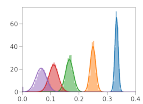
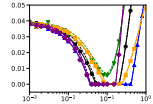
Theoretical Biology

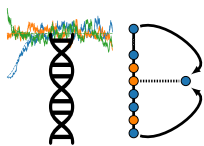
Pete Czuppon



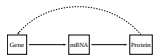
evolution



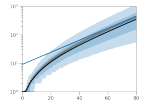
gene regulation



ecology



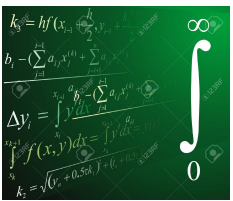
epidemiology



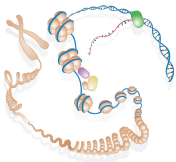
coding

&

maths



Potential projects



Transposable elements (TEs)

(self-replicating genetic elements)

→ study TE dynamics & silencing in the genome

Antibiotic resistance

→ study effect of constant antibiotic concentration assumption on the evolution of resistance

Population genetics

→ identification of the scaling of different selection regimes (quasi-neutral, weak, strong)

Alignment uncertainty

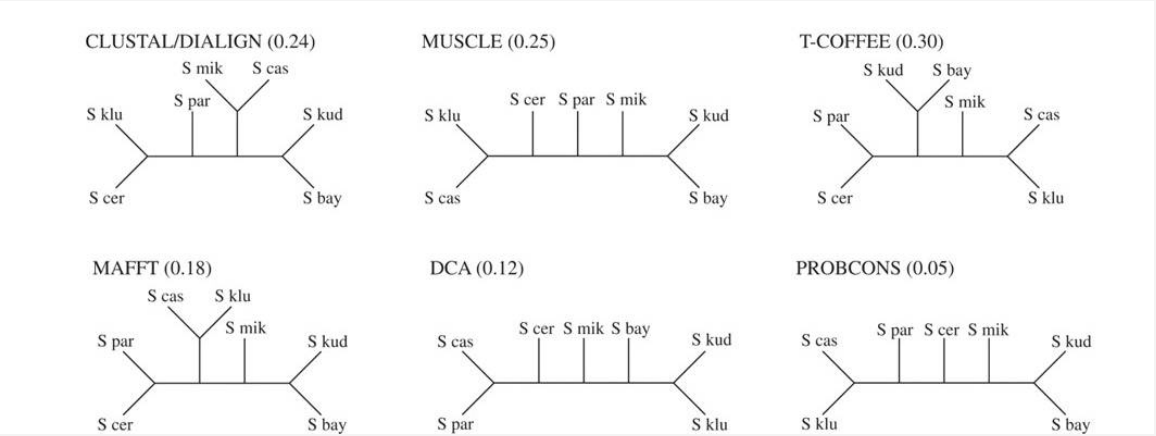
Supervisors: Carsten Kemena (c.kemena@uni-muenster.de)

Background: It is common to check trustworthiness of for example phylogenetic nodes based on bootstrap values. Those are dependent on the Alignment that is used. However multiple sequence alignments, building the basis of multiple analyses (phylogeny, selection, etc) usually taken as correct or are filtered based on gap patterns.

Aim: Perform a first step into taking alignment reliability into account for downstream analyses. Calculate a reliability score for a given alignment.

Methods: Create multiple sequence alignments using different methods and calculate the support for each alignment. Also measure influence of alignment method on downstream analyses.

Literature: Wong et al., Alignment uncertainty and genomic analysis, Science 2008
Sela et al, GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters, NAR 2015



Improving domain detection

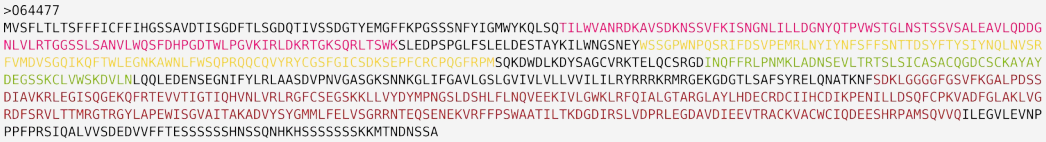
Supervisors: Carsten Kemena (c.kemena@uni-muenster.de)

Background: A good domain annotation is important for domain based analyses. We have started developing a tool to improve domain annotations based on domain context. Additionally a structure based component would be of interest.

Aim: Search for methods that can divide a sequence or 3d structure of unknown domains into different units that can then be compared based on its structure to known structures of protein domains.

Methods: The usage of new 3D structure prediction tools (alphafold and successors) to model 3D structure of sequence segments and comparison tools for 3D structures.

Literature Paysan-Lafosse et al., InterPro in 2022, NAR 2022
Terrapon et al., Detection of new protein domains using co-occurrence: application to Plasmodium falciparum, Bioinformatics 2009



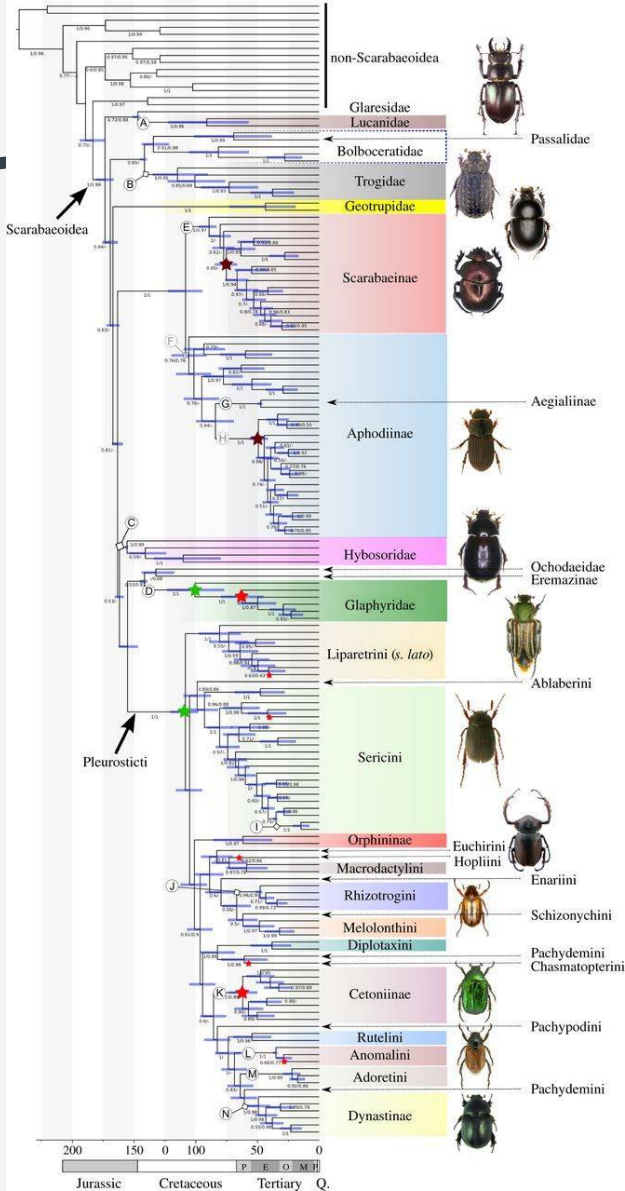
Genomic adaptations in subsocial beetles

Supervisors: Sarah Rinke-Stack (s.rinke@uni-muenster.de)

Background: 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.

Aim:
- Investigate genomic differences between non-social plant-feeders and social dung-beetles.
- Analyse gene family expansions and contractions related to nutrition and social phenotype.

Methods:
- Bash/Python for basic genomic comparisons
- Cafe for gene family expansion/contraction
- R/Python for statistical analysis



Literature Biedermann, P.H.W., Nuotclà, J.A. (2020). Social Beetles.
Philips et al. (2004). A phylogenetic analysis of dung beetles.

Universität Münster

The role of alternative splicing in caste differences in a higher termite

Supervisors: Mark Harrison & Alina Mikhailova

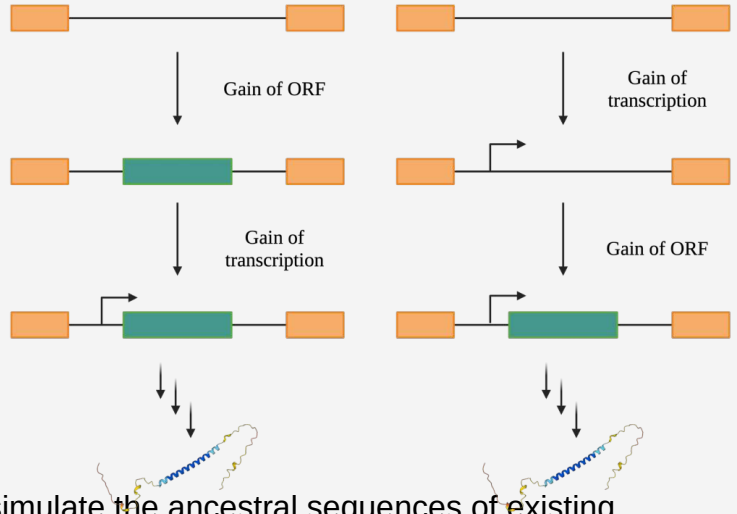
Background: In some termite species, incredibly divergent phenotypes (castes: queens, workers, soldiers) are produced from the same genome by differential regulation of gene expression. Little is known on how alternative splicing contributes to this polymorphism.

Aim: Investigate the importance of alternative splicing in caste differences in a higher termite.

Methods: data-analysis with, for example, hisat, htseq-count, DEXseq, R

Literature: Harrison, Jongepier, Robertson, et al. Hemimetabolous genomes reveal molecular basis of termite eusociality, Nature ecology & evolution, 2018; Harrison, Dohmen et al. Complex regulatory role of DNA methylation in caste-and age-specific expression of a termite, Open Biology, 2022

Purification of Ancestral Atlas protein



Supervisors: Lars Eicholt, Andreas Lange

Background:

De novo genes, lacking similarity to others, suggest a non-duplication origin. The hypothesis posits that many of these genes could evolve from non-coding sequences (de novo gene emergence), though as of now, there's no concrete proof. Ancestral sequence reconstruction (ASR) offers a powerful tool to simulate the ancestral sequences of existing proteins, aiding in understanding the emergence of putative de novo proteins.

Aim: Purifiy two ancestral Proteins of Atlas from Inclusion bodies
Methods: Protein expression & purification, maybe Circular dichroism

Literature:
Eicholt LA et al. Protein Science. 2022
Lange A et al. Nat Comm, 2021
Bornberg-Bauer, E, Hlouchova, K, Lange A COSB, 2020

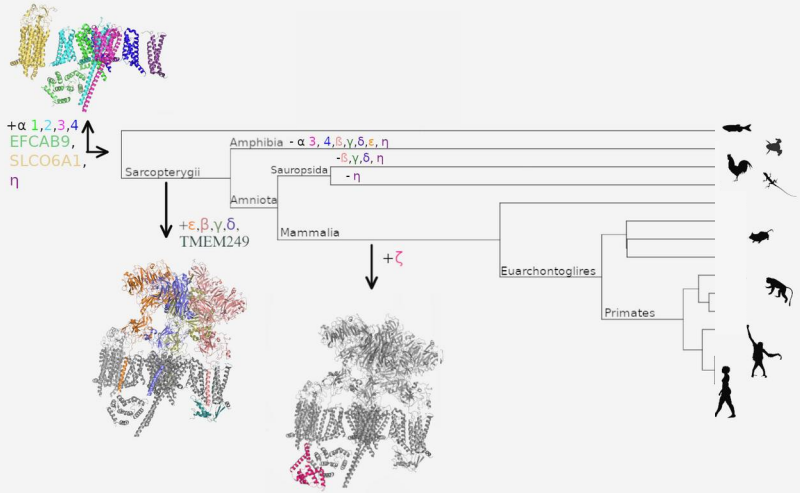
Emergence and Evolution of the CatSpermasome

Supervisors: Margaux Aubel

Background: Sperm specific ion channels (SSICs) are ion channels within spermatozoa, that cover crucial functions and mechanisms for the normal functionality of spermatozoa. The CatSpermasome, consisting of 14 subunits, is an SSIC in the flagella of spermatozoa.

Aim: Search and detect subunits in more species to expand the Tree and if possible find the origins of different subunits. Check if the composition of domains changed during evolution.

Methods: BLAST, multiple sequence alignment, pfamscan



Literature Lin, S. et al. (2021) ‘Structure of a mammalian sperm cation channel complex’, Cai, X. and Clapham, D.E. (2008) ‘Evolutionary genomics reveals lineage-specific gene loss and rapid evolution of a sperm-specific ion channel complex: CatSpers and CatSperβ’, Berger A. (2024) ‘Evolution of the CatSpermasome’

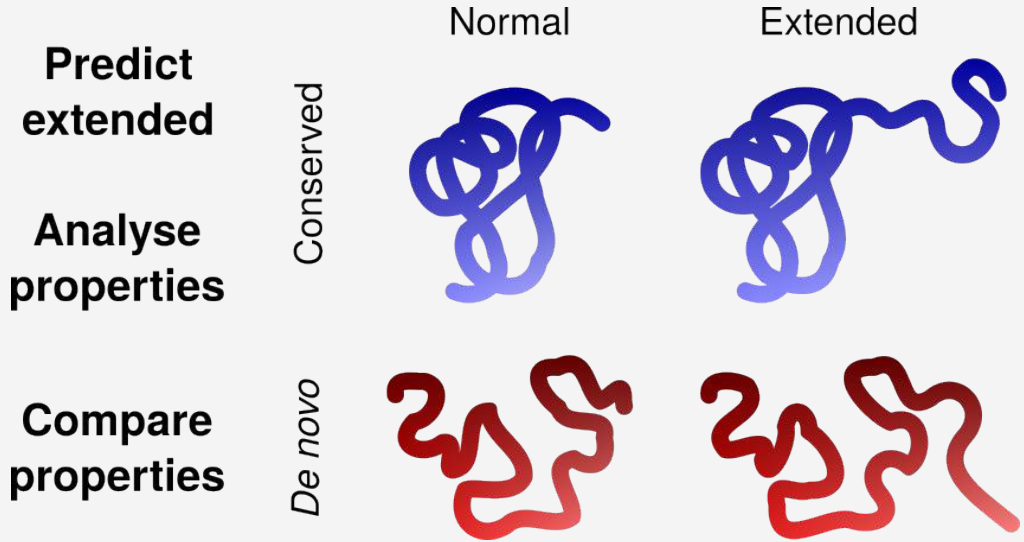
Comparing c-terminal extended vs non-extended proteins

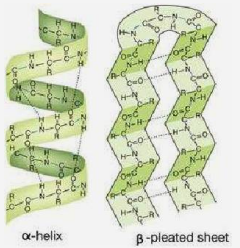
Universität Münster

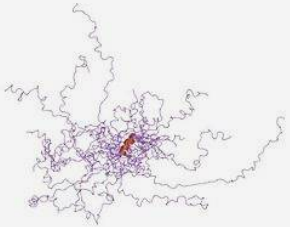
Supervisors: Bharat Ravi

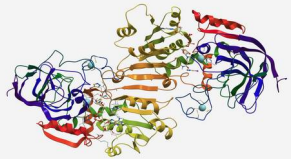
Aim:
Understand how C-terminal extensions can change protein properties

Methods:
1. Find next in-frame stop codon for every ORF.
2. Predict biophysical properties (software).
3. Compare *de novo* vs conserved, and extended vs normal
4. Perform stats









wissen.leben
bornberglab.org

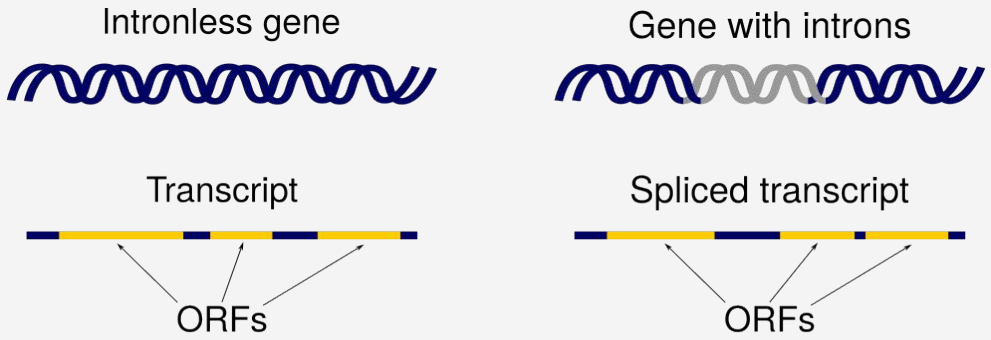
Comparing *de novo* intronless and spliced transcripts

Universität Münster

<u>Supervisors:</u> Bharat Ravi

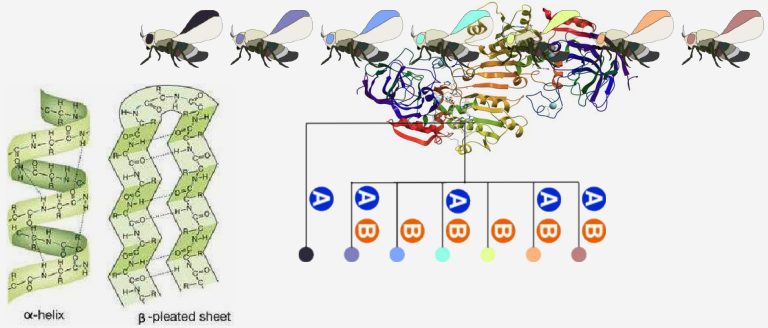
<u>Aim:</u>
To understand if splicing determines the evolution of *de novo* protein coding genes.

<u>Methods:</u>
1. Find intronless and spliced transcripts.
2. Find ORFs.
3. Compare properties of Transcript & ORF: (intronless vs spliced)
4. Perform stats



Transcript properties: Expression, Stability, Conservation

ORF (protein) properties: Disorder, 2°/3° Structure, Length



wissen.leben
bornberglab.org

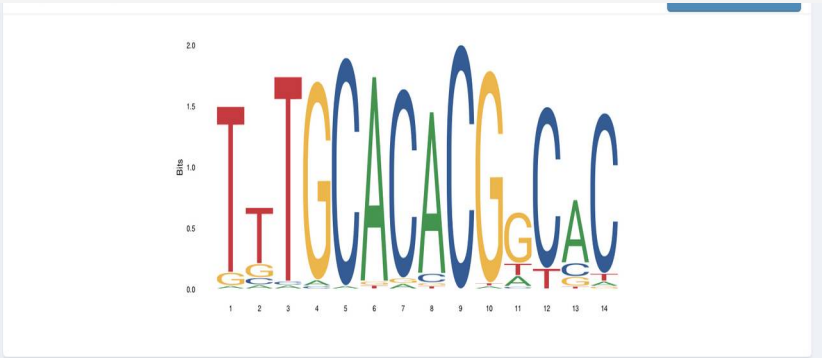
Transcription factor binding variability

Supervisors: Alun Jones

Background: Transcription factors are the key facilitators to producing those endle by regulating gene expression throughout the genomes. How transcription factors create biological variation varies greatly and in this project we aim to better understand of binding variability can lead to variation in regulation.

Aim: Look at variation in binding strength across binding motif sites for multiple transcription factors.

Methods: Using the JASPAR database and UNIPROBE we will use a sliding wind approach to look at escores across binding motif. Will use the biomotifs matrix module from biopython.



Frequency matrix

JASPAR | TRANSFAC | MEME | RAW PFM | Reverse comp.

A[	1	3	1	2	1	53	2	52	0	1	9	0	46	2	]
C[	0	5	1	1	71	0	69	3	67	0	2	63	10	62	]
G[	4	10	0	66	0	1	2	1	0	70	42	0	7	1	]
T[	50	40	53	0	0	1	0	1	0	1	10	8	2	3	]

Literature: Evolutionary Potential of Cis-Regulatory Mutations to Cause Rapid Changes in Transcription Factor Binding
https://pubmed.ncbi.nlm.nih.gov/30597011/

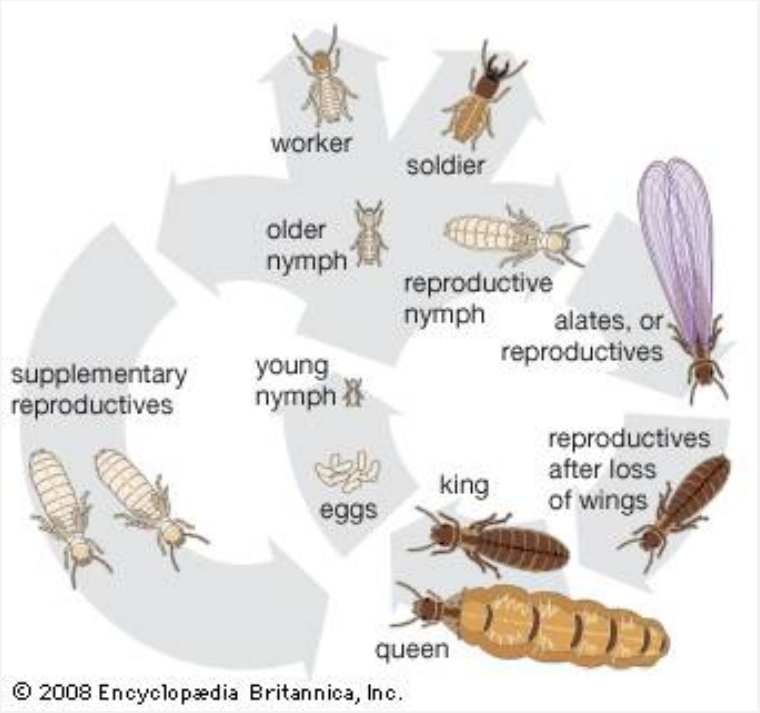
A Social Transcription Factor

Supervisors: Alun Jones

Background: The zinc finger kruppel homolog - 1 (Kr-h1). Kr-h1 expression is driven by juvinille hormone (JH) which has been shown to be key for caste determination in multiple eusocial species [1, 3, 2]. However, the evolution of Krh1 has not been looked at in regard to social evolution
across lineages and particularly within blattodea.

Aim: Identify common structures in Kruppel across social

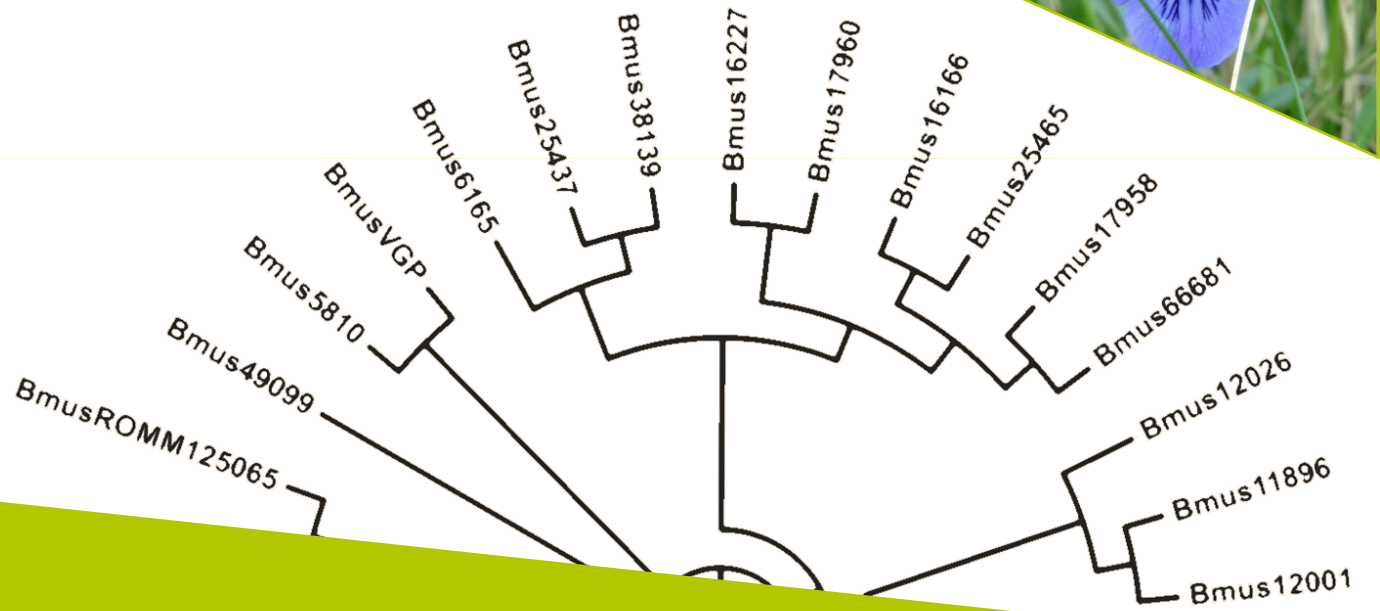
Methods: Extract sequences from a single eusocial origin in orthoDB. Using various python modules to investigate protein properties for each of the species' sequences and see differences between social and eusocial.



Universität Münster

Evolution & Biodiversity of Plants

Prof. Dr. Kai Müller

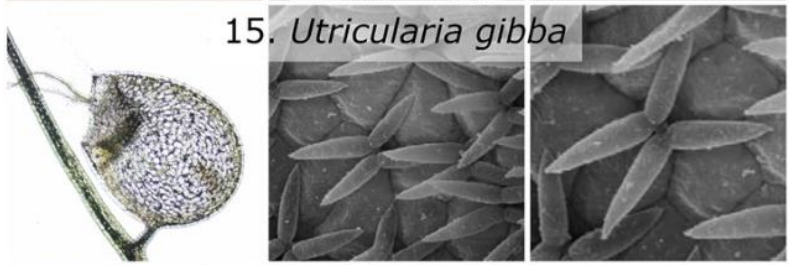


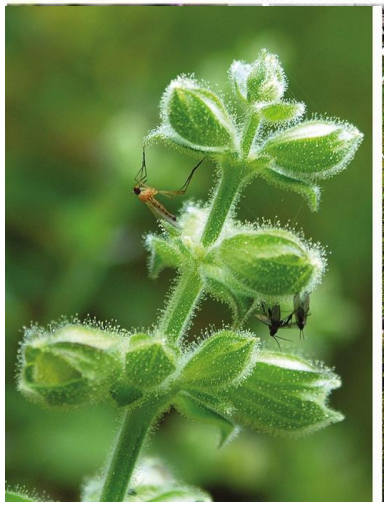
Lentibulariaceae

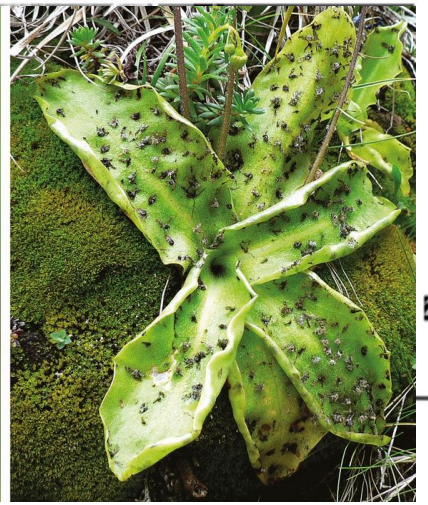
- Carnivorous plants
- Within the order Lamiales
- Diff. traps / shared glands
- Co-opting of defense

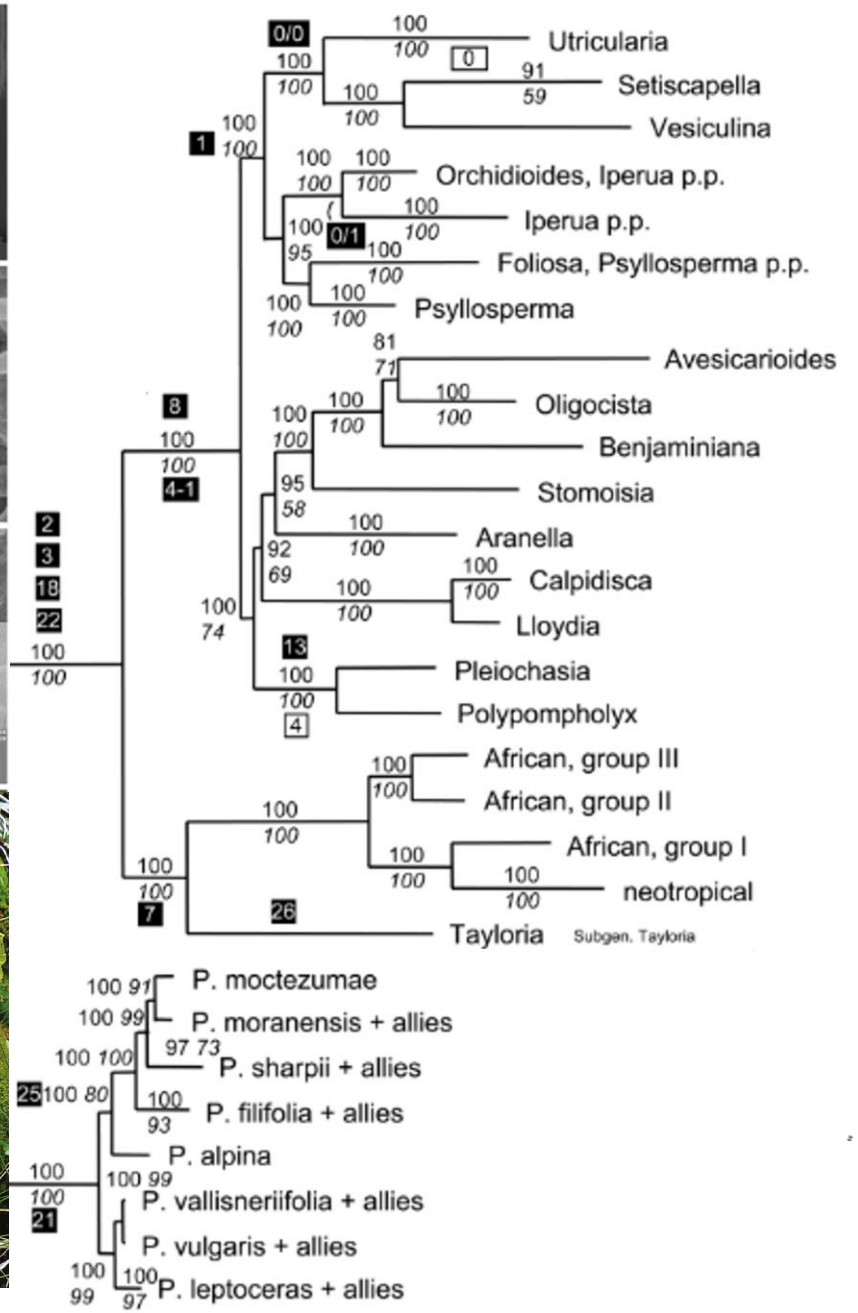
Unknowns

- Evolutionary origins
- 1 or multiple origins?
- Molecular mechanisms





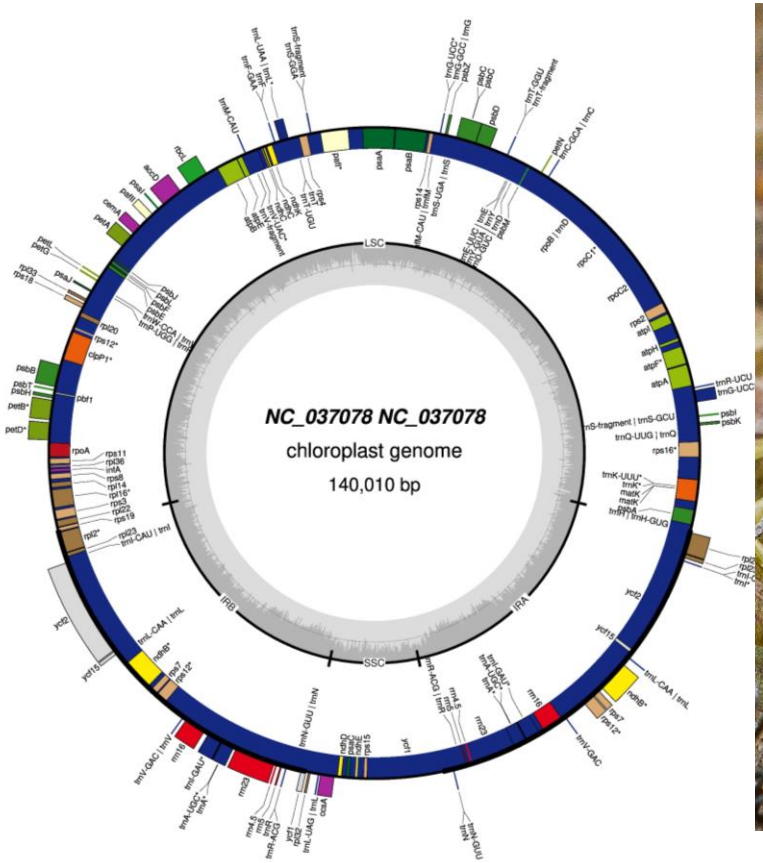




Goals

- Create a lab culture of *P. lusitanica / gigantea*.
- Construct a chloroplast genome.
- Construct a robust phylogeny of

with related Lamiales.



Methods

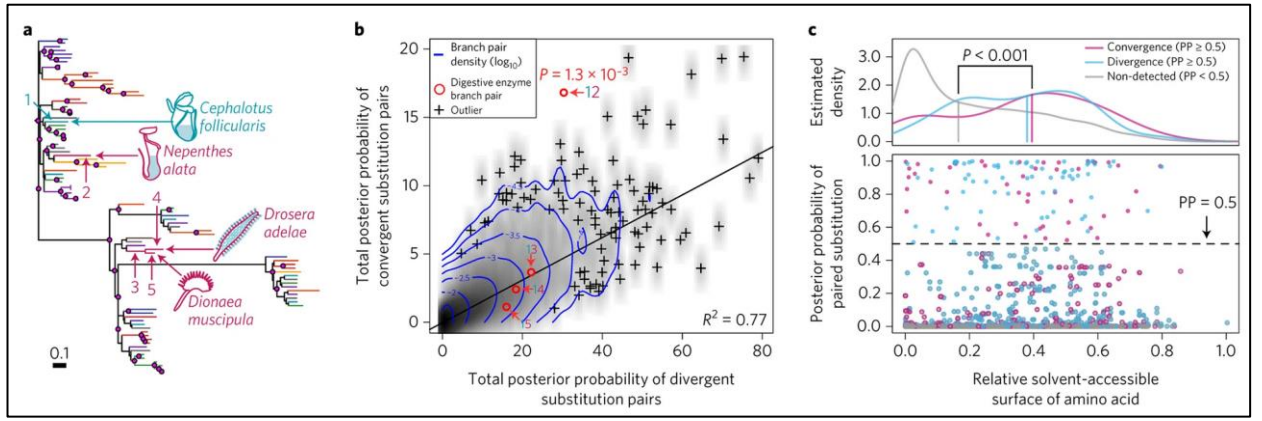
- Wet lab (culturing, DNA extr. etc.)
- High throughput sequencing
- Bioinformatics and phylogenetics

Pitfall traps and microorganisms

- Convergent evolution
- Complex digestion mechanisms
- Associated microorganisms

Unknowns

- Role in the evolution of carnivory
- Diff. between „old“ and „young“
- Changes in non-native habitats

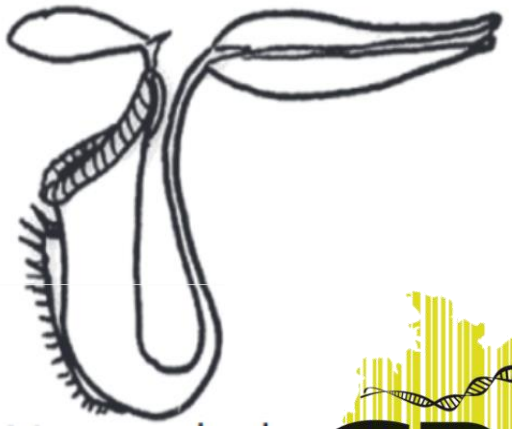


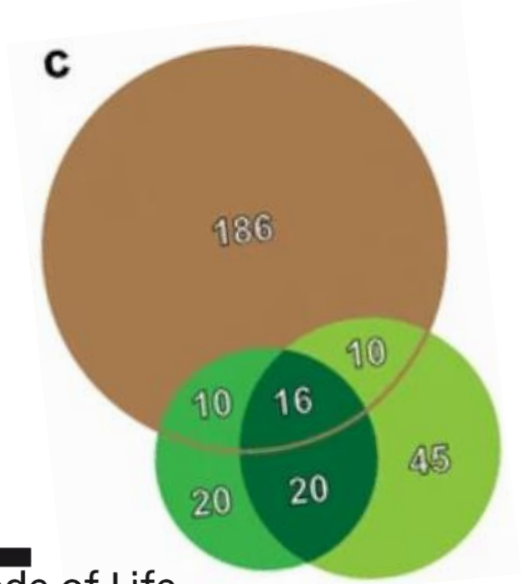
Goals

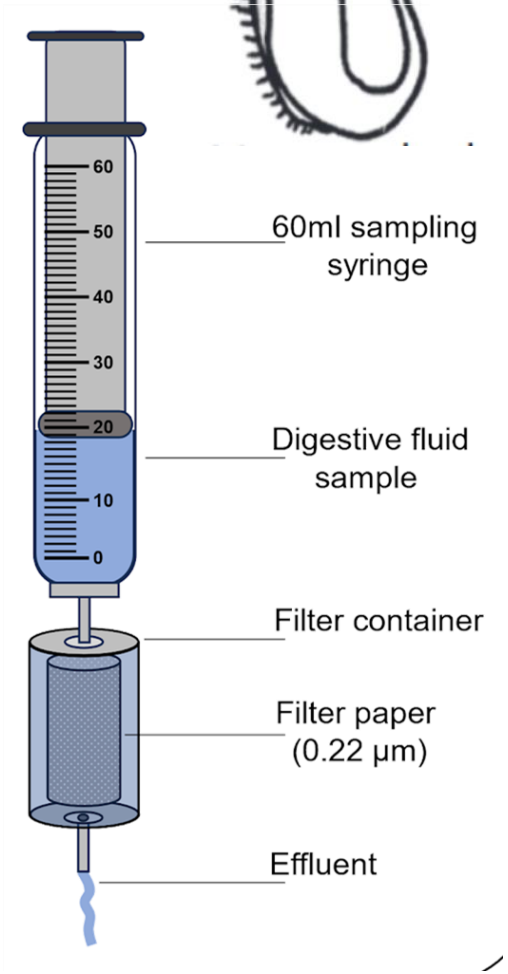
- Characterize differences and diversity of microorganisms.
- Establish workflow for field studies.

Methods

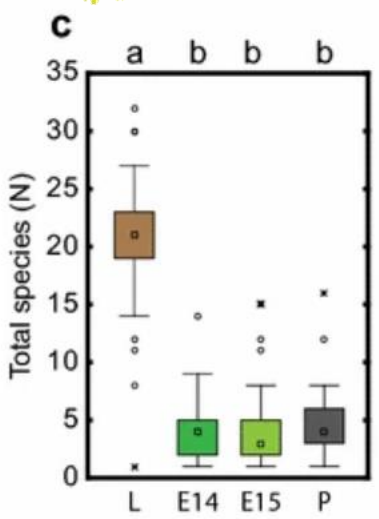
- Wet lab (sampling, DNA extr., PCR)
- High throughput sequencing
- Bioinformatics

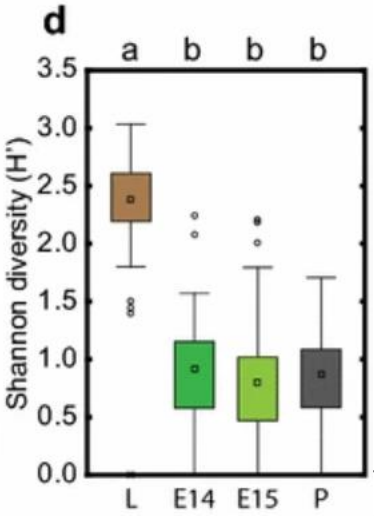






GBOL
German Barcode of Life





Universität Münster

Evolution & Biodiversity of Plants

Prof. Dr. Kai Müller