**Genetic variation and aquatic plant management: Key concepts and practical implications**

RYAN A. THUM*

**INTRODUCTION**

Evolution can occur rapidly, and over contemporary timescales (Hairston et al. 2005). Adaptation can occur during time frames that are relevant to natural resource management. Terrestrial weed scientists have learned this first hand by watching in real time the rapid evolution of resistance by many species to many different herbicides (see Heap 2017). The evolutionary potential of weeds is determined by genetic variation, which can be much greater than is recognized by some traditional weed scientists that do not have background or interest in genetics.

Genetic variation has not historically been a major focus of aquatic plant management research or practice. Yet, genetic studies reveal that managed aquatic plant taxa can exhibit genetic variation that can be relevant to their potential for growth, spread, impact, and control efficacy. For example, the genetic basis of fluridone resistance in hydrilla [Hydrilla verticillata (L.f) Royle] has been uncovered (Michel et al. 2004), and genetic screening of hydrilla populations can be used to predict fluridone efficacy (Benoit and Les 2013). Similarly, evidence for management-relevant genetic variation has been identified in watermilfoil (Myriophyllum spp.; e.g., Berger et al. 2012, Thum et al. 2012, LaRue et al. 2013), fanwort (Cabomba caroliniana Gray; Bultemeier et al. 2009), and flowering rush (Butomus umbellatus L.; Lui et al. 2005). Understanding where, when, and how genetic variation will be important for management decisions and outcomes is still unknown, and the study of genetic variation and its impacts on management therefore provides exciting prospects for the invasive aquatic plant management research agenda.

In this paper, I provide a conceptual overview of the types of genetic studies that could be areas of research focus for invasive aquatic plant management. It is not intended to be an exhaustive review of genetics, nor is it intended to review specific methodology and protocols, because those are ever-changing with rapid technological and computational advances in genetics. Rather, it is intended to introduce some of the salient features of the broad field of genetics that aquatic plant researchers could consider when developing and designing projects.

The term “genetics” broadly encompasses many different concepts and methods. Therefore, before undertaking a project about genetic variation, it is helpful to identify which specific aspects of genetics are of interest. The larger field of genetics can be broadly divided into numerous different subdisciplines, and I refer readers to any of the numerous college-level genetics textbooks available for further introduction to these subfields (e.g., Pierce 2014, Brooker 2015, Snustad and Simmons 2016). This review focuses on two specific subfields that are most immediately accessible to genetic studies of invasive aquatic plants: population genetics and ecological genetics.

**POPULATION GENETICS**

Population genetics is the subfield of genetics concerned with the patterns, causes, and consequences of genetic variation in populations, and how genetic variation varies over space and time. Population genetic variation is most commonly quantified by measuring allele frequencies for one to many molecular markers, as opposed to quantifying phenotypic variation (see Ecological Genetics section below). The molecular markers used for population genetic analyses are commonly considered to be “neutral,” and not directly associated with any phenotypic traits. However, this is not always the case, and many current population genetics studies are specifically designed to look for genetic loci that are under selection and/or related to particular phenotypic traits of interest (see “Detecting selection and adaptation” below).

Basic population genetic descriptions can be informative for aquatic plant managers. Understanding how much genetic diversity occurs within a population can provide insight into the potential for variation in ecological traits. For example, many aquatic plant species are capable of asexual propagation, and it is therefore possible for a population to contain anywhere from a single genetic clone to many genetic clones that might differ in traits that are of interest to managers. Molecular markers can be used to test these alternatives. Similarly, molecular markers can be used to determine whether two populations are genetically similar or different, which can provide insight into the potential for different populations to exhibit different ecological traits of interest to managers. Identifying genetically distinct populations can be helpful for designing experiments that test for significant variation in phenotypic traits of interest. Finally, genetic surveys of individual populations over time can determine whether populations are genetically stagnant versus dynamic, which might provide insight into the potential for ongoing and contemporary adaptation (e.g., herbicide-resistance evolution).

In addition to basic population genetic descriptions, I highlight below three prominent emphases of current population genetics for which there are very few, if any, published studies on invasive aquatic plants, but that could...
be pursued by interested students and researchers. The molecular techniques and statistical analyses used will depend on the specific questions being asked, and logistical constraints in any given study system. An exhaustive review of methods is beyond the scope of this paper, but I briefly highlight considerations for molecular markers and sampling schemes in each area.

**PHYLOGEOGRAPHY**

The term “phylogeography” refers to the study of historical processes that have influenced the contemporary geographic distribution of genetic variation across the landscape. For example, numerous studies have considered how Quaternary geology (e.g., glacial-interglacial cycles during the past 2 million yr) has influenced contemporary genetic variation (e.g., see Hewitt 1999 and Soltis et al. 2006 for extensive reviews of examples). Many of these studies have found genetic signatures that indicate the isolation of lineages into distinct geographic refugia during glacial maxima, followed by expansion of these lineages following glacial retreat. The signatures of these events are still identifiable in contemporary populations using certain molecular markers.

Phylogeographic studies of aquatic plants can be relevant to aquatic plant management. For example, biogeographic patterns of genetic variation in invasive aquatic plant taxa can be used to infer their geographic origins and introduction pathways and vectors. Information on geographic origins is of specific interest to researchers searching for biocontrol agents, which are often specific to geographic regions. Furthermore, inferences about pathways and vectors for introduction can be used to inform prevention efforts. In addition, because of its emphasis on history and species formation, phylogeographic studies can lead to the identification of genetically distinct groups that can represent cryptic taxa (e.g., Thum et al. 2011). Because genetically distinct cryptic taxa can exhibit ecological differences that are relevant to managers, subsequent field and laboratory study might wish to distinguish these groups (e.g., Tavalire et al. 2012).

The emphasis on the history and biogeography of populations distinguishes phylogeography from other areas of population genetics that focus on contemporary factors and processes. At its core, phylogeography depends on constructing gene genealogies (the evolutionary relationships of different alleles at one or more genes), and relating these to the geographic distributions of alleles. I direct interested readers to Avise (2000) and Knowles (2009). Because of its emphasis on genealogies, phylogeographic studies typically utilize DNA sequence data, because these data lend themselves well to genealogy construction.

However, other molecular markers, such as microsatellites and amplified fragment length polymorphisms (AFLPs) can be used. Similarly, because the emphasis is on geographic distributions of alleles, sampling strategies for phylogeographic studies should emphasize the number of sampling locations, and should seek to cover as much of the geographic range of the species being investigated as possible. For example, if a study only had resources to collect data on 500 individuals, it would be better to sample 10 individuals from each of 50 locations scattered across the geographic range of the focal species than it would be to sample 50 individuals from each of 10 locations across the species’ range. In addition, when attempting to infer geographic origins of invasive species, the geographic coverage and number of locations sampled throughout the native range should be maximized.

**LANDSCAPE GENETICS**

A related subfield of population genetics is landscape genetics. This subfield is similar to phylogeography in its emphasis on spatial patterns of genetic variation, but is distinct from phylogeography in its emphasis on present-day factors influencing genetic variation. At its core, landscape genetics combines genetic data on populations with landscape ecology methods to relate contemporary patterns of gene flow to the landscape. This is a rapidly developing field, and I refer readers to a recent text by Balkenhol et al. (2015) for further details.

Landscape genetics could be immediately relevant to aquatic plant managers in at least two ways. First, landscape genetic approaches could be used to identify features of the landscape that facilitate versus impede the movement of invasive aquatic plant taxa among water bodies, which could help inform management efforts to limit and prevent their spread. For example, landscape genetic studies of introduced red fox (*Vulpes vulpes*) in California have helped identify areas of introduction and routes of geographic spread, and have provided insight into the importance of localized reproduction versus immigration in introduced populations (Sacks et al. 2016). Second, landscape genetics studies increasingly attempt to identify genetic signatures that indicate adaptation to environmental conditions across the landscape. Such approaches could be used to inform questions about the extent to which invasive aquatic plant species adapt to environmental conditions, including management (see also next section).

Because landscape genetics focuses on a more contemporary timescale compared to phylogeography, landscape genetics studies often use molecular markers that evolve quickly and are highly diverse, such as microsatellites or single nucleotide polymorphisms (SNPs). These markers are generally better-suited to comparing genetic diversity among populations that share recent ancestry compared to more slowly-evolving DNA sequence based markers (e.g., mitochondrial or chloroplast DNA). In general, the more molecular markers that are used, the more statistical power the resulting dataset will have to detect genetic differences among populations, and to relate those differences to landscape features.

It is also important to consider the extent of geographic sampling, how to select populations for sampling, and how many individuals to sample per population. The number and locations of populations to sample will vary depending on the biology of the species, the size of the geographic area of interest, and the project constraints. However, a good rule of thumb is to sample across hypothesized barriers or environmental gradients, so that these landscape features...
are explicitly captured in analyses. Because landscape genetics analyses generally utilize allele frequency estimates within and among populations, it is important that a sufficient number of individuals are sampled per population to estimate allele frequencies. The exact number will vary based on the biology of the species and the constraints of the project, but generally greater than 10 individuals per population at a minimum, to upwards of 50 or more.

**DETECTING SELECTION AND ADAPTATION**

Evolutionary ecologists are frequently interested in detecting selection and adaptation with population genetic data. For example, population genetic data were used to demonstrate rapid adaptation in flowering time associated with establishment and spread of introduced Pyrenean rocket (*Sisymbrium austriacum* Jacq. subsp. *ehryanthum* Rouy & Foucaud) in parts of Europe (Vandepitte et al. 2014).

Population genetic methods for detecting putative selection and adaptation could therefore be relevant to aquatic plant managers. For example, the evolution of herbicide resistance is currently of great concern to weed scientists, and very little is currently known about how commonly invasive aquatic plant management practices lead to herbicide resistance. Utilizing population genetic methods to identify selection for herbicide resistance evolution is thus one possible application of genetics in aquatic plant management.

There is a burgeoning field of developing analyses aimed at detecting selection and adaptation with population genetic data. At their core, these methods rely on utilizing “genome scans” to identify “outlier loci” that exhibit unique genetic signatures relative to the majority of loci. Briefly, this involves collecting allele frequency data for many molecular markers scattered throughout the genome in two or more populations. The populations chosen for study would typically differ in a specific trait of interest. For example, genome scans can be used to compare populations that were known to exhibit different responses to an herbicide, and molecular markers that exhibit unusually large genetic differentiation relative to the bulk of molecular markers would be identified as markers that are near genes that are involved in herbicide resistance versus susceptibility. Similarly, genome scans in the same population conducted over different time points can identify molecular markers that exhibit unusually rapid genetic changes relative to other markers, which could indicate genes that are involved in ongoing adaptation to environmental conditions.

Because genome scans rely on estimating allele frequency differences among populations of interest, it is best to maximize the number of individuals to maximize the precision of allele frequency estimates. Similarly, because the number of genes under selection might be very small relative to the whole genome (“needles in a haystack”), it is best to maximize the number of molecular markers to maximize the genomic coverage. The specific molecular markers used can vary among studies, but next-generation genotyping-by-sequencing methods are increasingly used because they tend to have high coverage across the genome relative to molecular markers such as microsatellites. Finally, populations should be carefully chosen for genome scan studies, and should be known to differ significantly in the trait(s) of interest.

Despite the rapid ongoing methodological developments for detecting selection and adaptation in natural populations, it is important to note that distinguishing the exact evolutionary causes for genetic variation within and among populations is inherently difficult, because different evolutionary forces can lead to similar patterns of genetic variation. For example, selection and population bottlenecks can leave similar genetic signatures on individual molecular marker loci. For this reason, pairing population genetic data with other types of studies can be more powerful than relying solely on population genetic data to make inferences about the causes and consequences of genetic variation. For example, laboratory common garden and/or field reciprocal transplant experiments are powerful ways to determine whether populations exhibit genetic differences for specific traits of interest (e.g., herbicide tolerance versus susceptibility). Similarly, genetic mapping of quantitative traits (Quantitative Trait Locus mapping; QTL) provides an independent test of whether outlier loci identified by genome scan methods in natural populations are actually associated with traits of interest. Finally, functional genetic studies are ultimately needed to determine the mechanisms by which different loci and alleles lead to phenotypic differences among individuals and populations. Nevertheless, population genetic methods can provide powerful genetic descriptions to inform experimental comparisons.

**ECOLOGICAL GENETICS**

Ecological genetics focuses on the genetics of phenotypic traits that are ecologically important, and are related to organismal fitness (i.e., affect survival and reproduction). For aquatic plant management researchers, any phenotypic traits related to the establishment and spread potential, or to tolerance versus susceptibility to control methods (e.g., herbicide tolerance), would be of specific interest. Ecological genetics emphasizes genetics of natural populations, as opposed to laboratory strains that are commonly the focus of other subfields of genetics (e.g., classical genetics, molecular and biochemical genetics, etc.). Ecological genetics research commonly involves both field and laboratory studies, and commonly considers variation in focal traits across space and/or time. For all of these reasons, ecological genetics provides an accessible framework and opportunities for invasive aquatic plant studies.

Determining the relative influence of genetic versus environmental factors on phenotypic variation in focal traits within and among natural populations lies at the heart of ecological genetics. This concept is specifically relevant to aquatic plant managers, because they seek to alter environmental conditions in ways that decrease the fitness of target populations. For example, herbicide applications are specific environmental perturbations imposed to decrease survival, growth, and reproduction. However, the effect of environmental variables can differ among genotypes, and
genetic variation is the raw material for evolutionary change. The extent of genetic variation for fitness under specific herbicide conditions will determine the evolutionary potential for herbicide resistance.

Some phenotypic traits are discrete (e.g., Mendel’s classic examples of yellow versus green peas, or smooth versus wrinkled peas). Such traits are commonly determined by a single gene, and are little influenced by environmental conditions. However, many traits that influence organismal fitness are continuously distributed, and are called quantitative traits. Quantitative traits are commonly controlled by several to many different genes, and are also commonly influenced by the environment. For most growth and reproduction-related traits of interest to invasive aquatic plant managers, we do not know the underlying genetic basis. Therefore, the environmental and genetic sources of variation in these traits within and among populations must be determined statistically. This is done by measuring the phenotype(s) of interest in individuals of known relatedness in one or more environments, and using statistical models to parse out the effects of genetic and environmental variation. I briefly outline the main concepts of quantitative genetics in the remainder of this section. I refer readers to excellent texts on this subject for further reading (Falconer and Mackay 1996, Lynch and Walsh 1998, Conner and Hartl 2004).

The total phenotypic variation in a study population (Vp) is determined by the combination of environmental (Ve) and genetic variation (Vg): Vp = Vg + Ve. The proportion of the phenotypic variation in the study population that is due to genetic variation is known as the heritability of the trait. Heritability ranges from 0 to 1: a value of 0 indicates that the observed phenotypic variation in the study population is due entirely to environmental variation, whereas a value of 1 indicates that the observed phenotypic variation in the study population is due entirely to genetic variation. Most heritability estimates are somewhere in between 0 and 1, because most traits are influenced by both genetic and environmental variation. Note, too, that the influence of the environment on a phenotype can depend on the genotype (genotype-by-environment interaction), in which case the way in which an environmental gradient affects the phenotype (the “reaction norm”) is itself inheritable. For weeds specifically, there are numerous published studies on wild radish (Raphanus raphanistrum L. and R. sativus L.) that provide excellent illustrations of the concepts above.

Heritability of traits is of primary interest because it determines how much the phenotype can evolve, because genetic variation provides the raw material for evolution. If the phenotypic variation in a population is primarily determined by environmental variation, the population will have a limited capacity to evolve. However, if the phenotypic variation in a population is determined substantially by genetic variation for the trait, then the population will have high evolutionary potential. The evolutionary potential for herbicide resistance is of special concern to weed scientists. For example, heritability studies have demonstrated evolutionary potential for glyphosate resistance in rigid ryegrass (Lolium rigidum Gaudin; Busi and Powles 2009).

In the broadest sense, the heritability of a trait is the fraction of total phenotypic variation that is due to total genetic variation. But, note that genetic variation can be further broken down into different forms (see below). Heritability estimates based on total genetic variation are known as the broad sense heritability of a trait, and can be expressed as H² = Vg/Vp. Broad-sense heritability is an appropriate estimate of genetic variation for clonal organisms, because the entire genotype is passed down directly from parents to offspring.

However, sexually-reproducing organisms do not pass down entire genotypes to their offspring, but pass on one of their two alleles at random from each locus to each of their offspring. Alleles are additive in their effects when the impact of alleles on the phenotype is determined simply by adding up the individual effects of each allele (Va). However, different alleles at a given gene can interact nonadditively to determine a phenotype. If interaction occurs between alleles at the same genetic locus, it is known as dominance (Vd). If interaction occurs among alleles at different loci, it is known as epistasis (Ve). Thus, the total phenotypic variation can be expressed as Vp = Va + Vd + Ve, where Va is known as the additive genetic variance, and Vd and Ve are collectively referred to as nonadditive genetic variance. Because it is individual alleles that are passed on during sexual reproduction, heritability for sexual organisms is most appropriately determined using the additive genetic variance. This is known as narrow-sense heritability, and can be expressed as h² = Va/Vp.

The design of a heritability study depends on several important factors, including the mating system (asexual versus sexual), desire to split up different components of genetic variance (additive versus nonadditive), environmental factors of potential interest or importance, and logistics (e.g., some organisms are difficult to do controlled breeding with), among other things. For clonal organisms, the breeding design is very simple. Different clones of the same genotype (i.e., different ramets of the same genet) are used as replicates in one or more environments, and an appropriate analysis of variance is used to test for a significant effect of genotype (i.e., the among-clone variation) to indicate broad-sense heritability. This design is sufficient because broad-sense heritability is relevant for clonal organisms that pass down total genetic variation to their offspring. However, for sexually-reproducing organisms, narrow-sense heritability is more appropriate because individual alleles are passed down to offspring, not genotypes. In its essence, narrow-sense heritability methods rely on generating sets of offspring (families) of known relatedness (e.g., full siblings or half siblings) and using an appropriate statistical model to partition out the different kinds of genetic variance. For example, a significant effect of “family” indicates heritability, and the variance components can then be used to calculate the different kinds of genetic variance, depending on the study design. I refer readers to Falconer and Mackay (1996), Lynch and Walsh (1998), and Conner and Hartl (2004) for more details. In general, heritability studies should strive for 20 or more genotypes (for clonal organisms) or families (for sexual organisms). However, the exact number of genotypes or
families will vary according to the questions of interest, and logistical realities.

Heritability is a statistical property of a study population, and therefore there is no universal heritability for a trait. Heritability frequently varies among different populations, and among the same population in different environments or over time. This is true because populations frequently differ in the number and frequencies of different alleles at genes affecting phenotypic traits. For example, a population might be fixed for a single allele at a gene affecting a trait, whereas a second population might be fixed for a different allele at that same gene. Let’s assume that the different alleles lead to different phenotypes. The heritability for the trait in both populations is 0, because there is no genetic variation in either. However, the heritability for the trait if both populations are considered a single population would be positive, because there would then be phenotypic and genetic variation in the combined study population. Similarly, heritability for a phenotypic trait in a population will change over time along with any changes in the allele frequencies at genetic loci influencing the trait. Furthermore, the heritability of a trait for a given population can also differ among environments, because environmental variation contributes to phenotypic variation. For example, stressful environments can cause greater environmental variance, and estimates of heritability in these environments would tend to be lower than in less stressful environments.

Finally, note that the focus on phenotypic traits means that ecological genetics studies do not require the use of molecular genetic markers to determine the genetic basis for traits. However, such studies increasingly utilize molecular markers to genetically map traits of interest to particular area(s) of the genome (Quantitative Trait Locus Mapping; QTL). These studies can be used to estimate the number, effects, and locations of genomic regions associated with variation in phenotypic traits. Genetic mapping therefore represents an intersection between population and ecological genetics. For example, once genomic regions associated with specific phenotypes are identified, these regions can be further investigated to identify the genetic sources of variation within and among populations and across time.

PROSPECTS FOR INTEGRATING GENETIC VARIATION INTO AQUATIC PLANT MANAGEMENT

Ultimately, there are numerous opportunities to study genetic variation in ways that are relevant to aquatic plant management. Studies of genetic variation will lead to a deeper understanding of the number of genes affecting phenotypic traits of interest, their locations throughout the genome, and the effects of different alleles on those traits. This understanding, in turn, could lead to genetic methods that predict important phenotypic traits of interest, such as the potential for establishment and spread, or the predicted efficacy of a proposed control method. For example, the mutations that confer target-site herbicide resistance are now known for many herbicides and species (see Powles and Yu 2010), and genetic assays for these mutations can therefore identify herbicide-resistant populations, and track their spread across populations. An important example from aquatic plant management is genetic screening of hydrilla populations to predict fluridone efficacy (Benoit and Les 2013).

In an immediate sense, adaptive aquatic plant management programs could conceivably incorporate genetic data into their survey, monitoring, and evaluation efforts. Population genetic surveys of invasive aquatic plant species could be used to inform the design of genetic studies aimed at practical implications of phenotypic variation for important traits, such as growth and herbicide susceptibility. For example, surveys conducted across geographic locations with the target species could be used to determine whether there are any distinct genetic groups that occur within or among water bodies. If distinct groups are identified, they could be explicitly considered in laboratory- and field-based studies of growth and management to determine whether any differences in phenotypic traits related to growth, reproduction, or control efficacy correspond with the different genetic groups. Similarly, temporal monitoring of population genetic variation over time in managed and unmanaged lakes could be used to determine whether there is any evidence for evolutionary changes as a result of management, which could inform laboratory comparisons of populations to determine whether relevant phenotypic evolution is ongoing.

ACKNOWLEDGEMENTS

I am grateful to Jay Ferrell for the invitation and opportunity to contribute to this compendium. I thank Greg Chorak, Brian Ross, Jeff Pashnick, Paula Guastello, and an anonymous reviewer for helpful suggestions that improved the manuscript.

LITERATURE CITED


LaRue EA, Zuellig MP, Netherland MD, Heilman MA, Thum RA. 2013. Hybrid watermilfoil lineages are more invasive and less sensitive to a commonly used herbicide than their exotic parent (Eurasian watermilfoil). Evol. Appl. 6:462–471.


