royal Society: Biological Sciences* doi:10.1098/rspb.2010.0392

GENUS	ef1a	ITS	NDS	NADH SU1	Rps5	Rps2	18SrRNA	GAPDH	rRNA	CAD	cytb	MDH	28SrRNA	COI	Wingless	28SrRNA
Stratiomyidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Rhadinidae	YES	NO	NO	YES	YES	YES	NO	NO	NO	YES	YES	YES	NO	YES	YES	NO
Chathamiidae	YES	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	YES	NO	NO
Oeconeiidae	YES	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO	NO
Helicophidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Eupterotidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Hydrobiosidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Saturniidae	YES	NO	NO	YES	NO	NO	YES	NO	NO	YES	YES	NO	NO	YES	NO	NO
Glossosomatidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Kokiridae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Geometridae	YES	NO	NO	YES	YES	NO	NO	YES	NO	NO	NO	YES	YES	YES	YES	NO
Notodontidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Antheliidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Limnephilidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Autostichidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO
Drepanidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	NO
Arctiidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	YES	YES	YES	NO
Oecophoridae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO
Pieridae	YES	NO	YES	YES	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO
Cosmopterigidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Pagiorinidae	YES	NO	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO
Nymphalidae	YES	NO	YES	YES	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO
Noctuidae	YES	NO	YES	YES	NO	NO	YES	NO	NO	NO	YES	NO	YES	YES	YES	NO
Lymantriidae	YES	NO	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	NO	YES	YES	NO
Nolidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO	NO
Crambidae	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	YES	YES	NO
Conesucidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Fragilariaphyxidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO
Sphingidae	YES	NO	NO	YES	YES	NO	YES	NO	YES	NO	NO	YES	YES	YES	YES	NO
Leptoceridae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Polycentropodidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Lycaenidae	YES	YES	YES	YES	NO	NO	NO	NO	YES	YES	YES	NO	YES	YES	YES	NO
Philopotamidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Helicopsychidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Molannidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Hydropsychidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Dipseudopsidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Lasiocampidae	YES	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO
Lecithoceridae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO
Hedylidae	YES	NO	NO	YES	YES	YES	NO	YES	NO	YES	NO	YES	NO	YES	YES	NO
Hesperiidae	YES	NO	YES	YES	NO	NO	NO	NO	NO	NO	YES	NO	NO	YES	YES	NO
Pyralidae	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	YES	NO	NO
Sericostomatidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Berseidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Philorheithridae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Prodoxidae	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO	YES	NO	NO
Brahmaeidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Apatelodidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Hybliaeidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO
Lemoniidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Gracillariidae	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	YES	YES	YES
Calocidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Odontoceridae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Hepialidae	YES	NO	NO	NO	YES	NO	YES	YES	NO	NO	YES	YES	NO	YES	NO	NO
Micropterigidae	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO
Rhyacophilidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Sialidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	NO	YES	NO	NO
Ecnomidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Stenopsychidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Psychomyiidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Hydroptilidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Xiphocentronidae	YES	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Thyridae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO
Gelechiidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	YES	YES	NO
Carthaeidae	YES	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Elachistidae	NO	NO	NO	NO	YES	YES	NO	NO	YES	NO	YES	NO	YES	YES	YES	NO
Lepidostomatidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Phryganopsychidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Calamoceratidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Apataniidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Uenoidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Rhyganeidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Hydrosalpingidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Petrothrincidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Anomalopsychidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Atriplectididae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Helicopididae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Brachycentridae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Tasimiidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Neptiliculidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	YES	NO	NO
Plectrotarsidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	NO	NO	NO
Erioconaridae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Neopseustidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Agathiphagidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	YES	NO	NO
Mnesarchaeidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Opostegidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Acanthoperoctetidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Cecidoseidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO
Heterobathmiidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	YES	NO	NO
Tineidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Sesiidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	YES	NO	NO
Psychidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Tischeriidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	YES	NO	NO
Zygaenidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	YES	NO	NO
Palaephatidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Heterogynidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO
Tortricidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	YES	NO	YES	YES	YES	NO	NO
Cossidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	YES	NO	NO
Lacturidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO
Adelidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Himantopteridae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO
Somabrachyidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO
Momphidae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	YES	NO	NO
Pterophoridae	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO
Pisulidae	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	NO	YES	YES	NO	NO
Limnocentropodidae	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO	NO
Barbarichthonidae	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	YES	NO	NO
Goeridae	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	YES	NO	NO
Bombycidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Mimallonidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Endromidae	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO
Carposinidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	YES	NO	NO
Choreutidae	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	YES	NO	NO	YES	NO	NO
Yponomeutidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Blastobasidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Argyresthiidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Plutellidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Bucculatricidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Coleophoridae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Limacodidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Uraniidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Xylorictidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Batrachedridae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Coryalidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Xylomyidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO
Urodidae	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO

Figure 2: Insect families and gene clusters represented in supermatrix