***Ischnura* project research questions**

**Core aims and questions**

Step 1: neutral structure, descriptive statistics of the three transects, Tajima’s D (expansion v. contraction), landscape genetics with landscape limits to gene flow. Migration rates and direction? Maybe with dispersal or morph freqs?

1. Parallelism in clinal phenotypic, genetic and epigenetic variation along climate-induced range expansion transects
   1. How variable/predictable is evolution during range expansion?
   2. Do we see parallel phenotypic traits (dispersal, morphology, climate adaptation) evolve?
   3. What is the genomic basis of adaptation? Do we see selection on alternative alleles for polygenic traits (GWAS)? What is their linkage structure and has it changed during range expansion?
   4. What are the environmental drivers of selection (GEA)? Do we see a release from selection?
   5. How do candidate alleles turn over along the transect? Does this follow the shape of the gradient?
2. Interaction between neutral, deleterious and adaptive evolutionary changes during range expansion
   1. How do dispersal evolution, local adaptation, and expansion load interact?   
      (do dispersal and niche traits co-evolve along the gradient?)
   2. How do they interact with the expansion velocity and the steepness of the environmental gradient?
   3. How can we best distinguish between surfing and adaptive alleles?
3. What is the relative contribution of genetic and epigenetic variants to phenotypic clines and plasticity?
   1. How do genetic and epigenetic changes underpin dispersal and climate adaptations?
   2. How do genetic and epigenetic changes predict range expansion rates?
   3. What is the relative importance of adaptation and plasticity along the gradient?
   4. How does coding sequence variation and methylation status reflect regional variation in thermal reaction norms? What is the relative and total impact of (epi)alleles on plasticity?
4. How can we best forecast range expansion in the field based on landscape genomic data and specific evolutionary processes?
   1. How do different mechanisms (drift, spatial sorting, selection gradients) determine spatial (epi)genetic patterns? How do they feed back on the expansion velocity?
   2. How do IBMs and correlative forecasting approaches differ in input data needs, output prediction detail, prediction accuracy, and applicability?
5. How do biotic interactions with local species communities affect range expansion rates and evolution?
   1. What is the effect of the abundance and diversity of competitor species on range expansion?
   2. How does this affect the population density and genetic structure of *Ischnura*?
   3. How does spreading into a species-rich vs. species-poor landscape change environmental selection pressures, gene flow, and dispersal evolution?
6. Holobiont PhD project (plus some of the existing parasite data)
   1. How does the holobiont shift across latitudes and with organismal development?
   2. How do these patterns relate to changes in dispersal and population characteristics?
   3. What are microbial effects on host dispersal ability, mating success, and environmental tolerances?

**NERC grant: all research questions mentioned in the proposal**

What genetic and evolutionary processes promote arrive and survive phenotypes?

1. how do contemporary range shifts affect population structure and evolution?
2. what genetic and epigenetic factors predispose species to poleward invasions?
3. how can we best forecast future responses?

* What evolutionary processes permit species to advance at the leading edge of their range, allowing them to both arrive (via increased dispersal) and survive (adaptive niche shifts) in newly available regions?
* How do genetic and epigenetic changes underpin dispersal and climate adaptations?
* How do genetic and epigenetic changes predict range expansion rates?
* How do dispersal and niche traits co-evolve during range expansion?
* How does their genetic and epigenetic basis interact over real expansion gradients?
* What is the relative importance of adaptation and plasticity along the gradient?
* What is the relative importance of adaptive and neutral spatial evolutionary processes for spatial genetic structure, clinal variation in dispersal and climate adaptation, and overall rates of range shifting?

WP1

* What are the genetic and epigenetic signatures of range expansion?
* How do drift, gene flow, local adaptation, and plasticity shifts proceed along range expansion gradients?
* Is epigenetic or genetic clinal variation more important in explaining trait and demographic clines, generating spatial structure, and responding adaptively along environmental gradients?
* Do we see parallelism in these processes?
* LFMM: what are candidate SNPs and DMRs in environment/morphology x gene association analyses? What are the environmental drivers of selection? (GEA , GWAS for lcWGS)
* What is the linkage structure of underlying candidate loci? Has this changed during range expansion?
* GDM: what is the allelic turnover in candidate SNPs, neutral and overall variation along environmental gradients?
* What is the turnover in methylation presence/absence and CpG alleles?

H1: Patterns of SNP allele frequency turnover associated with dispersal and climate adaptation are shared similarly across geographically comparable transects or target common genetic pathways.

H2: CpG alleles increase in frequency, and methylation variability at DMRs increases, towards range limit populations across all four transects, reflecting a significant role of plasticity during range shift adaptation.

H3: Spatial epigenomic variability tracks environmental gradients more strongly than does spatial genomic

variability, particularly at the approach of the expanding range margin.

WP2

* Are the candidate SNPs/DMRs associated with dispersal, thermal tolerances, and thermal reaction norms in laboratory-reared individuals from the core and edge?
* What is the role of genetic vs epigenetic clines in explaining changes in plasticity during range shifts?
* How do thermal reaction norms of growth/survival/thermal tolerance/dispersal reflect (epi)genomic variation at candidate loci?
* How does coding sequence variation and methylation status reflect regional variation in thermal reaction norms? What is the relative and total impact of (epi)alleles on plasticity?

H1: Plasticity in thermal and dispersal traits increases at the range edge and is better predicted by epigenetic variation.

H2: SNPs and those CpG polymorphisms underlying DMRs, identified as associated with climate and dispersal phenotypes in WP1, will also covary with climate and dispersal thermal reaction norms.

H3: Methylation levels will vary according to thermal rearing environment, more so in the edge region than in the core.

WP3

* How does genetic and epigenetic variation predict colonization and spread? Which one predicts spread better?
* Can we correctly forecast the distribution of variants at the range edge 10 years later?
* Does this differ between a ‘climate tracking’ vs. a ‘climate adapting’ scenario?
* Does explicitly including genetic, ecological and evolutionary processes in an IBM improve prediction accuracy?
* How do different mechanisms (drift, spatial sorting, selection gradients) determine spatial (epi)genetic patterns?
* How do IBMs and correlative forecasting approaches differ in input data needs, output prediction detail, prediction accuracy, and applicability?

H1: Projections for future distributions involving both genetic and epigenetic variants are more accurate than those based on either type of variation alone.   
H2: The distribution of adaptive and neutral variants at the range front at a previous sampling time (2013/2014) can predict the new distribution of variants at the range front

H3: Genetically explicit individual-based models better predict range shifts than correlative approaches.

**Marie Curie questions**

* How do climatic gradients affect morphological and fitness clines during range expansion?
  + (H1): Spatial sorting leads to increased dispersal and fecundity at the range front, accelerating the expansion rate.
* How repeatable are patterns of neutral, deleterious and adaptive genetic variation across range

expansion transects in relation to climatic, dispersal and fecundity traits?

* + H1: Genomic signatures of selection and local adaptation are convergent along the four transects.
  + H2: Alleles associated with thermal tolerance and dispersal are correlated and evolve jointly.
  + H3: Genetic load increases towards the expansion front.
* How can adaptive, neutral, and deleterious genetic variation impact range expansion rate and success?
  + H1: Greater adaptive trait variation and steeper allelic turnover across environmental gradients are associated with faster spread
  + H2: Stronger spatial turnover rates of deleterious alleles slow expansion rates
  + however, H3: clines in H1 and H2 may exhibit complex interactions, so I expect emergent properties from the modelling

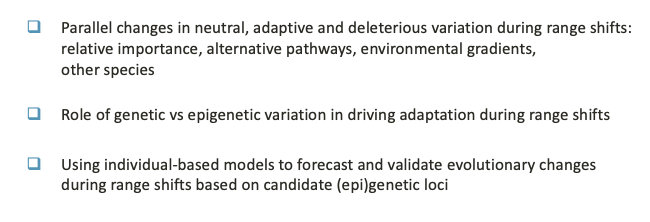
**Marie Curie target publications**

1. Clines in morphology and fitness along climate-mediated expansion transects, and how these relate to underpinning environmental gradients (target journal: Global Ecology and Biogeography)
2. The relative balance of neutral, deleterious, and adaptive genomic processes within and across transects and environmental gradients (target: Nature Ecology and Evolution)
3. A methods paper describing the integration of individual-based demographic and dispersal models with landscape genomic data (target: Genome Biology and Evolution)
4. our major results combining landscape genomic data with IBMs to understand which evolutionary processes best predict (variation in) expansion rate in the field (target: Science)

**NERC target publications (same colors indicate similar topics)**

1. Parallelism in clinal genetic and epigenetic variation associated with climate-induced range expansion in the wild
2. Contribution of genetic and epigenetic variants to clinal variation in phenotypes, plasticity, and coordination of dispersal and thermal traits
3. The role of neutral, adaptive, and deleterious genetic variation and epigenetic variation for rates of range shifting
4. Comparing and validating methods for range shift forecasting based on candidate (epi)genetically variable loci

**Major papers from brainstorming session**



**Meeting Lesley/Nicky discussing research questions on November 3**

* Local adaptation across the range; GEA / environmental drivers of selection; releases from selection
* Expansion load
* Gene flow across the gradient
* Genomic forecasting; allelic turnover: are shapes similar to the shape of the environmental gradient?
* Distinguishing surfing v adaptation
* Could spatial processes facilitate local adaptation along the gradient or harm it?
* Link with phenotypes: GWAS
* Interactions between dispersal evolution, local adaptation and expansion load
* A feeling this should all be easier with WGS...
* Do we really care about deviations from Hardy-Weinberg, demographic history, etc?
* Make sure we can test hypotheses; simulate things with SLIM?
* Problem: surfing alleles and adaptive alleles both increase in frequency along the gradient (so what do Fst outliers really mean?)

**A few notes on** **methods**

**lcWGS**

You *can* estimate individual heterozygosity, it is just averaged over many SNPs

Consider a minimum of 2x for diploid organisms

Trade-off: accuracy in allele frequency estimation (pop structure, outlier detection) vs. diversity statistics dependent on rare alleles (min 4x is better)

Relative values of LD still work! Could go down to 1x, need min 20x per pop