



Applications in the search for genomic selection signatures in fish

María E. López^{1,2*}, Roberto Neira¹ and José M. Yáñez^{2,3*}

¹ Faculty of Agricultural Sciences, University of Chile, Santiago, Chile

² Aquainnovo, Puerto Montt, Chile

³ Faculty of Veterinary and Animal Sciences, University of Chile, Santiago, Chile

Edited by:

Peng Xu, Chinese Academy of Fishery Sciences, China

Reviewed by:

Magadan Momo Susana, Institut National de la Recherche Agronomique Unité de Virologie et Immunologie Moleculaires, France
William S. Davidson, Simon Fraser University, Canada

*Correspondence:

María E. López and José M. Yáñez, Faculty of Agricultural Sciences and Faculty of Veterinary and Animal Sciences, University of Chile, Avenue Santa Rosa 11735, La Pintana, P. O. Box 8820808, Santiago, Chile
e-mail: me.lopez.dinamarca@gmail.com; jmayanez@uchile.cl

Selection signatures are genomic regions harboring DNA sequences functionally involved in the genetic variation of traits subject to selection. Selection signatures have been intensively studied in recent years because of their relevance to evolutionary biology and their potential association with genes that control phenotypes of interest in wild and domestic populations. Selection signature research in fish has been confined to a smaller scale, due in part to the relatively recent domestication of fish species and limited genomic resources such as molecular markers, genetic mapping, DNA sequences, and reference genomes. However, recent genomic technology advances are paving the way for more studies that may contribute to the knowledge of genomic regions underlying phenotypes of biological and productive interest in fish.

Keywords: fish, selection, domestication, single nucleotide polymorphisms, genome

INTRODUCTION

Selection signatures are genomic regions that harbor DNA sequences involved in genetic variation of traits subject to natural or artificial selection (Qanbari et al., 2012). Currently, due to advances in genomic technologies and statistical methods, such signatures can be identified in the genomes of various species.

Most studies in this field of genetics are based on the concept of *hitchhiking*, which suggests that selection affects the genome at a specific region, leaving “signatures” around the selected gene(s) (Smith and Haigh, 1974). Specifically, the hitchhiking theory focuses on the spread of new variants in a population due to selection for their favorable effects (Przeworski, 2002; Kim and Nielsen, 2004). Selection involving alleles from the population’s standing genetic variation produces specific and detectable DNA sequence patterns (Hermisson and Pennings, 2005).

The search for these molecular signatures has been the subject of intense research in recent years in both domesticated and wild populations of plants and animals, as well as in humans. These studies have been motivated by two main objectives: (1) a strong interest in the evolutionary past of the species and basic molecular mechanisms governing this evolution and (2) the expectation of an association between these genomic regions and biological functions or phenotypes of interest, since these regions should have some functional or adaptive importance underlying their selection (Nielsen et al., 2007). These studies are possible due to the development of various methods aimed at detecting selection at the molecular level in population samples. Information on allelic frequencies or haplotype patterns segregated in the population

can be used to identify signatures, since selection modifies the patterns of genetic variation expected under the neutral theory of molecular evolution.

Most studies in domesticated populations have focused on detecting relatively old selection signatures dating back hundreds or thousands of generations, e.g., (Flori et al., 2009), with few studies on genetic changes during early domestication stages (Trut et al., 2009).

Certain fish species provide unique models for studying the effects of selection and domestication, as their populations were domesticated recently and are available as both wild and domesticated populations simultaneously.

In this article we present different aspects involved in studying selection signatures at a genomic level in different species and discuss about the potential application of these studies in fish populations to unravel recent selection and domestication processes in these species.

IDENTIFICATION OF LOCI ASSOCIATED WITH TRAITS OF INTEREST

The search for genes controlling phenotypic variation can be performed in two different ways. First, the “top-down” approach which begins with knowledge of the phenotype of interest and uses genetic analysis to identify genes or causal regions. These approaches include candidate gene studies, identification of Quantitative trait loci (QTLs) and association mapping. These studies have certain limitations, including the need for an *a priori* hypothesis about which genes underlie the trait of interest, information about family relationship between individuals, as well as, access

to a large number of relatives with phenotypic records (Gu et al., 2009). Second, the “bottom-up” approach, in contrast, begins with genomic information and involves statistical evaluation of molecular information to identify regions subject to selection (Ross-Ibarra et al., 2007). This approach searches for patterns of linkage disequilibrium, genetic differentiation, or frequency spectrum that are inconsistent with the neutral evolution model to identify selection signatures (Qanbari et al., 2010). Recent advances in genomics provide a new paradigm for the “bottom-up” strategy concerning population genomics, a discipline that infers genetic and evolutionary parameters of a population based on datasets from the whole genome (Black et al., 2001).

In this context, population genomics relies on two basic principles or assumptions. First, neutral loci will be equally affected by demographic effects and by the evolutionary history of the population. Second, loci under selection will tend to behave distinctively, revealing atypical variation patterns (Luikart et al., 2003).

MODELS OF SELECTION

Natural selection can be defined as the differential contribution of genetic variation to future generations (Aquadro et al., 2001) due to differential reproduction of some phenotypes/genotypes over others under prevailing environmental conditions at a given time (Futuyma, 1998). It is the driving force behind Darwinian evolution and can be subdivided into different types, depending on the evolutionary outcome (Hurst, 2009).

Directional selection tends to decrease variation *within* a population but may increase or decrease variation *among* populations. Positive selection is a type of directional selection that favors alleles that increase fitness of individuals. When directional selection eliminates unfavorable mutations, it is called purifying selection (also known as negative selection).

Diversifying (or disruptive) selection favors variety and benefits individuals with extreme phenotypes over intermediate. In this type of selection, the propagation of an allele never reaches fixation, and therefore it may occur when an allele is initially subject to positive selection, and then negative selection when the frequency becomes too high (Nielsen, 2005).

Balanced selection, which helps to maintain an equilibrium point at which both alleles remain in the population, has several forms, including frequency-dependent selection and overdominance, which occurs when the heterozygote has the higher biological fitness, and therefore variability is maintained in the population (Nielsen, 2005).

SELECTION SIGNATURES

In the classic “hitchhiking” scenario, first described by Smith and Haigh (1974), a new allelic variant that represents a favorable adaptive substitution originates within the population as a new mutation, and its frequency increases as a result of constant selection pressure. When a favorable allele is selected, and its frequency increases to fixation in a population, genetic variation in the surrounding DNA segment is altered; that is, the increased frequency of the selected allele also produces increased frequency of closely-linked alleles (Pennings and Hermisson, 2006).

The ancestral variation, i.e., genetic variation present in a population prior to a selection process, is maintained only if

recombination during this phase disrupts the association between an adjacent locus and the selected site. The resulting pattern of such a selective event is a strong reduction in genetic variation around the selected site, known as a “*hard sweep*,” which corresponds to the classic selective sweep (Pritchard et al., 2010).

There is a second scenario in which an adaptive substitution involves multiple copies of a favorable allele in the population. This may occur for two reasons. First, when an adaptation arises from genetic variation, many copies of the favorable allele may be present in the population. Fixation of this allele may involve descendants of more than one of these copies. Second, a favorable allele can be introduced in the population by recurrent mutation or migration during a selection phase, and again, several descendants of independent origin may contribute to the allelic fixation. In both cases, different alleles of loci adjacent to any such favorable copies will be retained in the population, resulting in different haplotypes (Pennings and Hermisson, 2006).

Selection signatures involving descendants of more than one copy of the selected allele and, therefore with different haplotypes at closely-linked sites, are called “*soft sweeps*.” This type of selection signature results in different haplotype patterns than the “*hard sweeps*” described above and it is more difficult to detect as it only produces a slight reduction in the levels of adjacent polymorphisms (Cutter and Payseur, 2013).

Furthermore, when adaptation occurs by polygenic selection, it induces an increase in the allelic frequency of several loci which have a favorable effect on a particular phenotype; however, these polygenic alleles do not necessarily achieve fixation, and the resulting haplotype pattern corresponds to several partial selection signatures or multiple “*partial sweeps*” (Pritchard et al., 2010).

Finally, when purifying or negative selection reduces the frequency or eliminates a deleterious allele, the genetic diversity at linked loci also decreases, which is known as “*background selection*” (Charlesworth, 1993).

Figure 1 schematically summarizes the patterns caused by “*hard sweeps*,” “*soft sweeps*,” and “*partial sweeps*” that correspond to selective events for favorable variants in a population, as well as the pattern produced by “*background selection*.”

DOMESTICATION AND RECENT ARTIFICIAL SELECTION IN FISH

Domestication is the process by which various species have been adapted to a captive environment by humans. Such adaptation is accomplished through systematic breeding over generations and is characterized by changes in behavior, morphology, and physiology, as well as adaptive genetic changes caused by artificial and natural selection (Price, 1999).

In fish, domestication occurred very recently as compared to other land animals. One theory to explain the late domestication of aquatic species suggests that, due to the high fertility of these species, a small number of broodstock were required to obtain a sufficiently large progeny in subsequent generations. After a few generations, the inbreeding depression increases considerably; therefore, fitness and productive behavior decrease. As a result, fish farmers were forced to repeatedly take new broodstock from the wild environment, interrupting the continuity of domestication and breeding (Gjedrem et al., 2012). For this reason, aquaculture

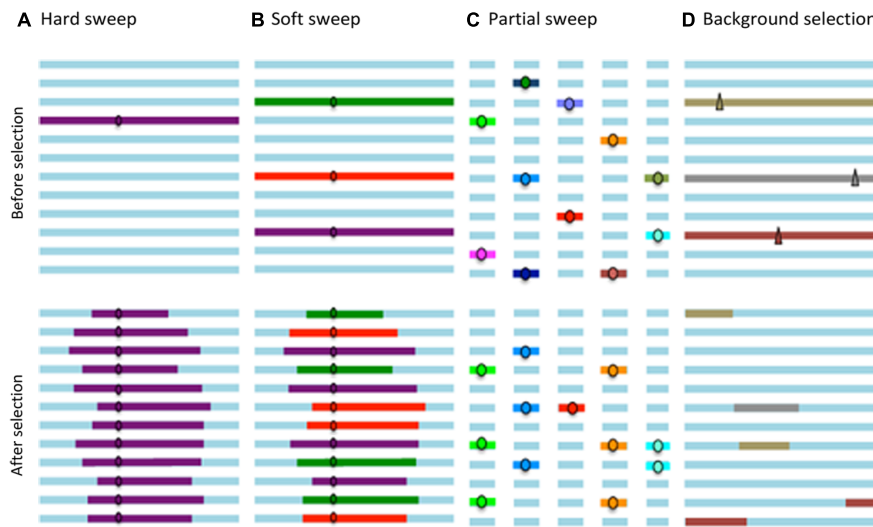


FIGURE 1 | The horizontal lines represent the haplotypes in a population; the circles favorable alleles; and the triangles deleterious alleles. **(A)** Hard sweeps involve fixation of a new favorable variant. **(B)** Soft sweeps involve fixation of equivalent alleles from different genetic

backgrounds and therefore result in different haplotypes. **(C)** Partial sweeps involve increased allele frequency at several loci, without reaching fixation. **(D)** Background selection eliminates deleterious variants and linked loci.

has lagged behind land animal and plant culture in the use of breeding to enhance biological production efficiency.

On the other hand, it is estimated that less than 10% of aquaculture production is based on genetically improved stocks (Gjedrem, 2012), although the annual genetic gains reported for aquaculture species are substantially higher than for land animals. For example, selection response reported for grow-related traits are even higher than 10% in fish populations, which can substantially enhance aquaculture production by selective breeding (Gjedrem et al., 2012). Recent high selective pressures in farmed fish populations may have shaped genome variation in regions harboring causative mutations of selected traits. The identification of these regions may help in the understanding of the effect of selection events and identification of genetic variants involved in phenotypic variation in fish populations.

EXAMPLES OF DOMESTICATION AND BREEDING PROGRAMS IN AQUATIC SPECIES

The fish belonging to the family *Ciprinidae* are most likely the first fish species to be domesticated. For instance, the goldfish (*Carrassius auratus*) is an ornamental fish, believed to have been domesticated in China before the XVI century and later taken to Japan and Europe (Purdom, 1993). Another important group of ornamental fish is the koi carp, a variety derived from common carp (*Cyprinus carpio*) and mainly cultivated in Japan. The large variety of colors and forms among koi carp resulted from directed selection and crossbreeding (Gjedrem, 2005).

There is evidence in common carp showing a large response to selection for furunculosis survival rates (Schaperclaus, 1962). In 1987, Ilyassov (1987) showed results from 4 to 5 generations of selection in this species for resistance to dropsy disease, which increased survival by 30–40% as compared to unselected carps.

On the other hand, selection responses reported for growth rate in rohu carp (*Labeo rohita*) have been particularly high, reaching almost a 30% per generation (Gjedrem, 2012).

Furthermore, salmonid species are the most intensively selected fish populations. In this regard, the rainbow trout (*Oncorhynchus mykiss*) has a long history of domestication and breeding in the United States, Norway, Finland, and Denmark (McAndrew and Napier, 2011). In 1932, investigators began to select individuals to improve growth rate, number of eggs, and characteristics of sexual maturity (Donaldson and Olson, 1957). Currently there are 13 breeding programs worldwide aimed at improving growth rate, age at sexual maturity, fillet quality, and disease resistance in this species (Rye et al., 2010).

In the case of Atlantic salmon (*Salmo salar*), breeding programs exist in Norway, Scotland, Ireland, Australia, and Chile (Norris et al., 1999; Metcalfe et al., 2003; Glover et al., 2009; Dominik et al., 2010; Rye et al., 2010; McAndrew and Napier, 2011). Several traits of commercial interest such as growth, sexual maturity, meat quality, and disease resistance have been incorporated into breeding objectives. Furthermore, findings from genomic technologies have been incorporated into these breeding programs, for example, the use of QTLs to assist selection for resistance against the viral disease named infectious pancreatic necrosis (Houston et al., 2008; Moen et al., 2009).

Among Pacific salmon, the chinook salmon (*Oncorhynchus tshawytscha*) originating in British Columbia (BC), Canada was one the first species of salmon to be domesticated (Kim et al., 2004). Currently, its farming is limited and there are two breeding programs in operations (Rye et al., 2010). Moreover, genetic improvement programs for coho salmon (*Oncorhynchus kisutch*) have been successful in selecting for harvest weight and early spawning, with selection responses of about 10% per generation (Neira et al., 2006).

Tilapias are the second-most important group of cultivated fish in the world. The dominant species is the Nile tilapia (*Oreochromis niloticus*); however, other species of the genus *Oreochromis* (Neira, 2010) are also cultivated. The GIFT (Genetic Improvement of Farmed Tilapias) program, begun in 1987 in the Philippines, systematically compared wild and commercial strains in various aquatic environments and established a family-based selection system to improve growth rate (Eknath et al., 1993). The program is currently managed by the World Fish Center in Malaysia and genetic gains for growth-related traits are among 10–15% (Ponzoni et al., 2011).

Breeding programs have recently been established for other important species such as, sea bass (*Dicentrarchus labrax*; Vandeputte et al., 2009), sea bream (*Sparus aurata*), turbot (*Scophthalmus maximus*), Atlantic cod (*Gadus morhua*; Glover et al., 2011), halibut (Glover et al., 2007), and tuna (Owen, 2011).

All of these domestication and artificial selection processes shape the genomes of cultured fish populations, resulting in selection signatures that could potentially be identified using molecular and statistical methods.

APPROACHES USED FOR DETECTING SELECTION SIGNATURES

When a new allelic variant that does not affect the fitness of individuals originates in a population, it is not affected by natural selection and is said to be neutral. Statistical tests aimed at testing a neutral evolution model can be divided into three main classes: (1) tests based on polymorphisms within species; (2) tests based on the differences between species; and (3) tests that use information within and between species. A description of these three approaches is given below.

TESTS BASED ON POLYMORPHISMS WITHIN SPECIES

Frequency spectrum

The frequency spectrum is defined as the allele frequency distribution of a large number of independent loci in a given sample (Nielsen, 2005; Vogl and Clemente, 2012). Deviations from expectations of the neutral model (no selection, recombination, population subdivision, or changes in the effective population size) could be indicative of selection: purifying or negative selection tends to increase the fraction of mutations segregating at low frequencies, while positive selection increases the number of alleles observed at high frequencies (Hurst, 2009).

Many tests for detecting selection signatures are based on information provided by the frequency spectrum obtained from DNA sequence data. One of the most commonly used is the Tajima's (1989) D test, which compares two measures of genetic variation (θ). The first is obtained from the average of nucleotide differences between pairs of sequences, and the second is the total number of segregating sites (Nielsen, 2005). If the difference between these two measures is greater than expected under neutral evolution, this model is rejected. Other tests have incorporated phylogenetic information in order to estimate the direction of change and increase power to detect deviations from the null hypothesis of the neutral model (Perfectti et al., 2009). One such test is that of Fu and Li (1993), which also calculates a statistic based on the comparison of two genetic variation estimates, adding phylogenetic

information. For example, a related species may be added as an outgroup, such as the inclusion of the chimpanzee in an analysis of human genetic variation (Nielsen, 2005). Likewise, Fay and Wu (2000) developed a test based on the concept that the frequency spectrum expected under neutrality must be enriched with mutations at low frequencies, and that therefore, mutations at high frequencies are atypical.

Researchers have used this approach to detect selection signatures in several species. In humans, for example, evidence of selection has been found in genes related to the immune system and social behavior (Sabeti et al., 2002; Williamson, 2007). In other species such as chickens, it has been possible to identify genomic regions related to production-related traits such as eggshell hardness and immune system characteristics (Qanbari et al., 2012).

Linkage disequilibrium (LD) and haplotype structure

Linkage disequilibrium (LD) refers to the non-random association of alleles at two or more loci. That is, if two alleles at two loci segregate together in greater proportion than expected by chance, it is said that these loci are in linkage disequilibrium. This measure has been widely used to study various demographic events and evolutionary processes in plants and animals, such as breeding systems, patterns of geographic subdivision, events of natural, and artificial selection, gene conversion, mutation, and other forces that can cause changes in gene frequency (Slatkin, 2008). The LD is affected by different evolutionary factors, including recombination, admixture, bottlenecks, gene flow, genetic drift, inbreeding, and selection (Slatkin, 2008). As a consequence, LD across the genome can vary within and between populations.

Thus, another approach to detect genomic selection signatures is based on statistical comparisons of atypical LD patterns at specific haplotypes of certain genomic regions that are inconsistent with the neutral evolution model (Mueller, 2004). This approach has been used in numerous studies to detect selection signatures in humans and in domesticated species (Sabeti et al., 2002; Voight et al., 2006; Hayes et al., 2008). These studies are based on the concept that in a large population, a neutral variant, which by definition is not under selection, will take many generations to become fixed or lost. Recombination and the passing of generations act with stronger intensity, and therefore, LD around these neutral alleles erodes quickly, leaving a smaller surrounding haplotype (Kimura, 1983; Nielsen et al., 2005b; Sabeti et al., 2006).

Conversely, alleles under positive or balanced selection carry other linked alleles with them, generating increased LD in the genomic region, as described for the hitchhiking effect (Smith and Haigh, 1974). LD between these alleles is slowly eroded, such that the adjacent haplotype is longer than expected by chance (Sabeti et al., 2002). Thus, large haplotypes reflect positive selection. This forms the basis of the EHH statistic ("extended haplotype homozygosity") suggested by Sabeti et al. (2002), which is defined as the probability that two randomly selected chromosomes carrying the core haplotype are identical by descent, and also measures the decay of haplotype homozygosity as a function of the distance. EHH allows for identification of regions with atypical frequencies of extended haplotypes and has been effectively used to detect

signatures of recent positive selection within a population (Tang et al., 2004; Walsh et al., 2006).

Voight et al. (2006) developed the statistic iHS or “integrated Haplotype Score” which allows to compare the area under the curve of EHH distribution between ancestral and derived alleles. This approach is based on the fact that the EHH area of an allele under selection will be greater than that of a neutral allele; therefore, the integral of EHH captures this effect. iHS corresponds to a standardized ratio between the areas under the curve of ancestral and derived alleles, which is equal to 0 when the EHH decay is similar for both types of alleles. A negative iHS value near -1 indicates extended haplotype around a derived allele, whereas positive values near one indicate extended haplotype around an ancestral allele.

The iHS statistic is more sensitive for detecting rapid increases in frequencies of the derived allele produced by selection. However, it cannot detect selection signatures resulting from complete or nearly complete fixation of a beneficial allele in the population, and therefore cannot detect a significant fraction of variants under positive selection (Qanbari et al., 2011). For this reason, Tang et al. (2007) reported a new method involving comparison of EHH at the same site, but between populations, i.e., an approach based on the genetic diversity among divergent populations. These statistics are called site-specific EHH ($EHHS$); the area under the $EHHS$ curve (iES); and the standardized ratio of iES between two populations (Hellmann et al., 2003), which reflect haplotype variation among populations. The search for selection signatures from EHH statistical derivatives has been performed in several species such as cattle (Qanbari et al., 2011), poultry (Li et al., 2012; Zhang et al., 2012), swine (Ai et al., 2013) and humans (Sabeti et al., 2007).

Index of population differentiation

The F_{ST} (Wright, 1951) is a statistical measure of genetic variation due to differences in allele frequencies between and within populations (Holsinger and Weir, 2009). The F_{ST} statistic has been one of the most widely used methods for detecting genomic regions that have been under selection (Gianola et al., 2010; Qanbari et al., 2011). The F_{ST} for a locus that has been selected in one population but not another will be higher than in other loci not affected by selection, where genetic diversity is mainly caused by genetic drift (Holsinger and Weir, 2009). Genetic drift affects all loci in the genome similarly; however, loci under selection often behave differently and therefore may present atypical patterns of variation. These atypical patterns can be determined by genotyping, for example, a large number of single nucleotide polymorphisms (SNP) throughout the whole genome, where loci influenced by selection may be identified by deviations from the empirical distribution of F_{ST} statistic (Cavalli-Sforza, 1966; Akey et al., 2002). That is, relative to a neutral model, outliers with value below a certain level suggest the effect of balanced selection, while outliers with values above a certain level are indicative of directional selection.

Various estimates of the F_{ST} statistic have been developed and applied in a number of studies to search for selection signatures (Akey et al., 2002; Hayes et al., 2009; Amaral et al., 2011). However, although the outlier approach may be effective in identifying genes

under selection, it poses several challenges, such as susceptibility to genotyping errors, population stratification, and false positives, as well as variations in mutation rate and low sensitivity (Narum and Hess, 2011). It is also well known that the outlier detection methods have limited power to detect disruptive selection (Beaumont and Balding, 2004) and weak forms of divergent selection (Wright and Gaut, 2005).

TESTS BASED ON DIFFERENCES BETWEEN SPECIES

The statistical methodology to detect selection signatures by comparing information between species relies on the fact that genomic substitutions in coding regions are present in two forms: non-synonymous mutations (d_n), which can lead to the replacement of amino acids in the resulting proteins, and synonymous mutations (d_s), which do not cause amino acid substitution because of the redundancy of the genetic code (Nielsen, 2005; Biswas and Akey, 2006).

The d_n/d_s ratio provides information about evolutionary forces acting upon a particular gene. For example, at loci under neutrality, the d_n/d_s ratio will be equal to 1. Genes subject to functional limitations, such that a non-synonymous substitution is detrimental, will tend to be eliminated from the population by negative selection; therefore, $d_n/d_s < 1$. Conversely, an excess of non-synonymous mutations over synonymous mutations ($d_n/d_s > 1$) provides evidence for the action of positive selection in favor of non-synonymous substitution, which could provide a comparative advantage at the protein level (Nielsen et al., 2005a).

Based on these concepts, several studies have detected selection in many genes and organisms, such as genes related to immune response (Endo et al., 1996; Hughes, 1997; Sawyer et al., 2004), viral receptor genes (Fitch et al., 1997; Nielsen and Yang, 1998; Bush et al., 1999), genes associated with fertility (Swanson et al., 2001, 2003), and genes involved in sensory perception and smell in humans (Gilad et al., 2000).

TESTS THAT USE INFORMATION WITHIN AND BETWEEN SPECIES

The neutral theory of molecular evolution indicates that genomic regions that evolve rapidly and, thus, have high divergence between species, will also show high levels of polymorphisms within species. The Hudson–Kreitman–Aguade (HKA) test compares the level of polymorphisms within each species and observed divergence between related species for two or more loci. The test can determine if it is likely that the observed difference is due to neutral or adaptive evolution (Hudson et al., 1987). The HKA test is the precursor to the McDonald–Kreitman test (Howe et al., 2013), which compares synonymous (P_S) and non-synonymous (P_N) mutations at a specific locus that are polymorphic within a species and synonymous (D_S) and non-synonymous (D_N) mutations that are fixed between species. Under neutrality, the ratios between P_N/P_S and D_N/D_S should be the same, while positive selection leads to increased divergence of synonymous substitutions ($D_N/D_S > P_N/P_S$; McDonald and Kreitman, 1991).

GENOMIC RESOURCES IN FISH

In recent decades, the development of DNA markers has greatly contributed to the study of animal genetics. DNA markers allow us

to observe and exploit variation across the genome of an individual (Liu and Cordes, 2004; Tier, 2010).

In fish, a wide range of DNA markers have been used, including amplified fragment length polymorphisms (AFLP), random amplified polymorphic DNA (RAPD), sequence tagged sites (STS), variable number of tandem repeats (VNTR), microsatellites or simple sequence repeats (SSR), SNP, and expressed sequence tags (EST; Liu, 2007). Currently, with the development of high-throughput sequencing technologies many gigabases of nucleotide sequences can be generated in a short period of time, and many SNP and other polymorphisms can be detected using bioinformatics methods (Liu, 2011). These techniques provide an affordable and reliable scale of DNA sequencing in several organisms (Mardis, 2008). They are extensively used in *de novo* sequencing, quantification of gene expression by RNA-seq (“RNA sequencing”; Wang et al., 2009), massive identification of SNP markers using RAD-sequencing (“restriction site associated DNA sequencing”; Rowe et al., 2011), and population genomics studies (Hohenlohe et al., 2010; De Wit et al., 2012).

Although teleost fish are the largest group of vertebrates (about 27,000 species), they are underrepresented in genome sequencing projects (Spaink et al., 2013). **Table 1** shows some of the species that have undergone genome sequencing projects to date.

Extracted and modified from Spaink et al. (2013). The terms scaffolds or contigs indicate that the genome of the species has been partially sequenced, and the term chromosome indicates that sequencing has been anchored to the existing physical map of the species.

SELECTION SIGNATURES IN FISH

In fish, studies aimed at detecting selection signatures are performed mainly in the context of molecular ecology disciplines. Most of them have been limited to a low level of resolution and restricted to specific genomic regions.

MODEL FISH SPECIES

Using SNP markers from ESTs, loci with outlier F_{ST} values were identified in wild populations of zebrafish (*Danio rerio*), suggesting directional selection in genes associated with energy metabolism, homeostasis regulation, and signal transduction, which could be associated with local adaptation among different populations. Further, evidence was found to suggest balanced selection of the gene encoding the receptor for the NS1A influenza virus (Whiteley et al., 2011). In the same study, outlier F_{ST} values were found for loci in laboratory strains related to oxidoreductase activity, chromatin condensation, immune response, and induction of apoptosis, among other processes, which could be associated with the domestication process of cultured strains (Whiteley et al., 2011).

CICHLIDS

Keller et al. (2013) detected outlier SNP patterns between five cichlid species from the Lake Victoria area in East Africa, identifying signatures of divergent selection between the two genera that include these species. These selection signals were associated with male color, depth distribution, feeding patterns, and

morphological traits that distinguish the genera. Moreover, evidence has been found to suggest selection in the homeobox genes (*dlx*) involved in the development of the nervous system, the craniofacial skeleton, and the formation of connective tissue and appendages (Diepeveen et al., 2013).

SALMONIDS

In lake whitefish (*Coregonus clupeaformis*), a fish of the salmon family distributed along northern Alaska and all of Canada, 24 loci were identified that revealed selection signatures associated with QTL of certain adaptive traits such as natatorium behavior, growth rate, morphology, and reproduction characters (Rogers and Bernatchez, 2007).

In Atlantic salmon, Vasemägi et al. (2012) used microsatellite markers and SNP to locate 10 genomic regions showing signatures of directional selection related to characteristics such as growth rate and morphology. Martinez et al. (2013) found strong evidence of selection in a microsatellite marker on chromosome 3, which harbored QTL for body weight. Furthermore, there is evidence that genes associated with immune response have been subject to greater selection pressure compared with other regions of the genome (Tonteri et al., 2010; Portnoy et al., 2014). Other studies in genera *Oncorhynchus*, *Salmo*, and *Salvelinus* have revealed signatures of balanced selection for genes of the major histocompatibility complex IIB (Aguilar and Garza, 2007; Limborg et al., 2012). In brown trout, analysis with markers linked to genes related to the immune response showed evidence of having been subjected to selection (Jensen et al., 2008). Finally, evidence was found in both brown trout and sockeye salmon of disruptive selection at two loci within the major histocompatibility complex IIB (Hansen et al., 2010; Gomez-Uchida et al., 2011; Meier et al., 2011).

OTHER FAMILIES

In guppies (*Poecilia reticulata*), outlier F_{ST} values suggest that between 3.5 and 6.5% of SNP markers are under directional selection. Some of these loci are near QTL associated with ornamental traits, and they are also in EST (Willing et al., 2010).

In Atlantic cod in 1960, Sick detected evidence of selection in the locus encoding Hemoglobin (*Hb I*) and the *Pan I* locus, which encodes a protein related to the neuroendocrine system and has recently been associated with vesicle transport in adipocytes (Pogson, 2001). In the same species, Moen et al. (2008) identified 29 SNP with outlier F_{ST} , suggesting that these loci are or have been under selection. These loci were found in genes involved in muscle contraction, immune response, and production of ribosomal proteins. Moreover, Nielsen (Nielsen et al., 2009) found evidence of directional selection for local adaptation to various environmental conditions, such as loci with outlier F_{ST} values associated with genes involved in the production of proteins for thermal shock (*Hsp90*), determination of sexual behavior (*Aromatasa*), and formation of photoreceptor cells for perception of light (*rhodopsin*).

In stickleback (*Gasterosteus aculeatus*), a fish of the *Gasterosteidae* family, studies using microsatellite markers to assess genetic diversity among marine and freshwater populations revealed evidence of directional selection that might be associated with

Table 1 | Sequenced genomes available in fish as of 2013.

Order	Family	Species	Common name	Size estimation (Mb)	Chromosomes	Scaffolds	Status	Reference
Anguilliformes	Anguillidae	<i>Anguilla anguilla</i>	European eel	1,000	–	–	Scaffolds or contigs	Henkel et al. (2012a)
		<i>Anguilla japonica</i>	Japanese eel	1,150	–	323,776	Scaffolds or contigs	Henkel et al. (2012b)
Beloniformes	Adrianchthyidae	<i>Oryzias latipes</i>	Japanese rice fish	585.33	24	82,496	Chromosomes	Kawahara et al. (2007)
Characiformes	Characidae	<i>Astyanax mexicanus</i>	Mexican Tetra	964.31	–	10,735	Scaffolds or contigs	Hinaux et al. (2013)
Coelacanthiformes	Latimeriidae	<i>Latimeria chalumnae</i>	African coelacanth	2,183.72	–	22,818	Scaffolds or contigs	Amemiya et al. (2013)
Cypriniformes	Cyprinidae	<i>Danio rerio</i>	Zebra fish	1,412.47	25	4,560	Chromosomes	Hinaux et al. (2013), Howe et al. (2013)
		<i>Cyprinus carpio</i>	Common carp	875	25	–	Chromosomes	Xu et al. (2014)
Cyprinodontiformes	Nothobranchiidae	<i>Nothobranchius furzeri</i>	Turquoise killifish	1,500	–	5,299	Scaffolds or contigs	Reichwald et al. (2009)
		<i>Nothobranchius kuhntae</i>	Beira killifish	–	–	5,934	Scaffolds or contigs	
Gadiformes	Poeciliidae	<i>Xiphophorus maculatus</i>	Southern platyfish	652.84	–	20,640	Scaffolds or contigs	Schartl et al. (2013)
		<i>Gadus morhua</i>	Atlantic cod	608.29	–	427,427	Scaffolds or contigs	Star et al. (2011)
Gasterosteiformes	Gasterosteidae	<i>Gasterosteus aculeatus</i>	Three-spined stickleback	446.62	–	–	Scaffolds or contigs	
Lepisosteiformes	Lepisosteidae	<i>Lepisosteus oculatus</i>	Spotted gar	945.86	29	2,105	Chromosomes	
Perciformes	Cichlidae	<i>Oreochromis niloticus</i>	Nile tilapia	816.12	–	5,901	Scaffolds or contigs	
		<i>Pundamilia nyererei</i>	Python island	698.8	–	7,236	Scaffolds or contigs	
		<i>Rhamphochromis esox</i>	–	1,100	–	55,751	Scaffolds or contigs	Loh et al. (2008)
		<i>Mchenga conophoros</i>	–	1,100	–	61,923	Scaffolds or contigs	Loh et al. (2008)
		<i>Melanochromis auratus</i>	Golden mbuna	1,100	–	63,297	Scaffolds or contigs	Loh et al. (2008)
		<i>Neolamprologus brichardi</i>	Princess Burundi	685.96	–	9,098	Scaffolds or contigs	
		<i>Haplochromis burtoni</i>	–	698.98	–	8,001	Scaffolds or contigs	
		<i>Labotropheus fuelleborni</i>	Blue mbuna	1,100	–	58,245	Scaffolds or contigs	Loh et al. (2008)
		<i>Maylandia zebra</i>	Zebra mbuna	1,100	–	65,094	Scaffolds or contigs	Loh et al. (2008)
Perciformes	Moronidae	<i>Dicentrarchus labrax</i>	European sea bass	98.25	–	–	Scaffolds or contigs	
Pleuronectiformes	Cynoglossidae	<i>Cynoglossus semilaevis</i>	Flatfish	477	21	80,677	Chromosomes	Chen et al. (2014)
Salmoniformes	Salmonidae	<i>Oncorhynchus mykiss</i>	Rainbow trout	1,877.5	29	79,941	Chromosomes	Berthelot et al. (2014)
		<i>Salmo salar</i>	Atlantic salmon	2,435.31	29	843,055	Chromosomes	Davidson et al. (2010)
Tetraodontiformes	Tetraodontidae	<i>Takifugu flavidus</i>	Sansai fugu	314.95	–	34,332	Scaffolds or contigs	
		<i>Takifugu rubripes</i>	Pufferfish	281.57	22	7091	Chromosomes	Aparicio et al. (2002)
		<i>Tetraodon nigroviridis</i>	Green spotted puffer	308.45	–	–	Scaffolds or contigs	

adaptation of certain populations to freshwater environments (Mäkinen et al., 2008).

FUTURE DIRECTIONS

The search for and detection of genomic signatures produced by selection has provided valuable information that contributes to the understanding of evolutionary forces affecting the genome and gene functions that control phenotypes of biological and economic interest (Nielsen, 2005; Nielsen et al., 2007).

Some fish species provide the great advantage of simultaneous availability as both a wild and a cultivated population. Additionally, these species have unique characteristics in terms of population structure and intra-specific adaptive divergence, mainly due to the diversity of environmental conditions that fish populations inhabit, resulting in populations that exhibit characteristics of strong local adaptation. Comparative studies among these populations would provide benefits in terms of elucidating the effects of selective processes and recent domestication events, which could improve the understanding regarding the impact of the interaction between domesticated and wild populations, the identification of genetic factors involved in economically important traits for aquaculture and unravelling the actual phenotypic variation within and between fish populations.

In domesticated species, the main motivation behind the search for selection signatures lies in the possibility of finding genes or genomic regions associated with traits of economic interest. The development of next-generation sequencing technologies and high-throughput genotyping has made it possible to investigate the effect of selective pressures on genome variation in several domesticated species. In cattle and sheep, researchers have detected selection signatures associated with carcass yield traits, tail fat deposition, dairy traits (Moradi et al., 2012; Rothhammer et al., 2013), reproductive traits (Gautier and Naves, 2011; Qanbari et al., 2011), immune response (Gautier and Naves, 2011), coat color, and horn development (Druet et al., 2013), among other characters of interest. Also, in swine, selection signatures have been identified in genomic regions associated with traits such as coat color, ear morphology, reproductive characteristics, and fat deposition (Wilkinson et al., 2013). In chickens, researchers have identified selection signatures associated with eggshell hardness and immune system characteristics (Qanbari et al., 2012), to mention just a few examples. These studies may provide a basis for conducting similar research that allows for investigation of the genomic regions affected by the processes of domestication and natural or artificial selection in fish populations, allowing for discovery of new genes that underlie phenotypic traits of interest and understanding processes relevant for conservation purposes.

There is currently little genomic information for fish species as compared to humans or domesticated animals. This is one of the reasons why selection signatures studies have been conducted in only a few species and generally limited to a low level of genomic coverage (Vasemägi et al., 2005, 2012). However, recent advances in genomic technologies, including high quality reference genome sequences, construction of genetic maps, and development of high-density SNP arrays are paving the way for systematic study of genetic variation in these species.

The development and application of next-generation sequencing approaches will represent a powerful strategy to improve the resolution and accuracy when detecting regions under selection in several species. This may lead to determination of the causative genetic factors involved in several biological aspects of aquaculture species. However, the application of these results in the aquaculture development requires further studies aiming at determining effective and practical applications of this technology. Candidate disciplines to be benefited from the discovery of selected regions using next-generation sequencing are, for example, genetic improvement, vaccine and pharmaceutical development and fish nutrition.

CONCLUSION

The development of genomic methodologies has contributed greatly to the study of genetic variation between and within species. High-resolution studies at the level of the whole genome can identify selection signatures explaining phenotypic variation between and within populations, and therefore potentially identify genetic variants underlying characteristics of biological and economic interest.

Although the application and utility of these techniques in aquaculture species has been limited by a lack of genomic information, there is a great potential for conducting such studies, especially in species for which there are genome sequencing projects and high-density molecular markers platforms availability.

ACKNOWLEDGMENTS

This work was supported by the Programa Formación Capital Humano Avanzado from CONICYT, Government of Chile, and partially funded by grants from CORFO (11IEI-12843 and 12PIE-17669), Government of Chile.

REFERENCES

- Aguilar, A., and Garza, J. (2007). Patterns of historical balancing selection on the salmonid major histocompatibility complex class II β gene. *J. Mol. Evol.* 65, 34–43. doi: 10.1007/s00239-006-0222-8
- Ai, H., Huang, L., and Ren, J. (2013). Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide SNP markers. *PLoS ONE* 8:e56001. doi: 10.1371/journal.pone.0056001
- Akey, J., Zhang, G., Zhang, K., Jin, L., and Shriver, M. (2002). Interrogating a high density SNP map for signatures of natural selection. *Genome Res.* 12, 1805–1814. doi: 10.1101/gr.631202
- Amaral, A. J., Ferretti, L., Megens, H.-J., Crooijmans, R. P. M. A., Nie, H., Ramos-Onsins, S. E., et al. (2011). Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLoS ONE* 6:e14782. doi: 10.1371/journal.pone.0014782
- Amemiya, C. T., Alfoldi, J., Lee, A. P., Fan, S., Philippe, H., Maccallum, I., et al. (2013). The African coelacanth genome provides insights into tetrapod evolution. *Nature* 496, 311–316. doi: 10.1038/nature12027
- Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.-M., Dehal, P., et al. (2002). Whole-genome shotgun assembly and analysis of the genome of *fugu rubripes*. *Science* 297, 1301–1310. doi: 10.1126/science.1072104
- Aquadro, C. F., Bauer Dumont, V., and Reed, F. A. (2001). Genome-wide variation in the human and fruitfly: a comparison. *Curr. Opin. Genet. Dev.* 11, 627–634. doi: 10.1016/S0959-437X(00)00245-248
- Beaumont, M. A., and Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13, 969–980. doi: 10.1111/j.1365-294X.2004.02125.x

- Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., et al. (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* 5:3657. doi: 10.1038/ncomms4657
- Biswas, S., and Akey, J. (2006). Genomic insights into positive selection. *Trends Genet.* 22, 437–436. doi: 10.1016/j.tig.2006.06.005
- Black, W. C. IV, Baer, C. F., Antolin, M. F., and Duteau, N. M. (2001). POPULATION GENOMICS: genome-wide sampling of insect populations. *Annu. Rev. Entomol.* 46, 441–469. doi: 10.1146/annurev.ento.46.1.441
- Bush, R. M., Fitch, W. M., Bender, C. A., and Cox, N. J. (1999). Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol. Biol. Evol.* 16, 1457–1465. doi: 10.1093/oxfordjournals.molbev.a026057
- Cavalli-Sforza, L. (1966). Population structure and human evolution. *Proc. R. Soc. Lond. Ser. B-Biol. Sci.* 164, 362–379. doi: 10.1098/rspb.1966.0038
- Charlesworth, B. (1993). The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet. Res.* 63, 213–227. doi: 10.1017/S0016672300032365
- Chen, S., Zhang, G., Shao, C., Huang, Q., Liu, G., Zhang, P., 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. doi: 10.1038/ng.2890
- Cutter, A. D., and Paysseur, B. A. (2013). Genomic signatures of selection at linked sites: unifying the disparity among species. *Nat. Rev. Genet.* 14, 262–274. doi: 10.1038/nrg3425
- Davidson, W. S., Koop, B. F., Jones, S. J., Iturra, P., Vidal, R., Maass, A., et al. (2010). Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol.* 11:403. doi: 10.1186/gb-2010-11-9-403
- De Wit, P., Pespeni, M. H., Ladner, J. T., Barshis, D. J., Seneca, F., Jaris, H., et al. (2012). The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Mol. Ecol. Resour.* 12, 1058–1067. doi: 10.1111/1755-0998.12003
- Diepeveen, E., Kim, F., and Salzburger, W. (2013). Sequence analyses of the distal-less homeobox gene family in East African cichlid fishes reveal signatures of positive selection. *BMC Evol. Biol.* 13:153. doi: 10.1186/1471-2148-13-153
- Dominik, S., Henshall, J. M., Kube, P. D., King, H., Lien, S., Kent, M. P., 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(Suppl. 1), S56–S61. doi: 10.1016/j.aquaculture.2010.05.038
- Donaldson, L. R., and Olson, P. R. (1957). Development of rainbow trout brood stock by selective breeding. *Trans. Am. Fish. Soc.* 85, 93–101. doi: 10.1577/1548-8659(1955)85[93:DORTBS]2.0.CO;2
- Druet, T., Pérez-Pardal, L., Charlier, C., and Gautier, M. (2013). Identification of large selective sweeps associated with major genes in cattle. *Anim. Genet.* 44, 758–762. doi: 10.1111/age.12073
- Eknath, A. E., Tayamen, M. M., Palada-De Vera, M. S., Danting, J. C., Reyes, R. A., Dionisio, E. E., et al. (1993). Genetic improvement of farmed tilapia: the growth performance of eight strains of *Oreochromis niloticus* tested in different farm environments. *Aquaculture* 111, 171–188. doi: 10.1016/0044-8486(93)90035-W
- Endo, T., Ikeo, K., and Gjobori, T. (1996). Large-scale search for genes on which positive selection may operate. *Mol. Biol. Evol.* 13, 685–690. doi: 10.1093/oxfordjournals.molbev.a025629
- Fay, J. C., and Wu, C.-I. (2000). Hitchhiking under positive darwinian selection. *Genetics* 155, 1405–1413.
- Fitch, W. M., Bush, R. M., Bender, C. A., and Cox, N. J. (1997). Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc. Natl. Acad. Sci. U.S.A.* 94, 7712–7718. doi: 10.1073/pnas.94.15.7712
- Flori, L., Fritz, S., Jaffre Zic, F., Boussaha, M., Gut, I., Heath, S., et al. (2009). The genome response to artificial selection: a case study in dairy cattle. *PLoS ONE* 4:e6595. doi: 10.1371/journal.pone.0006595
- Fu, Y. X., and Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* 133, 693–709.
- Futuyma, D. J. (1998). *Evolutionary Biology*. Sunderland, MA: Sinauer Associates Inc.
- Gautier, M., and Naves, M. (2011). Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Mol. Ecol.* 20, 3128–3143. doi: 10.1111/j.1365-294X.2011.05163.x
- Gianola, D., Simianer, H., and Qanbari, S. (2010). A two-step method for detecting selection signatures using genetic markers. *Genet. Res. (Camb.)* 92, 141–155. doi: 10.1017/S0016672310000121
- Gilad, Y., Segre, D., Skorecki, K., Nachman, M. W., Lancet, D., and Sharon, D. (2000). Dichotomy of single-nucleotide polymorphism haplotypes in olfactory receptor genes and pseudogenes. *Nat. Genet.* 26, 221–224. doi: 10.1038/79957
- Gjedrem, T. (2005). *Selection and Breeding Programs in Aquaculture*. Dordrecht: Springer. doi: 10.1007/1-4020-3342-7
- Gjedrem, T. (2012). Genetic improvement for the development of efficient global aquaculture: A personal opinion review. *Aquaculture* 344–349, 12–22. doi: 10.1016/j.aquaculture.2012.03.003
- Gjedrem, T., Robinson, N., and Rye, M. (2012). The importance of selective breeding in aquaculture to meet future demands for animal protein: a review. *Aquaculture* 350–353, 117–129. doi: 10.1016/j.aquaculture.2012.04.008
- Glover, K. A., Dahle, G., and Jørstad, K. E. (2011). Genetic identification of farmed and wild Atlantic cod, *Gadus morhua*, in coastal Norway. *ICES J. Mar. Sci.* 68, 901–910. doi: 10.1093/icesjms/fsr048
- Glover, K. A., Otterå, H., Olsen, R. E., Slinde, E., Taranger, G. L., and Skaala, Ø. (2009). A comparison of farmed, wild and hybrid Atlantic salmon (*Salmo salar* L.) reared under farming conditions. *Aquaculture* 286, 203–210. doi: 10.1016/j.aquaculture.2008.09.023
- Glover, K. A., Svåsand, T., Olesen, I., and Rye, M. (2007). *Atlantic Halibut – Hippoglossus Hippoglossus: Genetic Impact of Aquaculture Activities on Native Populations*. Genimpact Final Scientific Report (EU contract n. RICA-CT-2005-022802), 176. Available at: <http://genimpact.imr.no>
- Gomez-Uchida, D., Seeb, J., Smith, M., Habicht, C., Quinn, T., and Seeb, L. (2011). Single nucleotide polymorphisms unravel hierarchical divergence and signatures of selection among Alaskan sockeye salmon (*Oncorhynchus nerka*) populations. *BMC Evol. Biol.* 11:48. doi: 10.1186/1471-2148-11-48
- Gu, J., Orr, N., Park, S., Katz, L., Sulimova, G., Machugh, D., et al. (2009). A genome scan for positive selection in thoroughbred horses. *PLoS ONE* 4:e5767. doi: 10.1371/journal.pone.0005767
- Hansen, M. M., Meier, K., and Mensberg, K.-L. D. (2010). Identifying footprints of selection in stocked brown trout populations: a spatio-temporal approach. *Mol. Ecol.* 19, 1787–1800. doi: 10.1111/j.1365-294X.2010.04615.x
- Hayes, B., Chamberlain, A., Maceachern, S., Savin, K., Mcpartlan, H., Macleod, I., et al. (2009). A genome map of divergent artificial selection between Bos taurus dairy cattle and Bos taurus beef cattle. *Anim. Genet.* 40, 176–184. doi: 10.1111/j.1365-2052.2008.01815.x
- Hayes, B., Lien, S., Nilsen, H., Olsen, H., Berg, P., Maceachern, S., et al. (2008). The origin of selection signatures on bovine chromosome six. *Anim. Genet.* 39, 105–111. doi: 10.1111/j.1365-2052.2007.01683.x
- Hellmann, I., Ebersberger, I., Ptak, S. E., Paabo, S., and Przeworski, M. (2003). A neutral explanation for the correlation of diversity with recombination rates in humans. *Am. J. Hum. Genet.* 72, 1527–1535. doi: 10.1086/375657
- Henkel, C. V., Burgerhout, E., De Wijze, D. L., Dirks, R. P., Minegishi, Y., Jansen, H. J., et al. (2012a). Primitive duplicate hox clusters in the European Eel's genome. *PLoS ONE* 7:e32231. doi: 10.1371/journal.pone.0032231
- Henkel, C. V., Dirks, R. P., De Wijze, D. L., Minegishi, Y., Aoyama, J., Jansen, H. J., et al. (2012b). First draft genome sequence of the Japanese eel, *Anguilla japonica*. *Gene* 511, 195–201. doi: 10.1016/j.gene.2012.09.064
- Hermisson, J., and Pennings, P. S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169, 2335–2352. doi: 10.1534/genetics.104.036947
- Hinaux, H., Poulain, J., Da Silva, C., Noirot, C., Jeffery, W. R., Casane, D., et al. (2013). De novo sequencing of astyanax mexicanus surface fish and pachón cavefish transcriptomes reveals enrichment of mutations in cavefish putative eye genes. *PLoS ONE* 8:e53553. doi: 10.1371/journal.pone.0053553
- Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., and Cresko, W. A. (2010). Population genomics of parallel adaptation in three-spine stickleback using sequenced RAD tags. *PLoS Genetics* 6:e1000862. doi: 10.1371/journal.pgen.1000862
- Holsinger, K. E., and Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat. Rev. Genet.* 10, 639–650. doi: 10.1038/nrg2611
- Houston, R. D., Haley, C. S., Hamilton, A., Guy, D. R., Tinch, A. E., Taggart, J. B., et al. (2008). Major quantitative trait loci affect resistance to infectious

- pancreatic necrosis in Atlantic salmon (*Salmo salar*). *Genetics* 178, 1109–1115. doi: 10.1534/genetics.107.082974
- Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498–503. doi: 10.1038/nature12111
- Hudson, R. R., Kreitman, M., and Aguade, M. (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* 116, 153–159.
- Hughes, A. L. (1997). Rapid evolution of immunoglobulin superfamily C2 domains expressed in immune system cells. *Mol. Biol. Evol.* 14, 1–5. doi: 10.1093/oxfordjournals.molbev.a025694
- Hurst, L. D. (2009). Genetics and the understanding of selection. *Nat. Rev. Genet.* 10, 83–93. doi: 10.1038/nrg2506
- Ilyassov, Y. I. (1987). *Selection, Hybridization, and Genetic Engineering in Aquaculture*. Berlin: H. Heenemann GmbH & Company.
- Jensen, L. F., Hansen, M. M., Mensberg, K. L. D., and Loeschcke, V. (2008). Spatially and temporally fluctuating selection at non-MHC immune genes: evidence from TAP polymorphism in populations of brown trout (*Salmo trutta*, L.). *Heredity* 100, 79–91. doi: 10.1038/sj.hdy.6801067
- Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., et al. (2007). The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714–719. doi: 10.1038/nature05846
- Keller, I., Wagner, C. E., Greuter, L., Mwaiko, S., Selz, O. M., Sivasundar, A., 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. doi: 10.1111/mec.12083
- Kim, J., Withler, R., Ritland, C., and Cheng, K. (2004). “Genetic variation within and between domesticated chinook salmon, *Oncorhynchus tshawytscha* strains and their progenitor populations,” in *Genetics of Subpolar Fish and Invertebrates*, eds A. Gharrett, R. Gustafson, J. Nielsen, J. Seeb, L. Seeb, W. Smoker, et al. (Dordrecht: Springer), 371–378.
- Kim, Y., and Nielsen, R. (2004). Linkage disequilibrium as a signature of selective sweeps. *Genetics* 167, 1513–1524. doi: 10.1534/genetics.103.025387
- Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press. doi: 10.1038/nrg2473
- Li, D. F., Liu, W. B., Liu, J. F., Yi, G. Q., Lian, L., Qu, L. J., et al. (2012). Whole-genome scan for signatures of recent selection reveals loci associated with important traits in White Leghorn chickens. *Poult. Sci.* 91, 1804–1812. doi: 10.3382/ps.2012-02275
- Limborg, M. T., Blankenship, S. M., Young, S. F., Utter, F. M., Seeb, L. W., Hansen, M. H. H., et al. (2012). Signatures of natural selection among lineages and habitats in *Oncorhynchus mykiss*. *Ecol. Evol.* 2, 1–18. doi: 10.1002/ece3.59
- Liu, Z. E. (2011). *Next Generation Sequencing and Whole Genome Selection in Aquaculture*. Des Moines, IA: Blackwell Publishing Ltd. doi: 10.1002/9780470958964
- Liu, Z. J. (2007). *Aquaculture Genome Technologies*. Oxford, UK: Blackwell Publishing Ltd. doi: 10.1002/9780470277560
- Liu, Z. J., and Cordes, J. F. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238, 1–37. doi: 10.1016/j.aquaculture.2004.05.027
- Loh, Y.-H., Katz, L., Mims, M., Kocher, T., Yi, S., and Streebman, J. T. (2008). Comparative analysis reveals signatures of differentiation amid genomic polymorphism in Lake Malawi cichlids. *Genome Biol.* 9:R113. doi: 10.1186/gb-2008-9-7-r113
- Luikart, G., England, P. R., Tallmon, D., Jordan, S., and Taberlet, P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nat. Rev. Genet.* 4, 981–994. doi: 10.1038/nrg1226
- Mäkinen, H. S., Shikano, T., Cano, J. M., and Merilä, J. (2008). Hitchhiking mapping reveals a candidate genomic region for natural selection in three-spined stickleback chromosome VIII. *Genetics* 178, 453–465. doi: 10.1534/genetics.107.078782
- Mardis, E. (2008). Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* 9:15. doi: 10.1146/annurev.genom.9.081307.164359
- Martinez, V., Dettleff, P., Lopez, P., Fernandez, G., Jedlicki, A., Yañez, J. M., et al. (2013). Assessing footprints of selection in commercial Atlantic salmon populations using microsatellite data. *Anim. Genet.* 44, 223–226. doi: 10.1111/j.1365-2052.2012.02387.x
- McAndrew, B., and Napier, J. (2011). Application of genetics and genomics to aquaculture development: current and future directions. *J. Agric. Sci.* 149, 143–151. doi: 10.1017/S0021859610001152
- McDonald, J. H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351, 652–654. doi: 10.1038/351652a0
- Meier, K., Hansen, M. M., Bekkevold, D., Skaala, O., and Mensberg, K. L. D. (2011). An assessment of the spatial scale of local adaptation in brown trout (*Salmo trutta* L.): footprints of selection at microsatellite DNA loci. *Heredity* 106, 488–499. doi: 10.1038/hdy.2010.164
- Metcalfe, N. B., Valdimarsson, S. K., and Morgan, I. J. (2003). The relative roles of domestication, rearing environment, prior residence and body size in deciding territorial contests between hatchery and wild juvenile salmon. *J. Appl. Ecol.* 40, 535–544. doi: 10.1046/j.1365-2664.2003.00815.x
- Moen, T., Baranski, M., Sonesson, A. K., and Kjøglum, S. (2009). Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. *BMC Genomics* 10:368. doi: 10.1186/1471-2164-10-368
- Moen, T., Hayes, B., Nilsen, E., Delghandi, M., Fjalestad, K., Fevolden, S. -E., et al. (2008). Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *BMC Genet.* 9:18. doi: 10.1186/1471-2156-9-18
- Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K., and Mcewan, J. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genet.* 13:10. doi: 10.1186/1471-2156-13-10
- Mueller, J. C. (2004). Linkage disequilibrium for different scales and applications. *Brief. Bioinform.* 5, 355–364. doi: 10.1093/bib/5.4.355
- Narum, S. R., and Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Mol. Ecol. Resour.* 11, 184–194. doi: 10.1111/j.1755-0998.2011.02987.x
- Neira, R. (2010). “Breeding in aquaculture species: genetic improvement programs in developing countries,” in *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig, 8.
- Neira, R., Díaz, N. F., Gall, G. A. E., Gallardo, J. A., Lhorente, J. P., and Manteola, R. (2006). Genetic improvement in Coho salmon (*Oncorhynchus kisutch*). I: selection response and inbreeding depression on harvest weight. *Aquaculture* 257, 9–17. doi: 10.1016/j.aquaculture.2006.03.002
- Nielsen, E. E., Hemmer-Hansen, J., Larsen, P. F., and Bekkevold, D. (2009). Population genomics of marine fishes: identifying adaptive variation in space and time. *Mol. Ecol.* 18, 3128–3150. doi: 10.1111/j.1365-294X.2009.04272.x
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218. doi: 10.1146/annurev.genet.39.073003.112420
- Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., et al. (2005a). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* 3:e170. doi: 10.1371/journal.pbio.0030170
- Nielsen, R., Williamson, S., Kim, Y., Hubisz, M., Clark, A., and Bustamante, C. (2005b). Genomic scans for selective sweeps using SNP data. *Genome Res.* 15, 1566–1575. doi: 10.1101/gr.4252305
- Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C., and Clark, A. G. (2007). Recent and ongoing selection in the human genome. *Nat. Rev. Genet.* 8, 857–868. doi: 10.1038/nrg2187
- Nielsen, R., and Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148, 929–936.
- Norris, A. T., Bradley, D. G., and Cunningham, E. P. (1999). Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture* 180, 247–264. doi: 10.1016/S0044-8486(99)00212-214
- Owen, M. (2011). South Australian company completes bluefin tuna breeding program. *The Australian*.
- Pennings, P. S., and Hermisson, J. (2006). Soft sweeps II—molecular population genetics of adaptation from recurrent mutation or migration. *Mol. Biol. Evol.* 23, 1076–1084. doi: 10.1093/molbev/msj117
- Perfectti, F., Picó, F. X., and Gómez, J. M. (2009). La huella genética de la selección natural. *Ecosistemas* 18, 10–16.
- Pogson, G. H. (2001). Nucleotide polymorphism and natural selection at the panthophysin (Pan I) Locus in the Atlantic Cod, *Gadus morhua* (L.). *Genetics* 157, 317–330.
- Ponzone, R. W., Nguyen, N. H., Khaw, H. L., Hamzah, A., Bakar, K. R. A., and Yee, H. Y. (2011). Genetic improvement of Nile tilapia (*Oreochromis niloticus*) with special reference to the work conducted by the WorldFish Center with the GIFT strain. *Rev. Aquaculture* 3, 27–41. doi: 10.1111/j.1753-5131.2010.01041.x
- Portnoy, D. S., Hollenbeck, C. M., Vidal, R. R., and Gold, J. R. (2014). A comparison of neutral and immune genetic variation in Atlantic Salmon,

- Salmo salar* L. in Chilean aquaculture facilities. *PLoS ONE* 9:e99358. doi: 10.1371/journal.pone.0099358
- Price, E. O. (1999). Behavioral development in animals undergoing domestication. *Appl. Anim. Behav. Sci.* 65, 245–271. doi: 10.1016/S0168-1591(99)00087-8
- Pritchard, J. K., Pickrell, J. K., and Coop, G. (2010). The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* 20, R208–R215. doi: 10.1016/j.cub.2009.11.055
- Przeworski, M. (2002). The signature of positive selection at randomly chosen loci. *Genetics* 160, 1179–1189.
- Purdum, C. E. (1993). *Genetics and Fish Breeding*. London: Chapman and Hall, 277.
- Qanbari, S., Gianola, D., Hayes, B., Schenkel, F., Miller, S., Moore, S., et al. (2011). Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics* 12:318. doi: 10.1186/1471-2164-12-318
- Qanbari, S., Pimentel, E., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A., et al. (2010). A genome-wide scan for signatures of recent selection in Holstein cattle. *Anim. Genet.* 41, 377–389. doi: 10.1111/j.1365-2052.2009.02016.x
- Qanbari, S., Strom, T. M., Haberer, G., Weigend, S., Gheyas, A. A., Turner, F., et al. (2012). A high resolution genome-wide scan for significant selective sweeps: an application to pooled sequence data in laying chickens. *PLoS ONE* 7:e49525. doi: 10.1371/journal.pone.0049525
- Reichwald, K., Lauber, C., Nanda, I., Kirschner, J., Hartmann, N., Schories, S., et al. (2009). High tandem repeat content in the genome of the short-lived annual fish *Nothobranchius furzeri*: a new vertebrate model for aging research. *Genome Biol.* 10:R16. doi: 10.1186/gb-2009-10-2-r16
- Rogers, S., and Bernatchez, L. (2007). The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. *Salmonidae*) *Species Pairs*. *Mol. Biol. Evol.* 24, 1423–1438. doi: 10.1093/molbev/msm066
- Ross-Ibarra, J., Morrell, P., and Gaut, B. (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8641–8648. doi: 10.1073/pnas.0700643104
- Rothhammer, S., Seichter, D., Forster, M., and Medugorac, I. (2013). A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC Genomics* 14:908. doi: 10.1186/1471-2164-14-908
- Rowe, H. C., Renaut, S., and Guggisberg, A. (2011). RAD in the realm of next-generation sequencing technologies. *Mol. Ecol.* 20, 3499–3502. doi: 10.1111/j.1365-294X.2011.05197.x
- Rye, M., Gjerde, B., and Gjedrem, T. (2010). “Genetic improvement programs for aquaculture species in developed countries,” in *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig, 8.
- Sabeti, P., Reich, D., Higgins, J., Levine, H., Richter, D., Schaffner, S., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419, 832–837. doi: 10.1038/nature01140
- Sabeti, P., Schaffner, S., Fry, B., Lohmueller, J., Vailly, P., Shamovsky, O., et al. (2006). Positive natural selection in the human lineage. *Science* 3129, 1614–1620. doi: 10.1126/science.1124309
- Sabeti, P. C., Vailly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918. doi: 10.1038/nature06250
- Sawyer, S. L., Emerman, M., and Malik, H. S. (2004). Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol.* 2:e275. doi: 10.1371/journal.pbio.0020275
- Schaperclaus, W. (1962). *Trate de Pisciculture en Etang*. Paris: Vigot Freres, 208.
- Schartl, M., Walter, R. B., Shen, Y., Garcia, T., Catchen, J., Amores, A., et al. (2013). The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. *Nat. Genet.* 45, 567–572. doi: 10.1038/ng.2604
- Slatkin, M. (2008). Linkage disequilibrium [mdash] understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* 9, 477–485. doi: 10.1038/nrg2361
- Smith, J. M., and Haigh, J. (1974). The hitch-hiking effect of a favorable gene. *Genet. Res.* 23, 23–35. doi: 10.1017/S0016672300014634
- Spaink, H. P., Jansen, H. J., and Dirks, R. P. (2013). Advances in genomics of bony fish. *Brief. Funct. Genomics* 13, 144–156. doi: 10.1093/bfpg/elt046
- Star, B., Nederbragt, A. J., Jentoft, S., Grimholt, U., Malmstrom, M., Gregers, T. F., et al. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477, 207–210. doi: 10.1038/nature10342
- Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., Wolfner, M. F., and Aquadro, C. F. (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7375–7379. doi: 10.1073/pnas.131568198
- Swanson, W. J., Nielsen, R., and Yang, Q. (2003). Pervasive adaptive evolution in mammalian fertilization proteins. *Mol. Biol. Evol.* 20, 18–20. doi: 10.1093/oxfordjournals.molbev.a004233
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Tang, K., Thornton, K. R., and Stoneking, M. (2007). A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biol.* 5:e171. doi: 10.1371/journal.pbio.0050171
- Tang, K., Wong, L. P., Lee, E. J. D., Chong, S. S., and Lee, C. G. L. (2004). Genomic evidence for recent positive selection at the human MDR1 gene locus. *Hum. Mol. Genet.* 13, 783–797. doi: 10.1093/hmg/ddh099
- Tier, B. (2010). Genomic selection: promises and propriety. *J. Animal Breeding and Genetics* 127, 169–170. doi: 10.1111/j.1439-0388.2010.00876.x
- Tonteri, A., Vasemägi, A., Lumme, J., and Primmer, C. R. (2010). Beyond MHC: signals of elevated selection pressure on Atlantic salmon (*Salmo salar*) immune-relevant loci. *Mol. Ecol.* 19, 1273–1282. doi: 10.1111/j.1365-294X.2010.04573.x
- Trut, L., Oskina, I., and Kharlamova, A. (2009). Animal evolution during domestication: the domesticated fox as a model. *Bioessays* 31, 349–360. doi: 10.1002/bies.200800070
- Vandeputte, M., Dupont-Nivet, M., Haffray, P., Chavanne, H., Cenadelli, S., Parati, K., et al. (2009). Response to domestication and selection for growth in the European sea bass (*Dicentrarchus labrax*) in separate and mixed tanks. *Aquaculture* 286, 20–27. doi: 10.1016/j.aquaculture.2008.09.008
- Vasemägi, A., Nilsson, J., McGinnity, P., Cross, T., O’Reilly, P., Glebe, B., et al. (2012). Screen for footprints of selection during domestication/captive breeding of Atlantic Salmon. *Comp. Funct. Genomics* 2012:14. doi: 10.1155/2012/628204
- Vasemägi, A., Nilsson, J., and Primmer, C. R. (2005). Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Mol. Biol. Evol.* 22, 1067–1076. doi: 10.1093/molbev/msi093
- Vogl, C., and Clemente, F. (2012). The allele-frequency spectrum in a decoupled Moran model with mutation, drift, and directional selection, assuming small mutation rates. *Theor. Popul. Biol.* 81, 197–209. doi: 10.1016/j.tpb.2012.01.001
- Voight, B., Kudaravalli, S., Wen, X., and Pritchard, J. (2006). A map of recent positive selection in the human genome. *PLoS Biol.* 4:e72. doi: 10.1371/journal.pbio.0040072
- Walsh, E., Sabeti, P., Hutcheson, H., Fry, B., Schaffner, S., Bakker, P. W., et al. (2006). Searching for signals of evolutionary selection in 168 genes related to immune function. *Hum. Genet.* 119, 92–102. doi: 10.1007/s00439-005-0090-90
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg2484
- Whiteley, A. R., Bhat, A., Martins, E. P., Mayden, R. L., Arunachalam, M., Uusi-Heikkilä, S., et al. (2011). Population genomics of wild and laboratory zebrafish (*Danio rerio*). *Mol. Ecol.* 20, 4259–4276. doi: 10.1111/j.1365-294X.2011.05272.x
- Wilkinson, S., Lu, Z. H., Megens, H. -J., Archibald, A. L., Haley, C., Jackson, I. J., et al. (2013). Signatures of diversifying selection in European pig breeds. *PLoS Genet* 9:e1003453. doi: 10.1371/journal.pgen.1003453
- Williamson, S. H. (2007). Localizing recent adaptive evolution in the human genome. *PLoS Genet.* 3:e90. doi: 10.1371/journal.pgen.0030090
- Willing, E.-M., Bentzen, P., Van Oosterhout, C., Hoffmann, M., Cable, J., Breden, F., et al. (2010). Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. *Mol. Ecol.* 19, 968–984. doi: 10.1111/j.1365-294X.2010.04528.x
- Wright, S. (1951). The genetical structure of populations. *Ann. Eugen.* 15, 323–354. doi: 10.1111/j.1469-1809.1949.tb02451.x
- Wright, S. I., and Gaut, B. S. (2005). Molecular Population Genetics and the Search for Adaptive Evolution in Plants. *Mol. Biol. Evol.* 22, 506–519. doi: 10.1093/molbev/msi035
- Xu, P., Zhang, X., Wang, X., Li, J., Liu, G., Kuang, Y., et al. (2014). Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat. Genet.* 46, 1212–1219. doi: 10.1038/ng.3098

Zhang, H., Wang, S.-Z., Wang, Z.-P., Da, Y., Wang, N., Hu, X.-X., et al. (2012). A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. *BMC Genomics* 13:704. doi: 10.1186/1471-2164-13-704

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 October 2014; accepted: 15 December 2014; published online: 14 January 2015.

Citation: López ME, Neira R and Yáñez JM (2015) Applications in the search for genomic selection signatures in fish. *Front. Genet.* 5:458. doi: 10.3389/fgene.2014.00458
This article was submitted to *Livestock Genomics*, a section of the journal *Frontiers in Genetics*.

Copyright © 2015 López, Neira and Yáñez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.