Aquatic species represent a vast diversity of metazoans, provide humans with the most abundant animal protein source, and are of increasing conservation concern, yet landscape genomics is dominated by research in terrestrial systems. We provide researchers with a roadmap to plan aquatic landscape genomics projects by aggregating spatial and software resources and offering recommendations from sampling to data production and analyses, while cautioning against analytical pitfalls. Given the unique properties of water, we discuss the importance of considering freshwater system structure and marine abiotic properties when assessing genetic diversity, population connectivity, and signals of natural selection. When possible, genomic datasets should be parsed into neutral, adaptive, and sex-linked datasets to generate the most accurate inferences of eco-evolutionary processes.

Landscape Genomics and Aquatic Organisms

Aquatic species and their ecosystems play fundamental roles in sustaining global biodiversity and human populations [1]. Marine and freshwater ecosystems alike face numerous environmental challenges [2], which is an alarming fact considering that they harbor a tremendous amount of described metazoan flora and fauna. Environmental stressors are the greatest threat to freshwater habitats, which have caused a 83% decline in species abundances since 1970 [3]. Many marine fisheries are overexploited and on the brink of collapse [4]. Yet, little is known about how environmental changes are impacting the health and evolutionary potential of aquatic species, and under what conditions adaptation may occur. To address these needs, landscape genomics provides a powerful framework for understanding eco-evolutionary processes, assessing the viability of populations, and predicting the future health of species and aquatic ecosystems.

Landscape genetics emerged as a formal discipline 15 years ago as a powerful means to address problems of understanding how the interaction between ecological, evolutionary, and geographic factors influence population genetic structure (Box 1; [5]). More recently, the development of high-throughput genomic tools [6] made it possible to move from landscape genetics to landscape genomics – whereby genetic variation can be screened at the scale of the entire genome – offering greater power to disentangle adaptive from neutral genetic divergence and identify environmental factors acting as selective agents [7].

We define landscape genomics as ‘The use of genomic technologies to study genome-wide neutral and adaptive variation of ecologically diverse populations across heterogeneous landscapes to address novel or previously intractable questions’, such as forecasting of adaptive capacity (see Glossary) under environmental change [8]. Clearly, landscape genetic/genomic studies to date have been biased towards terrestrial ecosystems (Figure 1; [9]). Of all landscape
Trends in Ecology & Evolution

Box 1. A Perspective on the History of Landscape Genetics/Genomics

The roots of landscape genetics may be traced to 19th century biogeographers who noted variable community composition and species traits across the landscape [5]. The first theoretical articulation of spatial variation for neutral traits was Wright’s ‘isolation-by-distance’ (IBD) [84]. Cline analysis was an early analytical framework for landscape genetics because clines are associated with local adaptation and gene flow between populations [85]. Landscape genetics emerged as a discipline following: (i) development of methods to resolve genetic or protein (alcohol) variation at multiple loci in the 1960s and the realization that natural populations housed variation associated with environmental factors [86]; (ii) the DNA revolution beginning in the 1980s and associated growth of conservation genetics and molecular ecology; (iii) the realization that human alterations to habitats could impact genetic variation; and (iv) the subsequent founding of landscape ecology as a discipline in the 1990s [87]. All led to the notion that the concepts and tools of population genetics and landscape ecology could be combined to understand environmental heterogeneity and its impacts on genetic diversity, divergence, and microevolutionary processes. These ideas coalesced in Epperson’s Geographical Genetics [88] and the first definition of landscape genetics [5].

A novel aspect of landscape genetics was the use of individual-based approaches to assess fine-scale variation and more precisely localize barriers relative to population-level approaches. The landscape genetic approach initially focused on genetic assays and analytical methods available at the time with, understandably, little ability to draw broad inferences about pattern or process [5]. Lukart et al. [30] advocated a population genomic approach to studying associations between genetic and environmental variation, that is, simultaneously examining neutral and adaptive variation across putative selection gradients at thousands of loci across the genome. Later, landscape genetics was expanded by explicitly including adaptive and neutral variation and specifying the study of landscape composition and configuration, including matrix quality [89]. This idea was extended by calling for explicit quantification of landscape effects on genetic variation [47]. Other reviews highlighted: the formal recognition of a landscape genomic approach [8,67], landscape genetics in conservation [80], plants [91], infectious diseases [92], that neutral and selective factors impacting the genome may include species interactions, that is, landscape community genetics (Box 3; [27]), and the first textbook on the subject [87].

genetics papers published since 1991, only 13% were on aquatic systems (9% on freshwater and 4% in marine systems). This is partly because genomic resources are lacking for most aquatic species (see [10] for a marine–terrestrial comparison). Substantial differences exist between terrestrial and aquatic systems (see ‘Waterscape Characteristics’ below), questioning the translatability of terrestrial landscape genomics approaches to aquatic systems.

Waterscape Characteristics

Aquatic and terrestrial systems differ in fundamental ways relevant to landscape genomics. Water is often flowing with some current; therefore, most aquatic organisms need to spend more energy to stay in place than move. Marine and freshwater systems have many divergent properties, including differences in patterns of biodiversity, suggesting that processes generating biodiversity, and potentially tractable through landscape genomics, may differ between these realms. For instance, ~40% of all named fish taxa are found in fresh water, yet the percentage of the Earth’s surface that is fresh water is miniscule compared with the marine realm (0.8% vs 71%, respectively) [2].

The physical properties of water have created an environment that uniquely affects aquatic organisms and their eco-evolutionary dynamics. Water is ~800 times more dense than air and at least 40 times as viscous, but provides greater buoyancy. Water also has a higher thermal capacity (ability to maintain temperature) and conductivity (ability to transfer heat) than air. Oxygen solubility is inversely related to water temperature, with hypoxic conditions occurring for many aquatic organisms that experience warm temperatures [11]; thus, the coupling of temperature and oxygen has likely driven adaptations in aquatic ectotherms [12,13]. Aquatic environments also present particular physiological challenges for diadromous species that move between marine (hypertonic) and freshwater (hypotonic) environments (e.g., [14]).

Aquatic landscapes contain tremendous variation in habitat complexity and physical connectivity that distinguish them from terrestrial habitats. Both marine and freshwater environments
are highly dynamic with diel fluctuations in tides and currents in marine systems, or variation in daily discharge, water depth, and temperature in fresh water. Both aquatic environments also have seasonal fluctuations including upwelling in marine environments and flow rates in freshwater systems. In contrast to marine habitats, freshwater habitats are hierarchically organized by relative elevation and connected via headwater streams, reaches, and watersheds. Due to the dendritic nature of riverine systems, abiotic characteristics such as river branching extent and confluence position affect genetic variation and population structure (e.g., [15,16]). Furthermore, because predominant river currents are unidirectional, migration is expected to be asymmetric. In contrast, marine environments contain discrete yet connected habitat types such as the pelagic environment, near-shore (e.g., coral reefs and seagrasses), and estuaries (Figure 2). Terrestrial habitats, by contrast, are generally characterized by larger diel and annual fluctuations in temperature, particularly in polar and temperate regions, and are typically more structurally complex with steeper climatic gradients.

Due to the connected nature of aquatic systems, many aquatic organisms can encounter a broad range of habitats over their lifetime. For instance, reproductively mature adults of many species occupy dynamic intertidal and rocky near-shore habitats where temperature and solar radiation go through diel fluctuations, whereas their larval forms are often found in the more homogeneous pelagic zone (e.g., giant green anemone; Anthopleura xanthogrammica). Furthermore, in fresh water, some species may be adfluvial where juveniles born in streams move to lakes to mature before returning to streams as adults for spawning (e.g., bull trout; Salvelinus confluentus).

Landscape Connectivity and Gene Flow

Although aquatic systems have often been overlooked in favor of terrestrial systems for developing genetic connectivity model theory (e.g., least-cost path, circuitscape, etc.; Box 2), they provide a range of conditions and challenges to test methods and models [17–20]. Because of the physical properties of water, dispersal energetics are distinct in aquatic versus terrestrial environments. Consequently, aquatic organisms have evolved a myriad of behavioural, morphological, and life history traits that impact connectivity [10]. In freshwater systems, streams and rivers can often be represented in a one-dimensional cost surface. Conversely, marine environments often provide the ultimate challenge in connectivity modeling because partially/poorly defined barriers can lead to weak population structuring ($F_{ST}$ is typically <0.01), in a vast 3D environment where species interactions and survival are not well understood [21].

Measuring Genetic Connectivity in Aquatic Systems

In freshwater systems, migration can be easier to measure than in terrestrial ones, specifically in streams and rivers whose dendritic structure typically results in well-defined migratory paths. Link-based approaches are often applicable to these systems by using a least-cost path [17] or stream-tree approach where pairwise genetic distances ($F_{ST}$) are fitted to non-overlapping stream segments to assess population connectivity (Box 2; [19]). A challenge that remains, however, is determining resistance along migratory paths ([22], but see [20]).

Delineating discrete populations and connectivity among them pose distinct challenges in marine systems. A knowledge of physiological requirements such as thermal and osmoregulatory tolerance can be supplemented with predominant oceanic currents to generate Lagrangian dispersal models to inform circuitscape models [23]. In these cases, circuitscape methods can be applied to model gene flow, or an isolation-by-environment (IBE) type approach might also be suitable [24]. Along these lines, Xuereb et al. [25] determined that genetic connectivity of two populations of the giant California sea cucumber (Apostichopus californicus)

Glossary

Adaptive capacity: ability of a population to evolve in response to changing environments such that the mean population fitness is maintained or increases following the change.

Adfluvial: aquatic organisms that breed and develop in streams and subsequently enter nearby lakes to reach sexual maturity.

Anadromy/anadromous: migration strategy where an individual is born in fresh water, subsequently migrates to the marine environment where it develops as an adult, then returns to fresh water to spawn.

Association mapping: uses natural populations, as opposed to controlled breeding lines, to associate genomic regions with a trait (phenotypic or environmentally related) of interest (also known as LD mapping).

Catadromy/catadromous: migration strategy where an individual is born in the marine environment and subsequently migrates to fresh water to rear and develop, then returns to the marine environment to spawn.

Connectivity modeling: the application of a computational model (e.g., least-cost path, circuit theory, dispersal kernel, etc.) on a cost surface (resistance) map.

Cost surface: representation of the fitness cost associated with features of a landscape/waterscape for a given species as a set of spatially discrete weights (also known as a resistance map).

Dendritic network: the spatial arrangement of river basins in hierarchic units such as reaches, streams, subcatchments, and catchments, where two water segments join at confluence points and become a single segment.

Diadromy/diadromous: an organism that spends part of its life in fresh water and part in saltwater; see anadromy and catadromy for examples.

Environment-associated SNPs: SNPs with allele frequencies significantly associated with variation in one or more environmental variables of interest. Often identified via GEA analyses. Associated SNPs can be validated as adaptive SNPs.
was driven by local asymmetric currents as opposed to distance alone (isolation-by-distance; IBD). Duranton et al. [26] recently used haplotype length information in European sea bass (Dicentrarchus labrax) to estimate timing, directionality, and amount of gene flow. Finally, a landscape community genomics approach may help elucidate ecological and evolutionary processes important in structuring populations in particularly challenging systems ([27]; Box 3).

Defining Discrete Populations and Identifying Barriers to Gene Flow in Marine Species
Most barriers in the marine realm are porous or represent spatial clines (e.g., currents or thermal and haline gradients). Marine species are often assumed to have panmictic population structure (random mating resulting in high gene flow) due to the lack of potential barriers to movement. Recent studies, however, have demonstrated that high dispersal ability does not always mean that spatial genetic structure is unresolvable. Indeed, cryptic population structure exists within multiple marine species and is driven by environmental clines [28]. For instance, Benestan et al. [29] used a seascape genomics framework that allowed quantifying the relative importance of spatial distribution, ocean currents and sea temperature on connectivity among American lobster (Homarus americanus) populations.

The increased resolution of genomic data allows investigation of functionally distinct groups such as neutral, adaptive, and sex-linked (in genetically determined sex systems) loci. However, identifying sex-linked markers is difficult for many aquatic species because they are not sexually dimorphic and/or lack the genomic resources to do so. When possible, it is important to organize genomic data in this way because the relative strengths of mutation, migration, selection, and drift differ among these groups [30], which may lead to misleading patterns if analyzed in aggregate. For instance, Benestan et al. [31] showed that relatively few sex-linked markers (12 and 94, respectively), rather than genome-wide drift and gene flow, were driving genetic structure in both American lobster and Arctic char (Salvelinus alpinus). Similarly, adaptive markers associated with phenotype or particular environmental variables under selection often show a different pattern than neutral loci. In redband trout (Oncorhynchus mykiss gairdneri), Chen et al. [12] demonstrated that 5890 neutral loci revealed genetic differentiation as expected under IBD, whereas 13 outlier loci associated with cardiac and physiological function differentiated desert from montane populations irrespective of geographic distance.

Important advances of understanding gene flow and landscape connectivity could be made within the explicit incorporation of candidate adaptive markers into a landscape-resistance modeling framework (e.g., [16]). The addition of adaptive gene flow into connectivity modeling theory could improve understanding of adaptive capacity, as influenced by movement of adaptive alleles among populations [32], or by environmental variables driving selection along a migratory path [20]. Despite being computationally less challenging than terrestrial environments, freshwater systems have not been fully explored for theory purposes and in development of genetic connectivity models (but see [33]). For example, the influence of population topology (the spatial arrangement of populations throughout a landscape) on gene flow and population connectivity is often neglected in fresh waters, but could improve this type of research in terrestrial systems [34].

Genome Scans and Association Studies for Detecting Local Adaptation
Recent advances in sequencing technology, computational approaches, and genomic resources have enabled high-density genome scans to detect local adaptation, as well as genotype–environment associations (GEA) in natural populations [6,35]. In aquatic species, studies through experiments or mapped to genes of functional relevance. Evolutionary rescue: the recovery of a population from environmental perturbation via genetic adaptation. Genotype–environment association (GEA) analyses: univariate or multivariate analyses used to identify candidate adaptive SNPs by testing for direct associations between variation in allele frequencies and environmental variables. Isolation-by-distance (IBD): a pattern where genetic similarity decays with increasing geographic distance between individuals/populations. Isolation-by-environment (IBE): a pattern where genetic similarity decays with increasing ecological distance between individuals/populations. Linkage disequilibrium (LD): the non-random association between alleles at different regions in the genome, often caused by physical genomic proximity. Migration: dispersal of an individual followed by successful reproduction (also referred to as ‘effective migration’). Panmictic/panmixia: interbreeding between populations leading to no population genetic structure. Pelagic: open water in lakes, oceans and seas not near the bottom or shore. Quantitative trait locus (QTL): a genomic region associated with the variation of a quantitative (often phenotypic) trait. SNP: a variant position in the genome. Type I error: false-positive rate – the null hypothesis is rejected when it is actually true. Type II error: false-negative rate – failing to reject the null hypothesis when the alternative hypothesis is true.
have discovered the genetic basis for specific phenotypic traits [36], broad signals of local adaptation across landscapes [12,16,29], and candidate genes for conservation monitoring [37]. Genome scans and GEA tests have become routine and offer immense potential to investigate adaptive variation [38].

Researchers can now address critical questions related to evolutionary adaptation and resilience in aquatic ecosystems (e.g., [39]). Yet, study design for genome scans and GEA tests in aquatic systems requires careful consideration of many factors, some of which are distinct in marine versus freshwater systems. These include (i) sampling strategies; (ii) candidate environmental variables; (iii) marker density across the genome; and (iv) statistical approaches to detect drivers of selection, the type and strength of selection, and candidate genes involved. We focus on genome scans in an association mapping framework because non-model aquatic organisms are often not well-suited for quantitative trait locus (QTL) mapping, salmonid fishes being the exception rather than the rule [40].

Sampling strategy to adequately represent organisms across time and space (and to achieve statistical power) is a crucial component for both marine and freshwater landscape genomics studies, with temporal and spatial replicates needed to rigorously test the stability of selection...
In complex marine systems, additional layers of spatial dimensions must be considered [9,21]. For instance, many species are often broadly distributed across porous dispersal barriers, but population connectivity in the sea can be influenced by climatic gradients [28], spatially and temporarily variable recruitment associated to dynamic local oceanography [42], and multifarious ecological requirements of adults that utilize various niches across daily or seasonal timeframes [43]. Freshwater species show more limited dispersal, but often occupy different components of habitat based on temporal cycles and resource availability [44].

Anadromous or catadromous species that migrate between freshwater and marine environments are exposed to a broad range of conditions throughout their life cycle that may require additional sampling considerations to resolve adaptive variation related to each environment (e.g., [20]). Sampling at different life stages (e.g., larva vs adult) is crucial to confirm whether signals of selection reflect long-term local adaptation among genetically distinct populations (e.g., divergent selection), or short-term selection within the lifespan of individuals in a panmictic
Box 2. Measuring and Modeling Genetic Connectivity

Modeling gene flow within a connectivity framework is rooted in metapopulation and spatial ecology, where migration is described between habitat patches [93]; also see Box 1. Taylor et al. [94] advocated for the importance of understanding landscape connectivity as the degree to which the landscape facilitates or impedes movement among resource patches. This is the definition most often used for functional connectivity, which can further represent the response of individuals (physiological and behavioral) to the structural landscape and can disrupt or modify dispersal patterns that is realized through immediate or deferred mortality costs and risks [87]. Finally, functional connectivity is often measured in terms of the effective distance that represents the cost of a path between suitable habitat patches or across heterogeneous landscapes that is the Euclidean distance weighted by the cumulative resistance of all landscape types traversed [17].

A common approach to measure genetic connectivity in landscape genetics is to statistically compare the effective distance with some measure of genetic distance (often $F_{ST}$ or individual-based genetic differentiation metric, or allele frequencies). The most challenging part of connectivity modeling remains model selection, and there have been multiple simulation-based studies on model selection tests, with a perhaps overemphasis on Mantel tests (Table B: [95]). Despite extensive testing using simulation-based approaches, a consensus remains to be made on the most appropriate (or most correct) model selection test, and development of new approaches (and testing of older ones) is still ongoing.

Recent approaches to measure connectivity have used a mixed-model, maximum-likelihood, population-effects framework to identify linear water features (e.g., streams, canals, and ditches) as potentially important in the dispersal of a wetland bird [96]. Additionally, tools like StreamTree [19] might offer improved granularity in dendritic (or dendritic-like) systems where a specific $F_{ST}$ can be associated with each branch segment rather than each pair of populations; this could therefore be useful in identifying barriers to gene flow between populations. Brauer et al. [16] further used StreamTree with multiple matrix regression and randomization to integrate genetic connectivity model results into a GEA framework for rainbowfish (Melanotaenia fluviatilis).

Table 1. Model Selection Approaches for Assessing Population Connectivitya

<table>
<thead>
<tr>
<th>Statistical approach</th>
<th>Notes</th>
<th>Potential weaknesses</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mantel tests</td>
<td>Most common test for testing IBD in genetic structure</td>
<td>High type I error rates [96]</td>
<td>[97]</td>
</tr>
<tr>
<td>Partial Mantel tests</td>
<td>Common test for significance of ecological distance by partitioning Euclidean distance</td>
<td>High type I error rates [96]</td>
<td>[98]</td>
</tr>
<tr>
<td>Akaikes Information Criterion (AIC) and Bayesian Information Criterion (BIC)</td>
<td>Commonly used in many genetic analyses. BIC more heavily penalizes model overparameterization</td>
<td>Not appropriate for all mixed-model approaches, or multiple regression on distance matrices</td>
<td>[99]</td>
</tr>
<tr>
<td>Distance-based Moran’s eigenvector map (dbMEMs)</td>
<td>Capable of detecting spatial structure at several scales to help control for spatial correlation in tests of $y$-$x$ relationships</td>
<td>None yet determined</td>
<td>[29]</td>
</tr>
<tr>
<td>Mixed-model maximum-likelihood population-effects framework with (MLPE)</td>
<td>Can be used with AIC, BIC, or $R_{PE}^{2}$</td>
<td>None yet determined</td>
<td>[96]</td>
</tr>
<tr>
<td>Multiple matrix regression with randomization (MMRR)</td>
<td>Assesses the relative effects of IBD and IBE</td>
<td>Difficulties in estimating relative importance of correlated variables, as well as choosing best model selection method</td>
<td>[16]</td>
</tr>
</tbody>
</table>

aA non-exhaustive list of model selection approaches used to assess population connectivity in aquatic landscape genetics, including potential weaknesses for each method.
bThe $R_{PE}^{2}$ statistic measures the proportion of observed variation explained by the fixed effects of the model.

Population representing spatially varying balancing selection [45]. Additional sampling considerations include sex ratio of collections (when sex can be identified either phenotypically or genetically) because sex-linked variation could be falsely interpreted [31], and consideration of specific phenotypes within populations that may be controlled by genes of major effect [46]. Finally, detailed phenotyping (phenomics) may provide insight into specific morphology, behaviour, and development related to adaptive ecological processes [40].

A second factor to be considered relates to the choice of candidate environmental variables. Natural history provides the best source of information for developing a priori hypotheses about which variables might be ecologically relevant for the study species. Considerations about how environmental heterogeneity impacts habitat composition and structural and functional connectivity are nonetheless difficult to make ([47]; see Box 2). This is particularly true in marine systems.
Trends in Ecology & Evolution

Box 3. Landscape Community Genomics

Landscape genomics investigates interactions between population genetic diversity and environmental variation, whereas community genetics investigates interactions between genetic diversity and species interactions; landscape community genomics (LCG) is the merging of these two approaches [27]. To fully understand processes of eco-evolutionary change, researchers should consider simultaneously the effects of abiotic (environmental) and biotic (community) factors on demography, evolution, and genomic variation within and among populations.

To design a landscape community genomic study, researchers ideally include multiple strongly interacting species distributed across environmental (selection) gradients and both candidate adaptive and genome-wide (high-density) neutral loci. Here, we discuss three informative LCG examples: a terrestrial LCG study, an aquatic study lacking strong species interactions, and community environmental DNA (eDNA) studies lacking intraspecific population genetic markers.

An exemplary LCG study [100] involved the specialized Acon butterfly (Phengaris alcon), which is sensitive to grassland habitat configuration and requires the presence of the rare marsh gentian plant (Getiana pneumonanthe) and an ant species (Myrmica spp.). Restriction-site-associated (RAD) DNA sequencing was used to assess relations between genetic diversity, connectivity, habitat suitability, grazing (by livestock), and altitude. Climate warming and seasonal grazing abandonment strongly affected the distribution of the Acon butterfly because grazing and climate affect availability of the gentian host plant.

Raeymaekers et al. [101] used a comparative framework to test if two stickleback species differ in neutral and adaptive divergence along an environmental (salinity) gradient. Phenotypic and neutral marker differentiation along with genomic signatures of adaptation were stronger in the three-spined (G. aculeatus) than in the nine-spined (Pungitius pungitius) stickleback. Signatures of adaptation involved different genomic regions in the two species, and thus were non-parallel. Such multispecies studies provide insight into mechanisms underlying evolutionary change and adaptive strategies within landscapes. Future studies that include strongly interacting species (e.g., competitors, predator–prey, and host–pathogen) could prove to be especially informative.

eDNA metabarcoding will allow for genotyping or microhaplotyping of eDNA fragments from each of multiple species, simultaneously. It thus offers a potentially powerful means for population genetic/genomics studies, although few multilocus studies have been published (e.g., [102,103]). This approach will eventually allow for inferences about biotic and abiotic factors shaping population genetic structure and also community structure [27,104–106]. It is exciting to consider that future eDNA metabarcoding studies (including many neutral and adaptive loci) will eventually allow for LCG studies.

because of their asymmetric physical flows and dynamics, inherent non-stationarity, and size of habitats [21]. Landscape mapping that maximizes environmental variance is comparatively easier in freshwater than marine systems, where a large number of observational, modelled, and remotely sensed variables have recently become available for various scales [33]. Genome scans and GEA tests are bound to benefit from the increase in resolution and extent of spatial resources (examples in Table II ) driven by pressing human needs, such as fresh water availability for consumption and irrigation, fisheries resources through biophysical modelling, and tracking plastic in our seas through customizable simulations. These developments are expected to extend our options beyond the traditional candidate variables (e.g., temperature, salinity, and rainfall) and towards environmental mapping capable of informing on natural and anthropogenic disturbances, resource availability, range shifts, and biotic interactions.

Adjusting the density of markers to the research question, particularly in relation to linkage disequilibrium (LD), is a third important aspect when planning genome scan or GEA studies, with specific considerations for freshwater and marine species alike that often have limited genomic resources. It is ill-advised to draw strong inferences regarding candidate adaptive loci in cases where marker density is low and LD is high because adaptive loci can be mis-identified. As a reference point, LD estimates in wild fish populations have been reported from ~1 kb in zebrafish (Danio rerio) and threespine stickleback (Gasterosteus aculeatus) to 10–20 kb in the European eel (Anguilla anguilla) and up to 1 Mb in lake whitefish (Coregonus clupeaformis)
Trends in Ecology & Evolution

Table II. A List of Databases and Software for Researchers to Use in Aquatic Landscape Genomic Studies*

<table>
<thead>
<tr>
<th>Spatial Resource name</th>
<th>Description</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioClim 2.0</td>
<td>Global climate layers for mapping and spatial modeling</td>
<td><a href="http://worldclim.org/version2">http://worldclim.org/version2</a></td>
</tr>
<tr>
<td>BioSim</td>
<td>Simulation of climate-driven models to forecast future events</td>
<td><a href="https://cfs.nrcan.gc.ca/projects/133">https://cfs.nrcan.gc.ca/projects/133</a></td>
</tr>
<tr>
<td>Geoscience Australia</td>
<td>Geospatial datasets for Australia, including multiple online tools for data analysis</td>
<td><a href="http://www.ga.gov.au/">http://www.ga.gov.au/</a></td>
</tr>
<tr>
<td>Global Biodiversity Information Facility (GBIF)</td>
<td>Geographic distribution data for a multitude of species</td>
<td><a href="http://www.gbif.org">http://www.gbif.org</a></td>
</tr>
<tr>
<td>National Hydrographic Network</td>
<td>Geospatial data for Canada’s inland surface waters</td>
<td><a href="https://www.nrcan.gc.ca/earth-sciences/geochemistry/topographic-information/geobase-surface-water-program/21361">https://www.nrcan.gc.ca/earth-sciences/geochemistry/topographic-information/geobase-surface-water-program/21361</a></td>
</tr>
<tr>
<td>Marspec</td>
<td>High-resolution contemporary and paleo marine spatial ecology data</td>
<td><a href="http://www.marspec.org">http://www.marspec.org</a></td>
</tr>
<tr>
<td>Natural Earth Geographic</td>
<td>Public domain map dataset for map making and GIS usage</td>
<td><a href="http://www.naturalearthdata.com/">http://www.naturalearthdata.com/</a></td>
</tr>
<tr>
<td>NOAA WOD (World Ocean Database)</td>
<td>Oceanic datasets from 1 million-year-old sediment records to near real-time satellite images</td>
<td><a href="https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html">https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html</a></td>
</tr>
<tr>
<td>OceanParcels</td>
<td>Lagrangian framework to create customizable particle tracking simulations</td>
<td><a href="http://oceanparcels.org">http://oceanparcels.org</a></td>
</tr>
<tr>
<td>Ocean Surface Current Analyses Real-time (OSCAR)</td>
<td>Near-real-time global ocean surface data</td>
<td><a href="https://www.esr.org/research/oscar/">https://www.esr.org/research/oscar/</a></td>
</tr>
</tbody>
</table>

Genomics

<table>
<thead>
<tr>
<th>Software name</th>
<th>Description</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>BayEnv</td>
<td>Outlier loci and local adaptation identification via allelic frequencies and environmental variables</td>
<td>[72]</td>
</tr>
<tr>
<td>BayeSsEnv</td>
<td>Local adaptation detection via genotypic and environmental data</td>
<td>[73]</td>
</tr>
<tr>
<td>BayeScan</td>
<td>Outlier detection, no environmental data</td>
<td>[74]</td>
</tr>
<tr>
<td>gdm</td>
<td>Generalized dissimilarity modelling and gradient forests</td>
<td>[75]</td>
</tr>
<tr>
<td>Geneland</td>
<td>Identification of populations and their boundaries with genomic and geographic data</td>
<td>[76]</td>
</tr>
<tr>
<td>GESTE</td>
<td>Identification of environmental factors contributing to population structure</td>
<td>[77]</td>
</tr>
<tr>
<td>gILAnd</td>
<td>Univariate method for local adaptation identification via allelic frequencies and environmental variables</td>
<td>[78]</td>
</tr>
<tr>
<td>LEA* (LFMM)</td>
<td>Local adaptation detection via genotypic and ecological data</td>
<td>[79]</td>
</tr>
<tr>
<td>PCAdapt*</td>
<td>Outlier detection, no environmental data</td>
<td>[80]</td>
</tr>
<tr>
<td>PoolParty</td>
<td>Pipeline to identify genes associated with adaptation and phenotypic traits from whole genome resequencing</td>
<td>[81]</td>
</tr>
<tr>
<td>randomForest*</td>
<td>A powerful machine-learning algorithm to discern loci underlying phenotypic traits of environment association</td>
<td>[82]</td>
</tr>
<tr>
<td>vegan*</td>
<td>Implementation of RDA; local adaptation identification via allelic frequencies and environmental variables</td>
<td>[83]</td>
</tr>
</tbody>
</table>

*Denotes R packages.

[48–50]. For species with large effective population sizes, as is the case with many marine taxa, recombination may cause rapid linkage decay requiring high marker density to provide multiple SNP per linkage block to achieve sufficient power for detecting candidate adaptive genes [51]. In systems where LD is high, such as small, isolated freshwater populations, lower-density markers may be adequate to detect signals of adaptive variation, especially in inverted chromosomal regions with extended LD [52,53].
In either case, a quality reference genome assembly that is well annotated is a powerful resource to characterize the genomic architecture of adaptation that includes identification of candidate genes, genomic position, and putative biological function (see the Supplemental Information online; [38]). Although many aquatic species lack genomic resources, community efforts aimed at developing reference genomes across many taxa are expected to lead to tremendous improvements (e.g., Earth BioGenome Project seeks to sequence all known eukaryotic species [54]). Researchers can capitalize on these resources while also seeking to enhance them by contributing data to improve genome assemblies for target species (e.g., linkage maps, Hi-C libraries, and optical maps).

A fourth consideration for genome scans and GEA tests is the choice of statistical analyses that are best suited to address the study question and intricacies of aquatic systems. Genome scans are susceptible to detection of false-positive signals of adaptation, particularly in freshwater species comprising small, isolated populations prone to pronounced drift [55]. On the other hand, detecting local adaptation and genomic outliers can be a challenge in marine species with large and well-connected populations. Fortunately, several studies have provided guidance to balance Type I and II errors [35,41]. Statistical analyses that combine multiple approaches such as outlier tests, genome-wide association mapping, transcriptomics, and GEA offer corroborating evidence for local adaptation in aquatic systems [12]. Significance testing that accounts for multiple SNPs in LD provides stronger evidence than single-marker tests, as does multivariate testing for polygenic effects [40]. Recent simulations suggest that multivariate GEA methods such as redundancy analysis (RDA) provide the best balance of low false-positive and high true-positive rates across a range of demographic histories, sampling designs, sample sizes, and selection levels [35]. Current statistical models used for association mapping typically correct for population structure, but this may come with the caveat of reducing power to detect candidate loci if selection gradients follow the same direction as neutral structure [7]. Background selection combined with genetic hitchhiking can also generate correlation between local recombination rates and genetic diversity that could falsely be interpreted as a signal of divergent selection between populations [56].

Adaptive Capacity, Conservation, and Management of Wild Populations
Landscape genomics may advance conservation management and recovery of threatened and exploited populations by helping to understand their adaptive capacity to evolve under environmental change. Under climate change, ectothermic species face particular stresses to their preferred thermal niches, highlighting the importance of predicting adaptive capacities of aquatic populations [57]. Any intrinsic or extrinsic factors that will affect the strength of the four evolutionary forces can influence adaptive capacity. These include mutation rate and generation time, species life history, amount and architecture of genetic variation, effective population size and thus genetic drift, biotic and abiotic factors impacting the strength and mode of selection, and gene flow from ecologically distinct populations.

Using landscape genomic analyses to identify genotype–environment associations is an obvious first step for assessing selection in wild populations and integrating adaptive capacity into predictive models of vulnerability to environmental change [11,58]. At one end of the spectrum, landscape genomics can help assess adaptive potential of declining populations known to have persisted in variable and often degraded habitats, a topic of increasing importance and debate [45]. For example, in a range-wide study of a poorly dispersing and endangered fish, GEA tests that consider the effects of dendritic riverine structure recovered signals of adaptive diversity associated with a hydroclimatic gradient and human impacts [59]. The possibility that the small populations of this species are responding to selection was further
supported through comparative ecological transcriptomics [60]. Yet, other studies have suggested limited adaptation in small, geographically isolated populations that experience high inbreeding [55].

At the other end of the spectrum, landscape genomics can assess the influence of environmental heterogeneity and disturbance on local adaptation in abundant and exploited species with high gene flow. Such studies have indicated that heterogeneous environments may drive and maintain adaptive divergence among connected populations of marine [13,29], anadromous [20], and freshwater [61] species. These species may have the potential for tracking future environments because their individuals are capable of rapidly spreading alleles that affect fitness over vast distances.

The adaptive potential of a population is related to its ‘genomic vulnerability’, a metric defined within a landscape genomics framework as the ‘mismatch between current and predicted genomic variation based on genotype–environment correlations modelled upon contemporary populations’ [62]. Environment-associated SNPs can also be used to predict the putative environmental range for individuals with known genotypes [63]. This approach can help predict genetically mediated environmental limits across taxa and compare environmental ranges among multiple species over the same landscape [63]. Landscape genomics can also predict the spatio-temporal spread of adaptive alleles and resistance to spread of maladaptive alleles across space [32]. Frameworks for evidence-based genetic management decisions and policies exist (e.g., [64,65]), and in spite of the challenges associated with their implementation, genomic data have been used in many conservation-based decisions (see [66] for examples). When fueled with information about adaptive capacity, these frameworks should improve management plans targeting (i) the recovery of exploited populations; (ii) in situ and ex situ efforts of evolutionary rescue (e.g., captive breeding, translocations, and reintroductions); and (iii) the anticipated redesign of climate-ready populations.

**Concluding Remarks and Future Perspectives**

The lack of aquatic landscape genomics studies compared with the number of terrestrial landscape genomic studies is surprising (Figure 1). This is partly because genomic resources are lacking for aquatic species [10]. Therefore, a pressing need exists to develop resources to improve aquatic landscape genomics studies (e.g., reference genomes, transcriptomes, sex-linked markers, and large SNP catalogs). Freely available environmental databases are increasing for both marine and freshwater ecosystems, as well as geospatial tools and computer programs that help meet the particular challenges that aquatic landscape genomics studies face (Table 1).

Such challenges include more rigorously defining population structure and quantifying genetic and demographic connectivity in the marine realm, and gaining an understanding of landscape genomic patterns of species from understudied geographic regions (see Outstanding Questions). Along these lines, the inter-annual variability of abiotic conditions in many aquatic systems and their population-level effects, particularly in the marine realm, must be recognized; although field sampling is admittedly difficult, future studies would benefit from temporal replicates for understanding landscape genetic processes. Another major challenge is that strong inferences about GEAs may be constrained by false positives [67]. Arguably, associations provide indirect evidence of an actual functional relationship under the influence of natural selection. Consequently, future studies should rigorously test hypotheses derived from GEAs via gene functional analyses (e.g., comparative physiological studies), and perform experimental tests of natural selection [12].
Landscape genomics modeling can help predict population viability by facilitating modeling of complex interactions between biotic and abiotic factors that influence individual vital rates and control population distribution, abundance, growth rates, and species interactions (e.g., Box 3: [27,32]). Yet, this has been rarely conducted in aquatic ecosystems. Consequently, another potentially fruitful research area would be to apply recently developed landscape genomics and metamodels to test the reliability of models in forecasting changes of population sizes, connectivity, and community composition [32,68]. Aquatic landscape genomics research should also increasingly consider the role of differential gene expression and epigenetic inheritance as a mechanism for rapid adaptation [69,70], for instance, in the face of new stressors [12,60]. Similarly, genomics studies are revealing the important role of structural genetic variants (SGVs) in eco-evolutionary processes [53]. Catanach et al. [71] recently showed that in the Australasian snapper (Chrysophrys auratus), the number of base pairs affected by SGV variants was almost three times higher compared with other polymorphisms, such as SNPs, with a sizeable portion of these located in regions under putative selection.

In summary, although further work is needed to improve a quantitative and predictive theory of the genetic basis of adaptation and to validate recent approaches, knowledge derived from landscape genomics studies already provides a foundation to address real-world problems in the evolution and conservation management of aquatic biodiversity.

Acknowledgments
This work was supported in part by Genome Canada and Genome British Columbia (project code 242RTE), B.K.H. and G. L. were supported in part by funds provided by National Science Foundation grant DEB-1639014 and NASA grant NNK14AB84G, and we thank the Australian Research Council for a Future Fellowship (FT130101068) to L.B.B. We also thank C. Brauer and M. Whitlock for providing comments on an earlier version of the manuscript, and F. Alender for sending helpful publications and ideas on the origins of landscape genetic approaches.

Supplemental Information
Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tree.2019.02.013.

References
21. Rigon, C. et al. (2016) Navigating the currents of seascape genomics: how spatial analyses can augment population geno-
abilistic modeling tool for the multi-scale tracking of biotic and
abiotic variability in the ocean. Environ. Model. Softw. 42, 47–54
24. Wang, L.J. and Bradburd, G.S. (2014) Isolation by environ-
ment. Mol. Ecol. 23, 5649–5662
islands of differentiation in the European sea bass. Nat. Com-
mun. 9, 2518
27. Hand, B.H. et al. (2015) Landscape community genomics:
understanding eco-evolutionary processes in complex environ-
cryptic cline in the northwest Atlantic. Sci. Adv. 4, eaap2929
29. Benestan, L. et al. (2016) Seascape genomics provides evi-
dence for thermal adaptation and current-mediated population
25, 5073–5092
30. Luikart, G. et al. (2003) The power and promise of population
genomics: from genotyping to genome typing. Nat. Rev. Gen. 4,
981–984
sequencing: evidence for biases in genetic parameter estimation
and investigation of sex determination systems. Mol. Ecol. 26, 6767–6783
32. Landguth, E.L. et al. (2017) CDMetaPOP: an individual-based,
eco-evolutionary model for spatially explicit simulation of land-
scape demography. Methods Ecol. Evol. 8, 4–11
33. Davis, C.D. et al. (2018) Refining and defining riverscape genet-
ics: how rivers influence population genetic structure. Published
online January 31, 2018, https://doi.org/10.1002/wat2.1269
34. van Strien, M.J. (2017) Consequences of population topology
for studying gene flow using link-based landscape genetic
methods. Ecol. Evol. 7, 5070–5081
multilocus adaptation with multivariate genotype-environment
associations. Mol. Ecol. 27, 2215–2233
36. Prinos, D.J. et al. (2017) The evolutionary basis of premature
migration in Pacific salmon highlights the utility of genomics for
tion units: the genetic basis of adult migration timing in Pacific
salmonids. Evol. Appl. 11, 1518–1526
38. Manel, S. et al. (2016) Genomic resources and their influence
on the detection of the signal of positive selection in genome scans.
Mol. Ecol. 25, 170–184
leads to rapid loss of adaptive variation and restoration potential
177–186
evolution: problems, pitfalls, and promises. Trends Gen. 32,
155–164
41. Lotterhos, K.E. and Whitlock, M.C. (2016) The relative power of
genome scans to detect local differentiation depends on sampling
design and statistical method. Mol. Ecol. 24, 1001–1046
42. Teske, P.R. et al. (2015) On-shelf larval retention limits popula-
Prog. Ser. 532, 1–12
43. Rodriguez-Zarate, C.J. et al. (2018) Isolation by environment in
the highly mobile olive ridley turtle (Lepidochelys olivacea) in
the eastern Pacific. Proc. R. Soc. B, 285, 20180264
44. Fitzgerald, D.B. et al. (2017) Seasonal changes in the assembly
mechanisms structuring tropical fish communities. Ecology, 98,
21–31
and adaptation to environmental change: considerations from
population genomics in fishes. J. Fish Biol. 89, 2519–2556
life-History traits revealed by genome sequencing in Chinook
47. Storfer, A. et al. (2007) Putting the ‘landscape’ in landscape
genetics. Heredity, 98, 128–142
48. Whiteley, A.R. et al. (2011) Population genomics of wild and
landscape zebrafish (Danio rerio). Mol. Ecol. 20, 4259–4276
227, 117–132
threespine stickleback fish. Nat. Commun. 6, 8767
51. Hoban, S. et al. (2016) Finding the genomic basis of local adapta-
188, 379–397
52. Pearse, D.E. et al. (2014) Rapid parallel evolution of standing
variation in a single, complex, genomic region is associated with
281, 20140012
genomics of chromosomal inversions. Trends Ecol. Evol. 33,
427–440
55. Pointer, C. et al. (2017) Do tremendous genetic drift and accu-
mulation of deleterious mutations preclude adaptation to tem-
perature in a northern lacustrine fish? Mol. Ecol. 26, 6317–6335
56. Tine, M. et al. (2014) European sea bass genome and its varia-
tion provide insights into adaptation to euryhalinity and specia-
ion. Nat. Commun. 5, 5770
57. Pearson, R.G. et al. (2014) Life history and spatial traits predict
extinction risk due to climate change. Nat. Clim. Change, 4,
217–221
58. Harrison, K.A. et al. (2014) Using genomics to characterize
evolutionary potential for conservation of wild populations. Evol.
Appl. 7, 1008–1025
fish across a hydrodynamically heterogeneous river basin. Mol.
Ecol. 25, 5093–5113
60. Brauer, C.J. et al. (2017) Comparative ecological transcriptom-
ics and the contribution of gene expression to the evolutionary
potential of a threatened fish. Mol. Ecol. 26, 6941–6956
adaptive divergence despite high gene flow in golden perch
(Macquaria ambigua): implications for management and resil-
ience to climate change. Mol. Ecol. 27, 196–215
climate-driven population declines in a migratory bird. Science,
359, 83–86
63. Manel, S. et al. (2018) Predicting genotype-environment range from
genome-environment associations. Mol. Ecol. 27, 2823–2833
64. Bernatchez, L. et al. (2017) Harnessing the power of genomics
84. Wright, S. (1943) Isolation by distance. Genetics, 28, 114–128