Annotations on what makes this grant particularly great are in Post-Its. These Post-Its should be viewable in Preview on Mac or Adobe Acrobat Reader on any operating system.

The evolution of wing pattern mimicry in *Hypolimnas* butterflies

The identification of the genetic basis of ecologically relevant traits is a key goal in evolutionary biology. Wing pattern mimicry is an adaptive trait commonly found in Lepidoptera[1]. In the *Hypolimnas* genus of butterflies, two species, *H. misippus* and *H. bolina*, present female limited polymorphic Batesian mimicry[2,3]. While in Southeast Asia, *H. bolina* is a recognised Batesian mimic of multiple *Euploea* species, in the south Pacific Islands and Australia, it presents several non-mimetic forms. The genetic basis of its wing pattern was extensively studied by Clarke and Sheppard, who hypothesised that the polymorphism was determined by two autosomal loci[3]. Similarly, *H. misippus* presents detailed resemblances to the four morphs of the unpalatable *Danaus chrysippus* and its wing pattern is determined by three autosomal loci[4,5]. However, maladaptive intermediate morphs are also commonly found[6]. Altogether, this suggests that the current selection for mimicry in these species is weak and that other evolutionary forces might be at play[7,8].

One possibility that has been suggested for *H. misippus* is that selection for mimicry fluctuates over time following changes in population size[6,7]. Changes in the strength of selection lead to the maintenance of mimetic forms, but also allow for the presence of intermediates. Another possibility is that the genetic architecture of the trait itself leads to the maintenance of the polymorphism, such as in balanced lethal systems[9]. For example, in the tropical mimetic butterfly *Heliconius numata,* wing pattern polymorphism is controlled by a supergene containing three inversions, which has resulted in the accumulation of deleterious mutations and the reduction of fitness in homozygotes compared to heterozygotes[10]. This heterozygote advantage ensures the maintenance of the two alleles in the population. Therefore, improving our understanding of the genetic basis of mimicry in *Hypolimnas* butterflies will help elucidate the origin and maintenance of mimicry in this genus.

In my proposed project, I will clarify the evolutionary history of mimicry in *Hypolimnas* butterflies by achieving two goals.

Goal 1: Determine the origin of the loci controlling wing pattern in *Hypolimnas*.

Goal 2: Explore how mimicry is being maintained.

First, I will examine the evolution of mimicry within the genus, for which there are three hypotheses: 1) mimicry evolved only once and thus the genes controlling wing pattern in both species are the same; 2) mimicry evolved independently in each species and the wing pattern determining loci are different; and 3) mimicry evolved independently in each species and the wing pattern determining loci are the same. By identifying the genetic basis of wing pattern in *H. bolina* and *H. misippus* and comparing them, I will determine which of these scenarios is more plausible. Second, by comparing

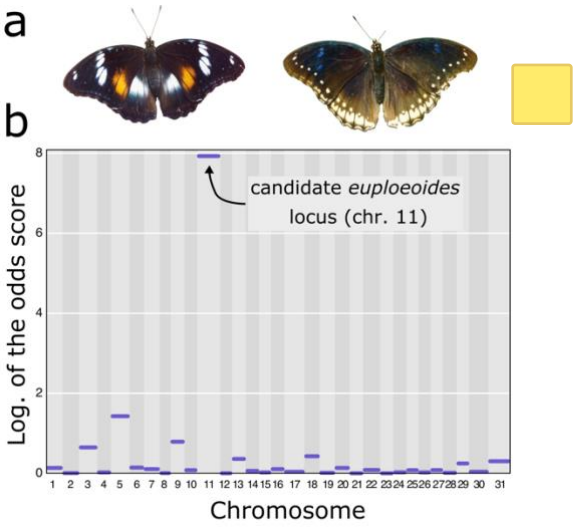

Figure 1: Wing pattern mimicry in *Hypolimnas bolina.* **a** Morphs *nerina* (left) and *euploeoides* (right). **b** Result of a linkage map showing that chromosome 11 contains the locus controlling the *euploeoides* morph.

the genetic basis of colour in *Hypolimnas* to other Lepidoptera species, I will clarify whether the same genes as in other species have been re-used once again to produce wing pattern phenotypes or whether it is determined by different loci. Finally, I will shed light on how the polymorphism is being maintained by checking for signatures of selection or inversions in the candidate loci.

To achieve my aims, I have produced chromosome level reference assemblies for *H. bolina* and *H. misippus,* and collected and sequenced 333 *H. misippus* adult females varying in wing phenotype. Also, I have sequenced two families of *H. bolina* individuals and located in the genome one of the loci controlling wing pattern (Fig. 1). Importantly, this locus is situated in the same chromosome as *cortex*, a gene controlling wing pattern in many other Lepidoptera species[11]. However, sequencing of wild samples is needed to narrow down the region and identify candidate genes. For that, I will sequence 70 individuals of each of the three main *H. bolina* morphs (*nerina*, *naresi* and *euploeoides*) with a total of 210 samples. I will produce low-coverage whole genome sequences of these samples and then, perform a Genome Wide Association Study using ANGSD[12] to identify the loci controlling wing pattern. Sequencing a large number of individuals at low coverage ensures a sample size large enough for the association study while minimising sequencing costs[13]. I have tested this approach in *H. misippus* with successful results (Fig. 2). Finally, I will compare my *H. bolina* and *H. misippus* results.

Once I have identified the candidate loci in both species, I will use the wild *H. bolina* and *H. misippus* samples to explore how the polymorphism is being maintained. First, I will check for evidence of balancing selection in the candidate loci, which would be expected if selection for mimicry fluctuates over time. To do that, I will look for excess of non-synonymous mutations, and explore statistics like Tajima's D and HKA, that indicate balancing selection[14]. Then, I will check for the presence of inversions in the candidate loci, as those are typically found in supergenes and would explain why multiple elements of the wing pattern are inherited together[15]. Also, the presence of a mechanism reducing recombination such as inversions could ensure the maintenance of polymorphism in the absence of strong selective pressures. Importantly, one of the loci controlling mimicry in *H. misippus* has been hypothesised to be a supergene[16].

Overall, by identifying genes controlling wing pattern in *H. bolina* and *H. misippus*, I will elucidate the evolutionary origin of mimicry in *Hypolimnas* and shed light onto the maintenance of mimicry in this genus.

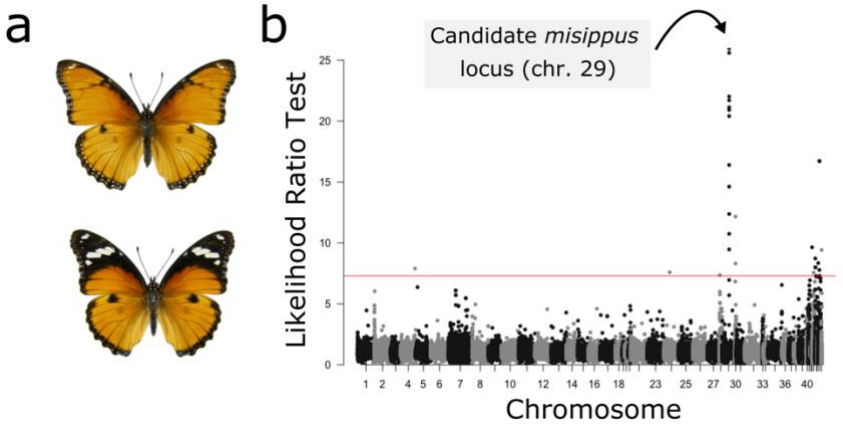

Figure 2: Wing pattern mimicry in *Hypolimnas misippus.* **a** Morphs *inaria* (top) and *misippus* (bottom). **b** Genome Wide Association Study of the locus controlling the misippus morph, located in Chromosome 29[5].

References

1. Quicke, D. L. J. *Mimicry, Crypsis, Masquerade and other Adaptive Resemblances*. (John Wiley & Sons, 2017).
2. Smith, D. A. S. Phenotypic diversity, mimicry and natural selection in the African butterfly Hypolimnas misippus L. (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society* **8**, 183–204 (1976).
3. Clarke, C. & Sheppard, P. M. The genetics of the mimetic butterfly Hypolimnas bolina (L.). *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* (1975) doi:10.1098/rstb.1975.0084.
4. Gordon, I. J. & Smith, D. A. S. Body size and colour-pattern genetics in the polymorphic mimetic butterfly Hypolimnas misippus (L.). *Heredity* **80**, 62–69 (1998).
5. VanKuren, N. W., Massardo, D., Nallu, S. & Kronforst, M. R. Butterfly Mimicry Polymorphisms Highlight Phylogenetic Limits of Gene Reuse in the Evolution of Diverse Adaptations. *Mol Biol Evol* **36**, 2842–2853 (2019).
6. Gordon, I. J. Natural-Selection For Rare and Mimetic Color Pattern Combinations in Wild Populations of the Diadem Butterfly, Hypolimnas-Misippus L. *Biological Journal of the Linnean Society* **31**, 1–23 (1987).
7. Gordon, I. J., Edmunds, M., EDgar, J. A., Lawrence, J. & Smith, D. A. S. Linkage disequilibrium and natural selection for mimicry in the Batesian mimic Hypolimnas misippus (L.) (Lepidoptera: Nymphalidae) in the Afrotropics. *Biological Journal of the Linnean Society* **100**, 180–194 (2010).
8. Clarke, C. A., Clarke, F. M. M., Gordon, I. J. & Marsh, N. A. Rule-breaking mimics: palatability of the butterflies Hypolimnas bolina and Hypolimnas misippus, a sister species pair. *Biological Journal of the Linnean Society* **37**, 359–365 (1989).
9. Muller, H. J. Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors. *Genetics* **3**, 422–499 (1918).
10. Jay, P. *et al.* Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nature Genetics* 1–6 (2021) doi:10.1038/s41588-020-00771-1.
11. Nadeau, N. J. *et al.* The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* **534**, 106–110 (2016).
12. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* (2014) doi:10.1186/s12859-014-0356-4.
13. Therkildsen, N. O. & Palumbi, S. R. Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. *Molecular Ecology Resources* **17**, 194–208 (2017).
14. Fijarczyk, A. & Babik, W. Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **24**, 3529–3545 (2015).
15. Thompson, M. J. & Jiggins, C. D. Supergenes and their role in evolution. *Heredity* **113**, 1–8 (2014).
16. Gordon, J. S. & Smit, D. a S. Genetics of the mimetic African butterfly Hypolimnas misippus: hindwing polymorphism. *Heredity* **63**, 409–425 (1989).

Budget and budget justification

The 210 *Hypolimnas bolina* samples that will be sequenced have already been collected. Thus, the only costs of the project are towards DNA sequencing. Sequencing will be done with Novogene UK, as it provides the most affordable option. Details of the samples, genome size, sequencing strategy and cost are stated below.

Number of samples	210 adult female *Hypolimnas bolina*
Genome Size	440 Mb
Coverage	8.6 X
Lanes	1 lane
Number of reads to be sequenced	2,667 million read pairs
Amount of data	800 Gb
Total cost	£5,200 ($7,110.25)*

*$1.37 exchange rate from GBP to USD (2 February, 2021).

Other sources of funding:

Costs of this project exceed the award from the society. Thus, to fully cover the costs of this project, I have applied and will apply to other sources of funding.

First, I have applied to the Lepidopterists' Society Ron Leuschner Memorial Fund for Research, which provides an award of $500.

Second, I will apply to the Varley-Gradwell Travelling Fellowship in Insect Ecology, which provides up to £2,500 ($3,418.39*) to fund research in insect ecology. Despite being a travelling fellowship, the Varley-Gradwell Fellowship can fund other activities, such as sequencing, as long as the research is focused on insect ecology.

Finally, for my PhD research, I was awarded a NERC research grant. The remaining costs of this project will be covered with it.