

Research module project (Forschungsmodul)

## Annotation and basic analyses of beetle genomes

### Background:

Genome annotation is an early step in the pipeline of genomic analyses. The quality of the annotation can have a great impact on the downstream analyses. To allow for uniform annotation across a dataset, using the same annotation pipeline for all genomes is important. In case of publicly available genomes and genome annotations, one can remove potential artificial differences due to different annotation pipelines by (re-)annotating genomes.

Subsocial beetles take care of their offspring by provisioning food and shelter, which is a rare trait in invertebrates. Most species of the Scarabaeidae family (scarab beetles) are non-social plant feeders, while all species of the Scarabaeinae subfamily (dung beetles) are subsocial, with varying degrees of parental care. So far it is unknown how subsociality evolved within this beetle family.

**Overall objective:** In this project, all publicly available Scarabaeidae genomes will be (re-)annotated with an established pipeline. The genome annotations will be used for basic genomic analyses and results will be correlated to sociality phenotypes.

**Potential hypotheses**, to be supplemented and further developed within the project

1. Genomic features differ between and correlate with differences in sociality
2. Differences in nutrition can also have an impact on observed genomic variation.

### Materials & methods

The student will use a set of publicly available Scarabaeidae genomes and a variety of programs for the annotation and analysis of genomic features. The correlation between genomic features and phenotype, i.e. brood care and nutrition are analysed. If time permits, selection analyses can be performed and the impact of the genome annotation pipeline on the selection analysis results can be investigated.

- Established Pipeline using BRAKER for genome annotation
- Bash/Python for basic genomic comparisons
- R/Python for statistical analysis

### Requirements:

- Some experience with the command line and learning the usage of different programs
- Interest in genome annotation and genomic analyses
- Ability to work independently

**Supervision:** Sarah Rinke-Stack, Room 17, s.rinke-stack@uni-muenster.de, Molecular Evolution and Bioinformatics Group (<http://bornberglab.org/>).

## References

- Biedermann, Nuotclà (2020). Social Beetles.
- Philips et al. (2004). A phylogenetic analysis of dung beetles.
- Gabriel et al (preprint). BRAKER3: Fully Automated Genome Annotation Using RNA-Seq and Protein Evidence with GeneMark-ETP, AUGUSTUS and TSEBRA
- Bell (2021). Transposable element annotation in non-model species - the benefits of species-specific repeat libraries using semi-automated EDTA and DeepTE de novo pipelines.