

DMP The evolution of caste plasticity and caste dimorphism in insect societies

DMP The evolution of caste plasticity and caste dimorphism in insect societies

ADMIN DETAILS

Project Name: DMP The evolution of caste plasticity and caste dimorphism in insect societies

Grant Title: G0F6622N

Principal Investigator / Researcher: Tom Wenseleers

Description: The evolution of sociality in ants, bees and wasps is a prime example of a major transition in evolution, where individuals went to live in societies characterised by the presence of an advanced reproductive division of labour. In primitively eusocial species, the allocation to reproductive and nonreproductive roles in the colony is highly plastic, while in more advanced eusocial species, the queen and worker castes are morphologically distinct. Recently, it was shown that the caste plasticity in a primitively eusocial ant species showed reversible changes in brain volume in function of reproductive role, with breeding females reducing their brain size to allow for an increased investment in egg production. In the present project, we will study the occurrence of such caste plasticity in function of reproductive role across a wider range of primitively and advanced eusocial temperate and Neotropical ants, bees and wasps using a combination of state-of-the-art genomic, light sheet fluorescence microscopy (LSFM) brain imaging and behavioural experiments. This will allow us to study shifting patterns in the evolution of behavioural plasticity in eusocial insects, including possible tissue trade-offs and changes in the modularity in brain gene expression patterns. In addition, we will study how pre-existing gene regulatory ground plans in primitively eusocial wasps could have laid the basis for the evolution of morphologically defined castes in advanced eusocial wasps.

Institution: KU Leuven

1. GENERAL INFORMATION

Name applicant

Tom Wenseleers

FWO Project Number & Title

Bilateral Scientific Cooperation State of São Paulo (Brazil – FAPESP) - G06622N Title of the application:
The evolution of caste plasticity and caste dimorphism in insect societies

Affiliation

- KU Leuven

2. DATA DESCRIPTION

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe the origin, type and format of the data (per dataset) and its (estimated) volume, ideally per objective or WP of the project. You might consider using the table in the guidance.

All **primary new data** (observational data, image recordings, RNA sequences, programming code, analysis scripts) will be **stored digitally** on the lab's two **QNAP NAS systems**, which communicate over a secure connection and are backed up daily. Final analysed data and code will be published in the Dryad database upon final publication. Final analysed data and code will be published in the Dryad database upon final publication and DNA sequences will be uploaded to the public NCBI database. Most datasets are relatively small except the genomics data. We currently count with 45 TB of storage in our lab's dedicated QNAP NAS systems that can still be upgraded with extra storage if necessary. RNA sequences will be uploaded to the public NCBI database.

WP1. The evolution of caste plasticity in function of reproductive role

WP 1.1 Sample collection

For bees, we will compare the advanced eusocial honeybee *Apis mellifera* with the primitively eusocial Orchard bee *Euglossa melanotricha*, while for ants, we will compare the highly eusocial black garden ant *Lasius niger* with the primitive ponerine Dinosaur ant *Dinoponera quadricaps*.

- Field sampling data of 20 individuals of each species *Apis mellifera*; *Euglossa melanotricha*; *Lasius niger* and *Dinoponera quadricaps*. Collection of age-matched workers, ovary activation status will be determined via dissection. Date and place of the samples also stored (stored in .csv format). Estimated volume: 0.1GB.

WP 1.2 Brain imaging

- One image of total brain size per species and caste (reproductive or non-reproductive), with detailed data on brain volume and structure volume (stored in .tiff format). Estimated volume: 10GB.
- Detailed data on brain volume and structure volume (stored in .csv format). Estimated volume: 0.1GB.
- R scripts of data analyses (stored in R. format). Estimated volume: 0.1GB.

WP 1.3 RNA-seq brain transcriptome analyses

- RNA sequences genotyping. The raw sequences will be uploaded at NCBI prior publication (stored in gff. format). Estimated volume: 30GB.

- Quality control and number of reads obtained by RNAseq analyses (stored in .txt format). Estimated volume: 0.1GB.
- R scripts of data analyses (stored in R. format). Estimated volume: 0.1GB.

WP 2. The evolutionary origin of morphologically specialized reproductive castes

WP 2.1 Collecting female larval samples

- Experimental microsatellite genotyping data of 10 larvae from 6 colonies (total = 60 samples) of *Vespula vulgaris* (.FASTA format). Estimated volume: 0.1GB.
- Experimental microsatellite genotyping data of 10 larvae from 6 colonies (total = 60 samples) of *Polistes dominula* (.FASTA format). Estimated volume: 0.1GB.

WP 2.2 RNA-seq transcriptome analyses

- RNA sequences genotyping. The raw sequences will be uploaded at NCBI prior publication (stored in gff. format) Estimated volume: 15GB.
- Quality control and number of reads obtained by RNAseq analyses (stored in .txt format). Estimated volume: 0.1GB.
- R scripts of data analyses (stored in R. format). Estimated volume: 0.1GB.

WP 3. Neuroethology and chemical ecology of task specialisation in ants

WP 3.1 Sampling

- Field sampling data of 20 individuals of each species. Collection of age-matched workers, ovary activation status will be determined via dissection. Date and place of the samples also stored (stored in .csv format). Estimated volume: 0.1GB.

WP 3.2 Dissection, brain imaging and body measurements

- One image of total brain size per species and caste (reproductive or non-reproductive), with detailed data on brain volume and structure volume (stored in .tiff format). Estimated volume: 10GB.
- Detailed data on brain volume and structure volume (stored in .csv format). Estimated volume: 0.1GB.
- R scripts of data analyses (stored in R. format). Estimated volume: 0.1GB.

WP 3.3 Chemical analyses

- Chromatograms and data of GC-MS (stored in .cdf and .csv format). Estimated volume (EV): 5 GB.

3. LEGAL & ETHICAL ISSUES

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference: -

Short description of the kind of personal data that will be used: -

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. DOCUMENTATION & METADATA

What documentation will be provided to enable reuse of the data collected/generated in this project?

All data will be kept in a separate folder per WP in our QNAP NAS servers containing the laboratory notebook of the researcher, including name of the researcher, date, sampling protocol, the experimental design in Word .docx format. Raw data will be stored according to specific file formats mentioned in data description (section 2). Analysis R scripts will be also kept. All processed data, programming code and analysis scripts will be made publicly available via the Dryad repository and/or public databases (NCBI for RNA sequences) upon final publication. An inventory of all biological material stored containing species, date, origin, number of samples, responsible researcher, and location of sample in the laboratory will be kept in excel .xlsx and updated throughout the project.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

Due to the nature of the project a metadata standard will not be used for the whole project but instead WP specific descriptive metadata files will be created in excel format containing relevant information such as title, version, author and file format for the identification and retrieval of specific files within a WP.

5. DATA STORAGE & BACK UP DURING THE FWO PROJECT

Where will the data be stored? How is back up of the data provided?

We count with two 25 TB QNAP NAS systems with RAID array that are backed up daily on each other. Hence the risks of losing data are minimum since we always have at least two copies of the files and

individual hard drive failures can easily be replaced and or/upgraded. The system is physically located at the Department central servers and are accessed via a secure University network.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

We currently have 50TB available on our dedicated servers and at the moment we have ca. of 40% of storage available. Nonetheless, the storage of the QNAP NAS servers can be further upgraded if necessary.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

We do not expect to have extra costs for data storage as the two dedicated QNAP NAS systems are already available in the group. Maintenance costs (e.g. to replace failing hard drives of the NAS systems) will be covered by this project.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Our QNAP NAS servers are only accessible via a secure University connection. In addition, password protected access to specific project folders is granted individually to authorized researchers in the project.

6. DATA PRESERVATION AFTER THE FWO PROJECT

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All digitally stored data will be kept for at least ten years but most likely indefinitely on both University and public data repositories. All scientific publications will be accompanied by the relevant raw data and datasets deposited in public repositories and DNA sequences will be uploaded to the public NCBI database.

Where will the data be archived (= stored for the longer term)?

Digital data will be stored in the laboratory QNAP NAS servers and published data will also be available on Dryad.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

We currently have sufficient storage space on our servers. Maintenance costs to replace faulty hard drives or to upgrade the storage, as well as one time submission costs to Dryad will be covered by this project.

7. DATA SHARING AND REUSE

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

All published data will be made fully available and unpublished data will be available upon specific collaboration requests.

Where/how will the data be made available for reuse?

- In an Open Access repository (Dryad and NCBI)
- Upon request by mail (brain images)

When will the data be made available?

- Upon publication of the research results

The full datasets will be available on Dryad immediately after publication and raw data will be made available upon request.

Who will be able to access the data and under what conditions?

Published data will be uploaded in a csv format in Dryad as an open access dataset under a CC-BY license. Therefore, it will be available to anyone for any purpose, provided that they give appropriate credit to the creators. Unpublished data will be made available upon request on a collaboration project basis.

What are the expected costs for data sharing? How will the costs be covered?

No extra costs are expected, except a one time submission-fee to Dryad (when not covered by the scientific journal) that will be covered by this project.

8. RESPONSIBILITIES

Who will be responsible for data documentation & metadata?

The PI, Prof. Tom Wenseleers

Who will be responsible for data storage & back up during the project?

The PI, Prof. Tom Wenseleers

Who will be responsible for ensuring data preservation and reuse ?

The PI, Prof. Tom Wenseleers

Who bears the end responsibility for updating & implementing this DMP?

The PI bears the end responsibility of updating & implementing this DMP.