

Project title

Genome duplication in parasitic plants: features and consequences

Supervisors with affiliations

Mark Blaxter, Institute of Evolutionary Biology.

Alex Twyford, Institute of Evolutionary Biology.

Primary supervisor contact email

Mark.Blaxter@ed.ac.uk

Project background



Parasitism is a ubiquitous feature of the natural world, found in every major organismal group, and across every ecosystem. In plants, parasitism is an extremely successful strategy that has evolved independently on at least a dozen occasions. There are over 4000 parasitic species. These parasitic plants include crop pests (*Orobanche*, *Striga*) and ecosystem engineers that affect community assembly (*Rhinanthus* and many hemiparasites). Studies of parasitic plants are essential for improving food security and in understanding ecosystem services, while also providing important insights into the origin of novel evolutionary strategies.

A major goal in parasitic plant biology is to understand the genetic basis of the parasitic habit. Comparative transcriptomic analyses identified ~180 genes upregulated in the haustorium (the parasite attachment organ) shared between three divergent taxa (Yang et al., 2015). Many of these genes are the products of gene duplication that has occurred since divergence from non-parasitic relatives. This raises the question of whether genome duplication plays an important role in the origin of plant parasitism.

This project will use genomic tools to investigate the presence and evolutionary fate of duplicate gene copies in parasitic eyebrights (*Euphrasia*). These generalist parasitic plant species are extremely ecologically diverse, and include a number of recently evolved UK endemics of conservation concern. They are particularly well-suited to genomic analysis because they have small genomes compared to other parasitic plants. The research project will involve:

- assembling the nuclear genome of a diploid *Euphrasia* species from next generation and single-molecule data
- assembling the genome of a tetraploid *Euphrasia*
- performing gene finding and rich functional annotation using RNA-seq data
- comparing diploid and tetraploid *Euphrasia* (and related non-parasitic taxa) to identify retained duplicates and explore their functions and evolutionary histories
- using low-coverage and resequencing data from all 19 British native taxa and new lines collected from European populations that may represent independent genome duplication events to explore pattern and process of duplicate evolution across the genus.

These new insights will be essential for understanding the evolution of plant parasitism, and may represent a step towards developing genetic tools for control of devastating crop parasites.

Key research questions

What are the genomic changes that have occurred between parasitic *Euphrasia* and related non-parasitic lineages?

Which duplicated gene copies have been retained in the genome of tetraploid *Euphrasia*?

Do comparative evolutionary analyses of parasitic plants point towards a common set of genes involved in plant parasitism?

Methodology

1. Generating single-molecule, long read data for assembly of *Euphrasia* genomes
2. Deploying high performance computing toolkits (virtual "cloud computing") to perform assembly and annotation of diploid and tetraploid *Euphrasia*
3. Building a rich interactive database of *Euphrasia* and comparator genomes to present assembly and annotation data
4. Comparing the gene content and structure of *Euphrasia* and the related non-parasitic *Mimulus guttatus*
5. Field collection and sequencing of European *Euphrasia*, and analysis of genomic diversity
6. Broad-scale comparative analysis of parasitic and non-parasitic genomes

Training

A comprehensive training programme will be provided comprising both specialist scientific training and generic transferable and professional skills. Training will include: (a) bioinformatics in python and statistical analysis in R, (b) specialist training in genome assembly and annotation, including advanced cloud computing, (c) field sampling, plant identification, DNA extraction and next generation sequencing.

Requirements

The applicant should have an interest in bioinformatics and biodiversity, and a willingness to learn comparative genomic analysis. Prior experience in bioinformatics and statistical analysis in R would be beneficial but not essential.

Further reading or any references referred to in the proposal

Westwood, J.H., Yoder, J.I. & Timko, M.P. (2010). The evolution of parasitism in plants. *Trends in plant science*, 15, 227-235.

Yang, Z., Wafula, E.K., Honaas, L.A., Zhang, H., Das, M., Fernandez-Aparicio, M. *et al.* (2015). Comparative Transcriptome Analyses Reveal Core Parasitism Genes and Suggest Gene Duplication and Repurposing as Sources of Structural Novelty. *Molecular Biology and Evolution*, 32, 767-790.

Wang, X., gussarova, G., Ruhsam, M., de Vere, N., Metherell, C., Hollingsworth, P.M. *et al.* (2017). DNA barcoding British *Euphrasia* reveals deeply divergent polyploids but lack of species-level resolution. *bioRxiv* <https://doi.org/10.1101/164681>

A project summary

What are the genomic consequences of adopting a parasitic lifestyle in plants? Genome sequencing of eyebrights (*Euphrasia*) will reveal the role of gene and genome duplication in parasite evolution.