

Advances of genotyping-by-sequencing in fisheries and aquaculture

Yan-He Li · Han-Ping Wang

Received: 16 May 2015 / Accepted: 29 March 2017 / Published online: 3 April 2017
© Springer International Publishing Switzerland 2017

Abstract The use of genotyping has enabled the characterization and mapping of genes and the study of stock identification, population genetics, evolution, ecological speciation, and invasion, as well as genomic evaluation, sex control and sex determination, nutrition, biomarkers for disease, and quantitative trait loci mapping for marker-assisted selection in fisheries and aquaculture. High-throughput variant discovery has been made possible in multiple species by the recent advent of next-generation DNA sequencing technologies. New genotyping methods that are high-throughput, accurate, and inexpensive are urgently needed for gaining full access to the abundant genetic variation of organisms. This approach is known as genotyping-by-sequencing (GBS), which holds great promise as a research tool because of its ability to allow simultaneous marker discovery and genotyping at low cost and with a simple molecular biology workflow for fisheries and aquaculture

studies. Since it was first developed for rice in 2009, GBS has been applied in over 50 species/studies by the end of 2014. It is also increasingly in use in fisheries and aquaculture and has been applied in nearly 40 species/studies from 2015 to present. This review summarizes the genotyping methodologies, recent advances in next-generation DNA sequencing technologies to achieve GBS, and the promises this approach holds as a genome-wide genotyping application in fisheries and aquaculture. Additionally, we discuss the potential of whole-genome sequencing (WGS) in GBS and present the advances of WGS in fisheries and aquaculture.

Keywords Fish · Genotyping-by-sequencing · Next-generation sequencing · Single nucleotide polymorphism markers · Whole-genome sequencing

Electronic supplementary material The online version of this article (doi:[10.1007/s11160-017-9473-2](https://doi.org/10.1007/s11160-017-9473-2)) contains supplementary material, which is available to authorized users.

Y.-H. Li · H.-P. Wang (✉)
Fish Genetics and Breeding Laboratory, The Ohio State University South Centers, 1864 Shyville Road, Piketon, OH 45661, USA
e-mail: wang.900@osu.edu

Y.-H. Li
College of Fisheries, Huazhong Agricultural University, Wuhan, Hubei 430070, People's Republic of China

Introduction

Understanding the relationship between genetic variation and biological function on a genomic scale is expected to provide fundamental new insights into the biology, evolution and pathophysiology of humans and other species (Syvänen 2001). Of all these insights, identifying genetic variation is the first and essential requisite. The use of genotyping has enabled the characterization and mapping of genes and the study of species diversity, invasion, evolution, and

disease, as well as biomarkers for stress, marker-assisted selection (MAS), genomic evaluation, and evolution in fisheries and aquaculture (Clark et al. 2011; Cortés et al. 2011; Kumar et al. 2012; O'Quin et al. 2012; Strandén and Christensen 2011). However, due to the financial cost, most researchers in fisheries and aquaculture have identified genetic markers or variations using low-cost conventional genotyping methods that only focus on a few specific DNA fragments (Hu and Vick 2003; Li et al. 2012). The genetic polymorphisms identified often afforded limited resolution to address the research question. Thus, new genotyping methods based on genome-wide sequencing of hundreds or thousands of loci that are high-throughput, accurate and inexpensive are urgently needed to gain full access to the abundant genetic variation of organisms.

Genotyping is the process of determining polymorphisms among individual's nucleotide sequences using biological assays and comparing it to another individual's sequence or a reference sequence (Levy et al. 2007; Etter et al. 2011a). Genotyping can reveal the alleles that an individual has inherited from their parents, as well as patterns of differentiation due to environmental conditions or experimental treatments. Therefore, genotyping is important in the research of genes and gene variants associated with genomic evaluation, nutrition, stress, disease, and evolution (Clark et al. 2011; Cortés et al. 2011; O'Quin et al. 2012; Strandén and Christensen 2011; Sreeparvathy 2013). In general, genotyping typically involves the generation of allele-specific genomic variants of interest followed by their detection for genotype determination. During the 1990's, RFLP (restriction fragment length polymorphism) markers (Lander and Botstein 1989) dominated the scene, and then PCR (polymerase chain reaction)-based markers, including RAPD (random amplified polymorphic detection) (Lynch and Milligan 1994) and AFLP (amplified fragment length polymorphism) (Vos et al. 1995; Mueller and Wolfenbarger 1999), and simple sequence repeat (SSR or microsatellite) amplification (Zietkiewicz et al. 1994; Powell et al. 1996a) have subsequently become increasingly popular. Meanwhile, SSR, sometimes in combination with higher-throughput PCR markers, such as AFLP (amplified fragment length polymorphism) and TRAP (target region amplified polymorphism) markers (Hu and Vick 2003; Alwala et al. 2006), has also been

frequently applied. Moreover, genotyping of SNP and SSR alleles based on hybridization to DNA microarrays or beads (Heller 2002) and high-throughput sequencing have potential to become more popular (Buerstmayr et al. 2009).

Regardless of the marker used, here we attribute the genotyping technologies used in research to five methodologies (Fig. 1):

1. Enzyme digestion-based genotyping identifies genotypes of the individuals based on analyzing the differences in DNA fragment lengths after treatment with restriction enzyme(s). RFLPs are a typical example of enzyme digestion-based genotyping and show sequence variation based on differences in DNA fragment lengths after treatment with a restriction enzyme and electrophoretic fragment analysis (Lander and Botstein 1989; Botstein et al. 1980; Poland et al. 2012; Powell et al. 1996b).
2. PCR-based genotyping is a kind of genotyping method mainly based on identifying the variations in DNA fragments by PCR. In fact, almost all current genotyping technologies, with only a few exceptions (Kwok and Xiao 2004), require the PCR amplification step, because in most technologies, PCR amplification of a desired allelic region is performed initially to introduce specificity and increase the number of molecules for allelic discrimination (Kim and Misra 2007). SSR and AFLP are examples of this technology (Mueller and Wolfenbarger 1999; Powell et al. 1996a, b; Tautz and Renz 1984; Tautz 1989).
3. Chip-based genotyping is based on hybridization of DNA to microarrays or beads (Tang et al. 1999). This approach is useful in association studies to identify genomic regions that may be associated with a particular phenotype. With this technology, a probe array can be used to query a predefined set of a large number of molecular markers located across the genome (Kim and Misra 2007). Although the methodology has achieved the goal of high-throughput, it is laborious, time-consuming, and expensive to design, produce, and process microarrays suited for specific mapping populations (Huang et al. 2009).
4. ESTs (Expressed Sequence Tags)-based genotyping, is the investigation of individual genotypes by identifying the variations in RNA fragments.

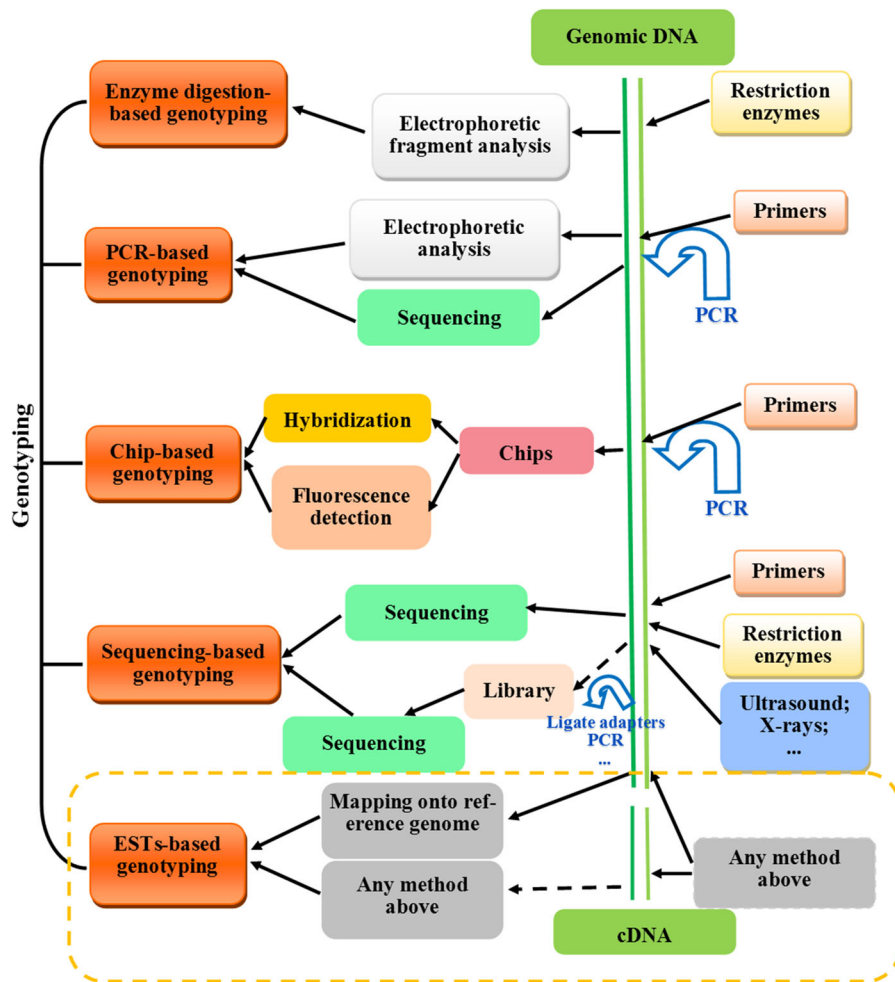


Fig. 1 Major genotyping methodologies and their characteristics. For the methodology based on enzyme digestion, genomic DNA is digested by restriction enzymes to get the different length fragments which are then detected by electrophoresis; for the PCR-based genotyping, primers are designed to amplify genomic DNA and get the predicted fragments which are then detected by electrophoresis or sequencing; for genotyping based on chip, through PCR amplification or other ways, chips contained DNA fragments are obtained, and then checked by fluorescence detection and hybridization; for genotyping based on sequencing, (1) the fragments from PCR are directly

sequenced, which is same as one way of PCR-based, (2) the libraries constructed using the fragments obtained from restriction enzymes, PCR, ultrasound, X-rays etc. are sequenced; for the ESTs-based genotyping (in the box with yellow dotted line), (1) the DNA fragments sequences can be obtained by any method mentioned, then mapped onto a reference genome and get the EST region sequences, (2) the cDNA (complementary DNA) fragments sequences obtained by any method and then detected by any method you use to get the ESTs sequences

Traditionally, EST-based genotyping is performed using the DNA sequences themselves to define biological populations through the use of molecular tools. However, evaluations of the germplasm with single nucleotide polymorphism markers (SNPs) or SSRs derived from coding regions or ESTs in complementary DNAs (cDNAs) may enhance the role of genetic markers

by assaying the variation in transcribed and known-function genes (Eujayl et al. 2002). EST-SSR discovery was initiated based on the single-pass sequencing of cDNAs derived from populations of messenger RNAs (mRNAs) from various tissues or specimens. It is also of interest to compare estimates of genetic diversity calculated from both genomic and EST-derived SSR markers

because the EST-SSR markers represent transcribed regions of the genome (Gupta and Rustgi 2004).

5. Sequencing-based genotyping, is generally centered on single-molecule DNA/cDNA sequencing, and sequencing-based high-throughput genotyping methods, which may provide SNP genotype and haplotype data with an unrivaled level of throughput (Kim and Misra 2007). Single-molecule DNA sequencing methods, such as SSR or mitochondrial gene amplification and sequencing, can provide the haplotype data and genotypic information for species populations (Doubleday et al. 2009; Khamnamtong et al. 2009; Li et al. 2012). Sequencing-based high-throughput genotyping methods, such as GBS and other genome-level sequencing, can simultaneously discover a large number of variants and obtain genotypic information at the genomic level. In most cases, these sequencing-based high throughput genotyping methods represent a combination of the above mentioned methodologies depending on the aim of the research study being conducted. For example, chip-based genotyping includes the PCR methodology, and GBS technology, the new method for genotyping, combines the enzyme digestion, PCR, and sequencing methodology.

Whole-genome sequencing (WGS), where the entire genome of an organism is sequenced remains prohibitively expensive for most applications. However, if we focus on specific regions for gene discovery, for example only the protein-coding portion of the genome, the cost is lower and therefore within reach for a greater number of projects (Majewski et al. 2011). Next-generation sequencing (NGS) makes the discovery of molecular markers based on specific regions of the whole genome cost-effective. Further, impressive progress has been made in the field of NGS (Kothiyal et al. 2009; McPherson 2009), and NGS has lowered the cost of sequencing to an affordable level through advancements in the fields of molecular biology and technical engineering (Shokralla et al. 2012; Buermans and den Dunnen 2014). Additionally, a number of genome complexity reduction strategies for discovery of SNPs such as genotyping by sequencing (GBS), have been developed by capitalizing on the strengths of NGS and its

relatively low cost (Huang et al. 2009; Davey et al. 2011).

SNPs discovery, whose downstream usefulness in linkage map construction (Limborg et al. 2014), genetic diversity analyses (Pujolar et al. 2013; Nunez et al. 2015; Underwood et al. 2016), association mapping (Yoshizawa et al. 2015), and MAS has been demonstrated in several species, is certainly one of the most common recently discovered applications of NGS (Cortés et al. 2011; Kumar et al. 2012). Due to SNPs abundance in the genome, and the relative ease in determining their frequency in a cost-effective and parallel manner in a given panel of individuals, the marker has emerged as one of the most widely-used genotyping markers (Deschamps et al. 2012). The GBS methodologies hold the greatest promise to serve the widest base of researchers because of the ability to allow simultaneous discovery of molecular markers including SNPs, and low-cost genotyping with a simple molecular biology workflow (Davey et al. 2011). Presently, we often see two terminologies, GBS and restriction-site-associated DNA sequencing (RAD-Seq), in publications. Both methods are a simple highly multiplexed system for constructing reduced representation libraries for the NGS platform, and involve digesting the genome of each individual in a population to be studied with a restriction enzyme (Elshire et al. 2011), except that the specific procedures for preparing the sequencing fragments are somewhat different. Thus, genotyping based on RAD-seq is included along GBS technology in this review.

The development of GBS, which is a sequencing-based high-throughput genotyping method, combines the advantages of time and cost effectiveness, dense marker coverage, and high-resolution mapping (Huang et al. 2009; Shokralla et al. 2012; Buermans and den Dunnen 2014). GBS, which was originally developed for crops (such as rice, maize and barley) (Huang et al. 2009; Elshire et al. 2011), is a simple, reproducible, highly multiplexed approach based on the Illumina or semiconductor sequencing platform (Elshire et al. 2011; Mascher et al. 2013; Wang 2015). The procedure of constructing GBS libraries based on reducing genome complexity with restriction enzymes is presented in Fig. 2. The major advantages of GBS are both technical simplicity (Davey et al. 2011) and the availability of informatics pipelines (Elshire et al. 2011; De Donato et al. 2013). GBS has the potential to

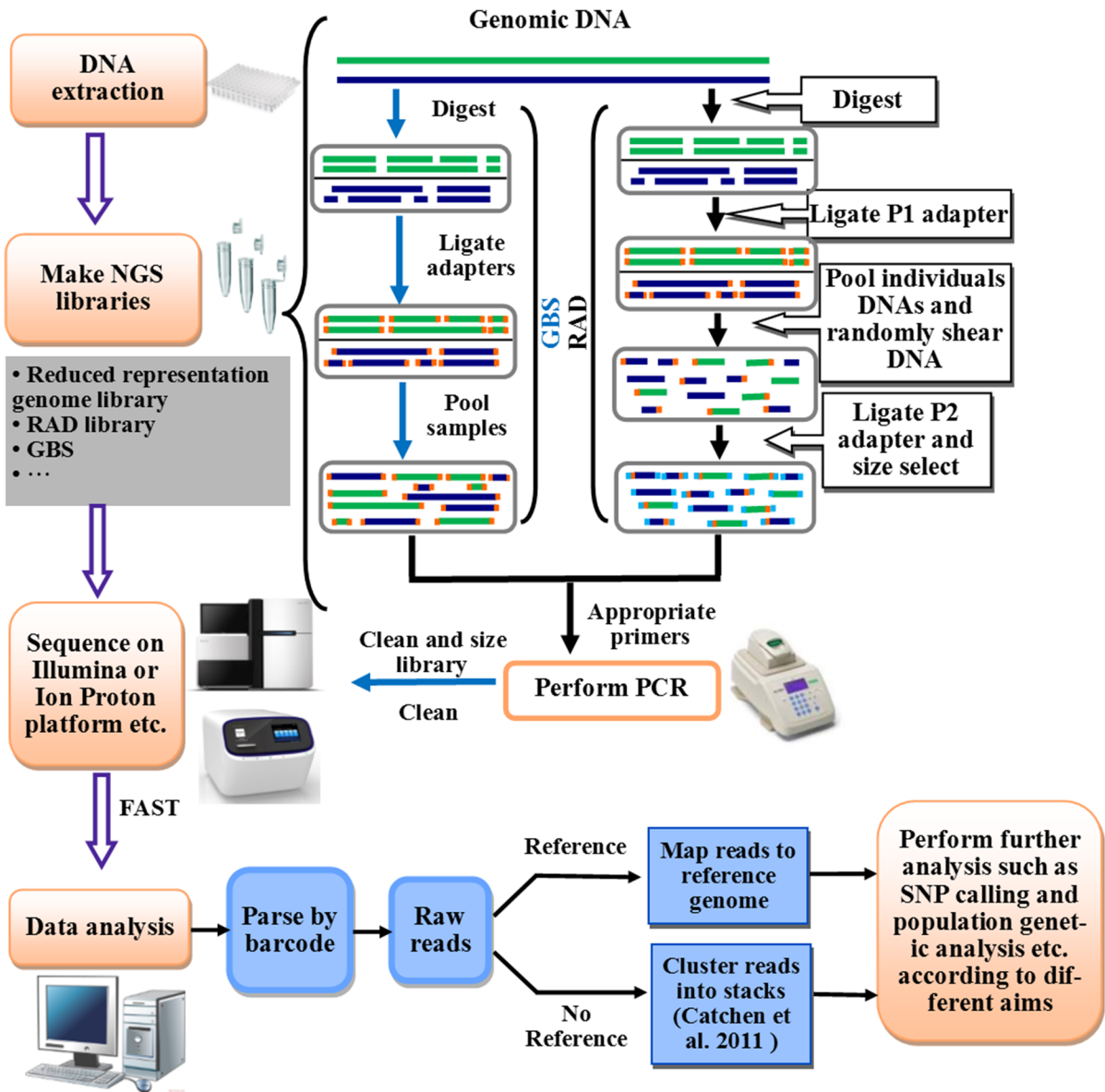


Fig. 2 A workflow of the genotyping-by-sequencing technology. Two reduced representation libraries (GBS and RAD library) are constructed to prepare the sequencing fragments for the NGS platform; the data for performing further analysis such

as SNP calling, population genetic analysis etc. after parsing barcode and mapping reads to reference genome (when there is reference genome) and clustering reads into stacks (when there is no reference genome)

be a truly revolutionary technology in the areas of fisheries, aquaculture, and related genomics. GBS allows a targeted fraction of the genome (a reduced representation library) to be sequenced with NGS technology rather than the entire genome, even in species with little or no previous genomic information and large genomes (Narum et al. 2013a). GBS is an efficient and cost-effective method for the

simultaneous discovery and genotyping of large numbers of molecular markers (SNPs, SSR, etc.) (Poland and Rife 2012; Poland et al. 2012; Dodds et al. 2015). The molecular markers discovered by GBS can be used to address questions related to population genomics, evolution (Lescak et al. 2015), selective breeding, genome-wide association studies (GWAS) (Dong et al. 2016), genetic characterization of

populations, disease prevention, etc. in a variety of diverse organisms in fisheries and aquaculture.

GBS has been applied in nearly 90 aquatic species/studies so far since it was first developed for rice in 2009 (Huang et al. 2009), and the number of GBS studies applied to fisheries and aquaculture research questions is increasing rapidly. Such trends indicate that GBS is becoming more widely used as a high-throughput genotyping system, compared with other whole-genome sequencing technologies. However, even though WGS remains prohibitively expensive for most research applications, because WGS can obtain more sequencing reads and provide more information about the genome, WGS is expected to push GBS forward to a whole genome level and provide more promises for applications in various research fields. This review summarizes genotyping, the application of NGS technologies toward GBS, the applications and advances in GBS in fisheries and aquaculture, and some advances in WGS in fisheries and aquaculture.

The application of NGS to high throughput SNP genotyping, and the advent of GBS

The types of markers and related methods have evolved rapidly with advances in molecular biology techniques since the first use of DNA-based markers laid the groundwork for gene discovery through forward and reverse genetics (Huang et al. 2009). Since NGS technologies appeared on the market ten years ago, tremendous progress has been made in terms of speed, read length, and throughput, along with a sharp reduction in the per-base cost, which democratized NGS and paved the way for the development of a large number of novel NGS applications. Genotyping based on new high-throughput sequencing holds promise to reduce costs and labor for traditional genotyping methods (Miller et al. 2007), and also for a methodological leap forward in genotyping and genetic mapping (Kanamori et al. 2016). NGS technology not only increased the sequencing throughput by several orders of magnitude, but also allows the simultaneous sequencing of a large number of samples using a multiplexed strategy (Craig et al. 2008; Cronn et al. 2008), and the ability to simultaneously genotype hundreds or thousands of SNPs or microsatellites in a single assay (Steemers and Gunderson 2007; Boussaha et al. 2012; Norrell et al. 2014).

In order to process this massive amount of data, bioinformatics research is constantly developing new software and algorithms, data storage approaches, and even new computer architectures to better meet the computation requirements (Dominik et al. 2010; Goddard et al. 2010).

Increasing the sequencing throughput can achieve the following advantages for genotyping: more individuals sequenced in a population, more markers sequenced per individual, and each marker genotyped at a greater depth and with greater accuracy. In addition, Illumina and semiconductor sequencing platforms have developed a genotyping system that combines high throughput and accuracy with low cost per SNP analyzed, since the involved instruments, kits and reagents output by Illumina or Life Tech become cheaper and cheaper (Oliphant et al. 2002; Mascher et al. 2013; Pukk et al. 2015; Wang 2015). To efficiently use the high sequencing capacity of NGS platforms, SNP assays have been designed for a high level of multiplexing. The PCR amplification step, which sometimes is regarded as a key technical obstacle, is required to gain sufficient sensitivity for genotyping SNPs in large, diploid genomes after the complexity of the genome is reduced by restriction enzyme digestion (Syvanen 2005). Presently, various approaches are available for obtaining sets or subsets of genomic restriction fragments for NGS which include restriction associated–associated DNA (RAD) (Baird et al. 2008; Etter et al. 2011b; Willing et al. 2011), diversity arrays technology (DArT) (Sansaloni et al. 2011), and complexity reduction of polymorphic sequences (CRoPS) (van Orsouw et al. 2007) are available for obtaining sets or subsets of genomic restriction fragments for NGS. With continuous advances and cost reduction in sequencing technologies, this genotyping based on genome-level sequencing may replace the conventional marker-based genotyping approach to provide a powerful tool for large-scale gene discovery and for addressing a wide range of biological questions (Huang et al. 2009).

GBS applications in fisheries

A better understanding of the natural genetic variation at the population level and the selective forces it is subjected to is needed for effective fisheries management (Bohling, et al. 2016; Messer et al. 2016). The increasing number of markers afforded by GBS

extends its application to many research questions in fisheries management. Applications of GBS range from genetic tagging to identify fish of unknown origin, refine estimates of the “real-time” migration rate and dispersal, and estimate the effective population size (Zhang et al. 2016). In this review, we focus on both advances and applications of GBS in fisheries management, conservation, and invasive species by bringing together the most important contributors to the field of fisheries (Fig. 3; Table S1).

Fisheries management

Genetic stock identification

Genetic stock identification (GSI) is an important tool for fisheries management to differentiate stocks of interest (Larson et al. 2014a). The genetic marker of choice for GSI has evolved from allozymes to microsatellites, and most recently to SNPs (Schlötterer 2004; Hauser and Seeb 2008). Recent genomic sequencing techniques have made it possible to screen thousands of putative SNPs in hundreds of individuals at one time (Allendorf et al. 2010; Narum et al. 2013a). The GBS approach has already been used to assess the

hybridization between two species of trout (Hohenlohe et al. 2011; Amish et al. 2012) and promises to be extremely applicable to the development of SNP panels for GSI (Storer et al. 2012). For example, approximately 1850 SNPs of western Alaskan Chinook salmon (*Oncorhynchus tshawytscha*) have been genotyped, of which 96 displaying high levels of among population differentiation have been converted into high throughput genotyping assays. This set of loci, called “RAD96”, has proven more useful for GSI of western Alaska Chinook salmon than those in CTC191 (191 SNPs), which was developed previously (Larson et al. 2014a). In addition, GBS technology can improve GSI in populations characterized by low genetic divergence and should be applicable to other populations and species. A comprehensive, open-access baseline of 45 SNPs was assembled from 172 Chinook salmon populations ranging from Russia to California for GSI (Templin et al. 2011). GBS of haploids was used to generate a high-density linkage map with 7035 loci and screened an existing panel of 8036 loci for signatures of selection (Limborg et al. 2014). The linkage map enabled the identification of novel genomic regions to display signatures of parallel selection shared between lineages of pink salmon (*Oncorhynchus gorbuscha*). The findings have

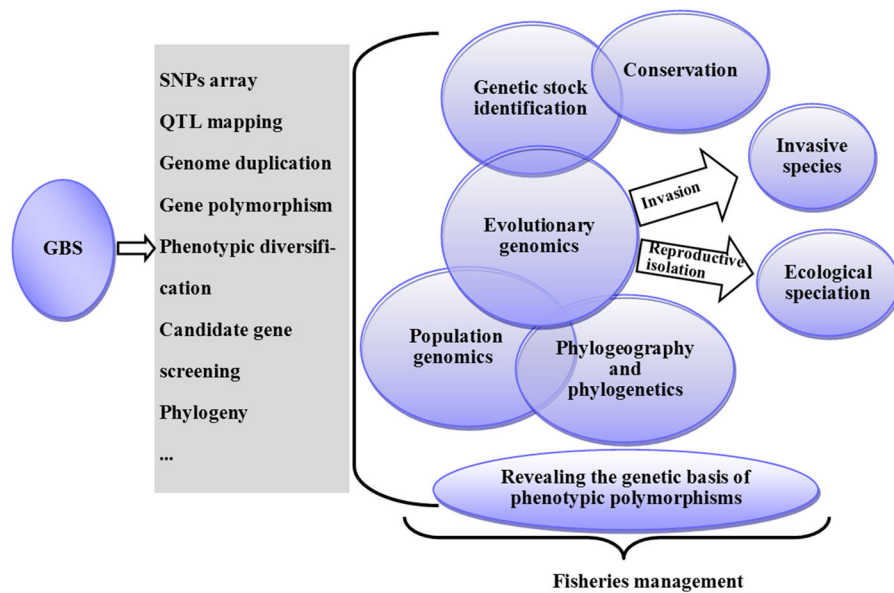


Fig. 3 Aspects and applications of genotyping-by-sequencing in fisheries. Genotyping-by-sequencing (GBS) can be used in applications for SNPs array, QTL mapping, genome duplication, gene polymorphism, phenotypic diversification, candidate

gene screening, phylogeny, etc. involved in fisheries research areas such as genetic stock identification, evolutionary genomics, and revealing the genetic basis of phenotypic polymorphisms

implications for understanding asynchronous population abundances, as well as predicting future ecosystem impacts from lineage-specific responses to climate change.

Revealing the genetic basis of phenotypic polymorphisms

Generally, it is difficult to determine the genetic basis of polymorphisms, although phenotypic polymorphisms can be applied to the study of the micro-evolutionary forces that maintain genetic variation and mediate speciation. However, RAD sequencing has recently attracted attention for revealing the genetic basis of phenotypic polymorphisms (Allgood et al. 2016; Carlson et al. 2015). RAD sequencing, combined with other techniques such as allele-specific PCR and Sanger sequencing, genotyping, quantitative trait loci (QTLs) mapping etc., could help to explore the genetic basis of phenotypic polymorphisms (Takahashi et al. 2013; Hecht et al. 2013; Guo et al. 2015). For examples, the genetic basis underlying male color dimorphism in the Lake Tanganyika cichlid fish *Cyprichromis leptosome* was determined with RAD sequencing, combined with allele-specific PCR and Sanger sequencing (Takahashi et al. 2013). A complex multi-genetic basis was suggested with several loci of small effect distributed throughout the genome that contributes to the migration of rainbow trout, based on that thousands of SNPs from RAD sequencing in two wild populations of migratory steelhead and resident rainbow trout (*Onchorhynchus mykiss*) were genotyped from the Pacific Northwest of the United States (Hecht et al. 2013). Significant QTLs were identified for body shape by geometric morphometric analyses of a cross between two ecologically and morphologically divergent, sympatric cichlid species (*Amphilophus zalius* and *Amphilophus astorquii*) (Franchini et al. 2014) and gnathopogon fishes (Kakioka et al. 2015) using double-digest restriction-site-associated DNA sequencing (ddRAD-seq). A system to study sleep and vibration attraction behavior in adult *Astyanax mexicanus* (blind Mexican cavefish) was established and the functional and evolutionary relationship between the QTL traits was investigated using high coverage QTL mapping based on RAD sequencing (Yoshizawa et al. 2015). QTL mapping can help identify candidate genes associated with

adaptation (Guo et al. 2015), and reveal the classic adaptive trait has a surprisingly modular and nonparallel genetic basis (Glazer et al. 2015).

Population and evolutionary genomics

Broadly, population genomics is used to understand the roles of evolutionary processes that influence variation across genomes and populations based on the simultaneous study of numerous loci or genome regions (Luikart et al. 2003). Understanding the genetics of how species adapt to changing environments across genomes and populations is a fundamental topic in evolutionary genomics (Du et al. 2015). Studies of genome-wide effects provide more reliable information regarding population demography and phylogenetic history (Longo and Bernardi 2015), while locus-specific effects help identify genes that are important for fitness and adaptation. Treating population genetic parameters as characters to be mapped onto phylogenies could be used to understand the evolutionary drivers of diversity within and differentiation between populations (Cutter 2013). The GBS technology provides novel opportunities for gathering genome-scale sequence data in natural or cultured populations, laying the empirical foundation for the evolving field of population genomics.

Population genomic analyses using high-confidence molecular markers developed from high-throughput sequencing can result in a highly resolved phylogeography which has not been achieved in previous studies using traditional markers (Reitzel et al. 2013). Using multiple loci to infer population and species histories has been used broadly in phylogeography and phylogenetics. Over the last decade, most practitioners of phylogeography and phylogenetics have developed and screened molecular markers suitable to their study system and appropriate to their evolutionary timescale of interest based on conventional laborious and often expensive methods (McCormack et al. 2013). The GBS technology may be a potential means to condense the many steps of multilocus data generation for model or even non-model organisms into a more time-efficient and cost-effective process. Genotyping many genome-wide markers in multiple divergent populations (generally ≥ 3 populations of multigenerational crosses or divergent natural populations) can allow the identification of markers that have extreme levels of differentiation

(outlier loci) and are likely to be within or close to genes involved in adaptation (Stapley et al. 2010). SNPs analysis generated by GBS technology can help develop a resource for analysis of genetic divergence, and has great potential to develop genomic resources for any aquatic organism in order to investigate the genetics of adaptation. RAD tags sequenced through NGS could be used to identify and genotype thousands of SNPs in individuals from multiple populations (Stapley et al. 2010; Araneda et al. 2016; Dayan et al. 2015).

Recent studies on population and evolutionary genomics in fisheries using GBS mostly focused on genetic diversity (Pujolar et al. 2013), phylogenetic relationships (Jones et al. 2013; Wagner et al. 2013), phylogeography (Hess et al. 2013), population structure (Xu et al. 2014a; Skovrind et al. 2016), population differentiation (Hohenlohe et al. 2010; Lah et al. 2014), and adaptation (Narum et al. 2013b; Seeb et al. 2014; Wagner et al. 2013; Ferchaud et al. 2014; Erickson et al. 2016). For example, using the ddRAD-seq method, Lah et al. (2014) estimated the population structure and differentiation and identified splits between porpoises from the North Sea and the Baltic, and within regions in the Baltic Sea (between the Belt Sea and the Inner Baltic Sea). Likewise, the applicability of RAD-tag genotyping for genome-wide scans of polymorphisms was evaluated across two cetacean species belonging to distinct families: the short-beaked common dolphin (*Delphinus delphis*) and the harbor porpoise (*Phocoena phocoena*) (Viricel et al. 2014).

Based on SNPs located within protein coding genes identified using high-throughput genotyping, GBS will be useful for population screening to search for adaptive variation (Bruneaux et al. 2013; Vera et al. 2013). For example, genome-wide tests of more than 10,000 SNPs in red band trout (*Oncorhynchus mykiss gairdneri*) found multiple SNPs that are significantly associated with survival under thermal stress, including Hsp47, which consistently appeared as a strong candidate gene for adaptation to desert climates (Narum et al. 2013b). Using RAD-tag sequencing, phylogenetic relationships were inferred among 16 species of cichlid fishes from a single rocky island community within Lake Victoria's cichlid adaptive radiation (Wagner et al. 2013). Generally, such studies have often been hampered by shallow genetic divergences, particularly in young adaptive radiations and

the lack of genome-scale molecular data. However, Wagner et al. (2013) demonstrated the power that NGS data sets hold to resolve even the most difficult of phylogenetic challenges and provided the first conclusive evidence for the monophyly of species in the Lake Victoria cichlid radiation.

The SNPs represent genomic regions that may be particularly informative for understanding adaptive evolution and exploring how differing genetic backgrounds within a species respond to selection from the same natural environment (Seeb et al. 2014) and in human-altered environments (Laporte et al. 2016). For example, the genomes of guppies (*Poecilia reticulata*) from natural and experimental populations were scanned using RAD-seq, and the results revealed several notable patterns of selection (Fraser et al. 2015).

Ecological speciation

Ecological speciation may occur when an adaptive radiation, which is recognized by a rapid burst of phenotypic, ecological, and species diversification results in reproductive isolation (Martin and Feinstein 2014). Reproductive isolation may be a function of the evolution of ecological specialization, and divergent selection of traits underlying adaptation to different environments (Futuyma and Moreno 1988; Hawthorne and Via 2001). As a consequence, the study of ecological speciation has become an extremely productive research area in modern evolutionary biology. Recent advances in sequencing technologies offer an unprecedented opportunity to study the mechanisms involved in ecological speciation (Faria et al. 2014). Phylogenetic approaches based on the high-throughput sequencing may be used to study the effect of specialization on lineage diversification (Vamosi et al. 2014).

The screening of thousands of loci throughout the genome (genome scans) by means of GBS technologies or even whole-genome sequencing, enables the identification of loci with exceptionally high levels of differentiation between populations or ecotypes. However, genome scans by means of GBS technologies still need to be combined with alternative lines of evidence, particularly regarding the effect of individual genes on adaptive phenotypes, in order to study ecological speciation. To address the questions of how consistent reproductive isolation is across species'

ranges, and to what extent reproductive isolation might remain polymorphic as species diverge, outcomes of hybridization could be compared using GBS. While GBS allows patterns of hybridization to be elucidated, a better understanding of the determinants of variation in reproductive isolation is needed, and the studies of reproductive isolation in hybrids should consider how the dynamics and mechanisms of reproductive isolation vary over ecological space and over evolutionary time (Mandeville et al. 2015). For example, patterns of genetic differentiation were compared between nascent species within an adaptive radiation of *Cyprinodon* pupfishes using GBS (Martin and Feinstein 2014), and the results showed that the scale-eater was further along the speciation-with-gene-flow continuum than the durophage, suggesting that different adaptive landscapes underlying these two niche environments drive variable progress toward speciation within the same habitat. This study provides an illustrative example of the power afforded by the analysis of large numbers of SNPs for estimating key parameters underlying evolutionary divergence.

Instances of parallel ecotypic divergence in which adaptation to similar conditions repeatedly cause similar phenotypic changes in closely related organisms are useful for studying the role of ecological selection in speciation (Roda et al. 2013). Very different evolutionary histories for different regions of the genome between *Mbipia* spp. and *Pundamilia* spp. were found using RAD tags suggesting at least two intergeneric hybridization events between these cichlid taxa led to the evolution of novel trait combinations and new species in Lake Victoria cichlids (Keller et al. 2013). The broad phenotypic variation of cichlids coupled with recent divergence also makes them an ideal model system for understanding speciation, adaptation, and phenotypic diversification using genomic information (Henning and Meyer 2014). For example, a color trait with a known adaptive function was mapped in cichlid fishes using the ddRAD-seq method, which are well-known textbook examples for rapid rates of speciation and astonishing phenotypic diversification (Henning et al. 2014).

Twelve eulachon (*Thaleichthys pacificus*, Osmeridae) populations ranging from Cook Inlet, Alaska and along the west coast of North America to the Columbia River were examined by RAD-seq to elucidate patterns of neutral and adaptive variation in this high

gene flow species, and this study highlighted the importance of integrating genome-wide neutral and adaptive genetic variation for applications of conservation and management (Candy et al. 2015). Similarly, to investigate patterns of neutral and adaptive genetic variation within the Chilean blue mussel (*Mytilus chilensis*), panels of adaptive SNPs obtained using GBS were used to identify genetic structure and geographic origin of Chilean Blue mussel populations in southern Chile, and will allow additional investigations into patterns of neutral and adaptive genetic variation related to population structure (Araneda 2015). However, sometimes functional reproductive isolation is maintained in sympatry even though genome-wide divergence is slight. For example, three sympatric species (the barred hamlet *Hypoplectrus puella*, the black hamlet *Hypoplectrus nigricans*, and the butter hamlet *Hypoplectrus unicolor*) in Belize, Honduras, and Panama are confirmed to be phenotypically distinct and reproductively isolated from transect surveys and spawning observations. However, a survey of approximately 100,000 restriction site-associated SNPs in 126 samples from the three species of coral reef fish across three replicate populations revealed extremely slight genome-wide divergence among the species ($F_{ST} = 0.0038$), which indicated that ecomorphological differences and functional reproductive isolation are maintained in sympatry in a backdrop of extraordinary genomic similarity (Puebla et al. 2014).

Conservation

Nowadays, with the current and anticipated impacts of anthropogenically induced climate change, understanding how organisms interact with their environments has become a key focus of conservation efforts with many species. For understanding this relationship, one method presently used is genetic maps and QTL mapping to detect genomic regions linked to phenotypic traits of importance for adaptation. Thus, research on population management and conservation urgently requires rapid and inexpensive methods for genome-wide SNP discovery and genotyping (Everett and Seeb 2014).

Presently, GBS coupled with the use of haploids would provide a robust pathway to rapidly develop genomic resources in non-model organisms. Using GBS to both detect and map thousands of SNPs in

haploid Chinook salmon (*Oncorhynchus tshawytscha*) and applying this map to detect QTL-related to temperature tolerance and body size in families of diploid Chinook salmon resulted in the successful detection of three QTLs for temperature tolerance, one QTL for body size at the experiment-wide level, and one QTL at the chromosome-wide level (Everett and Seeb 2014). These QTLs represent preliminary progress toward linking traits of conservation interest to regions in the Chinook salmon genome. Similarly, to explore the population structure, demography, and adaptive divergence in five populations of Chinook salmon from western Alaska, approximately 10,944 single-nucleotide polymorphisms were detected through RAD sequencing, and 733 loci and three genomic regions under putative selection were identified by outlier tests based on genetic differentiation (Larson et al. 2014b). These markers and genomic regions are considered excellent candidates for future research and can be used to create high-resolution panels for genetic monitoring and population assignment.

The high-throughput SNP assays based on GBS data provide an important resource to address critical uncertainties associated with the conservation and recovery of the imperiled species. An optimal combination of SNP assays was specially developed for conservation applications in Pacific lamprey (*Entosphenus tridentatus*) (Hess et al. 2015) and Delta smelt (*Hypomesus transpacificus*) (Lew et al. 2015). Using RAD-seq, 6145 SNPs were identified and genotyped in kokanee (*Oncorhynchus nerka*) from multiple spawning sites in Okanagan Lake (British Columbia, Canada), and the study provided preliminary evidence that RAD-seq may be an effective tool for conservation and management (Lemay and Rusello 2015).

To address that conservation management, traceability and enforcement suffer a lack of appropriate DNA markers for the genetic identification of sturgeon at the species, population and individual level, RAD sequencing was also employed to discover and characterize single nucleotide polymorphism (SNP) DNA markers for use in sturgeon conservation in the four tetraploid species (*Acipenser baerii*, *Acipenser naccarii*, *Acipenser persicus* and *Acipenser gueldenstaedtii*) (Ogden et al. 2013).

Additionally, hybridization and invasion are also important in conservation biology and are among the

most serious threats to the persistence of many native species (Sax et al. 2007). Interspecific hybridization can break up co-adapted gene complexes, disrupt local adaptation, and lead to genomic introgression or even genomic extinction (Allendorf et al. 2001; Bradbury et al. 2015). GBS also could be used to assess genomic patterns of introgression, detect candidate super invasive alleles driven to high frequency by selection, and estimate individual- and population-level admixture. Candidate adaptive super-invasive alleles, which show elevated rates of introgression above the genome-wide background, were identified from rainbow trout (*Oncorhynchus mykiss*) with excessively high admixture proportions in hybridized populations through overlapping paired-end RAD sequencing (Hohenlohe et al. 2013). RAD loci from rainbow trout mapped to a draft genome assembly could provide the marker density required to identify genes and chromosomal regions influencing selection in admixed populations of conservation concern and evolutionary interest (Hand et al. 2015).

Serious economic losses and environmental damages including biodiversity loss and ecosystem degradation may result from invasive species, which are an important element of global change and may result in impairment of ecosystem services worldwide (Pyšek and Richardson 2010). A species introduced into a new region must either possess sufficiently high levels of physiological tolerance and plasticity or must undergo genetic differentiation to achieve the required levels of fitness. There is growing realization that the integration of genetics and ecology is critical in the context of biological invasions because the two are explicitly linked. An ecological genetics approach for studying biological invasions refers to the ecology of particular genotypes and the role they play in adaptation to new environments, and ultimately, invasion success (Lawson Handley et al. 2011). It is crucial to integrate genetics with ecology in the context of biological invasions since the ecological conditions in a new environment may be distinctly different from the original range, which would present major adaptive challenges for an invasive population (Reznick and Ghalambor 2001; Ciosi et al. 2008). Understanding biological invasions, which have become regarded as “natural experiments” that offer unique insights into ecological and evolutionary processes occurring in real time (Lee 2002; Sax et al. 2007), is increasingly

considered crucial for implementing successful management policies.

GBS applications in aquaculture

Incorporating GBS into aquaculture related research is necessary for advancing progress on a variety of fronts including aquaculture genomic resources, genomic evaluation for breeding programs, nutrigenomics and so on. Here, we summarize the advances and application of GBS in aquaculture and related fields (Fig. 4; Table S2).

SNPs discovery and their use in unifying aquaculture genomic resources

SNPs are rapidly becoming the markers of choice in natural and aquacultural population genetics due to a variety of advantages relative to other factors, including higher genomic density, data quality, reproducibility and genotyping efficiency, as well as ease of portability between laboratories (Schunter et al. 2014). Ideally, a set of SNP markers randomly distributed throughout the genome would be developed for each

population studied. GBS moves us closer to this goal by incorporating the simultaneous discovery of SNPs and the genotyping of individuals. For example, utilizing GBS, SNP markers were characterized in 190 individuals from five wild and domesticated blue catfish (*Ictalurus furcatus*) populations, which showed that GBS could provide a rapid, low-cost approach to generate and genotype informative markers (Li et al. 2014).

GBS now enables the rapid discovery and genotyping of genome-wide SNP markers (especially suitable for the development of dense SNP linkage maps) that could serve as an important platform for improving genome assembly and unifying multiple genomic resources for aquaculture projects. For examples, using a RAD-sequencing approach, 99% of the SNPs identified for the rainbow trout (*Oncorhynchus mykiss*) are novel, allowing annotation to the positions of the SNPs in the working draft of the rainbow trout reference genome and some SNPs are likely to be in coding sequences (Palti et al. 2014). The 3569 genome-wide SNP markers were detected from the widely cultured GIFT (Genetically Improved Farmed Tilapia) strain of Nile tilapia (*Oreochromis*

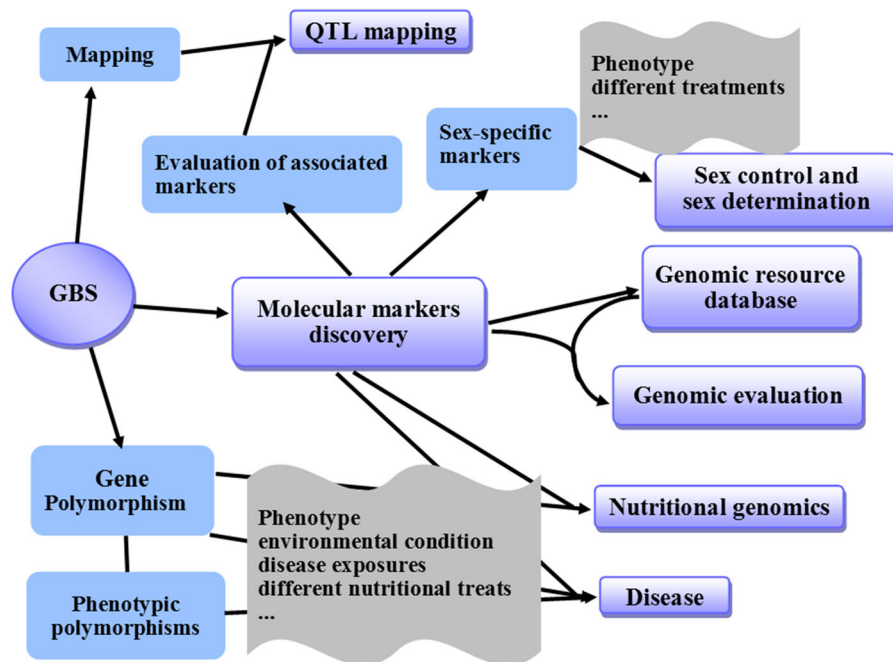


Fig. 4 Aspects of genotyping-by-sequencing application in aquaculture. Genotyping-by-sequencing (GBS) can be used in applications for molecular markers discovery, QTL mapping,

phenotypic polymorphism, gene polymorphism, etc. involved in aquaculture research areas such as sex control and sex determination, genomic evaluation, and nutritional genomics

niloticus), which allows the tracing of escapees from aquaculture and the monitoring of effects of introgression and hybridization (Van Bers et al. 2012). The novel SNPs from RAD-sequencing identified from Baltic Sea herring (*Clupea harengus*) (Corander et al. 2013) could be used to detect and estimate the degree of genetic differentiation among populations (Guo et al. 2016). The development and characterization of a high-density SNPs linkage map based on *SbfI* RAD-Seq SNP markers from two Atlantic salmon (*Salmo salar*) reference families were recently documented as a resource for salmonid genomics research and for estimation of phylogenetic relationships across closely and distantly related species (Gonen et al. 2014, 2015).

Genomic evaluation for breeding programs

Genomic evaluation is a blend of genomic predictions and traditional evaluations that include the mean genetic merit of parents (Wiggans et al. 2010). The accurate selection of breeding aquatic organisms can be enhanced through the use of genomic evaluation information (Reinhardt et al. 2009; Clark et al. 2011). Genomic evaluation systems and genome-wide studies require large numbers of individuals who are both densely genotyped for SNPs across the genome and intensively phenotyped to yield reliable results (Clark et al. 2011; Strandén and Christensen 2011). GBS is quickly expanding to fill these requirements. The developed imputation algorithm and software and the resulting single-stage genomic evaluation method based on the SNP markers obtained using GBS provide powerful new ways to exploit imputation and to obtain more accurate genetic evaluations (Hickey et al. 2012). The detection of conflicts in the stored genotypes, online access for problem resolution, and use of a single set of SNPs for genomic evaluations have simplified the tracking of genotypes and genomic evaluation as a routine and official process (Wiggans et al. 2009, 2010).

At present, some studies focus on the genomic evaluation and selection using GBS. Tens of thousands of high quality polymorphic SNPs in the Atlantic salmon (*Salmo salar*) genome were discovered successfully by RR-Seq, RAD-Seq and RNA-Seq methods (Houston et al. 2014). The Affymetrix Axiom SNP array based on this study (Houston et al. 2014) will facilitate high-resolution studies to determine the genetic architecture of traits of economic and

ecological importance, to study the structure of Atlantic salmon populations and to apply genomic selection in breeding programs. This study (Jones et al. 2014) identified putative QTL and marker associations, and confirmed genetic results that oyster (*Pinctada maxima*) growth traits (i.e., shell height, length, width and weight) are polygenic and provided additional evidence that these traits have sufficient genetic basis for improvements in selective breeding programs.

QTL mapping

QTLs are regions of the genome affecting variation in quantitative (phenotypic) traits. QTLs are identified from statistical associations between the genotypic markers and the phenotypic trait(s) of interest after genotypic markers in individuals which are generated from a known pedigree, and a linkage map that shows the order of the markers and the relative distance (in centimorgans; cM) between them are created (Mueller and Wolfenbarger 1999). The goal is to determine the genotype of each line in the mapping population relative to the parental genotypes. The type and number of markers applied depends on the equipment and resources available. The number and choice of markers should allow full coverage (e.g., no gaps >20 cm) of the genome and should include all suspected QTL regions. The number and design of the phenotyping experiments is crucial for successful QTL mapping (Buerstmayr et al. 2009). However, the application of QTL mapping remains difficult for many organisms because the construction of a detailed linkage map requires the identification of many polymorphic loci that are dispersed throughout the genome (Mueller and Wolfenbarger 1999) and is particularly difficult when the associations between QTLs and phenotypic traits are too weak and too unreliable to be useful in the selection (Goddard et al. 2010).

The genome-wide SNP genotypes that are used in genomic selection are also used in GWAS to map genes for complex traits (Hausmann et al. 2002; Goddard and Hayes 2009; Liu et al. 2015a). Linkage maps are a valuable tool for genome-wide applications, such as genome-wide association studies, QTL mapping and genome scans (Briec et al. 2014; Manousaki et al. 2016). The GBS holds great promise for the construction of high-density QTL-linked

markers linkage maps in organisms lacking extensive genomic resources (Jiao et al. 2014). Presently, using RAD-seq, QTL-linked markers linkage maps were constructed in (Houston et al. 2012), Lake Whitefish (*Coregonus clupeaformis*) (Gagnaire et al. 2013), wild resident rainbow trout and wild migratory steelhead trout (*Oncorhynchus mykiss*) (Hecht et al. 2012), *Pinctada fucata martensii* (Japanese pearl oyster) (Shi et al. 2014), and large yellow croaker (*Larimichthys crocea*) (Ao et al. 2015) and so on.

Sex determination and sex differentiation

The significant sexual dimorphism between some females and males makes a genetic test for the early identification of sex desirable (Rondeau et al. 2013). The development of molecular probes to identify sex or the establishment of molecular signatures characteristic of a given phenotype can have an immediate application in the aquaculture industry for broodstock selection, monosex production, and better performance (Piferrer et al. 2012). The recent application of genomic tools, including microarrays, NGS approaches, and epigenetics, can significantly contribute to the deciphering of the molecular mechanisms involved in the processes of sex determination and sex differentiation (Piferrer et al. 2012; Gamble 2016). GBS can help explore the region containing the main sex-determining locus and provides genetic markers that are tightly linked to this locus through an association that persists across the population (Wang et al. 2015).

Developing sex-associated markers and understanding the sex determination mechanism can shorten the time for the development of monosex female production, thus reducing the costs of grow-out. The sex-associated markers can be of use in refining the production of genetically monosex male or female stocks for aquaculture. For instance, through NGS techniques, sex-specific markers, gonadal-soma derived factor (*gsdf*) and the master sex determinant locus were identified in the medaka species *Oryzias luzonensis* (Rondeau et al. 2013). A major sex-determining region on linkage group was identified and explained nearly 96% of the phenotypic variance (Palaiokostas et al. 2013b). A combined sex linkage map was constructed based on SNP markers that resolved into 35 linkage groups (Brown et al. 2016). A major sex-determining locus to linkage group 13 was

mapped in Atlantic halibut (*Hippoglossus hippoglossus*) (Palaiokostas et al. 2013a), and two discrete levels of degeneration on the Y chromosome were identified (Roesti et al. 2013). Using sequence-based polymorphic restriction-site-associated markers, genetic loci associated with sex was identified by analyzing the F₂ offspring of reciprocal crosses between Oregon (AB) and Nadia (NA) wild-type zebrafish (*Danio rerio*) stocks (Anderson et al. 2012). A narrow genomic region strongly associated with sex determination in lake whitefish (*Coregonus clupeaformis*) was evidenced as well as phenotypic and expression QTL underlying ecologically important traits (Gagnaire et al. 2013). A publicly-available high-density SNP array created was used to explore the population structure of Atlantic salmon (*Salmo salar*) and predict the phenotypic sex, and will be used for genomic prediction and selection for resistance (Houston 2015).

Clarifying the interactions between genotype and nutritional status

Dietary interventions based on a knowledge of nutritional requirements, nutritional status, and genotype can be used to prevent, mitigate, or cure chronic disease (Sreeparvathy 2013). With the continuous development of genomics technology, nutriology scholars have begun to explore the interaction between genes and diets moving nutrition research into the genomics era. Thus, nutrition research has stepped into the genomic era. Studying the interaction between feeding styles and genetic makeups has established a new discipline within nutrition research called nutrigenomics (DellaPenna 1999; Walker and Blackburn 2004; Chen and Zhao 2013).

Nutrigenomics is mainly based on the use of molecular biology techniques to clarify the interactions between nutrients and genes, applying the technologies of DNA microarrays, NGS, proteomics, etc. Using in-depth and systemic research, scientists have established methods of dietary intervention and principles of nutrition and health care based on individual genomic structure features that could be applied to the development of individualized nutritional treatment strategies (Chen and Zhao 2013). The variety of diets available does mean that each of hundreds of types of fish may contain slightly different

quantities of key chemical ingredients. In some cases, different strains of fish may even produce different chemicals. Thus, to know what genes are expressed in the organisms which will be our food can have important health implications. The same is true for all of the fish foods we eat; for example, we may choose to eat fish with more abundant cancer-fighting proteins or healthier oils. Simply put, genomic analysis has a great potential role in our choices regarding culturing and modern aquaculture methods.

Aquaculture should be considered a full partner with human medicine and health. Instead of focusing on aquaculture as a yield per acre, the insertion or activation of genes within an aquatic organism could be pushed to the forefront of research. We believe that the application of genomics technologies to food technology will reduce research and development times, thus shrinking costs and shortening the time to market (Daniel 2007). Generally, reactions to common dietary ingredients are thought to be directly related to the presence of specific SNPs (Chen and Zhao 2013). SNPs are the most common form of DNA sequence variations and act as useful polymorphic markers for investigating genes. However, not all SNPs directly influence the quality and/or quantity of the gene product. As more information emerges regarding the links among SNPs, dietary components and phenotypes, it should become easier to predict those subjects who might benefit most from dietary interventions. The simultaneous examination of multiple SNPs may offer special advantages in defining the biological response to food components or drugs because multiple genes are likely to be involved in determining physiological processes and their ultimate influence on an individual's phenotype (Trujillo et al. 2006).

Gene polymorphism, which is also called genetic susceptibility due to changes in the genetic sequence and copy number, can also affect the absorption, metabolism and action sites (molecular target) of nutrients (van Der Werf et al. 2001; Milner 2004). Individuals carrying SNPs for specific sensitive genes might require certain nutrients and may possess greater dietary resistance to certain diseases (Liu-Stratton et al. 2004). GBS is a potentially powerful and useful tool for discovering plentiful SNPs for specific sensitive genes. Despite great progress being made in functional and mechanistic discoveries surrounding SNPs, true nutritional studies demonstrating the link

between diet, gene and phenotype are still quite rare. It is much more common to find links between genes and disease, and then later to make either a hypothetical or observational prediction as to the influence of nutrition.

If the precedent of genotyping in the dietary prevention of disease is to be established, the field of nutrition must become more proactive in demonstrating the mechanisms driving such connections between diets and phenotypes, according to specific genetic variations (Roberts et al. 2001). Research into these so-called nutrient gene interactions with genome-wide SNP arrays are complicated by the multiplicity of genes, dietary components and gene-nutrient interactions (Hunter 2005; Barnes 2008). Genome-wide SNP discoveries can help to identify the genes regulating these relevant traits, unique feeding behaviors and food preferences. To our knowledge, there is no document on identifying SNPs in nutrigenomics using GBS presently, although GBS provides promise in this field. However, there is a study on discovering approximately one-hundred times more potential SNPs which were found in feeders than in non-feeders using transcriptome sequencing and digital gene expression profiling of mandarin fish (He et al. 2013), which indicated the potential applications of transcriptome sequencing and other high-throughput sequencing technologies including GBS in fish nutrition.

Identifying mutations associated with disease

In disease genetics studies utilizing SNPs, the primary aim is to identify SNPs that cause changes in cellular biological processes of inducing diseased states (Kwok and Gu 1999; Schorka et al. 2000; Emahazion et al. 2001; Tost and Gut 2005). Most genetic variations, including SNPs, insertions or repeats, have been identified by sequencing genes that code for enzymes or transporters related to the disease of interest (Afman and Muller 2006). In both disease genetics and pharmacogenomics studies, a large number of SNPs must be genotyped in equally large sample populations (Judson et al. 2002; Gunderson et al. 2006). Hence, achieving a high level of throughput has been a critical factor in recently developed genotyping assays (Gut 2001; Kwok 2001; Syvänen 2001; Jenkins and Gibson 2002).

A typical approach for improving the throughput of genotyping is to increase the number of SNPs that are

analyzed simultaneously; in other words, the multiplexing level of the assay is expanded (Kim and Misra 2007). The relationship between a specific genotype and a phenotype can then be used to characterize the susceptibility genes associated with a disease and to identify the encoded proteins for the prevention and treatment of the disease (Kim and Misra 2007). Complex diseases and their susceptibility variants are generally investigated by population association studies on several candidate genes or by genome-wide linkage analyses of familial samples (Kim and Misra 2007). The fine mapping of genes associated with disease in large genomic regions previously defined by linkage analysis also requires the genotyping of hundreds or thousands of SNPs.

Efficient SNP genotyping technologies will be required for routine clinical application once disease-predisposing genes have been identified and once the allelic variants that predict disease or improve diagnostics have been specified. Moreover, many of these complex diseases and traits may be caused by rare alleles that can only be detected by re-sequencing complete genomic regions in multiple individuals, first for the identification of variants (Evans et al. 2004) and later for disease diagnostics. To this end, there are many calls for the genotyping of SNPs on a large scale (Syvänen 2001; Syvanen 2005). GBS can then be selectively filtered into the few required for a target experiment. Thus, GBS allows the identification of certain common alleles that contribute to complex disease susceptibilities, and can help identify mutations within genes affecting various diseases, such as bacterial cold water disease (BCWD) (Liu et al. 2015b) and *Vibrio anguillarum* disease (Shao et al. 2015). For example, the thickness of individual retinal layers among 115 F2 hybrid progeny from the eyed \times eyeless forms of *Astyanax mexicanus* (blind cavefish) were quantified, QTLs affecting retinal thickness were scanned based on RAD-seq data, and candidate genes within these QTL regions were identified. (O'Quin et al. 2013). This would serve as a starting point for future studies on the genetics of retinal degeneration and eye disease using the evolutionary mutant model *Astyanax*.

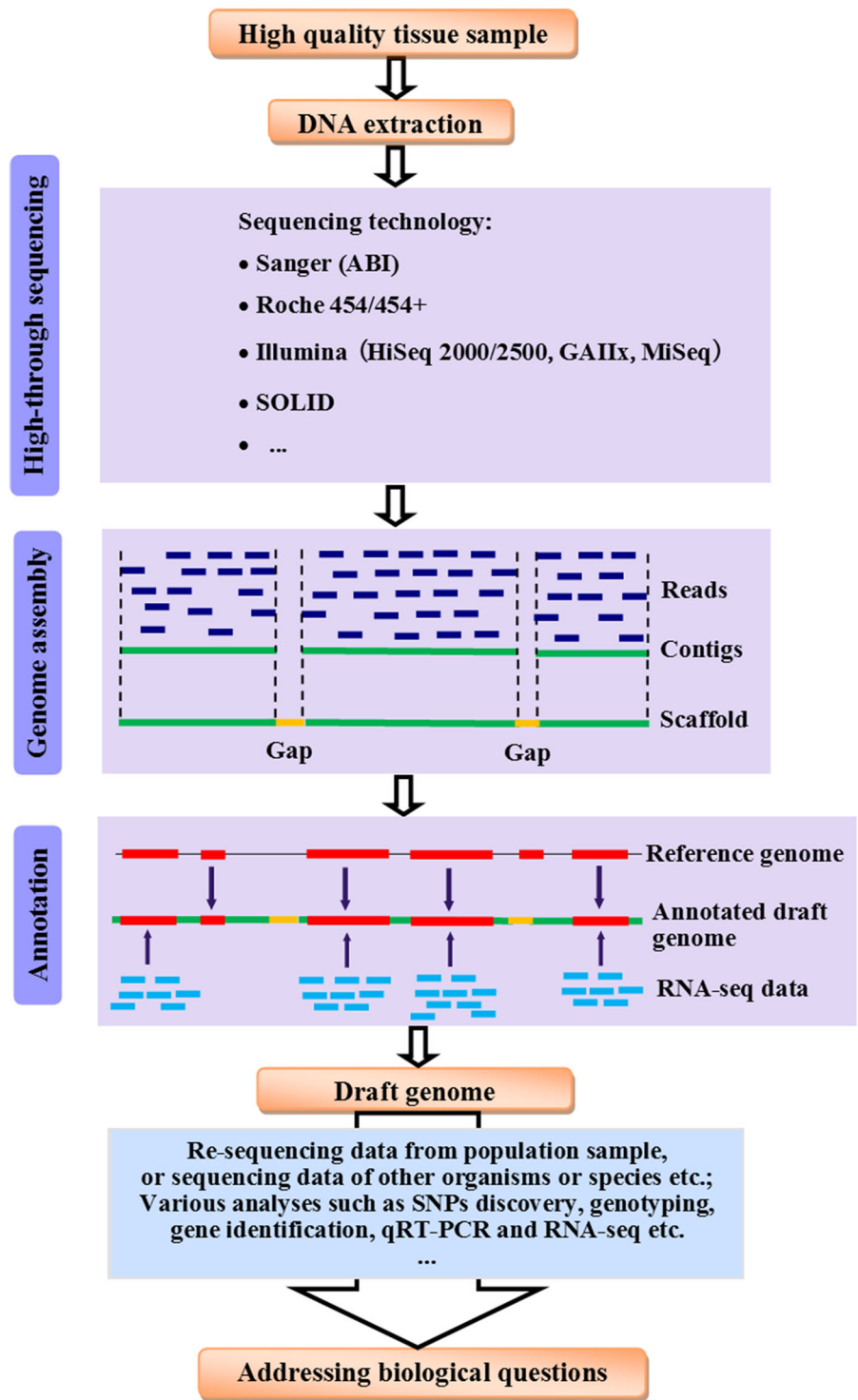
WGS in fisheries and aquaculture

Most applications of phylogenomics, population genomics, evolutionary genomics and nutritional

genomics, must rely on the analysis of hundreds of individuals; thus, methods that reduce the genome to a subset of loci should be more cost-effective for some time to come. Thus, the present GBS undeniably shows the promising prospects for application in such fields. However, WGS would be extensively used over a broad range of research fields if it is as cost-effective as the present GBS technology and the bioinformatic, computational, and data storage challenges inherent in WGS could be addressed. WGS could provide reference to whole-genome so that it could help characterize those unidentified tags from GBS, identify more molecular markers, and more accurately analyze the information from GBS. Nevertheless, despite the presence of inefficiencies in the targeting process, it has become feasible to sequence all targeted regions or sequences, such as exons, non-coding RNAs, fragments with adaptor primers or probes, and highly conserved regions of the genome. Such targeted sequencing approaches have the general advantage of increasing the sequence coverage of regions of interest at a lower cost and higher throughput. As the cost of WGS continues to decrease, this modality should become increasingly popular due to its ability to survey most of the genome alongside additional classes of mutations (Majewski et al. 2011). WGS is expected to trigger a major shift in a broad number of research areas and should push GBS technologies toward a true “whole genome level” with manifold promising applications in various research fields.

Presently, because of the cost and considerable effort involved in WGS, the most important first step is to thoroughly consider whether a complete genome sequence is necessary to address the biological question at hand; once this decision is made, a genome project requires careful planning with respect to the organism involved and the intended quality of the genome draft (Ekblom and Wolf 2014). A brief workflow of a typical whole-genome sequencing project is presented in Fig. 5, and there have been some applications of WGS in aquaculture and fisheries (Table S3). For example, since the genomic evolution and diversification of the teleosts are important subjects for understanding vertebrate evolution, the freshwater puffer fish *Tetraodon nigroviridis* genome, which has the smallest known vertebrate genome, was sequenced and compared with human genomes in 2004 (Jaillon et al. 2004). Additionally, the genome of the small, egg-laying, freshwater teleost, the medaka

Fig. 5 Brief workflow of a whole-genome sequencing project. Whole-genome sequencing (WGS) projects generally contains the following steps: DNA sample preparation, high-throughput sequencing, genome assembly, annotation, and further data analysis



(*Oryzias latipes*), was sequenced by Kasahara et al. (2007), who found a strict genetic separation of 4 million years between the two populations

investigated. The result suggested that differential selective pressures acted on specific gene categories, based on the dense SNP information. Most

applications of WGS in aquaculture and fisheries have focused on characterizing the genomic mechanisms, genetic and evolutionary basis of phenotypic diversity (Star et al. 2011; Jones et al. 2012; Nakamura et al. 2013; Wang et al. 2013; Brawand et al. 2014; Yim et al. 2014). WGS applications could offer a unique model to better understand the molecular biology of several traits, such as complex behavior, live births and malignant diseases, including cancer (Schartl et al. 2013), and suggest a hypothesis for the evolution of the behavioral complexity of fish (Howe et al. 2013; Berthelot et al. 2014), sex-determination mechanisms (Chen et al. 2014). The applications of WGS could also help pave the way for better molecular research and improvements in the genome-assisted breeding of fish in combination with a high-resolution genetic map (Xu et al. 2014b).

Conclusion

Substantial improvements in the base calling software and sequencing chemistries have allowed NGS technologies to deliver higher sequencing throughputs per run. In turn, this capacity has enabled deeper multiplexing for a fixed average sequencing depth per sample. At the same time, bioinformatics techniques and training are becoming more important skills as the technology continues to advance. The GBS approach has several advantages, including that preliminary sequence information is not required and that all newly discovered markers originate from the population being genotyped, making GBS an especially useful tool for examining non-model organisms common in fisheries and aquaculture. Indeed, we have highlighted a number of important GBS studies focused within this review demonstrating its applicability. However, challenges remain, including, but not limited to, the following: (1) the need for training in bioinformatics skills, (2) acquiring access to powerful computers to launch analyses and (3) integrating new concepts into fisheries management and aquaculture. Additionally, a new study showed that RADseq is prone to a number of sources of error that may have consequential effects for population genetic inferences that have received only limited attention in terms of the estimation and reporting of genotyping error rates (Mastretta-Yanes et al. 2015). WGS will propel GBS technologies toward a real “whole genome level” and provide more promising

applications to various research fields related to fisheries and aquaculture, although at present, WGS remains relatively expensive for most of these applications.

Acknowledgements This study was supported by the National Institute of Food and Agriculture (NIFA), U.S. Department of Agriculture, under Agreement No. 2010-38,879-20946. Salaries and research support were provided by state and federal funds appropriated to The Ohio State University, Ohio Agricultural Research and Development Center, and Huazhong Agricultural University. We thank Joy Bauman for her comments on the manuscript.

References

- Afman L, Muller M (2006) Nutrigenomics: from molecular nutrition to prevention of disease. *J Am Diet Assoc* 106:569–576
- Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids—setting conservation guidelines. *Trends Ecol Evol* 16:613–622
- Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nat Rev Genet* 11:697–709
- Allgood KS, Lescak EA, Bassham SL, Catchen JM, Von Hippel FA, Cresko WA (2016) Linking phenotypic and genomic evolution in very young populations of threespine stickleback. *Integr Comp Biol* 56:E4
- Alwala S, Suman A, Arro JA, Veremis JC, Kimbeng CA (2006) Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. *Crop Sci* 46:448
- Amish SJ, Hohenlohe PA, Painter S et al (2012) RAD sequencing yields a high success rate for westslope cutthroat and rainbow trout species-diagnostic SNP assays. *Mol Ecol Resour* 12:653–660
- Anderson JL, Rodriguez Mari A, Braasch I, Amores A, Hohenlohe P, Batzel P, Postlethwait JH (2012) Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. *PLoS ONE* 7:e40701. doi:10.1371/journal.pone.0040701
- Ao J, Li J, You X et al (2015) Construction of the high-density genetic linkage map and chromosome map of large yellow croaker (*Larimichthys crocea*). *Int J Mol Sci* 16:26237–26248
- Araneda C (2015) Contrasting patterns of neutral and adaptive genetic variation of Chilean blue mussel (*Mytilus chilensis*) due to local adaptation and aquaculture. In: Plant and animal genome XXIII conference. Plant and animal genome
- Araneda C, Angelica Larrain M, Hecht B, Narum S (2016) Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecol Evol* 6:3632–3644
- Baird NA, Etter PD, Atwood TS et al (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376. doi:10.1371/journal.pone.0003376

- Barnes S (2008) Nutritional genomics, polyphenols, diets, and their impact on dietetics. *J Am Diet Assoc* 108:1888–1895
- Berthelot C, Brunet F, Chalopin D et al (2014) The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat Commun* 5:3657
- Bohling J, Haffray P, Berrebi P (2016) Genetic diversity and population structure of domestic brown trout (*Salmo trutta*) in France. *Aquaculture* 462:1–9
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331
- Boussaha M, Guyomard R, Cabau C, Esquerr D, Quillet E (2012) Development and characterisation of an expressed sequence tags (EST)-derived single nucleotide polymorphisms (SNPs) resource in rainbow trout. *BMC Genom* 13(1):238
- Bradbury IR, Hamilton LC, Dempson B, Robertson MJ, Bourret V, Bernatchez L, Verspoor E (2015) Transatlantic secondary contact in Atlantic salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Mol Ecol* 24:5130–5144
- Brawand D, Wagner CE, Li YI et al (2014) The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513:375–381
- Brieuc MS, Waters CD, Seeb JE, Naish KA (2014) A dense linkage map for Chinook salmon (*Oncorhynchus tshawytscha*) reveals variable chromosomal divergence after an ancestral whole genome duplication event. *G3 (Bethesda)* 4:447–460
- Brown JK, Taggart JB, Bekaert M, Wehner S, Palaiokostas C, Setiawan AN, Symonds JE, Penman DJ (2016) Mapping the sex determination locus in the hapuku (*Polyprion oxygeneios*) using ddRAD sequencing. *BMC Genom* 17:448
- Bruneaux M, Johnston SE, Herczeg G, Merila J, Primmer CR, Vasemagi A (2013) Molecular evolutionary and population genomic analysis of the nine-spined stickleback using a modified restriction-site-associated DNA tag approach. *Mol Ecol* 22:565–582
- Buermans HP, den Dunnen JT (2014) Next generation sequencing technology: advances and applications. *Biochim Biophys Acta* 1842(10):1932–1941
- Buerstmayr H, Ban T, Anderson JA (2009) QTL mapping and marker-assisted selection for Fusarium head blight resistance. *Plant Breed* 128:1–26
- Candy JR, Campbell NR, Grinnell MH, Beacham TD, Larson WA, Narum SR (2015) Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. *Mol Ecol Resour*. doi:10.1111/1755-0998.12400
- Carlson BM, Onusko SW, Gross JB (2015) A high-density linkage map for *astyanax mexicanus* using genotyping-by-sequencing technology. *G3-Genes Genom Genet* 5:241–251
- Chen L, Zhao E (2013) Advances of nutrigenomics: concept, content, technology and development. *J Exp Biol Agric Sci* 1(3):152–158
- Chen S, Zhang G, Shao C et al (2014) Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet* 46:253–260
- Ciosi M, Miller NJ, Kim KS, Giordano R, Estoup A, Guillemaud T (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Mol Ecol* 17:3614–3627
- Clark SA, Hickey JM, van der Werf JH (2011) Different models of genetic variation and their effect on genomic evaluation. *Genet Sel Evol* 43:18
- Corander J, Majander KK, Cheng L, Merila J (2013) High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. *Mol Ecol* 22:2931–2940
- Cortés AJ, Chavarro MC, Blair MW (2011) SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 123:827–845
- Craig DW, Pearson JV, Szelinger S et al (2008) Identification of genetic variants using bar-coded multiplexed sequencing. *Nat Methods* 5:887–893
- Cronn R, Liston A, Parks M, Gernandt DS, Shen R, Mockler T (2008) Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Res* 36:e122
- Cutter AD (2013) Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Mol Phylogenet Evol* 69:1172–1185
- Daniel H (2007) Genomics and proteomics: importance for the future of nutrition research. *Br J Nutr* 87:S305
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499–510
- Dayan DI, Crawford DL, Oleksiak MF (2015) Population genomics of rapid adaptation in *Fundulus heteroclitus* exposed to power station thermal effluents. *Integr Comp Biol* 55:E42
- De Donato M, Peters SO, Mitchell SE, Hussain T, Imumorin IG (2013) Genotyping-by-sequencing (GBS): a novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PLoS ONE* 8:e62137. doi:10.1371/journal.pone.0062137
- DellaPenna D (1999) Nutritional genomics: manipulating plant micronutrients to improve human health. *Science* 285:375–379
- Deschamps S, Llaca V, May GD (2012) Genotyping-by-sequencing in plants. *Biology (Basel)* 1:460–483
- Dodds KG, McEwan JC, Brauning R, Anderson RM, van Stijn TC, Kristjansson T, Clarke SM (2015) Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genom* 16(1):1047
- Dominik S, Henshall JM, Kube PD et al (2010) Evaluation of an Atlantic salmon SNP chip as a genomic tool for the application in a Tasmanian Atlantic salmon (*Salmo salar*) breeding population. *Aquaculture* 308:S56–S61
- Dong L, Xiao S, Wang Q, Wang Z (2016) Comparative analysis of the GBLUP, emBayesB, and GWAS algorithms to predict genetic values in large yellow croaker (*Larimichthys crocea*). *BMC Genom* 17:460

- Doubleday ZA, Semmens JM, Smolenski AJ, Shaw PW (2009) Microsatellite DNA markers and morphometrics reveal a complex population structure in a merobenthic octopus species (*Octopus maorum*) in south-east Australia and New Zealand. *Mar Biol* 156:1183–1192
- Du X, Oleksiak MF, Crawford DL (2015) A genotyping by sequencing study of natural populations of *Fundulus heteroclitus* inhabiting a strong pollution cline. *Integr Comp Biol* 55:E49
- Eklblom R, Wolf JBW (2014) A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl*. doi:10.1111/eva.12178
- Elshire RJ, Glaubitz JC, Sun Q et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi:10.1371/journal.pone.0019379
- Emahazion T, Feuk L, Jobs M et al (2001) SNP association studies in Alzheimer's disease highlight problems for complex disease analysis. *Trends Genet* 17:407–413
- Erickson PA, Glazer AM, Killingbeck EE, Agoglia RM, Baek J, Carsanaro SM, Lee AM, Cleves PA, Schluter D, Miller CT (2016) Partially repeatable genetic basis of benthic adaptation in threespine sticklebacks. *Evol Int J Org Evol* 70:887–902
- Etter P, Bassham S, Hohenlohe P, Johnson E, Cresko W (2011a) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods Mol Biol* 772:157–178
- Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA (2011b) Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE* 6:e18561. doi:10.1371/journal.pone.0018561
- Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor Appl Genet* 104:399–407
- Evans DM, Cardon LR, Morris AP (2004) Genotype prediction using a dense map of SNPs. *Genet Epidemiol* 27:375–384
- Everett MV, Seeb JE (2014) Detection and mapping of QTL for temperature tolerance and body size in Chinook salmon (*Oncorhynchus tshawytscha*) using genotyping by sequencing. *Evol Appl* 7:480–492
- Faria R, Renaut S, Galindo J et al (2014) Advances in ecological speciation: an integrative approach. *Mol Ecol Notes* 23:513–521
- Ferchaud AL, Pedersen SH, Bekkevold D, Jian J, Niu Y, Hansen MM (2014) A low-density SNP array for analyzing differential selection in freshwater and marine populations of threespine stickleback (*Gasterosteus aculeatus*). *BMC Genom* 15:867. doi:10.1186/1471-2164-15-867
- Franchini P, Fruciano C, Spreitzer ML et al (2014) Genomic architecture of ecologically divergent body shape in a pair of sympatric crater lake cichlid fishes. *Mol Ecol* 23:1828–1845
- Fraser BA, Kunstner A, Reznick DN, Dreyer C, Weigel D (2015) Population genomics of natural and experimental populations of guppies (*Poecilia reticulata*). *Mol Ecol* 24:389–408
- Futuyama DJ, Moreno G (1988) The evolution of ecological specialization. *Annu Rev Ecol Syst* 19:207–233
- Gagnaire PA, Normandeau E, Pavey SA, Bernatchez L (2013) Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish (*Coregonus clupeaformis*). *Mol Ecol* 22:3036–3048
- Gamble T (2016) Using RAD-seq to recognize sex-specific markers and sex chromosome systems. *Mol Ecol* 25:2114–2116
- Glazer AM, Killingbeck EE, Mitros T, Rokhsar DS, Miller CT (2015) Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. *G3-Genes Genomes Genet* 5:1463–1472
- Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet* 10:381–391
- Goddard ME, Hayes BJ, Meuwissen TH (2010) Genomic selection in livestock populations. *Genet Res (Camb)* 92:413–421
- Gonen S, Lowe NR, Cezard T, Gharbi K, Bishop SC, Houston RD (2014) Linkage maps of the Atlantic salmon (*Salmo salar*) genome derived from RAD sequencing. *BMC Genom* 15:166. doi:10.1186/1471-2164-15-166
- Gonen S, Bishop SC, Houston RD (2015) Exploring the utility of cross-laboratory RAD-sequencing datasets for phylogenetic analysis. *BMC Res Notes* 8:299
- Gunderson KL, Kuhn KM, Steemers FJ, Ng P, Murray SS, Shen R (2006) Whole-genome genotyping of haplotype tag single nucleotide polymorphisms. *Pharmacogenomics* 7:641–648
- Guo B, DeFaveri J, Sotelo G, Nair A, Merila J (2015) Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biol* 13(1):1–18
- Guo B, Li Z, Merila J (2016) Population genomic evidence for adaptive differentiation in the Baltic Sea herring. *Mol Ecol* 25:2833–2852
- Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. *Funct Integr Genom* 4:139–162
- Gut IG (2001) Automation in genotyping of single nucleotide polymorphisms. *Hum Mutat* 17:475–492
- Hand BK, Tyler DH, Kovach RP et al (2015) Genomics and introgression: discovery and mapping of thousands of species-diagnostic SNPs using RAD sequencing. *Curr Zool* 61(1):146–154
- Handley LJJ, Estoup A, Evans DM et al (2011) Ecological genetics of invasive alien species. *Biocontrol* 56:409–428
- Hauser L, Seeb JE (2008) Advances in molecular technology and their impact on fisheries genetics. *Fish Fish* 9:473–486
- Hausmann BI, Mahalakshmi V, Reddy BV, Seetharama N, Hash CT, Geiger HH (2002) QTL mapping of stay-green in two sorghum recombinant inbred populations. *Theor Appl Genet* 106:133–142
- Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412:904–907
- He S, Liang XF, Sun J et al (2013) Insights into food preference in hybrid F1 of *Siniperca chuatsi* (♀) × *Siniperca scherzeri* (♂) mandarin fish through transcriptome analysis. *BMC Genom* 14:601. doi:10.1186/1471-2164-14-601
- Hecht BC, Thrower FP, Hale MC, Miller, Nichols KM (2012) Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. *G3 (Bethesda)* 2:1113–1127

- Hecht BC, Campbell NR, Holecek DE, Narum SR (2013) Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Mol Ecol* 22:3061–3076
- Heller MJ (2002) DNA microarray technology: devices, systems, and applications. *Annu Rev Biomed Eng* 4:129–153
- Henning F, Meyer A (2014) The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. *Annu Rev Genom Hum Genet* 15:417–441
- Henning F, Lee HJ, Franchini P, Meyer A (2014) Genetic mapping of horizontal stripes in Lake Victoria cichlid fishes: benefits and pitfalls of using RAD markers for dense linkage mapping. *Mol Ecol*. doi:10.1111/mec.12860
- Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2013) Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Mol Ecol* 22:2898–2916
- Hess JE, Campbell NR, Docker MF et al (2015) Use of genotyping by sequencing data to develop a high-throughput and multifunctional SNP panel for conservation applications in Pacific lamprey. *Mol Ecol Resour* 15:187–202
- Hickey JM, Kinghorn BP, Tier B, van der Werf JH, Cleveland MA (2012) A phasing and imputation method for pedigreed populations that results in a single-stage genomic evaluation. *Genet Sel Evol* 44:9. doi:10.1186/1297-9686-44-9
- Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet* 6:e1000862. doi:10.1371/journal.pgen.1000862
- Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Mol Ecol Resour* 11(Suppl 1):117–122
- Hohenlohe PA, Day MD, Amish SJ et al (2013) Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Mol Ecol* 22:3002–3013
- Houston RD (2015) Application of genomics to selective breeding of Atlantic salmon. In: Plant and animal genome XXIII conference. Plant and animal genome
- Houston RD, Davey JW, Bishop SC et al (2012) Characterisation of QTL-linked and genome-wide restriction site-associated DNA (RAD) markers in farmed Atlantic salmon. *BMC Genom* 13:244. doi:10.1186/1471-2164-13-244
- Houston RD, Taggart JB, Cézard T et al (2014) Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genom* 15:90. doi:10.1186/1471-2164-15-90
- Howe K, Clark MD, Torroja CF et al (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496:498–503
- Hu J, Vick BA (2003) Target region amplified polymorphism: a novel marker technique for plant genotyping. *Plant Mol Biol Report* 21:289–294
- Huang X, Feng Q, Qian Q et al (2009) High-throughput genotyping by whole-genome resequencing. *Genome Res* 19:1068–1076
- Hunter DJ (2005) Gene-environment interactions in human diseases. *Nat Rev Genet* 6:287–298
- Jaillon O, Aury JM, Brunet F et al (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431:946–957
- Jenkins S, Gibson N (2002) High-throughput SNP genotyping. *Comp Funct Genomics* 3:57–66
- Jiao W, Fu X, Dou J et al (2014) High-resolution linkage and quantitative trait locus mapping aided by genome survey sequencing: building up an integrative genomic framework for a bivalve mollusc. *DNA Res* 21:85–101
- Jones FC, Grabherr MG, Chan YF et al (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61
- Jones JC, Fan S, Franchini P, Scharl M, Meyer A (2013) The evolutionary history of Xiphophorus fish and their sexually selected sword: a genome-wide approach using restriction site-associated DNA sequencing. *Mol Ecol* 22:2986–3001
- Jones DB, Jerry DR, Khatkar MS et al (2014) Determining genetic contributions to host oyster shell growth: quantitative trait loci and genetic association analysis for the silver-lipped pearl oyster, *Pinctada maxima*. *Aquaculture* 434:367–375
- Judson R, Salisbury B, Schneider J, Windemuth A, Stephens JC (2002) How many SNPs does a genome-wide haplotype map require? *Pharmacogenomics* 3:279–391
- Kakioka R, Kokita T, Kumada H, Watanabe K, Okuda N (2015) Genomic architecture of habitat-related divergence and signature of directional selection in the body shapes of *Gnathopogon* fishes. *Mol Ecol* 24:4159–4174
- Kanamori A, Sugita Y, Yuasa Y, Suzuki T, Kawamura K, Uno Y, Kamimura K, Matsuda Y, Wilson CA, Amores A, Postlethwait JH, Suga K, Sakakura Y (2016) A genetic map for the only self-fertilizing vertebrate. *G3-Genes Genomes Genet* 6:1095–1106
- Kasahara M, Naruse K, Sasaki S et al (2007) The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447:714–719
- Keller I, Wagner CE, Greuter L et al (2013) Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol Ecol* 22:2848–2863
- Khamnamtong B, Klinbunga S, Menasveta P (2009) Genetic diversity and geographic differentiation of the giant tiger shrimp (*Penaeus monodon*) in Thailand analyzed by mitochondrial COI sequences. *Biochem Genet* 47:42–55
- Kim S, Misra A (2007) SNP genotyping: technologies and biomedical applications. *Annu Rev Biomed Eng* 9:289–320
- Kothiyal P, Cox S, Ebert J, Aronow BJ, Greinwald JH, Rehm HL (2009) An overview of custom array sequencing. *Current protocols in human genetics*, Chapter 7: Unit 7.17. doi: 10.1002/0471142905.hg0717s61
- Kumar S, Banks TW, Cloutier S (2012) SNP discovery through next-generation sequencing and its applications. *Int J Plant Genomics* 2012:831460
- Kwok PY (2001) Methods for genotyping single nucleotide polymorphisms. *Annu Rev Genomics Hum Genet* 2:235–258
- Kwok PY, Gu Z (1999) Single nucleotide polymorphism libraries: Why and how are we building them? *Mol Med Today* 5:538–543

- Kwok PY, Xiao M (2004) Single-molecule analysis for molecular haplotyping. *Hum Mutat* 23:442–446
- Lah L, Benke H, Berggren P et al (2014) Investigating harbor porpoise (*Phocoena phocoena*) population differentiation using RAD-tag genotyping by sequencing. IWC Meeting Portal, Scientific Committee Annual Meeting 2014 (SC65B), <https://events.iwc.int/index.php/scientific/SC65B/paper/view/825>
- Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199
- Laporte M, Pavey SA, Rougeux C, Pierron F, Lauzent M, Budzinski H, Labadie P, Geneste E, Couture P, Baudrimont M, Bernatchez L (2016) RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Mol Ecol* 25:219–237
- Larson WA, Seeb JE, Pascal CE, Templin WD, Seeb LW, Taylor E (2014a) Single-nucleotide polymorphisms (SNPs) identified through genotyping-by-sequencing improve genetic stock identification of Chinook salmon (*Oncorhynchus tshawytscha*) from western Alaska. *Can J Fish Aquat Sci* 71:698–708
- Larson WA, Seeb LW, Everett MV, Waples RK, Templin WD, Seeb JE (2014b) Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evol Appl* 7:355–369
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends Ecol Evol* 17:386–391
- Lemay MA, Russello MA (2015) Genetic evidence for ecological divergence in kokanee salmon. *Mol Ecol* 24:798–811
- Lescak EA, Bassham SL, Catchen J, Gelmond O, Sherbick ML, von Hippel FA, Cresko WA (2015) Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proc Natl Acad Sci USA* 112:E7204–E7212
- Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP et al (2007) The diploid genome sequence of an individual human. *PLoS Biol* 5(10):e254. doi:10.1371/journal.pbio.0050254
- Lew RM, Finger AJ, Baerwald MR, Goodbla A, May B, Meek MH (2015) Using next-generation sequencing to assist a conservation hatchery: a single-nucleotide polymorphism panel for the genetic management of endangered delta smelt. *Trans Am Fish Soc* 144:767–779
- Li Y, Guo X, Cao X, Deng W, Luo W, Wang W (2012) Population genetic structure and post-establishment dispersal patterns of the red swamp crayfish *Procambarus clarkii* in China. *PLoS ONE* 7:e40652. doi:10.1371/journal.pone.0040652
- Li C, Waldbieser G, Bosworth B, Beck BH, Thongda W, Peatman E (2014) SNP discovery in wild and domesticated populations of blue catfish, *Ictalurus furcatus*, using genotyping-by-sequencing and subsequent SNP validation. *Mol Ecol Resour*. doi:10.1111/1755-0998.12272
- Limborg MT, Waples RK, Seeb JE, Seeb LW (2014) Temporally isolated lineages of pink salmon reveal unique signatures of selection on distinct pools of standing genetic variation. *J Hered*. doi:10.1093/jhered/esu063
- Liu S, Vallejo RL, Gao G, Palti Y, Weber GM, Hernandez A, Rexroad CE III (2015a) Identification of single-nucleotide polymorphism markers associated with cortisol response to crowding in rainbow trout. *Mar Biotechnol* 17:328–337
- Liu S, Vallejo RL, Palti Y, Gao G, Marancik DP, Hernandez AG, Wiens GD (2015b) Identification of single nucleotide polymorphism markers associated with bacterial cold water disease resistance and spleen size in rainbow trout. *Front Genet* 6:298
- Liu-Stratton Y, Roy S, Sen CK (2004) DNA microarray technology in nutraceutical and food safety. *Toxicol Lett* 150:29–42
- Longo G, Bernardi G (2015) The evolutionary history of the embiotocid surrperch radiation based on genome-wide RAD sequence data. *Mol Phylogenet Evol* 88:55–63
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nat Rev Genet* 4:981–994
- Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Mol Ecol* 3:91–99
- Majewski J, Schwartzenruber J, Lalonde E, Montpetit A, Jabado N (2011) What can exome sequencing do for you? *J Med Genet* 48:580–589
- Mandeville EG, Parchman TL, McDonald DB, Buerkle CA (2015) Highly variable reproductive isolation among pairs of *Catostomus* species. *Mol Ecol*. doi:10.1111/mec.13118
- Manousaki T, Tsakogiannis A, Taggart JB, Palaiokostas C, Tsaparis D, Lagnel J, Chatziplis D, Magoulas A, Papanoulakis N, Mylonas CC, Tsigenopoulos CS (2016) Exploring a nonmodel teleost genome through RAD sequencing-linkage mapping in common pandora, *Pagellus erythrinus* and comparative genomic analysis. *G3-Genes Genomes Genet* 6:509–519
- Martin CH, Feinstein LC (2014) Novel trophic niches drive variable progress towards ecological speciation within an adaptive radiation of pupfishes. *Mol Ecol* 23:1846–1862
- Mascher M, Wu S, Amand PS, Stein N, Poland J (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PLoS ONE* 8:e76925. doi:10.1371/journal.pone.0076925
- Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Pinero D, Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol Ecol Resour* 15:28–41
- McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol Phylogenet Evol* 66:526–538
- McPherson JD (2009) Next-generation gap. *Nat Methods* 6:S2–S5
- Messer PW, Ellner SP, Hairston NG Jr (2016) Can population genetics adapt to rapid evolution? *Trends Genet* 32(7):408–418
- Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res* 17:240–248
- Milner JA (2004) Molecular targets for bioactive food components. *J Nutr* 134:2492S–2498S

- Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. *Tree* 14:389–394
- Nakamura Y, Mori K, Saitoh K et al (2013) Evolutionary changes of multiple visual pigment genes in the complete genome of Pacific bluefin tuna. *Proc Natl Acad Sci* 110:11061–11066
- Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA (2013a) Genotyping-by-sequencing in ecological and conservation genomics. *Mol Ecol* 22:2841–2847
- Narum SR, Campbell NR, Meyer KA, Miller MR, Hardy RW (2013b) Thermal adaptation and acclimation of ectotherms from differing aquatic climates. *Mol Ecol* 22:3090–3097
- Norrell AE, Crawley D, Jones KL, Saillant EA (2014) Development and characterization of eighty-four microsatellite markers for the red snapper (*Lutjanus campechanus*) using illumina paired-end sequencing. *Aquaculture* 430:128–132
- Nunez JCB, Seale TP, Fraser MA, Burton TL, Fortson TN, Hoover D, Travis J, Oleksiak MF, Crawford DL (2015) Population Genomics of the Euryhaline Teleost *Poecilia latipinna*. *PLoS ONE*. doi:10.1371/journal.pone.0137077
- Ogden R, Gharbi K, Mugue N et al (2013) Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. *Mol Ecol* 22:3112–3123
- Oliphant A, Barker DL, Stuelpnagel JR, Chee MS (2002) BeadArray technology-enabling an accurate. *Biotechniques* 32:S56–S61
- O'Quin KE, Schulte JE, Patel Z et al (2012) Evolution of cichlid vision via trans-regulatory divergence. *BMC Evol Biol* 12:251. doi:10.1186/1471-2148-12-251
- O'Quin KE, Yoshizawa M, Doshi P, Jeffery WR (2013) Quantitative genetic analysis of retinal degeneration in the blind cavefish *Astyanax mexicanus*. *PLoS ONE* 8:e57281. doi:10.1371/journal.pone.0057281
- Palaiokostas C, Bekaert M, Davie A et al (2013a) Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genom* 14:566. doi:10.1186/1471-2164-14-566
- Palaiokostas C, Bekaert M, Khan MG et al (2013b) Mapping and validation of the major sex-determining region in Nile tilapia (*Oreochromis niloticus* L.) using RAD sequencing. *PLoS ONE* 8:e68389. doi:10.1371/journal.pone.0068389
- Palti Y, Gao G, Miller MR et al (2014) A resource of single-nucleotide polymorphisms for rainbow trout generated by restriction-site associated DNA sequencing of doubled haploids. *Mol Ecol Resour* 14:588–596
- Piferrer F, Ribas L, Diaz N (2012) Genomic approaches to study genetic and environmental influences on fish sex determination and differentiation. *Mar Biotechnol* (NY) 14:591–604
- Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5:92
- Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7:e32253. doi:10.1371/journal.pone.0032253
- Powell W, Machray GC, Provan J (1996a) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1:215–222
- Powell W, Morgante M, Andre C et al (1996b) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2:225–238
- Puebla O, Bermingham E, McMillan WO (2014) Genomic atolls of differentiation in coral reef fishes (*Hypoplectrus* spp., *Serranidae*). *Mol Ecol* doi:10.1111/mec.12926
- Pujolar JM, Jacobsen MW, Frydenberg J et al (2013) A resource of genome-wide single-nucleotide polymorphisms generated by RAD tag sequencing in the critically endangered European eel. *Mol Ecol Resour* 13:706–714
- Pukk L, Ahmad F, Hasan S, Kisand V, Gross R, Vasemagi A (2015) Less is more: extreme genome complexity reduction with ddRAD using ion torrent semiconductor technology. *Mol Ecol Resour*. doi:10.1111/1755-0998.12392
- Pyšek P, Richardson DM (2010) Invasive species, environmental change and management, and health. *Annu Rev Environ Resour* 35:25–55
- Reinhardt F, Liu Z, Seefried F, Thaller G (2009) Implementation of genomic evaluation in German Holsteins. *Interbull Bull* 40:219–226
- Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Mol Ecol* 22:2953–2970
- Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112–113:183–198
- Roberts MA, Mutch DM, German JB (2001) Genomics: food and nutrition. *Curr Opin Biotechnol* 12:516–522
- Roda F, Ambrose L, Walter GM et al (2013) Genomic evidence for the parallel evolution of coastal forms in the *Senecio lautus* complex. *Mol Ecol* 22:2941–2952
- Roesti M, Moser D, Berner D (2013) Recombination in the threespine stickleback genome—patterns and consequences. *Mol Ecol* 22:3014–3027
- Rondeau EB, Messmer AM, Sanderson DS et al (2013) Genomics of sablefish (*Anoplopoma fimbria*): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene. *BMC Genom* 14:452. doi:10.1186/1471-2164-14-452
- Sansaloni C, Petroli C, Jaccoud D et al (2011) Diversity arrays technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proc* 5:P54
- Sax DF, Stachowicz JJ, Brown JH et al (2007) Ecological and evolutionary insights from species invasions. *Trends Ecol Evol* 22:465–471
- Schartl M, Walter RB, Shen Y et al (2013) The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. *Nat Genet* 45:567–572
- Schlötterer C (2004) The evolution of molecular markers—Just a matter of fashion? *Nat Rev Genet* 5(1):63–69
- Schorika N, Fallina D, Lanchbury J (2000) Single nucleotide polymorphisms and the future of genetic epidemiology. *Clin Genet* 58:250–264
- Schunter C, Garza JC, Macpherson E, Pascual M (2014) SNP development from RNA-seq data in a nonmodel fish: How many individuals are needed for accurate allele frequency prediction? *Mol Ecol Resour* 14:157–165

- Seeb LW, Waples RK, Limborg MT, Warheit KI, Pascal CE, Seeb JE (2014) Parallel signatures of selection in temporally isolated lineages of pink salmon. *Mol Ecol* 23:2473–2485
- Shao C, Niu Y, Rastas P, Liu Y, Xie Z, Li H, Wang L, Jiang Y, Tai S, Tian Y, Sakamoto T, Chen S (2015) Genome-wide SNP identification for the construction of a high-resolution genetic map of Japanese flounder (*Paralichthys olivaceus*): applications to QTL mapping of *Vibrio anguillarum* disease resistance and comparative genomic analysis. *DNA Res* 22:161–170
- Shi Y, Wang S, Gu Z et al (2014) High-density single nucleotide polymorphisms linkage and quantitative trait locus mapping of the pearl oyster, *Pinctada fucata martensii* Dunker. *Aquaculture* 434:376–384
- Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* 21:1794–1805
- Skovrind M, Olsen MT, Vieira FG, Pacheco G, Carl H, Gilbert MTP, Moller PR (2016) Genomic population structure of freshwater-resident and anadromous ide (*Leuciscus idus*) in north-western Europe. *Ecol Evol* 6:1064–1074
- Sreeparvathy M (2013) Nutritional genomics. *Int J Biol Biol Sci* 2:150–153
- Stapley J, Reger J, Feulner PG et al (2010) Adaptation genomics: the next generation. *Trends Ecol Evol* 25:705–712
- Star B, Nederbragt AJ, Jentoft S et al (2011) The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477:207–210
- Stemers FJ, Gunderson KL (2007) Whole genome genotyping technologies on the BeadArray platform. *Biotechnology* 2(1):41–49
- Storer CG, Pascal CE, Roberts SB, Templin WD, Seeb LW, Seeb JE (2012) Rank and order: evaluating the performance of SNPs for individual assignment in a non-model organism. *PLoS ONE* 7:e49018. doi:10.1371/journal.pone.0049018
- Stranden I, Christensen OF (2011) Allele coding in genomic evaluation. *Genet Sel Evol* 43:25
- Syvänen AC (2001) Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet* 2:930–942
- Syvanen AC (2005) Toward genome-wide SNP genotyping. *Nat Genet* 37(Suppl):S5–S10
- Takahashi T, Sota T, Hori M (2013) Genetic basis of male colour dimorphism in a Lake Tanganyika cichlid fish. *Mol Ecol* 22:3049–3060
- Tang K, Fu DJ, Julien D, Braun A, Cantor CR, Köster H (1999) Chip-based genotyping by mass spectrometry. *Proc Natl Acad Sci* 96:10016–10020
- Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17:6463–6471
- Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res* 12:4127–4138
- Templin WD, Seeb JE, Jasper JR, Barclay AW, Seeb LW (2011) Genetic differentiation of Alaska Chinook salmon: the missing link for migratory studies. *Mol Ecol Resour* 11(Suppl 1):226–246
- Tost J, Gut IG (2005) Genotyping single nucleotide polymorphisms by MALDI mass spectrometry in clinical applications. *Clin Biochem* 38:335–350
- Trujillo E, Davis C, Milner J (2006) Nutrigenomics, proteomics, metabolomics, and the practice of dietetics. *J Am Diet Assoc* 106:403–413
- Underwood ZE, Mandeville EG, Walters AW (2016) Population connectivity and genetic structure of burbot (*Lota lota*) populations in the Wind River Basin, Wyoming. *Hydrobiologia* 765:329–342
- Vamosi JC, Armbruster WS, Renner SS (2014) Evolutionary ecology of specialization: insights from phylogenetic analysis. *Proc R Soc B Biol Sci*. doi:10.1098/rspb.2014.2004
- Van Bers NE, Crooijmans RP, Groenen MA, Dibbitts BW, Komen J (2012) SNP marker detection and genotyping in tilapia. *Mol Ecol Resour* 12:932–941
- van Der WMJ, Schuren FHJ, Bijlsma S, Tas AC, van Omen B (2001) Nutrigenomics: application of genomics technologies in nutritional sciences and food technology. *J Food Sci* 66:772–780
- van Orsouw NJ, Hogers RC, Janssen A et al (2007) Complexity reduction of polymorphic sequences (CRoPS): a novel approach for large-scale polymorphism discovery in complex genomes. *PLoS ONE* 2:e1172. doi:10.1371/journal.pone.0001172
- Vera M, Alvarez-Dios JA, Fernandez C, Bouza C, Vilas R, Martinez P (2013) Development and validation of single nucleotide polymorphisms (SNPs) markers from two transcriptome 454-runs of turbot (*Scophthalmus maximus*) using high-throughput genotyping. *Int J Mol Sci* 14:5694–5711
- Viricel A, Pante E, Dabin W, Simon-Bouhet B (2014) Applicability of RAD-tag genotyping for interfamilial comparisons: empirical data from two cetaceans. *Mol Ecol Resour* 14:597–605
- Vos P, Hogers R, Bleeker M et al (1995) AFLP—a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414
- Wagner CE, Keller I, Wittwer S et al (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Mol Ecol* 22:787–798
- Walker WA, Blackburn G (2004) Symposium introduction: nutrition and gene regulation. *J Nutr* 134:2434S–2436S
- Wang S (2015) Genotyping-by-sequencing (GBS) using a semiconductor sequencing platform. In: Plant and animal genome XXIII conference. Plant and animal genome
- Wang Z, Pascual-Anaya J, Zadissa A et al (2013) The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nat Genet* 45:701–706
- Wang W, Hu Y, Ma Y, Xu L, Guan J, Kong J (2015) High-density genetic linkage mapping in turbot (*Scophthalmus maximus* L.) based on SNP markers and major sex- and growth-related regions detection. *PLoS ONE*. doi:10.1371/journal.pone.0120410
- Wiggans GR, Sonstegard TS, VanRaden PM et al (2009) Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J Dairy Sci* 92:3431–3436

- Wiggans GR, VanRaden PM, Bacheller LR et al (2010) Selection and management of DNA markers for use in genomic evaluation. *J Dairy Sci* 93:2287–2292
- Willing EM, Hoffmann M, Klein JD, Weigel D, Dreyer C (2011) Paired-end RAD-seq for de novo assembly and marker design without available reference. *Bioinformatics* 27:2187–2193
- Xu J, Zhao Z, Zhang X et al (2014a) Development and evaluation of the first high-throughput SNP array for common carp (*Cyprinus carpio*). *BMC Genom* 15:307 <http://www.biomedcentral.com/1471-2164/1415/1307>
- Xu P, Zhang X, Wang X et al (2014b) Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat Genet*. doi:10.1038/ng.3098
- Yim HS, Cho YS, Guang X et al (2014) Minke whale genome and aquatic adaptation in cetaceans. *Nat Genet* 46:88–92
- Yoshizawa M, Robinson BG, Duboué ER et al (2015) Distinct genetic architecture underlies the emergence of sleep loss and prey-seeking behavior in the Mexican cavefish. *BMC Biol* 13:15. doi:10.1186/s12915-015-0119-3
- Zhang B-D, Xue D-X, Wang J, Li Y-L, Liu B-J, Liu J-X (2016) Development and preliminary evaluation of a genomewide single nucleotide polymorphisms resource generated by RAD-seq for the small yellow croaker (*Larimichthys polyactis*). *Mol Ecol Resour* 16:755–768
- Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20:176–183