CONTRIBUTIONS FROM POPULATION GENETICS TO ECOTOXICOLOGY AND STRESS ECOLOGY IN LIGHT OF TRANSFORMATION TO THE POPULATION GENOMIC ERA

C. PERTOLDI1, L. A. BACH2, J.-C. SVENNING2, C. DAMGAARD3 and M. BAYLEY3

1 Department of Ecology and Genetics, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
2 Ecoinformatics & Biodiversity group, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
3 Department of Teri, NERI University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
4 Department of Zoophysiology, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark

Abstract - With the advent of the genomic era, which has partly been driven by advances in stress ecology, there is enormous growth in molecular and computer simulation techniques. Here we propose combining some of these techniques to give more elaborate risk assessments that include the effects of population variation in genotypes, phenotypes, and the way they link to aspects of life history and adaptive potential. We focused on ways to ascertain whether phenotypic plasticity or evolutionary responses constitute the basis for observed stress responses, as well as on the extrapolation problem, i.e. how do responses under controlled conditions correspond to those observed in natural ecological populations or in evolutionary end-points of interest? Additionally, we discuss the ways to integrate environmental variability into risk analysis and pest control predictions that include gene-environment interactions, focusing also on the importance of erosion of genetic diversity by toxic stressors to the risk of population extinction.

Key words: Genetic variability, environmental stress, toxic stress, extinction

INTRODUCTION

Since the establishment of ecotoxicology as an independent discipline, the prefix has been a source of debate (Van Straalen, 2003; Relyea and Hoverman, 2006; Van den Brink, 2008). As defined by Truhaut (1977), the discipline from the outset included study of how chemicals partition and are transported in both abiotic and biotic environmental compartments, how they affect individual organisms and how these two subsets of information translate into effects at the ecosystem level. The input of ecologically relevant end-points into the regulatory framework for chemicals, generated by ecotoxicology, has been minor (Van Straalen, 2003) and the need for extrapolation from single species information to this has been overcome using uncertainty factors and species sensitivity distributions (Van den Brink, 2008). Van Straalen (2003) argued that the ecological side of ecotoxicology could be increasingly seen as a subdiscipline of the much wider field of stress ecology. With the advent of the genomic era, which has partly been driven by advances in stress ecology, there is enormous growth in molecular and computer simulation techniques. Here we propose combining some of these techniques to give more elaborate risk assessments that include the effects of population variation in genotypes, phenotypes, and the way they link to aspects of life history and adaptive potential. We focus on:

1. Ways to ascertain whether phenotypic plasticity or evolutionary responses constitute the basis for observed stress responses.
2. The extrapolation problem: how do responses under controlled conditions correspond to those seen in natural ecological populations or in evolutionary end-points of interest?

3. Ways to integrate environmental variability into risk analysis and pest control predictions that include gene-environment interactions.

4. The importance of the erosion of genetic diversity by toxic stressors to the risk of population extinction.

*Environmental and genetic variability and the problem of extrapolation*

Although environmental variation is not necessarily reflected in altered vital rates such as growth rate, the interplay between environmental variation and population dynamics has been shown in a variety of species (Stenseth et al., 1998; Bjornstad and Grenfell, 2001; de Little et al., 2007; Rotella et al., 2009). In terms of toxicological effects, it is therefore necessary to obtain considerable knowledge of the consequences of toxicants on life history parameters such as growth, survival and reproduction, as well as the consequences for the loss of genetic diversity (Bickham et al., 2000; Gardeström et al., 2008). Ultimately, we are interested in predicting the effects of single or mixed toxicants through risk analyses on the higher hierarchical levels of complexity and under a variety of environmental scenarios. Estimation of risk requires the applied methodology to be probabilistically based as for example seen in stochastic computer simulations.

Interestingly, in toxicology related disciplines, genetic and ecological variability have often been perceived as confounding parameters that reduce experimental replicability. However, recent efforts have been made to incorporate genetic techniques into toxicological investigations to exploit the variation for useful interpretations (Kristensen et al., 2004; Phillips and Hickey, 2010). Similarly, evolutionary biologists and ecologists have turned increasingly towards molecular genetics to study the demographic and genetic consequences of environmental stress on populations (Frankham, 2010). Although it is inherently difficult to quantify and sort out the respective significances of genetic variability vs demographic factors, the challenge is being addressed (Spielman et al., 2004; Frankham, 2010). Genetic variability may be irrelevant for a population facing short-term extinction due to demographic problems. It is generally accepted, however, that conservation of biodiversity and adaptability to anthropogenic stressors such as environmental toxins, ultimately depends on the sustained presence of genetic diversity (Brown et al., 2009). It is also accepted that populations can only persist if the rate of adaptive evolution at least matches the rate of environmental change since the evolutionary response of quantitative traits to selection requires the presence of genetic variability (Burger and Lynch, 1995). In fact, this is the case even in the presence significant plastic response capability, including adaptations in behavior, physiology, morphology, growth, life history and demography. The interest in genetic disciplines has contributed to our understanding of the effects of genetic erosion on extinction risk and the dynamics of adaptation of species to new environmental conditions (Lande and Shannon, 1996; Frankham, 2010). This in turn can be exploited in the toxicological fields by interdisciplinary efforts and, indeed, attempts to correlate genetic, demographic and phenotypic properties of the same populations will become more frequent in the toxicological literature (Van Straalen and Timmermans, 2002; Brown et al., 2009).

Environmental factors and their changes are largely mirrored in the genetic composition of the affected populations, which in turn affect the potential for adaptation to future selective forces such as toxic substances. Even small alterations of environmental conditions can influence the genetic composition of populations, both via demographic and selective responses (Lande and Shannon, 1996; Björklund et al., 2009). Understanding the consequences of demographic stochasticity in populations requires detailed knowledge of local fluctuations in population size, extinction probability and
colonization potential, as well as reproductive success, which can be gained from population dynamics analyses. DNA analyses are progressively used to estimate the extent and organization of genetic diversity in populations in order to infer the causes of spatio-temporal dynamics (Schwartz et al., 2007). Such assessment is performed by investigating the degree of neutral genetic variation, which is informative in inferring ancient or recent historical dynamics of populations. Information on the genetic composition of a population prior to environmental perturbation is now accessible thanks to the recent progress in biostatistics and mathematics (e.g., theory of coalescence, Bayesian statistics, individual-based population dynamics, algorithms for efficient simulation and sampling of complex processes), which have greatly improved the possibility to infer population genetic processes through the development of theoretical models (Stephens and Balding, 2009). Going beyond plain parameter estimation is possible in applying a Bayesian approach, which can integrate both genetic and non-genetic data and hence test hypotheses about the factors that control demographic and genetic changes. In particular, the development of Bayesian models aimed to infer historical population dynamics and population parameters are extremely promising (Riebler et al., 2008; Guillot et al., 2009). A principal notion in forecasting a population's ability to adapt and survive under changing environmental conditions is the effective population size ($N_e$) (Pertoldi et al., 2007; Pertoldi et al., 2008; Björklund et al., 2010). Bayesian models can assess both the historical $N_e$ and current $N_e$, but can also estimate the degree of genetic isolation and rates of gene flow (Guillot et al., 2009).

The complexity and problems associated with the use of molecular tools may partly explain why most toxicological investigations have largely been confined to controlled laboratory conditions. However, laboratory experiments do not typically adopt a multifactorial approach, but instead vary one parameter while holding others constant, which may constrain the ecological relevance of such studies. The next step, however, consists of extrapolating the ecological relevance of many laboratory experiments and linking laboratory findings to real-world situations. In laboratory experiments however, it is possible to conduct integrative studies using neutral molecular markers (e.g., microsatellites and sequencing) and correlate the neutral variability with the genetic variability detected in quantitative and fitness-related traits. Correlation between identified markers and fitness traits can be useful for embarking on toxicological studies since such correlations are sparsely established in natural populations.

**Exploiting population variation and molecular techniques**

In laboratory conditions, it is in fact possible to conduct quantitative genetic analyses, which are of importance in the assessment of the extinction risk both at the individual and population level, since this approach can give information on the amount of non-neutral genetic variability present for a given trait. This information allows us to scrutinize fitness components on various genetic and environmental backgrounds, producing information on the fate of genetic diversity and the strength of selection acting on the populations. This will in turn permit us to quantify the importance of a given environmental stress in the expression of functional genes. Note however, that in practice we are thus far limited to manageable organisms with short generation times. Nevertheless, our ultimate aim is to determine how much a response of a given trait to environmental change is due to plastic and/or evolutionary response. Such information is becoming extremely relevant for the toxicological field, as there is a need for detailed studies on how variation at the level of genes translates, through developmental and physiological processes, into phenotypic variation for ecologically important traits (Coulson et al., 2006). In combination with ecological genomics and quantitative genetics, these investigations will promote a great increase in our understanding of ecological responses, starting from genetic variation in natural populations to the description of shifts in phenotypes because of evolutionary responses to environmental changes (Luikart et al., 2003).
As mentioned above, quantitative genetic investigations have often been limited to laboratory conditions and the neutral molecular markers in natural populations are not necessarily relevant to understand the evolution of functional genes subject to selection, which are essential for assessing the potential adaptability of a population to environmental changes. In natural populations, it is difficult to show selection (let alone to quantify). However, genome scans and association studies are increasingly promising due to new statistical methods with improved power (Slate et al., 2008; Nielsen et al., 2009; Stephens and Balding, 2009). Although identifying selected and functionally important genes is no easy task, genome scans offer the possibility of finding genomic domains with selective value, which in turn is a first step in separating selection from the background of random genetic drift. This would make way for describing how changing environments (and fragmentation) can affect different domains of the genome. Hence, finding genomic domains under selection may be at least as useful as gene finding per se.

The causal relationship between molecular genetic variation and phenotype-based measures of success are associated with some debate. Part of this incongruity stems from confusing the levels of organization at which genetic variation and phenotypic accomplishment have been conceptualized (Coulson et al., 2006). Furthermore, molecular markers cannot identify the likelihood of loss of genetic variance in traits of ecological significance, as the correlation between molecular diversity (which is by definition neutral) and ecologically relevant traits (which are by definition non-neutral) is weak and becomes even weaker in expanding or declining populations. However, the attempt to correlate neutral and non-neutral variability can be made by using a promising new tool in conservation genetics consisting of single nucleotide polymorphisms (SNPs). It is at present viewed as the richest polymorphic genetic marker in many genomes and may circumvent some of the problems related to microsatellites because of the enhanced resolution of genetic variation. In natural populations, SNPs hold the potential to expand our ability to survey both neutral (non-coding region) variation as well as genes under selection (coding region), while also providing broader genome coverage compared to microsatellites (Morin et al., 2004).

Furthermore, moving the genomic methodology from lab-model organisms to non-model organisms is now becoming possible, allowing genomic analysis in a population- and species-wide fashion (Mitchell-Olds et al., 2008). Until recently, genomic tools and resources have unfortunately been limited when it came to key ecological species as opposed to models species with plenty of genomic approaches readily available. Relocating greenhouse or lab experiments into a reality of genetic and environmental variation poses the challenge of separating environmental forcing from effects associated with the genetic variation (Kristensen et al., 2004). The challenge posed by the absence of nucleotide sequence information in key ecological species can be addressed by high throughput sequencing of a collection of mRNA samples. A set of overlapping DNA segments derived from a single genetic source (contigs) obtained by de novo assembly of known sequences can be blasted against model organisms within close phylogenetic distance and ultimately reveal the existence of SNPs (Ouborg et al., 2010).

Until now, geneticists have generally focused on changes in amino acid sequences that alter the kinetic function of proteins, without considering other possible alterations of the DNA structure with evolutionary consequences such as altered epigenetic stress response regulation. The question of the importance of structural genetic variation (i.e. proteins) versus regulatory genetic variation should be kept in mind when applying molecular techniques. cDNA microarray technology has emerged as a powerful tool to monitor the gene expression of thousands of genes simultaneously (Ouborg et al., 2010). Recent identification of functional genes and genes linked to quantitative traits are opening the way to the analysis of functional genes and components of genetic control of physiological processes, and are therefore expected to contribute to the un-
understanding of local adaptation (Riebler et al., 2008; Marsano et al., 2010). Population genomics will very soon add important contributions to these issues, delivering large amounts of data on regulatory polymorphisms on a genomic scale. Moreover, we may address the question of whether the regulatory variation per se causes adaptation to local conditions and whether it is able to alter significantly lifetime reproductive success.

Finally, population transcriptomics is a recent effort to go beyond the analyses of sequence variation per se by looking at the actual gene expression and patterns of gene regulation in a population. This will allow a description of variation and population changes closer to the realized phenotypic level (Marsano et al., 2010; Ouborg et al., 2010).

**Theoretical approaches**

The development of theoretical models and the use of computer simulations has also contributed significantly to ecotoxicological field through, for example, the integration of genetics into metapopulation frameworks and the development of predictive models which incorporate both environmental and genetic data sets. These models include stochastic environmental effects, allowing us to make probabilistic predictions that can be reasonably precise when we consider averages over large scales. Considerable progress has been achieved in incorporating age- or stage-structure into population genetic models, mostly in the context of life history evolution and estimation of $N_e$ of large and stable populations (Engen et al., 2010). However, knowledge on the interaction between age- or stage-structure and other factors, such as variance in reproductive success, temporal fluctuations in population size, is still quite limited.

Deterministic simulations are based on algebraic equations that predict the likely outcome of sampling, while stochastic (Monte Carlo) simulation models mimic random processes. Although being transparent and analytically tractable, deterministic predictions cannot deal with the same level of complexity over many generations as stochastic simulations. One advantage of combining these approaches is apparent from simulation used to verify the accuracy when prediction equations are developed. Stochastic simulations are highly relevant for the design of risk estimates and there are no inherent limitations excluding representation of the genetic level.

Genetic and ecotoxicological data may be integrated into a modeling framework either by using an individual-based modeling (IBM) approach or by describing the selective and ecotoxicological process in a Bayesian network. The study objects, such as populations or individuals, do not necessarily comply with the mean field assumptions that all units must be organized as uniform masses and interactions are unconditioned and can be averaged. In such cases, the IBM or agent-based approaches can be appropriate ways to allow variation in many aspects of the individual’s characteristics as well as variable and complicated conditional interactions (Travis et al., 2009). Likewise, the geospatial implementations of IBM can account for specific spatial effects. This approach can be especially relevant for heterogeneous populations of higher animals in spatiotemporally heterogeneous environments with conditional behavior depending on its own state, the state of kin and conspecifics, or the specific states of the environment (Bach et al., 2006). In other words, the individual in an IBM does not perceive and interact with ‘the average individual’ of an abstract population according to ‘an average encounter rate’ and it does not experience ‘the average environment’. However, as entities, interactions and environment can be freely defined, it follows that the extreme flexibility can become a serious challenge in designing simulations to answer simple questions. In terms of genetics, another advantage of IBM is the straightforward implementation of genotypes, representing either neutral or selected genes, where the latter can allow the agents to adapt to changing environments – these are sometimes referred to as complex adaptive systems (CAS) (DeAngelis and Mooij, 2005). Also, the fact that events in IBM simulations are inherently stochastic may prove an advantage when yielding probabilities is the goal. Finally, much depends on the specific
question and available data. Empirical data can be entered in the models at several levels, but the real strength of this methodology is the bottom-up design. Here data is typically included on the lower levels (as knowledge of individual physiology or life history traits) and patterns on higher levels can be observed as emergent properties of the system (spatiotemporal population dynamics, evolutionary trajectories, community structure).

In a Bayesian network (also known as state-space modeling or structural equation modeling), the selective and ecotoxicological processes are assumed to operate on latent variables that model the relevant genetic and phenotypic states of the populations using standard population genetic models of selection and genetic drift (Lynch and Walsh, 1998). The data are integrated into the population genetic-modeling framework by linking the latent variables to the genetic and ecotoxicological data by likelihood functions. The resulting Bayesian network is then parameterized by MCMC methods (Carlin and Louis, 2000), which allows the testing of different genetic and ecotoxicological hypotheses using the joint posterior distribution of the parameters. One of the advantages of the Bayesian network approach is that the process error, e.g. genetic drift, is separated from the measurement error. This separation of the different sources of variation allows predictions to be made where the uncertainty due to process error alone, is included in the prediction.

Stochastic simulation models can also accommodate various global change scenarios, which may not be readily accomplished by mathematical analysis. Stochastic genetic models may mimic events at individual loci, so-called finite loci or allelic models, or may be parameter-based, describing average genetic effects according to quantitative genetics theory. Hence, the use of computer simulations will have an important role in the immediate future and will be utilized for: (i) modeling alternative scenarios for the dynamics of genetic diversity within and among populations exposed to different environmental regimes and evaluation of short- and long-term risks; (ii) linking the genotype with phenotype, for example, modeling how a given trait would develop in a given scenario (e.g. life-history or morphological traits (Pertoldi et al., 2003)). As factors or estimates can be manipulated at almost all levels of the modeling, effects from ecological changes can be predicted, especially those related to spatially and temporally dynamic environments. If the information obtained can be combined with empirical and molecular data, the model will provide a powerful tool for understanding real-world dynamics; (iii) assessment of how different environmental scenarios affect both genetic and demographic parameters; (iv) understanding how differences in life history between ecologically similar species lead to substantial differences in \( N_e \) and genetic variability (Vindenes et al., 2010). Investigating to which extent fluctuations in vital rate parameters induced by environmental change alter \( N_e \); and (v) quantification of the interactions of each particular life history parameter with other factors.

**Developments in geographical ecology with relevance for ecotoxicology**

Finally, another area of potential progress in stress ecology and ecotoxicology is the inclusion of new developments in geographical ecology towards much improved quantification of the determinants of species distributions and diversity patterns (Guisan and Zimmermann, 2000; Allen et al., 2002; Franklin, 2010). Notably the role of geographic variation in environmental factors, such as climate, creates an important basis for predicting responses to future climate change (e.g. Thomas et al., 2004; Morueta-Holme et al., 2010). Toxic impacts can be a crucial part of a population’s surroundings and therefore should be accounted for along with other relevant environmental factors (Holmstrup et al., 2010): but they have hitherto been largely ignored in geographical ecology. Furthermore, geographical ecology may tell us where species are most vulnerable (e.g., subject to marginal environmental conditions) and “geotoxicology” may tell us where emissions and depositions are largest. Indeed emissions and concentrations of ecotoxicological stressors show pronounced large-scale geographic patterns.
CONCLUSION

Multidisciplinary approaches that facilitate the implementation and development of new and recent genomic and theoretical tools may advance the ecotoxicological field. Further scientific progress could be accelerated by merging and complementing current efforts in evolutionary and ecological genetics by: (a) merging taxonomic, geographic, ecological and genetic databases; (b) using molecular data in synergy with quantitative traits and environmental data; (c) unraveling the distribution of variation at functional versus non-coding sequences in natural populations, and (d) estimating fitness in changing and stressful environments.

REFERENCES


