PROJECT TITLE: How have gene family duplications shaped the disparity and diversity of mammal clades?

DTP Research Theme(s): Living World

Lead Institution: The University of Bath

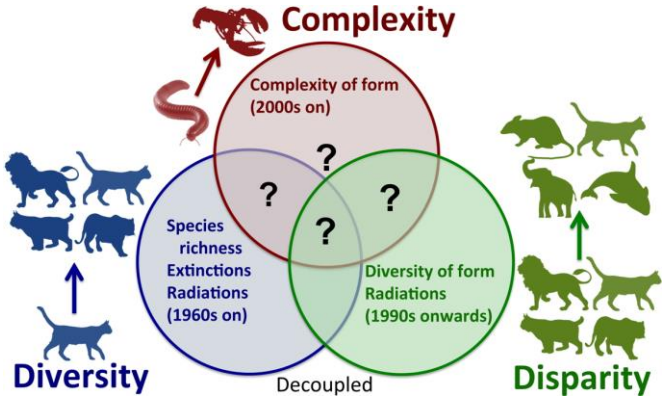
Lead Supervisor: Professor Matthew Wills, University of Bath, Milner Centre for Evolution

Co-Supervisor: Dr Araxi Urrutia, University of Bath, Milner Centre for Evolution
Co-Supervisor: Professor Martin Genner, University of Bristol, School of Biological Sciences

Project Enquiries: m.a.wills@bath.ac.uk

Project keywords: (please look at the list of FindaPhD subject areas provided and indicate which ones are relevant to your project)
Earth Sciences: Palaeobiology
Biological and Medical Sciences: Evoluiton, Zoology, Bioinformatics, Ecology and Conservation

The red or lesser panda (*Ailurus fulgens*) is a living fossil that is actually more closely related to weasels, otters and other mustelids than it is to the giant panda (*Ailuropoda melanoleuca*).



Concepts of **complexity, diversity and disparity**. ? = Areas targeted by this project.

Project Background

Understanding the forces that shape global biodiversity patterns was identified as one of the 25 greatest challenges for Science in the 21st Century (1, 2). Some mammal clades are enormously diverse, while their sister groups (originating at the same time) are far less so (e.g., 1,500 species of rodents versus 80 species of rabbits, hares and their allies). Why is this, and what might these differences tell us about the likely responses of groups to the present biodiversity crisis? In this strongly inter-disciplinary project, we will investigate the possible role of Small Scale Duplications of genes (SSDs) in shaping patters of diversity, anatomical complexity and morphological disparity of mammals in deep time.

As noted by Ernst Haeckel, developmental trajectories tend to become more complex with macroevolutionary time. Increasing interdependencies between genes and systems, coupled with their co-option for multiple functions (pleiotropy), result in more deleterious collateral consequences of mutations. This predicts that aspects of bodyplan design may become arbitrarily 'locked down' (e.g., seven neck vertebrae in most mammals), and that evolutionary innovation will be commonest when genetic redundancy is highest.

Small Scale Duplications of genes (SSDs) increase the number of genes within a given gene family (gene family size, GFS), and are one way in which mammals may circumvent such pleiotropic constraints and facilitate innovation.

Project Aims and Methods

The overarching objective is to understand the factors shaping the striking asymmetry of diversity within the tree of life, and within the mammals in particular. We will investigate the role of gene family duplicatons in this regard. However, the precise objectives and focus will depend upon the interests and aptitude of the student.

We will use annotated genome data from *Ensembl* and the Comparative Genomics database to map GFS onto large phylogenies. *OrthoMCL* groups will be used to investigate species with no *Ensembl* annotation. We will compile trait data from *Dryad* and *Morphobank*. We will identify 'shifts' in speciation rate and test for correlations between the rate of gene duplications, discrete morphological trait evolution and speciation rate.

We will differentiate single-branch rate shifts (e.g. rapid shifts in individual taxa) from shifts within clades (e.g. jumps to adaptive optima). We will model speciation, extinction, and trait evolution to account for variable sampling, evaluate different models of change (e.g. passive diffusion, early bursts, punctuation) and assess their fit to phylogenies via information criteria (e.g. Akaike; Bayesian), using a suite of *R* packages (e.g. *Auteur*; *Bayou*; *Ape*; *Gieger*; *Motmot*; *BAMM*; *Phytools*). If gene duplications underpin morphological evolution, branches with WGDs or higher GFS will have higher diversification rates, more morphological novelties and higher disparity than their sister clades.

Disparity will be quantified using discrete morphological data sampled from all aspects of anatomy (published repositories) and our own R scripts. Bodyplan complexity will be quantified using a variety of information statistics applied to both the serial differentiation of vertebrae and limb elements. Morphometric data will be collected where more nuanced shape quantification is needed (e.g. details of vertebrae, ribs, girdles and limbs.

Candidate Requirements

The project is suitable for anyone interested in macroevolutionary patterns, and wishing to pursue a career in bioinformatics, phylogenomics and big biological data.

+Training

The student with receive diverse training in bioinformatics, genomics, phylogenetics and statistical palaeobiology, as well as in methods for quantifying form and for mapping rates of trait changes across trees. They will have full access to the generic research and project specific training opportunities available through the DTP, in addition to a training budget the will allow the student to identify other, external training particular to their needs and interests.

The strongly interdisciplinary project combines genomic, bioinformatics (Urrutia and Genner) and palaeontological methods and approaches (Wills), and is a collaboration between the Milner Centre for Evolution (Bath) and the Department of Life Sciences (Bristol).

We will meet as a supervisory team every other month, and the student will be encouraged to attend the thriving evolution discussion groups in both Bristol and Bath. The physical proximity of the labs makes this eminently achievable.

References / Background reading list

1. Kennedy D & Norman C (2005) *Science* 309:75.
2. Pennisi E (2005) *Science* 309:90.
3. Hughes M, Gerber S, & Wills MA (2013) *PNAS* 110(34):13875-13879.
4. Oyston J, Hughes M, Gerber S, & Wills MA (2016) *Ann Bot* 117:859-879.
5. McShea D & Brandon RN (2010) *Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems* (University of Chicago Press, Chicago) p 170.
6. Adamowicz SJ, Purvis A, & Wills MA (2008). *PNAS* 105(12):4786-4791.
7. Wills MA, Briggs DEG, & Fortey RA (1998) *Systemat Assoc Spec Vol* 55:57-65.
8. Wills MA, Briggs DEG, & Fortey RA (1994). *Paleobiology* 20(2):93-130.
9. Ruta M, Angielczyk KD, Froebisch J, & Benton MJ (2013) *Proc R Soc B* 280(1768).
10. Davis KE, Hill J, Astrop TI, & Wills MA (2016) *Nature Comm* 7:13003.

Useful links

Enquiries relating to the project should be directed to the lead supervisor (see email address above for Project Enquiries). Enquiries relating to the application process should be directed to doctoraladmissions@bath.ac.uk

In order to apply, you should select the relevant University of Bath PhD online application form found here: https://www.bath.ac.uk/study/pg/applications.pl. When completing the form, please state in the 'Finance' section that you wish to be considered for GW4+ DTP funding and quote the project title and lead supervisor's name in the 'Your research interests' section.

Further information about the application process may be found here: http://www.bath.ac.uk/topics/postgraduate-research/

The application deadline is 1600 hours GMT Monday 7 January 2019 and interviews will take place between 4 and 15 February 2019. For more information about the NERC GW4+ DTP, please visit https://nercgw4plus.ac.uk.