Generation of genomics resources for the meadow spittlebug *Philaenus spumarius*

Roberto Biello*[,1], Thomas Mathers[1], Qun Liu[1], Ana S. B. Rodrigues[2], Ana Carina Neto[3], Maria Teresa Rebelo[3], Octávio S. Paulo[2], Sam Mugford[1], Sofia Seabra[2], Saskia Hogenhout[1]

[1]*Department of Crop Genetics, John Innes Centre, Norwich Research Park, Norwich, United Kingdom*

[2]*Centre for Ecology, Evolution and Environmental Changes (cE3c), Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal*

[3]*Centre for Environmental and Marine Studies (CESAM), Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal*

Philaenus spumarius (Linnaeus, 1758) and other spittlebug species (order Hemiptera) are xylem-feeding insects that vector the bacterial pathogen *Xylella fastidiosa*, which has caused dramatic declines of a wide variety of plant species and most recently destroyed the olive production industry in southern Italy. *P. spumarius* is native to the Palearctic regions, and was unintentionally introduced in other regions, e.g. USA and New Zealand. The insect quickly adapts to new environments, and may be considered invasive. As a first step to improve our knowledge of the evolution and genomics of this species, we generated a draft assembly of the 2.7 Gb genome of *P. spumarius*. The assembly has a scaffold N50 length of 120Kb with the longest scaffold of 4 Mb and a high (2.3%) heterozygosity level. Annotation of the genome revealed a highly repetitive genome and presence of 97% of conserved genes among arthropods suggesting a complete genome assembly. The assembly of the genome of a second froghopper species, *Aphrophora salicina*, revealed a genome size of 2.3 Gb. This assembly had a scaffold N50 length of 625Kb with the longest scaffold of 5 Mb and a lower (0.7%) heterozygosity level.

We started to generate genome assemblies of another 16 potential vectors of *X. fastidiosa* as well. In the future these will form the basis of comparative genome analyses to assess differences between phloem and xylem feedings insects within Hemiptera. They will also allow assessing the population structure and putative dispersal of these species between host plants. This research is part of a larger UK-government funded collaborative project (named BRIGIT; https://www.jic.ac.uk/brigit) and all genome sequence data and annotations generated in this project will be made freely available via http://sapfeederhub.jic.ac.uk/.

Detecting hybridisation from whole genome data in eyebrights

Max Brown*(S), Hannes Becher, Alex Twyford

University of Edinburgh

Hybridisation is a common phenomenon in plants and has the potential for a large range of evolutionary outcomes. In the UK, hybridisation across a ploidy level in Eyebrights has led to at least two hybrid species. These hybrid species have been postulated based on morphology, ecology and some limited early DNA work. Here, we use whole genome data of both the parental species and the hybrid species to detect signatures of hybridisation across the genome.

Genomic introgression in *Oreochromis* tilapias

Adam Ciezarek*, Antonia Ford, Tarang Mehta, Will Nash, Luca Penso-Dolfin, George Turner, Martin Genner, Federica Di-Palma and Wilfried Haerty

Earlham Institute

The *Oreochromis* tilapias are an economically important group of fish for aquaculture, whose production has expanded dramatically in the last two decades. A direct consequence of this success has been the introduction of exotic species in Tanzania, a hotspot for *Oreochromis* diversity. Introductions have had significant negative ecological effects on indigenous *Oreochromis* species, including species displacement and loss of population structure through hybridisation, having potential implications for local adaptations. This history of introgression has made untangling the population history of *Oreochromis* difficult. We address this issue using genome-wide sequencing data across 600 individuals from 29 *Oreochromis* species from across Tanzania and east Africa, identifying strong signatures of introgression between species at both the population and individual level. In addition to ancestral introgression, many hybrid individuals were found demonstrating ongoing gene flow in water bodies where species are sympatric, in some cases due to deliberate introductions. We quantified the impact of both modern gene flow and ancestral introgression on genomic diversity and differentiation, and investigated whether some genomic regions are resistant to introgression. Such regions may maintain adaptive variation, and therefore be important for population viability. We further report a new *Oreochromis* species from the Ruaha basin, previously identified as a population of *O. rukwaensis*. We anticipate that our results will have important implications when managing the translocation of species for food production.

The transcriptomic basis of cold acclimation across the *Drosophila* genus

N Cook*, Parker DJ, Smith TK and Ritchie MG

The University of St Andrews

Species of the *Drosophila* genus are ideal to examine both phenotypic evolution and its mechanistic basis due to their broad ecological range, including climatically diverse environments. In addition, the phylogenetic history of many species is well understood, facilitating genotype-phenotype mapping on an evolutionary scale. The ability to acclimate to seasonal cold has important fitness consequences for all organisms allowing them to maintain physiological function for longer as the temperature drops. The physiology of cold adaption is well understood but we know very little about the underlying mechanisms of this trait and how those mechanisms vary with species. Here, using an RNA-seq approach combined with metabolite profiling, we compare the changes associated with cold acclimation in nine *Drosophila* species in a phylogenetic framework. This data allows us to explore the evolutionary consequences of this adaptation.

Genetic background of a lethal posthitis disease in the European bison

Sylwia D. Czarnomska*, Shazia Kunvar, Małgorzata Tokarska

Centre of New Technologies, University of Warsaw, Stefana Banacha 2c, 02-097 Warsaw, Poland

European bison (*Bison bonasus*), the largest terrestrial mammal in Europe, went through a severe bottleneck and became extinct in the wild in early 1919. The lowland line (Bialowieza line) originates from only seven founders with an extremely varying genetic contribution (approximately 80% of the genes in contemporary populations come from just two individuals). European bison has been the subject of extensive studies and dramatic impoverishment of genetic variability has been confirmed using a variety of methods (mtDNA, MHC, microsatellites, SNPs). Suprisingly, population from Białowieza Forest does not show typical inbreeding depression symptoms (lowered reproduction parameters, health or viability problems), contradicting the notion that the European bison is in a precarious genetic situation. The only severe disease observed in the population is posthitis (necrotic inflamation of the prepuce), which affects approx. 6 % of males a year but does not show correlation with inbreeding level. The etiology of posthitis is unknown and latest hypotheses suggests mainly environmental and endemic factors. However, recent studies on European bison using Illumina Bovine HD 777 K microarray indicated significant associations between certain SNP markers and the occurence of the disease.

In order to further investigate genetic background of the disease we used RADseq technology and target sequencing of candidate genes, previously identified as associated with incidence of posthitis, in European bison samples acquired over several decades by Mammal Research Institute PAS in Bialowieza. The scientific objectives of the current project are: to identify species-specific SNPs associated with postithis, to verify the potential role of sex chromosomes in the pathogenic process and to design an informative and cost-effective SNP-chip based genetic test that allows to verify susceptibility to posthitis, as well as carriers of unfavorable variants.

Darwin Tree of Life Project: Genomes for all

Richard Durbin* on behalf of the Darwin Tree of Life Project

University of Cambridge

DNA sequencing is a transformative technology for biology, but until recently high quality genome reference sequences have been limited more or less to human, model organisms and microbes. This is changing rapidly. Advances in sequencing platforms and related technologies over the last few years now enable the generation of highly complete and accurate chromosomal reference genomes at increasing scale and reducing cost. It is possible to extrapolate these advances to conceive sequencing genomes for all accessible eukaryotic species in the next 10-15 years. To advance this goal the Darwin Tree of Life project aims to sequence the genomes of all 60,000 species of eukaryotic organisms in the north-west European islands of Britain and Ireland, collaborating with others working on other biota. A consortium of sequencing centres, museums, botanical gardens, institutes and universities has now been funded to start this venture. We have started collecting samples, sequencing, and releasing assemblies. This poster will summarise our structure, plans and progress, and provide a focal point at the meeting to collect input on taxa to prioritise and other strategic and practical issues. We can't do everything that everyone would like, but we want the Darwin project to deliver for the community, and we hope it will have community buy-in and engagement.

https://www.darwintreeoflife.org/

The effect of genetic architecture and selfing on the capacity of a population to expand its range

Martin Eriksson* (S), Marina Rafajlović

University of Gothenburg

Previous theoretical work on range expansions over heterogeneous environmental conditions showed that there is a critical environmental gradient where range expansion stops. For populations with freely recombining loci underlying the trait under selection (hereafter adaptive loci), the critical gradient, for one-dimensional habitats, depends on the fitness cost of dispersal, and the strength of selection relative to genetic drift. Here, we extend the previous work and ask: What is the role of the genetic architecture of adaptive loci during range expansions in populations with and without selfing? To answer this question, we use computer simulations of range expansions over steepening environmental gradients. We demonstrate that, while tighter genetic architectures slow down range expansions due to poor purging of locally deleterious alleles at the expansion front, they may also allow a species to occupy a greater range. In fact, for some parameter values, we find that a population with freely recombining adaptive loci experiences global extinction, whereas tighter genetic architectures allow for a successful expansion over a wide geographic range. In addition, we find that allowance of selfing may improve the ability of populations to expand their ranges. We discuss the mechanisms underlying these results.

Applying genomic analysis to the conservation of mountain gorillas (*Gorilla b. beringei*)

Ettore Fedele* (S), Jon Wetton, Mark Jobling

The University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom

Non-human great ape species are key to the maintenance of healthy ecosystems that support millions of people around the world. Despite their ecological importance, great apes are globally under threat due to habitat loss and fragmentation, hunting, and illegal wildlife trade. Therefore, monitoring changes in population size is of primary importance in order to develop effective action plans and assess the effectiveness of current conservation measures. The development of novel techniques for individual identification in population studies represent a corner stone in ecological research. In this context, field genomics, a relatively new discipline, has the potential to effectively support the conservation and monitoring of the species around the world.

The current project aims at using genetic markers (STRs and SNPs) for individual identification of great apes, with a particular focus on mountain gorillas (*G. b. beringei*), currently numbering 1000 individuals occurring in two protected areas in East Africa (i.e. Virunga National Park and Bwindi Impenetrable National Park). The Illumina® ForenSeq™ DNA Signature Prep Kit targeting human autosomal, Y- and X-chromosomal STRs and SNP markers was tested on 60 samples belonging to the two genera *Pan* and *Gorilla*. The data were used to assess the molecular evolution of STRs in different species, as well as power of discrimination and heterozygosity, central to the understanding of the resilience and adaptability of the remnant populations to various pressures and threats. The second part of this project is focused on the development of a SNP panel for species and individual identification to be used in non-invasive DNA samples (e.g. hair and faeces), exploiting the ONT™ MinION™ for in situ sequencing thus greatly reducing the time and costs of analysis while equipping local authorities with a powerful tool in the fight against poaching and illegal wildlife trade.

The Genetic Diversity of the Malaria Parasite *Plasmodium falciparum* in Two Sympatric Mosquito Species

A. J. Forster*(S), S. M. O'Loughlin & A. Burt

Imperial College London, London UK

Malaria remains a significant public health issue in Sub-saharan Africa (SSA). Burkina Faso and Mali are two countries with high malaria incidence; certain regions in Mali, particularly in the south where it shares a border with Burkina Faso, reach an average of more than 300 cases per 1000 individuals. In Burkina Faso this is true across the whole country. Two members of the *Anopheles gambiae s.l.* species complex are responsible for the majority of malaria transmission in SSA; *An. gambiae s.s.* and *An. coluzzii*. In Burkina Faso and Mali their distributions overlap however, although capable, they rarely interbreed and significant ecological and behavioural differences differentiate the sister species. A significant amount of research has been invested in the coevolutionary interplay between humans and *Plasmodium spp.* however the relationship between *P. falciparum* and its vector and how this influences transmission is less clearly defined.

Assessing the genetic diversity of *P. falciparum* indicates local transmission intensity and can signify the efficacy of control methods. In the past various loci have been used to assess genetic diversity in infections carried by humans. However, its assessment in vector-borne isolates has not, to our knowledge, been attempted. This not only disregards a key stage of the life of the parasite but may be a less invasive measure to employ in the field. In this study we will attempt to assess the genetic diversity of *P. falciparum* carried by wild *An. gambiae s.s.* and *An. coluzzii* collected in Mali and Burkina Faso by sequencing two Merozoite Surface Protein (MSP) gene loci. The results will be used to infer the transmission intensity in each location and whether differences in genetic diversity between two neighbouring high transmission countries are detectable. We will also compare parasites carried by each vector species to assess any differences.

Molecular and Bioinformatic Investigations into the Mating Behaviour of the Malaria spreading *Anopheles gambiae* Mosquito

A. J. Forster*(S), S. M. O'Loughlin & A. Burt

Imperial College London, London UK

The mating status of mosquitos has important implications for malaria. Methods of genetically modifying *Anopheles* mosquitoes as a means of reducing the impact of malaria are highly sought after and gene drive, which positively biases the inheritance of certain genes, is a primary focus of this research. To facilitate the implementation of gene drive it is important to understand mosquito mating behaviour as changes in mating frequency, number of mates and levels of interspecific mating between members of the *An. gambiae s.l*. complex all have implications for its spread. In addition to this it was recently found that females are more susceptible to infection by the malaria parasite *Plasmodium falciparum* when they are mated due to physiological changes induced by hormones transferred in sperm.

In the past mating frequency has been determined by an arduous process of dissecting females and checking for the presence of sperm. In 2016 the basic structure of the *An. gambiae* Y chromosome was characterised providing scope for a molecular technique. A polymerase chain reaction (PCR) based method of determining mating status in whole DNA mosquito extractions was developed. Although male specific DNA is rare, mini- and microsatellites have also been found to be strongly Y-linked in *An. gambiae* and differences in coverage of these regions in the sequenced genomes of wild females might indicate whether the female is mated. The coverage of certain minisatellites were found to be significantly greater in females deemed to be mated by PCR indicating that a bioinformatically based method of determining mating frequency in large scale genomic data sets may be a possible. This and a PCR based method of determining mating status in females will facilitate investigations into the mating status of mosquitos in the wild to shed light on its effects on malaria transmission.

Genome-wide site freqency spectra and reproduction in the high-fecundity Atlantic cod

Katrín Halldórsdóttir*, Bjarki Eldon, Jere Koskela, Einar Árnason

University of Iceland

The question about the mode of reproduction of an organism is essentially an ecoogical question. However, for an organism like the Atlantic cod, a highly fecund organism, its mode of reproduction cannot be determined using field research. Instead, a genomics approach is called for. We hypothesize that Atlantic cod reproduces by sweepstakes. To understand this mode of reproduction we turn to genomics. But what theoretical models describe its reproduction? Here we estimate parameters in two reproductive models, the Fisher-Wright and the Schweinsberg's model of reproduction by estimating parameters using coalescent methods. Does the Atlantic cod fit the classical Fisher-Wright model of reproduction, the Fisher-Wright model which includes population growth, the beta coalescent, the Bolthausen-Sniztman coalescent or something else? New results based on whole genome sequencing of 79 individuals of cod from a single locality in South-East Iceland and estimation of the parameter α in the various coalescent models will be discussed.

Evidence for a young social chromosome in a socially polymorphic ant

Max John* (S), Ben Braim, Bethan Hill, Jason Stevenson, Rob Hammond

Department of Genetics and Genome Biology, University of Leicester

Ants are notable for their ecological success, dominating terrestrial ecosystems on every continent on Earth. The >12,000 described species of ants have evolved a remarkable diversity of lifestyles and modes of social organisation. While much of this variation is interspecific, many species display polymorphic social organisation - the number of queens in a nest varying between one (monogyny) and many (polygyny). The techniques now exist to untangle the genetic and genomic basis of these complex behavioural polymorphisms. The work presented here describes evidence for a "social chromosome" underpinning a social polymorphism between "functional" monogyny vs polygyny in the ant *Leptothorax acervorum*. A social chromosome is an autosome bearing two heteromorphic, canonically non-recombining, supergene haplotypes, each containing many individual genes (similar to the non-recombining region of the Y chromosome). The inheritance of these supergenes determines the social phenotype of an ant colony. Using genome assemblies produced with 10X Genomics linked reads, together with pooled RADseq data from multiple populations, we are able to show that a large contiguous region (>10Mb) of *L. acervorum*'s genome contains most of the allelic variation associated with this ant's polymorphic social organisation. Between social phenotypes, this region shows elevated diversity when compared with the rest of the genome, consistent with selection-driven divergence. Within social phenotypes, this region shows reduced diversity when compared with the rest of the genome, again consistent with positive selection. Comparing between social phenotypes, microsatellite genotyping shows divergent patterns of allelic diversity within the social region, with significantly reduced allelic richness in the social region of individuals from polygynous populations. RNAseq-derived predictions of gene and regulatory sequences will facilitate an in-depth analysis of population-scale, individual-level whole-genome resequencing data. Two examples of similar genomic architecture being convergently evolved have been described previously in other species of ant (*Solenopsis invicta* and *Formica selysi*), making *L. acervorum* the third example known.

RNA Editing in Eusocial Hymenoptera

Alun Jones*(S) and Eamonn Mallon

University of Leicester

The allocation of reproductive work is a hall mark of eusocial societies. The task of reproduction is often attributed to a few or even a single individual within a colony. These individuals can have specific phenotypes that help with reproduction, such as increased size compared to non-reproductive members, but have to produce these phenotypes from the same genome as their non-reproductive counterparts. The ability to be reproductive is, more often than not, a flexible trait with differing levels of flexibility associated with the complexity of the eusocial society. How reproduction is controlled is unclear. Whilst changes in gene expression, novel genes and methylation differences have been associated with reproductive castes, it is still unknown if there is a conserved underlying mechanism or, if varying regulatory mechanisms occur in different lineages.

RNA editing is the conversion of one RNA base to another, the most common being A to I editing by adenosine deaminase. RNA editing has the potential to change protein function, by altering coding regions, but also altering expression through 3 prime UTR modifications. This flexibility means RNA editing is key for highly variable processes such as neurological regulation and immune responses. The allocation of reproduction in eusocial insects, with its variation in flexibility, has been shown to involve RNA editing in leaf cutter ants but it is unknown in other species.

Our aim is to use NGS data sets, containing DNA and RNA, to identify RNA editing in different eusocial Hymenoptera. By identifying sites where editing occurs and how it differs between reproductive castes we aim to identify 1) If editing is conserved as a mechanism for reproductive division of labour 2) If the genes targeted for editing are conserved across species 3) Whether the target of editing shifts depending upon reproductive flexibility.

Intercontinental genomic parallelism in multiple adaptive radiations

Isabel S. Magalhaes*, James R. Whiting, Andrew D.C. MacColl

University of Roehampton, Dept. of Life Sciences, Whitelands College , London , SW15 4JD

Parallelism, the evolution of identical traits in populations diversifying in similar conditions, provides good evidence of adaptation by natural selection. Adaptive radiations are celebrated as outbursts of biological diversity on the tree of life and are useful in the study of parallelism because of often substantial evolutionary replication. Many studies of parallelism have focused on comparisons of strongly different ecotypes or sharply contrasting environments, defined a priori, which could upwardly bias the apparent prevalence of parallelism. In our study we used a RADseq approach to test the extent of genomic parallelism associated with continuous variation in both environments and phenotypic traits, across four independent adaptive radiations of three-spined stickleback fish (*Gasterosteus aculeatus*) spanning the northern hemisphere. We used quantitative characterizations of environmental and phenotypic variation, which are often lacking from such studies, and quantified convergence of these axes across continents.

We found substantial evidence that evolution within radiations is associated with even relatively modest environmental change. Genomic parallelism is significantly greater than expected by chance for several variables and similar in its extent for phenotypic traits and environmental variables. Genetic similarity appeared to be the best predictor of the extent of parallelism between adaptive radiations with intra-continental pairs of radiations the strongest sources of parallelism. However, we also found evidence of common environments promoting genomic parallelism coming from the common acid-alkali axis experienced by all radiations, and the observations that Calcium and pH were associated with strong patterns of gene-reuse. Overall, our results show that genome-wide evolution continues to be repeatable at intercontinental scales and after hundreds of thousands of years of divergence, and provide useful insight into factors likely to influence genomic divergence at large geographical scales.

Evolutionary genetics of wing pattern mimicry in Hypolimnas butterflies

Anna Orteu*(S), Chris Jiggins

Department of Zoology, The University of Cambridge, Downing Street, Cambridge CB2 3EJ

The identification of the genetic basis of ecologically relevant traits is a key goal in evolutionary biology. Wing pattern in Lepidoptera is a classic example of adaptive evolution. Recent genomic analyses have identified the genetic basis of wing pattern in many butterfly and moth species. Surprisingly, on many occasions the same genes control wing pattern formation in multiple species, such as the gene *cortex* in butterflies and moths, and *doublesex* in the *Papilio* genus. The diadem butterfly, *Hypolimnas misippus*, is a complex example of Batesian mimicry and provides an exceptional opportunity to explore the role of convergence in wing pattern mimicry in Lepidoptera. Through female-limited polymorphic mimicry, the diadem butterfly mimics the four morphs of the African queen, *Danaus crysippus*. However, the mismatch of morph frequencies between the model and the mimic across Africa suggest that other selective forces act on the wing phenotypes. Wing pattern in *H. misippus* is controlled by 3 autosomal loci with epistatic relationships between them, and the combination of alleles produce multiple wing pattern morphs, of which many are non-mimetic and maladaptive. Clarifying the genetic basis of wing pattern in *H. misippus* will shed light on its complex evolutionary history. The sister species *H. bolina*, presents many similarities to *H. misippus*, such as female-limited polymorphisms and mimicry. Crucially, in *H. bolina*, *doublesex* has been identified as a suppressor of male-killing *Wolbachia*, evidence of the extreme flexibility of such a vital developmental gene. By identifying genes controlling wing pattern in *H. bolina* and *H. misippus*, I can elucidate which evolutionary forces act upon colour pattern in these species and explore the role of genetic convergence in the evolution of mimicry in butterflies.

Heritability of intra-individual variation in body temperature in fasted heterothermic rodents.

Rohan Raval*(S), Jan Boratyński, Maria Luisa Martin Cerezo, Marek Kucka, Yingguang Frank Chan, Karol Zub and Jarosław Bryk

University of Huddersfield, Queensgate, Huddersfield HD1 3DH

Minimising energetic burdens when food availability is severely limited is suggested as an adaptive response to starvation. Heterotherms can decrease body temperature (T_b) and enter into a state of torpor to reduce energy expenditure when food is scarce. While intra-individual variability in T_b has been shown as repeatable, its heritability in wild populations is not yet established. We are undertaking a large-scale study in a population of wild yellow-necked mice, *Apodemus flavicollis*, to uncover the genetic basis and heritability of thermal and metabolic responses to fasting. We have measured T_b, basal metabolic rate (BMR) and torpor (variable n ~150-500) in multiple generations of animals from a single population in the primeval Białowieża forest, north-eastern Poland. Each captured individual was genotyped (n ~ 700) using a modified ddRAD-seq protocol to reliably determine pedigrees and therefore reproductive success in the population over time, giving a direct measure of the fitness of each individual and phenotype. Subsequently, the genome-wide, high density markers allow us to conduct genotype-phenotype associations. We aim to uncover whether T_b in fasted mice reduces their energy expenditure, whether the observed intra-individual variability is heritable, and how both affect fitness in the wild. This study is one of the first to investigate the heritability of thermal and metabolic traits on a large scale and integrate genomic, evolutionary and ecological information about complex phenotypes on a whole population level.

Using doubleton variants and their ages to investigate demography and selection in the *Anopheles gambiae* species complex

Josh Reynolds*(S), Vassiliki Koufopanou, Austin Burt

Department of Life Sciences, Imperial College London, Silwood Park, Ascot, Berkshire, United Kingdom, SL5 7PY

Rare variants are known to be more differentiated than common variants in populations, and thus more revealing about underlying population structure. Doubletons – alleles which only occur in two individuals in a sample population – are among the most informative of all rare variants, and their ages have been found to provide detailed measures of population structure. Mosquitoes of the *Anopheles gambiae* species complex represent the main vectors of malaria in Africa, and their effective management requires detailed knowledge of population interactions. Here, doubleton sharing at sites under different selection pressures is compared, using data from the *Anopheles gambiae* 1000 Genomes Project. Adapting the framework of Mathieson and McVean (2014), the age of doubletons from different degenerate sites and populations are approximated to investigate population stratification and selection. We find that doubleton counts are elevated in 0-fold and private sites, suggestive of the effect of selection and demography on population differentiation, and show that the ages of these mutations are younger. Finally, we find that population structure inferred from differences in the ages of doubletons from different populations supports other studies and allows for assessment of both interactions of groups and individual population demographic histories. This shows that this method provides a valuable alternative to traditional measures of population structure in a non-human organism and that further work with allele ages can help inform future management of the Sub-Saharan malaria vectors.

Impact of air pollution on buff-tailed bumblebee (*Bombus terrestris*) and their gut microbiome

Hannah Sampson*(S), Prof. Julian Ketley, Dr. Eamonn Mallon, Dr. Julie Morrissey

University of Leicester

Bumblebees play a major role in global pollination. Consequently, their health is of high importance for food security worldwide. Yet, recent population estimates show that their numbers are declining. This decline has been attributed to habitat loss, infection and use of pesticides. An important factor for bee health that contributes to population survival is the gut microbiome composition. The bee gut microbiome is functionally comparable to the human gut microbiome as they both provide protection from pathogens, are specific to the host and help break down food. Without a balanced gut microbiome, the health of the bee is threatened through increased infection and mortality. The bee gut microbiome is relatively simple, being dominated by 8 core bacterial species providing a convenient study system. Previous published data shows that air pollution has an impact on bacteria. Therefore, our hypothesis is exposure to air pollution causes an imbalance in the bee gut microbiome. To test this, we exposed bees to black carbon (BC), a major component of air pollution particulate matter. We assessed the effects on bee behaviour, microbiome composition and gut bacteria treated *in vitro*. Bees treated with BC showed a significant reduction in viable bacterial cells in their faecal community. This supports the hypothesis that air pollution can cause an imbalance in the bee gut microbiome, and may adversely influence bee health and pollinator populations.

Evolution in action: using museum DNA to understand the evolution of disease resistance in island birds

Eleanor C Sheppard*(S), Lewis G Spurgin, Brent C Emerson, Matt Clark, and David S Richardson

University of East Anglia

Many natural populations are facing a "double threat" of reduced genetic diversity and new pathogen challenges. Therefore, understanding how hosts adapt to pathogens in small and fragmented populations has important ramifications for conservation, epidemiology and evolution. However, we still lack fundamental understanding of the mechanisms that enable natural populations to respond and adapt to changing pathogen pressures. Next-generation sequencing methods combined with temporal sampling of wild populations now make it possible to undertake an exciting, powerful and timely investigation of these questions. We use Berthelot's pipit (*Anthus berthelotii*), a small passerine endemic to the Canary Islands, Madeiran and Selvagens archipelagos, as a model system to examine spatio-temporal genomic variation. Populations of Berthelot's pipit possess a number of features that make them well-suited for studying adaptive evolution in the wild, including spatially varying selection pressures, such as different pathogen regimes, and abundant museum samples for temporal study. Furthermore, previous work has shown how evolution has shaped patterns of genetic variation across the populations, giving us insight into the colonisation and demographic history of the species and identifying potential genomic candidates for selection across the species range. We aim to determine (i) the rate of immunogenetic change over time (ii) how different mutational processes facilitate it, and (iii) the role of specific pathogen-mediated mechanisms.

Effect of recurrent selective sweeps on the polymorphism of HA and non-HA segments in influenza virus A/H3N2

Nahyeon Song*(S), Kangchon Kim, and Yuseob Kim

Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, Republic of Korea

Influenza virus subtype A/H3N2 persists in human hosts by fixing new antigenic variants that allow them to evade host immune responses. Recurrent fixation of new variants or selective sweeps at an antigenic gene segment, particularly at the hemagglutinin (HA) gene segment, reduce its level of neutral (synonymous) sequence diversity relative to that of non-antigenic (i.e. non-HA) segments. This difference results from co-infection of viruses followed by reassortment (equivalent to meiotic recombination in sexual reproducing organisms) between viral segments, each of which however does not undergo homologous recombination. In this study, we investigate this relative levels of sequence diversity using individual-based simulation that also incorporate complex meta-population dynamics of seasonal influenza viruses. Parameters in the model, particularly reassortment rate and sweep rate, are estimated by comparing simulation results and H3N2 sequence polymorphism. In addition, as a first step in deriving analytic approximation for inter-segmental polymorphism, the pattern of coalescent processes on the HA segment was investigated. We find that the classical theory for recurrent selective sweep is not adequate for H3N2, which evolves very fast in the regime of clonal interference.

European distribution of *Borrelia*, hantaviruses, ticks and other *Apodemus* mice-associated organisms based on carryover DNA sequences in RAD-seq data

Haeyam Taiy*(S), Rohan Raval, Maria Luisa Martin Cerezo and Jaroslaw Bryk

University of Huddersfield

DNA sequences from 400 samples of wood and yellow-necked mice *Apodemus sylvaticus* and *A. flavicollis* from 55 locations across Europe were genotyped using ddRAD-seq. We mapped the reads to two low-quality *Apodemus spp* genomes and used the unmapped reads to identify carryover sequences of various *Apodemus* mice-associated organisms, in particular Lyme disease agent *Borrelia*, hemorrhagic fever with renal syndrome agent *Hantaviridae* and *Ixodes ricinus* ticks. We aim to demonstrate the feasibility of recovery of their sequences from carryover RAD-seq data and construct a distribution of these pathogenic agents across Europe.

Demography and natural selection have shaped genetic variation in the widely distributed conifer Norway Spruce (*Picea abies*)

Xi Wang[1,2,*(S)], Carolina Bernhardsson[1,2], Pär K. Ingvarsson[2]

1Umeå Plant Science Centre, Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden;

2Linnean Centre for Plant Biology, Department of Plant Biology, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Under the neutral theory, species with larger effective population size are expected to harbour higher genetic diversity. However, across a wide variety of organisms the range of genetic diversity are orders of magnitude more narrow than the range of effective population size. This observation has become known as Lewontin's paradox and although aspects of this phenomenon have been extensively studied, the underlying causes for the paradox remain unclear. Norway spruce (*Picea abies*) is a widely distributed conifer species across the northern hemisphere and it consequently plays a major role in European forestry. Here, we use whole-genome re-sequencing data from 35 individuals to perform population genomic analyses in *P. abies* in an effort to understand what drives genome-wide patterns of variation in this species. Despite having a very wide geographic distribution and an enormous current population size, our analyses find that genetic diversity of *P.abies* is low across a number of populations (π=0.005-0.006). To assess the reasons for the low levels of genetic diversity, we infer the demographic history of the species and find that it is characterised by several re-occurring bottlenecks with concomitant decreases in effective population size can, at least partly, provide an explanation for low polymorphism we observe in *P. abies*. Further analyses suggest that recurrent natural selection, both purifying and positive selection, can also contribute to the loss of genetic diversity in Norway spruce by reducing genetic diversity at linked sites. Finally, the overall low mutation rates seen in conifers can also help explain the low genetic diversity maintained in Norway spruce.

Does multidimensional selection drive local adaptation? Evidence from experimental evolution

White NJ*(S), Eyres I, Butlin RK

University of Sheffield

Divergent selection is a key driver of local adaptation and ecological speciation. The number of different selection pressures that a population must adapt to, termed the "dimensionality" of selection, is predicted to strongly impact the build-up of local adaptation among populations connected by gene flow. Theory predicts that local adaptation should build rapidly under unidimensional selection but more slowly under multidimensional selection (if at all), as selection may struggle to overcome gene flow. Contrary to these predictions, we have found that data from experimental evolution shows stronger and more rapid local adaptation when divergent selection is multidimensional, rather than unidimensional.

Investigating the genetic basis of dominant Bt resistance in the African maize stalk borer (Lepidoptera: Noctuidae)

Samuel Whiteford*, Carl Yung, Janet Vorster, Annemie Erasmus, Alistair C. Darby, Johnnie van den Berg, Pascal Campagne, Ilik J. Saccheri

University of Liverpool

Studies of the Maize stalk-borer (*Busseola fusca*) in South Africa have indicated the existence of a field-evolved dominant Bt-resistance trait localised to the western edge of the Maize triangle. Here we utilise a pooled next generation sequencing approach aiming to determine candidate genomic sequences associated with this trait. A chromosome-level assembly of a complex and highly repetitive ~900Mbp genome was produced by partitioning a long-read dataset of a female F1 hybrid using parental short reads (Trio-Canu). In order to optimise haplotype partitioning, crosses were highly outbred and many were not productive, suggesting possible hybrid inviability between populations separated by the Great Escarpment. We subsequently scaffolded the data using in-vitro and in-vivo HiC libraries derived from siblings of the target F1 genome. The allele frequency data was initially characterised with Φst statistics calculated from geographically distinct pairs of Bt and non-Bt Maize fields using the AMOVA framework, however we were unable to obtain a signature of a highly differentiated single locus. In order to overcome problems caused by high variance in allele frequency estimates caused by sub-optimal sequencing coverage we explored various in-silico pooling strategies and obtain a set of candidate regions associated with the dominant Bt resistance trait.

Deciphering the developmental basis of the gene cortex expression in Lepidoptera

Charlotte Wright (S), Luca Livraghi, Ian Warren, Tomas Generalovic, Owen McMillan, Chris Jiggins

Department of Zoology, Downing Street, Cambridge, CB2 3EJ

Heliconius are a genus of neotropical butterflies that display a remarkable diversity in wing colour pattern and colour. Closely related species often display mimicry. This makes *Heliconius* an excellent study system for exploring the mechanisms and genetic basis of wing patterning. The species *H. melpomene* and *H. erato* exhibit wide variation in wing pattern across their ranges, but within each area, they display extremely similar wing patterns. The majority of patterning in *Heliconius* is dictated by three loci, involving a limited number of genes. *Cortex* is the least well understood of these genes. The spatial expression profile of *cortex* dynamically changes during development suggesting a complex role in patterning. Moreover, *cortex* appears to play different roles in different species. Recent work has attempted to unpick how *cortex* is differentially regulated in the two species. Reverse-transcription PCR and short-read RNA sequencing suggest that transcripts of *cortex* are subjected to alternative splicing, resulting in a diversity of spliceoforms. However, the techniques used to observe splicing so far have provided insufficient resolution to identify these different spliceoforms. One objective of my project is to characterise the temporal and spatial changes in alternative splicing of *cortex* during the development of *H. erato*. Another objective is to compare the splice forms expressed in different species to understand how *cortex* is differentially employed by different species. Insights into alternative splicing will be obtained by transcriptome-wide long-read RNA sequencing which allows complete transcripts to be detected. The generated data will also be important for improving genome-wide annotation. Observed differences in splice forms will be confirmed by reverse transcription quantitative PCR and further probed by in situ hybridisation.

Population genomic variation in the wood tiger moth across the European geographic range

Eugenie Charley Yen*(S), Joana Meier, Chris Jiggins

Department of Zoology, University of Cambridge

The wood tiger moth (*Arctia plantaginis*) uses conspicuous hindwing colouration to warn potential, avian predators of its unpalatability. Such aposematic signals are typically expected to be under purifying selection due to predator avoidance learning. Despite this, *A. plantaginis* males exhibit a discrete, hindwing colour polymorphism, with white and yellow morphs maintained under balancing selection within local populations. Morph frequencies further vary between populations across the wide Holarctic species distribution, with additional red morphs only found in the Caucasus, and black morphs found in some Russian and North American populations. *A. plantaginis* therefore provides a great, natural system to study the evolutionary forces that promote colour signal diversity on local and global scales, even in the face of gene flow and opposing selection towards signal uniformity. Here, we present preliminary results from the first application of whole genomic tools to this study system, facilitated by a novel, high-quality reference genome assembled through trio binning. We resequenced 51 wild individuals to explore population genomic variation across 5 European populations from the species range, by characterising genome-wide population structure and phylogenomic relationships between polymorphic Southern and Central Finnish populations, a white-only Estonian population, a yellow-only Scottish population and a divergent red Georgian population.

Can we harness disease resistance by association directly in wild sea beet?

Hélène Yvanne[1]*(S), Kumar Gaurav[2], Sanu Arora[2], Isabelle De Cauwer[3], Neil Hall[1], Brande Wulff[2] & Mark McMullan[2]

Earlham Institute[1], John Innes Centre[2], Université de Lille[3]

Plant host and pathogen co-evolution can maintain genetic polymorphism at resistance genes and effectors, respectively. However, crop domestication has reduced genetic variation and this is believed to be important for their ability to resist pathogens that freely evolve. Until now, modern agricultural practices have introduced novel resistance genes by crossing crop lines. More recently, our attention has turned toward exploring the genetic diversity of wild crop relatives.

My PhD project aims to identify and study resistance genes in wild populations with a goal to improve crops resistance. I will use the beet system to identify associations between resistance gene polymorphism and fungal pathogens resistance. Direct wild associations are difficult for a number of reasons, not least because many wild hosts, crop progenitors, may not share their pathogens with crops. I will use the sugar beet (*Beta vulgaris*) system, because its recent domestication from sea beet (*Beta vulgaris* subsp. *maritima*) means that numerous pathogens infect both wild and agricultural plants. Moreover, these plants can still be crossed.

To do this, I will use an association genetics method combined with genome sequencing, allowing the identification of specific sequences (k-mers) present in sea beet genes resistant to rust. The association genetics method will rely on an understanding of the population genetics of European sea beet. Moreover, I will explore genetic diversity and divergence, of associated resistance genes, across populations. This project includes a collaboration with the KWS company, which aims to clone the identified genes in the sugar beet crop.