2016

Role of genomics and transcriptomics in selection of reintroduction source populations

Xiaoping He
University of Windsor

Mattias L. Johansson
University of Windsor

Daniel D. Heath
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/glierpub

Part of the Genomics Commons

Recommended Citation

This Article is brought to you for free and open access by the Great Lakes Institute for Environmental Research at Scholarship at UWindsor. It has been accepted for inclusion in Great Lakes Institute for Environmental Research Publications by an authorized administrator of Scholarship at UWindsor. For more information, please contact scholarship@uwindsor.ca.
Role of genomics and transcriptomics in selection of reintroduction source populations

Running title: Genomics and reintroduction

Xiaoping He¹, Mattias L. Johansson¹, Daniel D. Heath¹*

¹Great Lakes Institute for Environmental Research, University of Windsor, Windsor, ON N9B 3P4, Canada

*Address correspondence to D. D. Heath, email dheath@uwindsor.ca

Keywords: reintroduction biology, conservation genomics, adaptation, acclimation, SNP, genetic diversity, stress

Abstract

The use and importance of reintroduction as a conservation tool to return a species to its historical range where it has become extirpated will only increase as climate change and human development accelerate habitat loss and population extinctions. Although the number of reintroduction attempts has rapidly increased over the past two decades, the success rate is generally low. As a result of population differences in fitness-related traits and divergent responses to environmental stresses, there is a high likelihood for differential performance among potential source populations upon reintroduction. It is well known that population performance upon reintroduction is highly variable and it is generally agreed that selecting an appropriate source population is a critical component of a successful reintroduction.
Conservation Genomics is an emerging field that addresses long-standing challenges in conservation biology, and the potential for using novel molecular genetic approaches to inform and improve conservation efforts is high. As the successful establishment and persistence of reintroduced populations is highly dependent on the standing genetic variation and environmental stress tolerance of the source population, we propose the application of conservation genomics/transcriptomics to guide reintroduction practices. Specifically, we propose estimating genetic diversity (standing genetic variation) for source populations using genome-wide functional loci to predict the potential for adaptation and using transcriptional profiling to measure expression response of fitness-related genes to environmental stresses as a proxy for acclimation (tolerance) capacity. Appropriate application of conservation genomics/transcriptomics has the potential to dramatically enhance reintroduction success in a time of rapidly declining biodiversity and accelerating environmental change.

**Introduction**

Reintroduction is the intentional release of a species into its historical range where it has become extirpated, and is distinct from reinforcement, where organisms are translocated to existing populations of the same species (IUCN/SSC 2013). Due to sharp declines in global biodiversity caused by human activities (Butchart et al. 2010), reintroduction has become an important conservation tool and is likely to increase in its application as extirpation rates increase (IUCN/SSC 2013). With 218 animal species reported as reintroduced by 1998, and double that number by 2005, the frequency of reintroduction attempts is already increasing rapidly (Fischer & Lindenmayer 2000; Seddon et al. 2007). While the number of reintroduction projects is increasing worldwide, the success rate of reintroduction is generally low (Armstrong & Seddon 2008). Fischer & Lindenmayer (2000) found that 49% of animal reintroductions with known outcomes could be considered successful while Cochran-
Biederman et al. (2015) found that the success rate for freshwater fish was as high as 58%. In fact, the real success rate may actually be much lower due to publication biases that drive higher publication rates for successful reintroduction reports relative to failed reintroductions and the high percentage of reintroductions with uncertain outcomes (Fischer & Lindenmayer 2000; Miller et al. 2014).

Given suitable receiving habitat, the selection of an appropriate source population is an essential and critical factor affecting reintroduction success (Schneider 2011; Forsman 2014; IUCN/SSC 2013). Populations can differ in reintroduction potential due to their genetic architecture (genome content and epistatic interactions) which may limit their scope for adapting to novel environments and narrow their tolerance range for environmental stressors encountered in the initial acclimation phase of reintroduction. In general, genetic architecture is not only closely related to fitness in a locally adapted population, but it is also a key factor in determining whether organisms can survive and thrive in novel or changing environments (Lande & Shannon 1996; Ouborg et al. 2010). The importance of the genetic architecture of source populations is reflected in the IUCN Guidelines for Reintroductions (IUCN/SSC 2013) which state that if a translocation consists of many individuals with high genetic diversity then source genetic architecture may not limit reintroduction success.

Conservation genomics is a new field that applies genomic technologies to address questions in conservation biology (Ouborg et al. 2010; Hoffmann et al. 2015). Some applications of conservation genomics simply increase the power of traditional population genetics approaches by using larger numbers of DNA markers, covering more of the genome (Kohn et al. 2006). However, conservation genomics and transcriptomics can also address more challenging long-standing issues in conservation biology, including quantifying fitness-
related genetic variation, measuring how environmental stress affects gene activity and
determining the molecular mechanisms of tolerance to environmental fluctuation and
pollutants (Ouborg et al. 2010; Whitehead et al. 2010; Harrisson et al. 2014). Here, we first
argue that the selection of source population is critical for successful species reintroduction
and must go beyond simple inbreeding avoidance (i.e., maximize neutral genetic diversity).
We then explain how genomic tools can facilitate selection of the optimal source populations
based on two criteria: 1) maximizing functional standing genetic variation to foster adaptive
potential and 2) maximizing potential adaptive plasticity to foster acclimation, or breadth of
tolerance. We argue that the appropriate application of conservation
genomics/transcriptomics has the potential to dramatically improve the success rate of
reintroduction, a critical tool for maintaining biodiversity in the face of rapid environmental
change.

**Source population selection**

**Difference in response and tolerance of environmental stresses**

Due to divergent evolutionary processes, populations of the same species showing differences
in morphology, behavior, life history and physiology have been observed in almost all
kingdoms. Population differences are not only observed in static traits, but differences also
exist in how individuals from different populations respond to environmental change or
stressors. Zhang et al. (2004) compared the response of *Populus davidiana* trees from
populations that experience different levels of drought stress and found differences in their
strategies to survive drought. Côte et al. (2012) compared embryo incubation time, body size
and survival of four Atlantic salmon (*Salmo salar*) populations under normoxic and hypoxic
conditions and observed significant population by environment interactions and parental
effects for all traits. DeKoning et al. (2004) measured acute and chronic handling stress
response in northern and southern populations of killifish (*Fundulus heteroclitus*) and found the southern population fish exhibited higher plasma cortisol response relative to pre-handling states. These examples highlight the broad range of responses to environmental stresses among presumably locally adapted populations, and hence differences in adaptive potential to new environments after reintroduction.

In addition to differential stress responses, populations also exhibit differing tolerance of environmental stress and resistance to pathogens. Marchin et al. (2008) found population differences in cold tolerance, growth and survival among 44 white ash (*Fraxinus americana*) populations, even after they had been grown in a common garden for 30 years. Fangue et al. (2006) compared thermal tolerance in six populations of the common killifish (*Fundulus heteroclitus*) and found that the three southern populations had a higher critical thermal maximum than the three northern populations while the latter had a lower critical thermal minimum. Finally, Uller et al. (2003) found that southern populations of the common lizard (*Lacerta vivipara*) were more resistant to viral eye disease, with higher survival after infection compared to northern populations in Sweden. In general, differences in environmental stress tolerance and pathogen resistance are reported across taxa, indicating a high likelihood that potential source populations differ in response traits that can affect their expected survival and performance after reintroduction.

**Reintroduction performance variation**

Successful reintroductions require that the released organisms are able to establish and persist in the target habitat. Establishment refers to survival and successful reproduction (Seddon et al. 2012), while persistence refers to increased numbers and density of reintroduced species in the target habitat (Armstrong & Seddon 2008). When individuals are reintroduced, they...
face novel selective pressures imposed by the new environment. Thus the survival of reintroduced individuals depends on how closely their phenotype matches locally favoured phenotypes (Ghalambor et al. 2007) or more generally, how well their phenotypes match the current conditions in the target habitat. Ideally, researchers should evaluate potential source populations for matching habitat environmental characteristics and the associated genetic architecture and phenotypic variation to choose the best possible source for reintroduction (see Sork et al. 2013). However, such an exhaustive evaluation of candidate source populations is often not feasible.

Generally, there are two main mechanisms by which organisms can cope with environmental stress: phenotypic plasticity and adaptation (Hansen et al. 2012; Harrisson et al. 2014). Phenotypic plasticity is the ability of individual genotypes to exhibit alternative phenotypes in response to different environments (Pigliucci et al. 2006). Genetic adaptation occurs when the genetic background of individuals within a population changes over time to maximize fitness in the new environment (Hendry et al. 2011; Hansen et al. 2012). In the short term, reintroduced individuals may survive through environmental acclimation via phenotypic plasticity (Schlichting & Smith 2002; Ghalambor et al. 2007). In the long term, genetic adaptation may be the key mechanism for introduced populations to survive and thrive (Hansen et al. 2012).

Since populations show substantial phenotypic differences at rest and in response to stress, it is not surprising that those populations may exhibit variation in reintroduction performance. For example, Olsson (2007) compared reintroduction performance of White Storks (Ciconia ciconia) between populations originating from North Africa and those from their native northeastern Europe, and found that the birds with native ancestry had higher reproductive
success and more offspring per pair than those with North African ancestry. Schneider (2011) tested several Atlantic salmon strains for reintroduction into the Rhine River, and concluded that only the Swedish Atran strain was suitable for reintroduction as its spawning time matched the original population of the Rhine River. The variation in performance among potential source populations is likely explained by differences in their genetic architecture shaped by historic evolutionary pressures.

Whether reintroduced organisms can survive and thrive in new environments depends on their acclimation and adaptation potential, both determined by the evolutionary history of the source population. It is thus important to characterise potential source populations for both their current standing variation (adaptive potential) as well as their scope for response to ecologically relevant stresses (acclimation). However, characterizing functional trait differences can be both technically and logistically difficult, while neutral genetic variation, even with large numbers of loci, may not accurately reflect genome-wide standing genetic variation. Here we propose a new paradigm for reintroduction source population selection: conservation genomics/transcriptomics (Fig. 1).

**Standing genetic variation and genome-wide SNPs**

**Genetic similarity and genetic diversity**

The standing genetic variation of source populations can significantly impact the outcome of reintroduction efforts. Two main aspects of standing genetic variation must be taken into account when selecting source populations: genetic similarity and genetic diversity. Ideally, donor populations that are genetically similar to the historical (extirpated) population in the targeted habitat should be selected for reintroduction. After environmental remediation or natural habitat recovery, the species could re-establish after reintroduction due to pre-existing
adaptations to the target habitat. For example, aurora trout (*Salvelinus fontinalis timagamiensis*) were extirpated in Whitepine Lake and Whirligig Lake in the 1960s due to acidification of the lakes (Snucins et al. 1995). In 1990, when the lakes recovered, captive aurora trout originating from wild individuals collected from the same lakes were reintroduced and natural reproduction was observed in Whirligig Lake in 1992 (Snucins et al. 1995). However, in most cases individuals from the original population are not available, and there are usually no DNA samples from the original population that can be used to compare genetic similarity between the original and potential source populations (Schwartz 2005). Furthermore, while an historic habitat may appear to be restored, it is likely that current conditions are changed, and thus even the original genetic stock may fail to re-establish. Thus, in most situations, we must select from extant populations of unknown genetic relatedness to the original population.

Using genetic diversity as a criterion for selecting source populations for reintroduction (Earnhardt 1999; IUCN/SSC 2013) will ensure substantial genetic variation for natural selection to act upon in the novel environment, maximizing adaptive potential (Lande & Shannon 1996). Avoiding low genetic diversity resulting from past genetic bottlenecks and inbreeding is also important for reintroduction success (Frankham 1995). Indeed this is the theoretical basis for using measures of genetic variation as surrogates for fitness in conservation (e.g., Reed & Frankham 2003). However, although high genetic diversity is important for population fitness, not all genetic variation is related to fitness. While neutral genetic marker applications in conservation have increased dramatically (Ouborg et al. 2010), the vast majority of those applications rely on small numbers of loci that may not reflect genome-wide diversity (Fig. 2). Furthermore, while one may assume that neutral marker diversity is correlated with functional standing genetic variation, this may not be correct.
(Hedrick 2001; Reed & Frankham 2001). Ideally, genome-wide coverage based on functional marker loci should be used to achieve more complex conservation goals than inbreeding assessment and genetic isolation quantification.

**Single nucleotide polymorphism**

Single nucleotide polymorphisms (SNPs) are DNA sequence differences at a single nucleotide where the less frequent allele of the polymorphism is 1% or higher (Vignal et al. 2002). SNPs are generally biallelic and distributed throughout the genome with high density (e.g., one SNP every 116 base pairs in the genome of Channel Catfish, *Ictalurus punctatus*; Sun et al. 2014). SNPs can be located in the coding region of genes, in introns and between the genes (Jukema & Agema 2001). Coding region SNPs (cSNPs) can be further divided into synonymous and nonsynonymous SNPs. Nonsynonymous SNPs are associated with changes in amino acid sequence and are thus most likely to represent functional genetic variation, although synonymous SNPs may be in linkage disequilibrium with unrecognized functional variation.

Variation in phenotype among and within populations is partially explained by variation in DNA sequence (Botstein & Risch 2003) and our understanding of how variation at specific gene loci affects phenotype is growing rapidly. Thus identifying genetic markers (e.g., SNPs) for variation in specific traits that are critical for reintroduction success will facilitate effective selection of source populations and individuals for reintroduction. For example, Johnston et al. (2014) used 4353 SNPs to conduct a genome-wide association study between SNPs and the age at which Atlantic salmon return from the sea to spawn, and identified 10 SNPs that were associated with maturation age. Such studies show the power of genomic approaches to identify functional DNA markers which can be used to evaluate source
populations for reintroduction. However, despite the promise of conservation genomics for improving the selection of source populations for successful reintroductions and a rapidly decreasing cost to characterize genome-wide SNPs, very few such studies have been reported, and no studies report reintroduction success. Thus despite the rapid growth of conservation genetics and the broad acceptance of the concepts of conservation genomics, to date genomics and transcriptomics have rarely been used to assist in species reintroduction efforts (Fig. 2).

Application of genome-wide SNPs to reintroduction

SNPs have been proposed to replace microsatellite markers for applications in conservation genetics (Morin et al. 2004). Genotyping of genome-wide SNPs has been reported for many species, including farm animals (Muir et al. 2008; Petersen et al. 2013), fish (Willing et al. 2010; Jones et al. 2012) and plants (Grattapaglia et al. 2011; Plomion et al. 2014).

Considering the importance of functional genetic variation and logistical convenience of genotyping genome-wide SNPs, we propose the use of SNP-based genome scans to estimate genetic diversity for reintroduction source population selection (Fig. 1). We further propose that two types of SNP genome scans are useful for reintroduction efforts: nonsynonymous SNPs and SNPs associated with fitness-related traits. Nonsynonymous SNPs change protein sequences, thus may reflect variation in protein function. SNPs already known to be associated with fitness-related traits may have either direct effects on phenotype, or be in linkage disequilibrium with genetic variation underlying phenotype variation. In either case, the estimation of genetic diversity at such SNP markers can effectively provide estimates of functional standing genetic variation (Fig. 1). Our proposed approach to known functional SNP scanning is appropriate for species with abundant genomic information and well-characterised gene function. However, many species of conservation concern have little
genomic information and the genetic bases for fitness-related traits are barely studied.

Therefore, we propose genotyping anonymous genome-wide cSNPs to estimate genetic diversity for species with poor genomic characterization. Although not all cSNPs are functionally important, they are much more likely to be associated with functional genetic variation than known neutral DNA markers. Once appropriate functional SNPs are identified, candidate populations showing the highest observed heterozygosity should be selected as source populations. This focus on maximizing heterozygosity in cSNPs will not only maximize standing genetic variation for functional traits and hence the likelihood of reintroduction success through adaptation, but also provide useful information regardless of source population size, as it provides objective functional criteria for choosing the source population with maximal adaptive potential for reintroduction.

**Gene expression and application of transcriptomics**

It has long been recognized that gene expression variation is the primary mechanism that leads to phenotypic variation within and among populations (Rifkin et al. 2003; Storey et al. 2007) as gene expression is influenced by both genetic and environmental factors (Petretto et al. 2006; López-Maury et al. 2008) specifically through regulation of expression and/or selective expression or silencing of genes (Schlichting & Smith 2002). Therefore, variation in gene expression is not only a physiological process, but also an important adaptive mechanism allowing organisms to respond to novel habitats or environmental change/stress (Schulte 2004).

**Population difference in stress response via gene transcription**

With the rapid development of transcriptomic tools, it has become simpler to simultaneously quantify transcription at thousands of gene loci, even in non-model species. Differentially-
expressed genes (at rest or in response to a challenge) represent potentially adaptive genetic variation among populations (Whitehead & Crawford 2006; Larsen et al. 2007; Giger et al. 2008; He et al. 2015). Thus transcriptional profiling can lead to quantitative estimates of relative environmental stress response among populations. Wellband & Heath (2013) compared transcriptional response to temperature and immune challenges in six wild rainbow trout populations and found that populations showed different transcriptional patterns in response to the challenges, despite the fact that the six populations were separated by less than 250km. Whitehead et al. (2010) compared transcriptome differences in response to polychlorinated biphenyl exposure between naturally tolerant and sensitive killifish populations and revealed that low expression of genes involving in the aryl hydrocarbon receptor signalling pathway may be a mechanism of pollution tolerance in killifish. While examples such as these underscore a large body of literature that demonstrates that populations generally respond to environmental stressors differently at the gene transcription level, the application of that knowledge to the selection of source populations for reintroductions is practically non-existent. The only published example of transcriptomics applied for reintroduction purposes focussed on the reintroduction of extirpated Atlantic salmon into Lake Ontario, Canada. Using a custom microarray, He et al. (2015) showed significant gene transcription differences at 21 genes between two possible source populations, demonstrated that differences were likely the result of selection, and recommended one source population for reintroduction based on those differences. Differences in gene expression patterns underlie the mechanisms of differential tolerance to environmental stress, and transcriptional profiling is thus an ideal, but underutilized, tool for selecting source populations for reintroduction.

**Application of transcriptional profiling to reintroduction**
Since variation in gene expression is directly linked to phenotypic variation, gene expression is potentially a powerful tool for the prediction of phenotypes (Oellrich et al. 2014). For example, Tung et al. (2012) compared gene expression among 10 rhesus macaque (Macaca mulatta) groups with different social status and found that the identified differentially-expressed genes could be used to predict social rank class with 80% accuracy. Miller et al. (2011) applied a nonlethal biopsy method to collect tissue from wild migrating Sockeye salmon (Oncorhynchus nerka), and used genome-wide gene transcription to identify a genomic signature that could be used to predict migration and spawning success (Miller et al. 2011). Although this kind of application of transcriptomics is still in its infancy and few examples exist in the literature, those examples demonstrate the possibilities.

Due to the plastic nature of gene expression and its important role in the adaptive response to environmental stressors (acclimation), we propose the application of transcriptional profiling to quantify acclimation potential among potential source populations (Fig. 1). Because many species of conservation concern are found only in small, highly fragmented populations, this focus on flexibility avoids the problem that small populations may be dominated by genetic drift and thus may no longer be locally adapted (Leimu and Fischer 2008, Willi et al. 2006). For species whose stress response genes are well characterised, transcriptional profiling could target a selected panel of candidate genes associated with coping with ecologically-relevant stressors and environmental fluctuation. For species whose stress response genes are poorly characterised, whole transcriptome analysis should be used to quantify acclimation potential. In such cases, ecologically relevant and physical environmental challenges would have to be applied to individuals from the candidate populations and their gene transcription profile assessed before and after the challenge. Based on their transcriptional response, we can evaluate and predict population performance upon reintroduction in two ways: (1) for cases
where we know a gene’s function in dealing with environmental stress, we select populations exhibiting adaptive responses; (2) for cases where we do not know whether up-regulation or down-regulation of a gene is beneficial, we select populations which show high plasticity.

**Limitations of Conservation Genomics/Transcriptomics**

Species reintroduction will become more common as habitats are altered and lost due to human activities and climate change. However, reintroduction efforts are costly, potentially environmentally risky and subject to complex regulatory requirements (IUCN/SSC 2013). Thus reintroductions are only feasible when reintroduction is central to species conservation and the species is of high priority – we propose that conservation genomics/transcriptomics are realistic possibilities for improving the likelihood of reintroduction success in key high risk situations. While the costs of genome-wide cSNP application and transcriptome profiling have come down substantially, it is still a major hurdle for often financially limited conservation efforts. We do not propose that genomics or transcriptomics should be the first line of response in a reintroduction effort, but rather they are promising tools for which the cost is dropping rapidly. More importantly, the effectiveness of the conservation genomics and transcriptomics applications that we propose have not yet been tested in any reintroduction that we are aware of, thus no empirical data exist that shows it materially improves reintroduction success. Ideally, translocation experiments in controlled systems should be used to test whether functional standing genetic variation is a good predictor for long-term introduction success, or if transcriptional profiling can predict short term acclimation and survival. As it is likely not feasible to apply conservation genomic/transcriptomic methods to reintroduction efforts for true species at risk, we propose an empirical test of the application of genomic and transcriptomic techniques in an artificial reintroduction experiment. Using short-lived, genomically well-characterised species as
models, groups would be introduced into controlled environments that range in environmental conditions such that they represent benign to potentially lethal levels of environmental stress. The putative “source” populations would be characterised as having either high or low levels of functional standing genetic variation and adaptive or non-adaptive transcriptional response to the environmental stressors in the artificial target habitats. Groups from the contrasting source populations would be introduced into the range of target habitats, and population performance would be monitored as reproduction in the short-term and as population size and habitat-use expansion in the medium-term. The prediction is that high functional standing genetic variation and high plasticity in transcriptional scope would drive increased short- and medium-term performance. Although the primary purpose of this essay is to make the argument that conservation genomics/transcriptomics has great promise and should be explored as a valuable tool in addressing the growing biodiversity conservation crisis, we cannot yet provide concrete evidence for its value in conservation efforts. However, our growing understanding of how the genome and transcriptome is shaped by interactions with the environment provides compelling evidence for conservation genomics/transcriptomics as emerging and valuable tools for effectively managing the world’s biodiversity.

Conclusions

Functional SNP genotyping and transcriptional profiling are potentially powerful tools for reintroduction in particular, and conservation in general. Conservation genomics and transcriptomics can not only answer long standing questions in conservation biology, but also provide important applications in reintroduction biology, specifically in selecting the appropriate source population. Characterizing genome-wide functional SNPs can provide quantitative estimates of fitness-related standing genetic variation and transcriptional
profiling can provide data on how individuals respond to environmental stresses. Such data would have immediate practical applications in reintroduction biology as metrics for source population selection. Since successful reintroduction requires both short term acclimation and long term adaptation to the targeted habitat, we strongly urge conservation biologists and managers to consider employing functional SNP scans to measure genetic diversity and transcriptional profiling to measure the response of candidate and anonymous genes as part of the selection process for source populations for reintroduction. When based on genomic and transcriptomic measurements of adaptation and acclimation, the selection of source population will be more effective and will increase the success rate of reintroductions globally.

References
Identifying correlates of success and failure of native freshwater fish reintroductions.
Conservation Biology 29:175–186.

response to hypoxic stress in Atlantic salmon. Journal of Evolutionary Biology

DeKoning, A. B. L., D. J. Picard, S. R. Bond, and P. M. Schulte. 2004. Stress and
Interpopulation Variation in Glycolytic Enzyme Activity and Expression in a Teleost


Fangue, N. A., M. Hofmeister, and P. M. Schulte. 2006. Intraspecific variation in thermal
tolerance and heat shock protein gene expression in common killifish, Fundulus


Forsman, A. 2014. Effects of genotypic and phenotypic variation on establishment are
important for conservation, invasion, and infection biology. Proceedings of the


Cambridge University Press.

non-adaptive phenotypic plasticity and the potential for contemporary adaptation in


Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. Tree Genetics & Genomes 9:901–911.


Fig. 1 Diagram showing how conservation genomics can be applied to source population selection for species reintroduction. Selected functional SNPs and genome-wide coding region SNPs (cSNP) can be used to measure standing genetic variation available for natural selection. Transcriptional profiling of candidate genes and/or whole transcriptome analysis can be used to quantify the population’s scope for acclimation in response to environmental stress.
Fig. 2 Numbers of papers published that used DNA-based markers for conservation applications from 2005 to 2014. The solid line with filled circles are all publications that used microsatellite or SNP markers with a conservation application ({conservation} AND {microsatellite OR SNP}). The dashed line with open circles are only those publications that used a genome-wide marker approach ({conservation} AND {whole genome}, followed by inspection for relevant studies).