Sun et al. BMC Genomics (2020) 21:674
https://doi.org/10.1186/s12864-020-07095-8

BMC Genomics

RESEARCH ARTICLE Open Access

Selection signatures of Fuzhong Buffalo based on whole-genome sequences

Ting Sun[1,2], Guang-yun Huang[1], Zi-hao Wang[1], Shao-hua Teng[1], Yan-hong Cao[1], Jun-li Sun[1], Quratulain Hanif[3,4], Ning-bo Chen[2], Chu-zhao Lei[2*] and Yu-ying Liao[1*]

Abstract

Background: Fuzhong buffalo, a native breed of Guangxi Zhuang Autonomous Region, is traditionally used as a draft animal to provide farm power in the rice cultivation. In addition, the Fuzhong buffalo also prepared for the bullfighting festival organized by the locals. The detection of the selective signatures in its genome can help in elucidating the selection mechanisms in its stamina and muscle development of a draft animal.

Results: In this study, we analyzed 27 whole genomes of buffalo (including 15 Fuzhong buffalo genomes and 12 published buffalo genomes from Upper Yangtze region). The ZHp, ZFst, π-Ratio, and XP-EHH statistics were used to identify the candidate signatures of positive selection in Fuzhong buffalo. Our results detected a set of candidate genes involving in the pathways and GO terms associated with the response to exercise (e.g., *ALDOA*, *STAT3*, *AKT2*, *EIF4E2*, *CACNA2D2*, *TCF4*, *CDH2*), immunity (e.g., *PTPN22*, *NKX2-3*, *PIK3R1*, *ITK*, *TMEM173*), nervous system (e.g., *PTPN21*, *ROBO1*, *HOMER1*, *MAGI2*, *SLC1A3*, *NRG3*, *SNAP47*, *CTNNA2*, *ADGRL3*). In addition, we also identified several genes related to production and growth traits (e.g., *PHLPP1*, *PRKN*, *MACF1*, *UCN3*, *RALGAPA1*, *PHKB*, *PKD1L*). Our results depicted several pathways, GO terms, and candidate genes to be associated with response to exercise, immunity, nervous system, and growth traits.

Conclusions: The selective sweep analysis of the Fuzhong buffalo demonstrated positive selection pressure on potential target genes involved in behavior, immunity, and growth traits, etc. Our findings provided a valuable resource for future research on buffalo breeding and an insight into the mechanisms of artificial selection.

Keywords: Buffalo, Resequencing, Selection signatures

Background

Buffalo has been an important livestock animal used as a source of food and draught power in tropical and subtropical regions. It can be divided into two types: river ($2n = 50$) and swamp buffalo ($2n = 48$) [1]. China is one of the countries with the largest population of swamp buffalo. A recent study showed that the swamp buffalo can be divided into two divergent groups: Southeast Asian buffalo (including buffaloes from Southeast Asia and Southwest China) and South China buffalo (including the buffalo from Upper Yangtze and Middle-Lower Yangtze) [2]. In particular, the buffalo distributed in Upper Yangtze exhibited weak gene flow from the Southeast Asian buffalo [2]. The Fuzhong (FZ) buffalo is a native livestock from Guangxi Zhuang Autonomous Region, one of the hottest and most humid region in the Southwest of China [3]. Except for its use as a draft animal like other swamp buffaloes, the FZ buffalo has also been used in bullfighting as a folk custom with long history in Guangxi Zhuang Autonomous Region [4]. The FZ buffaloes have been adaptive to the local climate with

* Correspondence: leichuzhao1118@126.com; liaoyuying@126.com
[2]College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi, China
[1]Animal Husbandry Institute of Guangxi Zhuang Autonomous Region, Guangxi Key Laboratory of Livestock Genetic Improvement, Nanning 530001, China
Full list of author information is available at the end of the article

© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

enhanced disease resistance and heat tolerance. To date, its genomic value and potential are yet to be discovered.

Recent developments of high-throughput sequencing and genotyping technologies allow the construction of detailed selection signature maps for human and various domestic animals. With the growing importance of buffalo industry, the sequencing and genotyping platforms have been used to investigate the genetic diversity, demographic history, and selective signatures, etc. Recently, a 90 K SNPChip for buffalo (Axiom® Buffalo 90 K Genotyping Array) has been developed, which includes about 90 K SNP loci covering the river buffalo genome [5]. Up to now, most studies mainly focused on the dairy buffalo. A previous investigation revealed clear divergence of two native Iranian buffalo breeds (Azeri and Khuzestani) and pointed to candidate genes for milk production, growth, immunity and nervous system, etc. [6]. Another study detected genes associated with the body size, immunity and coat color by analyzing the ROH of the Azeri and Khuzestani buffalo breeds using the same dataset [7]. The same chip has also been used for the GWAS analysis of water buffalo, revealing a set of candidate genes associated with milk traits (e.g., *MFSD14A*, *SLC35A3*, *PALMD*, *RGS22*) [8]. Furthermore, there were few studies performed at the whole genome level. Whitacre et al. detected 13 genes potentially involved in the development of the hindlimbs based on the whole genome sequence of buffalo [9]. A recent study described the genetic history and population structure of swamp and river buffalo by analyzing 121 whole genomes from 25 breeds with different geographical origins [2], intercepting the candidate genes associated with nervous system and muscle development in swamp buffalo, while genes related to heat-stress and immunity were detected in the river buffalo [2].

To date, several methods are employed to detect the selective sweeps in various livestock genomes. In the current study, the fixation index (Fst) was used to measure the genetic differentiation between populations [10]. The larger Fst value indicates the difference in the two populations. π-Ratio was used to identify the differences in nucleotide divergence between populations. Rubin et al. defined and applied a Z-score test for heterozygosity depression (ZHp) on chicken genome sequence, which basically expresses how much the expected heterozygosity in chromosome windows deviate from the average genome heterozygosity [11]. The extremely low ZHp scores indicate putative selective sweeps because of excess homozygosity. Moreover, the cross-population extended haplotype homozygosity test (XP-EHH) is used to detect ongoing or nearly fixed selective sweeps by comparing haplotypes between the two populations [12].

In the present study, 27 whole genomes of buffalo (including 15 newly sequenced FZ buffalo genomes and 12 published buffalo genomes from regions of Upper Yangtze (UY) were analyzed. The above methods (Fst, ZHp, π-Ratio, and XP-EHH) were used to search for the candidate signatures of positive selection for FZ buffalo.

Results

Genome resequencing of Fuzhong buffalo

A total of 15 FZ buffaloes were selected for genome resequencing, while 12 UY buffaloes genomic DNA sequences were obtained from previously published data [2] (Additional file 1: Table S1). The clean reads were aligned to the reference genome of buffalo (GCA_003121395.1) with an average alignment rate of 99.06%. The genome resequencing achieved an average depth of 10.54×. In total, 18,216,884 autosomal single nucleotide polymorphisms (SNPs) were identified, including 63,557 nonsynonymous and 128,420 synonymous coding variants (Additional file 2: Table S2). Principal components analysis (PCA) revealed the genetic separation between the FZ and UY buffaloes (Fig. 1).

Detection of selective sweeps in Fuzhong buffalo

Due to the genetic separation between FZ and UY buffalo, we performed selective sweep analysis to detect the selection signatures in FZ buffaloes. We used the Fst, Hp, π-Ratio, and XP-EHH tests to search for the genomic regions in FZ buffaloes. Fst is a descriptive statistic and a measure of population genetic differentiation between FZ and UY buffalo. We then calculated the Hp which estimated genetic polymorphism data and was used in population genetic analysis. The Fst and Hp values were normalized (ZFst, ZHp) using the Z-transformation method. Totally, we identified 1519 and 841 candidate genes from ZFst ($P < 0.005$, ZFst > 2.576) and ZHp ($P < 0.005$, ZHp < − 2.576), respectively (Fig. 2, Additional files 3-4: Table S3 and Table S4). In addition, the π-Ratio ($P < 0.005$, π-Ratio > 0.592) and XP-EHH ($P < 0.005$, XP-EHH > 2.180) analysis detected 826 and 675 candidate genes in FZ buffalo, respectively (Additional files 5-6: Table S5 and Table S6). The selective sweeps detected by at least two approaches were defined as the putative selective sweeps. Finally, a total of 599 genes were identified as the candidate genes for FZ buffalo. Among the 599 candidate genes, 21 genes detected by these four different statistics (Fig. 3a). Some of these genes were associated with production and growth traits (*PHLPP1*, *PRKN*, *MACF1*, *UCN3*, *RALGAPA1*, *PHKB*, *PKD1L2*) (Fig. 2). In addition, the selection signatures in UY buffalo were also detected, resulting in 1519, 949, 1065, and 669 candidate genes from ZFst, ZHp, π-Ratio and XP-EHH analysis, respectively. Totally, 707 overlapping genes were identified as the candidate genes for UY buffalo, while 58 candidate genes were shared between FZ and UY buffalo.

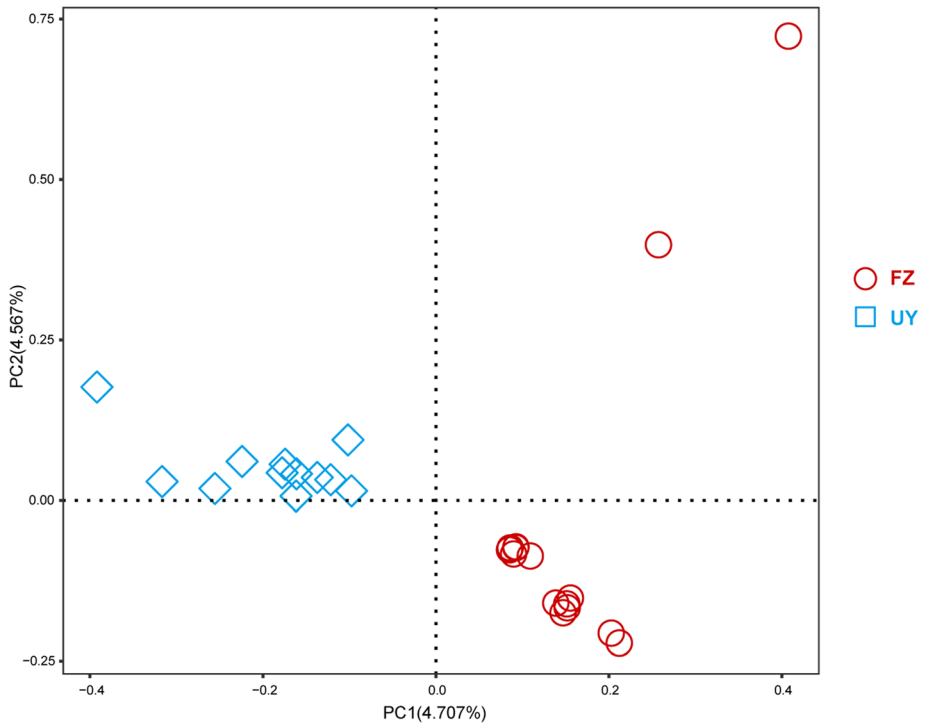

Fig. 1 Principal component analyses (PCA) of the FZ and UY buffaloes. FZ: Fuzhong buffalo; UY: buffaloes from Upper Yangtze region. Details of the sample information are explained in Additional file 1: Table S1

KEGG pathways and GO enrichments

To obtain a broad overview of the molecular functions of these identified candidate genes for FZ buffalo, KEGG pathways and GO enrichment analysis were performed (Fig. 3b, Additional files 7-8: Table S7 and Table S8). Some of the significant KEGG pathways (corrected *P*-value < 0.05) were associated with the cardiovascular system (arrhythmogenic right ventricular cardiomyopathy (ARVC), corrected *P*-value = 0.0102) and oxidative stress (HIF-1 signaling pathway, corrected *P*-value = 0.0030) (Fig. 3b, Additional files 7-8: Table S7 and Table S8). Several candidate genes (e.g., *ALDOA*, *STAT3*, *AKT2*, *EIF4E2*, *CACNA2D2*, *TCF4*, *CDH2*) were also found to be related to the above pathways.

We also detected significant KEGG pathways (chemokine signaling pathway, corrected *P*-value = 0.0044) and GO terms responsible for immunity (acute myeloid leukemia, corrected *P*-value = 0.0233; positive regulation of defense response to virus by host, corrected *P*-value = 0.0391; leukocyte activation, corrected *P*-value = 0.0442; immune system process, corrected *P*-value = 0.0122) involving immunity related genes (e.g., *NKX2-3*, *PIK3R1*, *ITK*, *TMEM173*, *MTSS1*). In addition, the KEGG pathways (e.g., parkinson disease, corrected *P*-value = 0.2241; glutamatergic synapse, corrected *P*-value = 0.0359; cholinergic synapse, corrected *P*-value = 0.0339) and GO terms (e.g., neuron development, corrected *P*-value = 0.0009; neurogenesis, corrected *P*-value = 0.0079; forebrain neuron differentiation, corrected *P*-value = 0.0455; central nervous system neuron axonogenesis, corrected *P*-value = 0.0454) also enriched genes associated with the nervous system in the FZ buffalo. Several neural-related genes (e.g., *PTPN21*, *ROBO1*, *HOMER1*, *MAGI2*, *SLC1A3*, *NRG3*, *SNAP47*, *CTNNA2*, *ADGRL3*) were also found to be enriched in FZ buffalo.

Discussion

To date, there have been several studies using the high-throughput sequencing and genotyping technologies to search for candidate genes associated with milk production, growth traits, immune and nervous system of buffaloes [6–9]. However, most of these studies were performed for the dairy river buffalo. The FZ buffalo has been used as a draft animal to provide farm power in rice cultivation. It has also been selected by local people to participate in the bullfight (a folk custom with long history in Guangxi Zhuang Autonomous Region) because of its developed muscles, high endurance, and disease resistance capacity [4]. According to the geographical location, the FZ buffalo belongs to the Southwest China. Our results confirmed the genetic separation between the buffaloes from Southwest China and Upper Yangtze region (Fig. 1) [2].

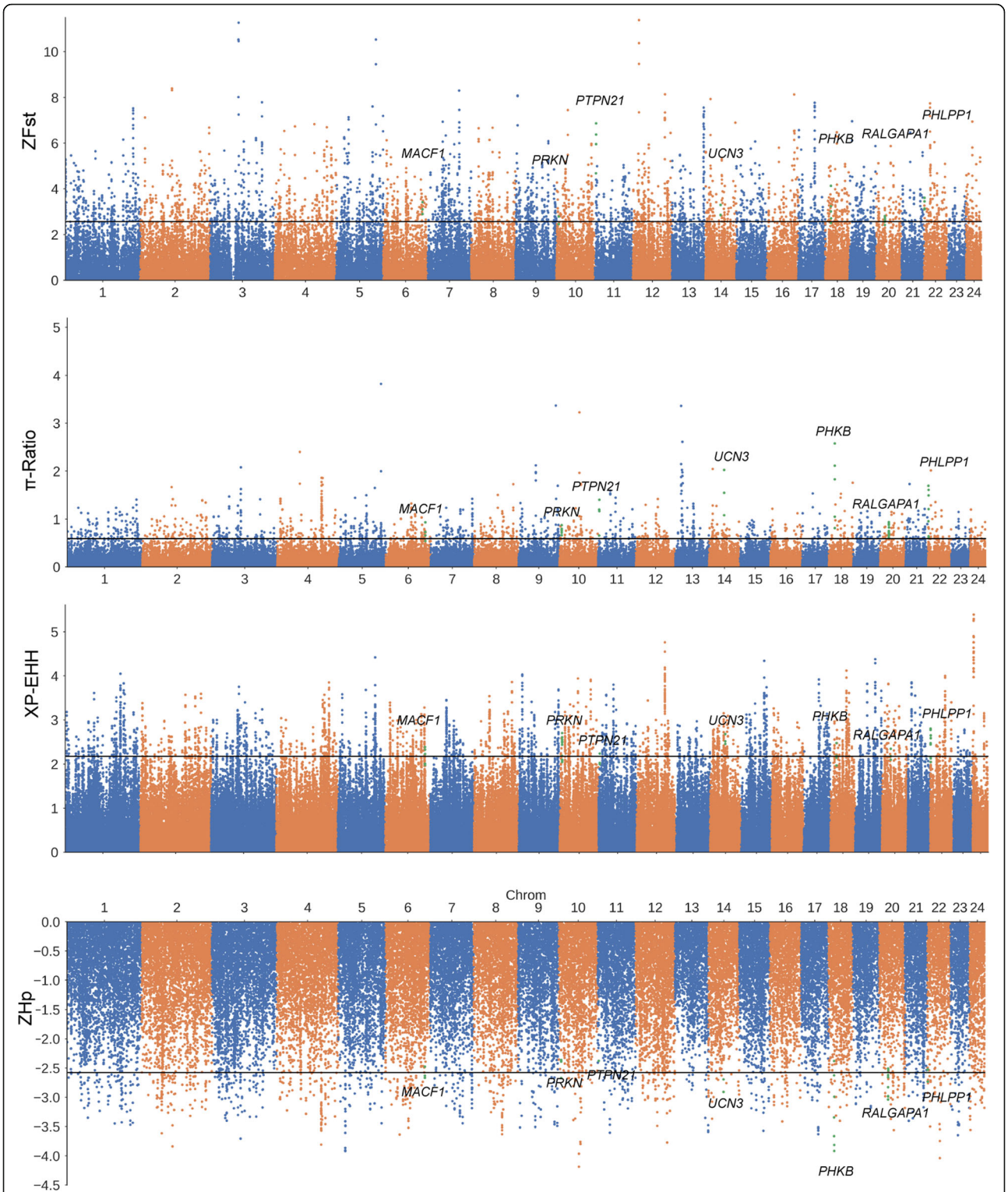

Fig. 2 Manhattan plot of selective sweeps in Fuzhong (FZ) buffalo. ZFst, π-Ratio, XP-EHH, and ZHp values were calculated in 50-kb sliding window with 20-kb step across all autosomes in FZ buffalo. The Manhattan plots are pointed the value for each window of ZFst, π-Ratio, XP-EHH and ZHp. The black horizontal line indicates the cut-off (*P*-value < 0.005) setted for extracting outlier for FZ buffalo (ZFst > 2.576; π-Ratio $_{UY/FZ}$ > 0.592, XP-EHH$_{FZ\text{-}to\text{-}UY}$ > 2.180, and ZH*p* < − 2.576). The green dots represent the regions that contain the highlight gene

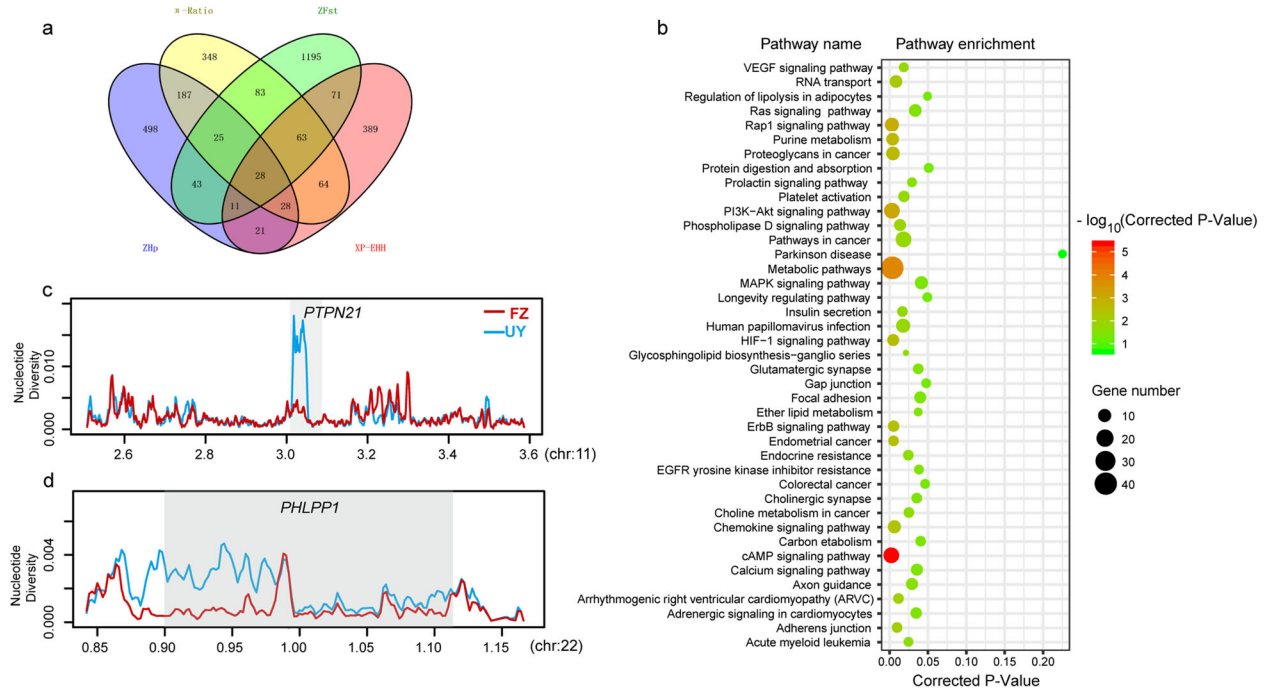

Fig. 3 Genomic regions with strong selective signals in FZ buffalo. **a** Number of candidate genes identified in FZ buffalo by the four methods listed in each of the Venn diagram components. **b** KEGG enrichment analysis for the identified candidate genes. Example of genes (**c**, **d**) with selection sweep signals in FZ buffalo. The nucleotide diversity was plotted using a 5-kb sliding window. The gene region was noted by gray shadow in (**c** & **d**)

In this study, we used four methods (ZFst, ZHp, π-Ratio, and XP-EHH) to detect the selective sweeps in FZ buffalo. These four approaches largely belong to two different types. The ZFst, ZHp, π-Ratio were based on allele frequency, while the XP-EHH was based on linkage disequilibrium patterns across genomes. Combining different detection approaches can provide complementary information, which was considered as an optimal strategy in detecting selection signatures [13]. Our results showed that ZHp and π-Ratio detected the most shared genes, while the XP-EHH and ZHp detected the lowest number of shared genes (Fig. 3a). It seems that the number of shared gene between different methods may be not associated with type of method.

FZ buffalo is with strong muscle, able to endure the strength to pull a plough through muddy rice paddies as well as bullfighting. Our results depicted the association of FZ buffalo with oxidative stress, cardiovascular system, and immunity in were several KEGG pathways and GO terms, including HIF-1 signaling pathway, adherens junction, insulin secretion, arrhythmogenic right ventricular cardiomyopathy (ARVC), which were reported to be involved in the reaction to exercise. HIF-1 signaling pathway plays an important role in the regulation of glycolysis during exercise and endurance performance [14]. During endurance training, the endurance exercise can result in oxidative stress and antioxidant defense, and the skeletal muscle experiences severe and repetitive oxygen stress [15]. Studies showed that the metabolic action of insulin was enhanced in skeletal muscle after exercise [16, 17]. Previous reports also indicated that the adherens junction pathway may be related to the exercise in human and horse [18, 19]. ARVC is a rare inherited heart-muscle disease, which could lead to sudden death in young people, athletes, and horse [20, 21]. Previous research has shown the severity of ARVC to be associated with strenuous endurance exercise [22]. In our current study, we identified several candidate genes (e.g., *ALDOA*, *STAT3*, *AKT2*, *EIF4E2*, *CACNA2D2*, *TCF4*, *CDH2*) involved in these two pathways. *ALDOA* (aldolase A, fructose-bisphosphate) is involved in a variety of cellular functions and biological processes, including muscle maintenance, regulation of cell morphology and migration, striated muscle contraction, actin cytoskeleton and actin polymerization and ATP biosynthesis [23–30]. The hypermethylation of *ALDOA* was found to be involved in the anaerobic metabolism in slow muscle fibers [31]. *STAT3* may contribute to the adaptation of skeletal muscle after the acute resistance exercise [32]. *AKT2* is a critical regulator for cardiomyocyte survival and metabolism [33]. *CACNA2D2* plays an important role in heart rate [34], and the *CACNA2D2*-knockout mouse showed lower heart rate [35]. *TCF4* contributes to muscle fiber and basement membrane recovery

following the muscle fiber damage induced by exercise [36]. Furthermore, *CDH2* was also identified as a candidate gene involved in cardiomyopathy [37]. All these aforementioned pathways, GO terms, and involved candidate genes directly or indirectly play an important role in oxidative stress and the overall body endurance. These results provided strong evidence that FZ buffalo bear endurance, strong muscular stature and powerful oxidative muscle strength.

Noteworthy KEGG pathway and GO terms (e.g., chemokine signaling pathway, immune system process, leukocyte activation) were over-represented, involving immunity related genes (e.g., *NKX2–3*, *PIK3R1*, *ITK* and *TMEM173*, *MTSS1*). Chemokines are small chemoattractant peptides that provide directional cues for the cell trafficking and thus are vital for protective host response. *PTPN22* plays an important role in autoimmune diseases [38]. *NKX2–3* plays a substantial role in the correct association of lymphocytes and splenic stromal elements [39]. *PIK3R1*, *ITK* and *TMEM173* were identified as potential positional candidate genes associated with infection of *Mycobacterium avium* subspecies paratuberculosis in cattle [40–42]. A study reported that the *MTSS1* was related to metritis in cattle [43]. Guangxi Zhuang Autonomous Region undergoes characteristic subtropical monsoon climate with hot and humid summers and dry and mild winters, which provides an environment for the spread of disease. Although the temperature and humidity are significantly higher than the Yangtze region [3], the native FZ buffalo is with high disease resistance. Our results provided the genetic evidence to illustrate the capability of the local buffalo adaptation to the local environment bearing enhanced disease resistance.

Moreover, the KEGG pathways (e.g., parkinson disease, cholinergic synapse, glutamatergic synapse) and GO terms (e.g., positive regulation of neuron differentiation, forebrain neuron differentiation, central nervous system neuron axonogenesis, regulation of neurogenesis) associated with nerve were detected, involving a set of genes (e.g., *PTPN21*, *ROBO1*, *HOMER1*, *MAGI2*, *SLC1A3*, *NRG3*, *SNAP47*, *CTNNA2*, *ADGRL3*). Moreover, it was found that *PAFAH1B1* was associated with learning, memory, and motor behavior [44]. *ROBO1* was reported to be related to reading disability in humans [45]. A study reported that *ROBO1* was significantly associated with breed-specific accomplishments of dog in competitive obstacle course events [46]. *HOMER1*, *MAGI2*, *SLC1A3*, *NRG3* were associated with schizophrenia [47–50]. *HOMER1* was identified as a candidate gene involving in the domestication of swamp buffalo [2]. *SNAP47* was proved to be involved in postsynaptic and presynaptic function [51]. The *CTNNA2* was associated with excitement-seeking and risk-taking, and were relevant to hyperactivity, substance use, antisocial and bipolar disorders [52]. *ADGRL3* was associated with attention-deficit/hyperactivity disorder [53]. *PTPN21* (protein tyrosine phosphatase, non-receptor type 21), a protein-coding gene which can positively influence cortical neuronal survival and enhance neuritic length [54], was identified as a potential risk gene for schizophrenia [55], showing a reduction in nucleotide diversity in FZ buffalo (Fig. 3c). In addition, we detected a SNP (g. 3, 035,884 C > T) in the intron 1 of *PTPN21*. In this locus, the C allele was dominant in FZ buffalo with the frequency of 87.5%, while in buffaloes from Upper Yangtze was rather rare (33.3%). The buffalo is used as a draft animal, which is very gentle and easy to handle and train. In addition, the FZ buffalo was also selected to participate in the traditional folk bullfighting in Guangxi Zhuang Autonomous Region. These results indicated that the nerve of FZ buffalo must have experienced selection, and these genes may act as the selective genes associated with the nerve of FZ buffalo.

Comparing 21 candidate genes detected by 4 methods with the published literature, some of these genes were found to be involved in production and growth traits. *PHLPP1* (PH domain and leucine-rich repeat protein phosphatase 1) was highlighted as functionally plausible candidate gene for pig growth and fatness traits [56, 57], which was also reported to be associated with birth growth trait in cattle [58, 59]. The selective sweep containing *PHLPP1* showed a reduction in nucleotide diversity in FZ buffalo (Fig. 3d). In addition, several genes were associated with meat quality (*PRKN*), feed efficiency (*MACF1*) and carcass traits (*ZNF280B* and *UCN3*) in cattle [60–63]. *RALGAPA1* was relative to reproductive traits in chicken [64]. *PHKB* was as a candidate gene for the feed-conversion ratio in pig [65]. Our results indicated that FZ buffalo experienced artificial breeding, and the growth traits were improved with the artificial selection.

Moreover, 58 candidate genes were found to be shared between FZ buffalo and UY buffalo. In particular, the *TIAM1*, *HOMER1* were identified as the neural-related genes involved in the domestication of swamp buffalo [2]. As we all known, almost all the swamp buffalo is used as the draft animal to provide farm power in rice cultivation, which is very docile and easy to handle and train. Our results further conformed the importance of nervous system for the domestication of swamp buffalo.

Conclusions

FZ buffalo is a native breed from Guangxi Zhuang Autonomous Region, which is mainly used as a draft animal to plow or level land, puddle rice fields, and bullfight. In this study, we re-sequenced the whole genomes of 15 FZ buffaloes, combining with 12 published buffalo genomes

from Upper Yangtze region to detect the selective signatures in FZ buffalo using ZHp, ZFst, π-Ratio, and XP-EHH statistics. Our results depicted several pathways, GO terms, and candidate genes to be associated with response to exercise, immunity, nervous system, and growth traits, which aid us towards a better understanding of the adaptive traits in FZ buffalo. In addition, our results indicated that the nervous system plays an important role in the domestication of swamp buffalo.

Methods

Sample collection and sequencing

We sampled a total of 15 ear tissues of FZ buffaloes from Guangxi Zhuang Autonomous Region China. The samples were collected from local farmers, who were interviewed in detail to ensure unrelatedness among the sampled individuals. Following sampling, the animals were returned to their owners. Genomic DNA was extracted from the ear tissue samples using the standard phenol-chloroform protocol. For each sample, 1-15 μg of DNA was used to construct the library with an average insert size of 500 bps. Sequencing was performed to generate 150-bp paired-end reads on an Illumina HiSeq 2000 according to the manufacturer's protocol. Moreover, 12 publically available buffalo genome sequences [2] from regions of the Upper Yangtze were used as a control group. Detailed information about all samples analyzed in this study was provided in the Additional file 1: Table S1. Raw FASTQ sequences have been deposited to NCBI Short Read Archive under the BioProject accession number PRJNA566371. A study has suggested that around 10 individuals can gain the power for the analysis of selection sweep [66]. Totally, 27 whole genomes of buffalo were used for the further analysis.

Alignments and variant identification

The clean reads were mapped to the reference genome (GCA_003121395.1) [67] using BWA-MEM with default settings [68]. Furthermore, Samtools, Picard tools, Genome Analysis Toolkit (GATK, version 3.6–0-g89b7209) were used to detect the single nucleotide polymorphisms (SNPs) [69]. All SNPs were filtered using the "Variant Filtration" module of GATK with the standard parameters as below: Variants with quality depth (QD) < 2; FS (Phred-scaled *P*-value using Fisher's exact test to detect strand bias) > 60; MQRankSum < − 12.5; ReadPosRankSum < − 8; MQ < 40.0; the mean sequencing depth of variants (containing all individuals) < 1/3× and > 3×; SOR > 3.0; maximum missing rate < 0.1; and SNPs were restricted to the two alleles.

Population structure analyses

The genetic relationships between FZ and UY buffaloes was performed by the principal component analysis (PCA). The SNPs were filtered using the MAF (--maf 0.05) using VCFtools [70]. Moreover, extended parameters (--indep-pairwise 50 5 0.2) of PLINK were used to remove one of the pairs of SNPs if the linkage disequilibrium (LD) with a squared correlation greater than 0.2, in windows of 50 variants and shifting by 5 variants. The SNPs that pass the pruning were used to perform PCA analysis. The PCA was performed using SmartPCA program in the package EIGENSOFT v5.0 [71].

Genome-wide selective sweep test

To identify the selective sweep regions, the Fst, Hp, π-Ratio, and XP-EHH tests were performed with 50 kb sliding window and 20 kb step. The SNPs were filtered with parameters (--maf 0.05 -max-missing 0.90) using PLINK 1.9 [72]. The Fst [73] was calculated using VCFtools with parameter "--weir-fst-pop group1 --weir-fst-pop group2 --fst-window-size 50000 --fst-window-step 20000 --maf 0.05 --max-missing 0.90" [70], while Hp was calculated as described previously [74]. The Hp and Fst values were converted to a standard normal distribution, denoted by ZHp and ZFst [11]. The genetic diversity (π-Ratio) was calculated using VCFtools with parameters as follows: "--keep gropu1/gropu2 --window-pi 50000 --window-pi-step 20000 --maf 0.05 --max-missing 0.90" [70] and house python scripts. The XP-EHH was performed for every SNP using the default settings by selscan v1.1 [75], and genotypes were phased using Beagle [76] with default parameters. The test statistic was the average normalized XP-EHH score in each 50-kb region. In the π-Ratio, and XP-EHH tests, the FZ buffalo were used as the target population, and the UY buffalo act as the reference population. The *P*-values were estimated based on values of the test using the normal distribution. Significant genomic regions for each method were identified by *P*-value < 0.005. The power of each test was different, any set of candidate genes may contain some false positives [77]. If there is any given signal consistently supported by other methods, it may be considered as a strong evidence that the locus has been under selection. Combining multiple tests can improve the power of detecting selection signatures [78]. To date, almost all the genome-scan-related studies used multiple analyses to further confirm the selected candidate regions to make the results more reliable. In the current study, two or more methods showed outlier signals (*P*-value < 0.005) in overlapping regions and were therefore considered as the candidate selective regions and were subsequently examined for the candidate genes [79]. The candidate regions were separated by a distance with less than 50 kb using the house script. In order to gain better understanding of the biological functions and pathway of the identified candidate genes, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) were performed using KOBAS 3.0 [80], considering a significant threshold for corrected *P*-value < 0.05.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12864-020-07095-8.

Additional file 1: Table S1. Overview of sample information and sequencing statistics.

Additional file 2: Table S2. Distribution of SNPs within various genomic regions.

Additional file 3: Table S3. A summary of genes from ZFst.

Additional file 4: Table S4. A summary of genes from ZHp.

Additional file 5: Table S5. A summary of genes from π-Ratio.

Additional file 6: Table S6. A summary of genes from XP-EHH.

Additional file 7: Table S7. KEGG pathway analysis of candidate genes in FZ buffalo.

Additional file 8: Table S8. GO enrichment analysis of candidate genes in FZ buffalo.

Abbreviations

FZ: Fuzhong; UY: Upper Yangtze; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; Fst: The genetic differentiation; Hp: Heterozygosity; π-Ratio: Genetic diversity; XP-EHH: Cross-population extended haplotype homozygosity; PCA: Principal component analysis

Acknowledgments

We would like to thank Dr. Qiuming Chen for his good suggestions.

Authors' contributions

Y.Y.L. and C.Z.L. conceived and designed the study. T.S., G.Y.H., Y.H.C., N.B.C. and J.L.S. performed statistical analysis. T.S. wrote the manuscript. Z.H.W., S.H.T., Q.H. revised the manuscript. Y.H.C., J.L.S., N.B.C., and C.Z.L. collected the samples. All authors reviewed the manuscript and gave final approval for publication.

Funding

The work was supported by the Guangxi special project for innovation-driven development (AA17204024) and the National Beef Cattle and Yak Industrial Technology System (CARS-37).

Availability of data and materials

Raw FASTQ sequences have been deposited to NCBI Short Read Archive under the BioProject accession number PRJNA566371.

Ethics approval and consent to participate

The ear tissues of buffaloes were collected with verbal agreement of the famers who owned the animals. All experimental procedures were performed in accordance with the Regulations for the Administration of Affairs Concerning Experimental Animals approved by the State Council of the People's Republic of China. The study was approved by the Institutional Animal Care and Use Committee of Northwest A&F University. This study was approved by the Institutional Animal Care and Use Committee of Northwest A&F University (Permit number: NWAFAC1019).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Animal Husbandry Institute of Guangxi Zhuang Autonomous Region, Guangxi Key Laboratory of Livestock Genetic Improvement, Nanning 530001, China. [2]College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi, China. [3]Computational Biology Laboratory, Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan. [4]Department of Biotechnology, Pakistan Institute of Engineering and Applied Sciences, Nilore, Islamabad, Pakistan.

Received: 1 May 2020 Accepted: 23 September 2020
Published online: 29 September 2020

References

1. Fischer H, Ulbrich F. Chromosomes of the Murrah buffalo and its crossbreds with the Asiatic swamp buffalo (Bubalus bubalis). Z Tierzücht Züchtungsbiol. 1967;84(1–4):110–4.
2. Sun T, Shen J, Achilli A, Chen N, Chen Q, Dang R, Zheng Z, Zhang H, Zhang X, Wang S, et al. Genomic analyses reveal distinct genetic architectures and selective pressures in buffaloes. GigaScience. 2020;9(2).
3. Zeng L, Chen N, Ning Q, Yao Y, Chen H, Dang R, Zhang H, Lei C. PRLH and SOD1 gene variations associated with heat tolerance in Chinese cattle. Anim Genet. 2018;49(5):447–51.
4. Resources CNCoAG: Animal genetic RESOURCES in China-bovines: China agriculture press; 2011.
5. Iamartino D, Nicolazzi EL, Van Tassell CP, Reecy JM, Fritz-Waters ER, Koltes JE, Biffani S, Sonstegard TS, Schroeder SG, Ajmone-Marsan P. Design and validation of a 90K SNP genotyping assay for the water buffalo (Bubalus bubalis). PLoS One. 2017;12(10):e0185220.
6. Mokhber M, Moradi-Shahrbabak M, Sadeghi M, Moradi-Shahrbabak H, Stella A, Nicolzzi E, Rahmaninia J, Williams JL. A genome-wide scan for signatures of selection in Azeri and Khuzestani buffalo breeds. BMC Genomics. 2018; 19(1):449.
7. Ghoreishifar SM, Moradi-Shahrbabak H, Fallahi MH, Jalil Sarghale A, Moradi-Shahrbabak M, Abdollahi-Arpanahi R, Khansefid M. Genomic measures of inbreeding coefficients and genome-wide scan for runs of homozygosity islands in Iranian river buffalo, *Bubalus bubalis*. BMC Genet. 2020;21(1):16.
8. Liu JJ, Liang AX, Campanile G, Plastow G, Zhang C, Wang Z, Salzano A, Gasparrini B, Cassandro M, Yang LG. Genome-wide association studies to identify quantitative trait loci affecting milk production traits in water buffalo. J Dairy Sci. 2018;101(1):433–44.
9. Whitacre LK, Hoff JL, Schnabel RD, Albarella S, Ciotola F, Peretti V, Strozzi F, Ferrandi C, Ramunno L, Sonstegard TS, et al. Elucidating the genetic basis of an oligogenic birth defect using whole genome sequence data in a non-model organism, *Bubalus bubalis*. Sci Rep. 2017;7:39719.
10. Holsinger KE, Weir BS. Genetics in geographically structured populations: defining, estimating and interpreting FST. Nat Rev Genet. 2009;10(9):639–50.
11. Rubin C-J, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, et al. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature. 2010; 464(7288):587–91.
12. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. Genome-wide detection and characterization of positive selection in human populations. Nature. 2007; 449:913.
13. Grossman SR, Shlyakhter I, Karlsson EK, Byrne EH, Morales S, Frieden G, Hostetter E, Angelino E, Garber M, Zuk O, et al. A composite of multiple signals distinguishes causal variants in regions of positive selection. Science. 2010;327(5967):883–6.
14. Mason SD, Howlett RA, Kim MJ, Olfert IM, Hogan MC, McNulty W, Hickey RP, Wagner PD, Kahn CR, Giordano FJ, et al. Loss of skeletal muscle HIF-1alpha results in altered exercise endurance. PLoS Biol. 2004;2(10):e288.
15. Kinnunen S, Atalay M, Hyyppä S, Lehmuskero A, Hänninen O, Oksala N. Effects of prolonged exercise on oxidative stress and antioxidant defense in endurance horse. J Sports Sci Med. 2005;4(4):415–21.
16. Mikines KJ, Sonne B, Farrell PA, Tronier B, Galbo H. Effect of physical exercise on sensitivity and responsiveness to insulin in humans. Am J Physiol Endocrinol Metab. 1988;254(3):E248–59.
17. Richter EA, Derave W, Wojtaszewski JFP. Glucose, exercise and insulin: emerging concepts. J Physiol. 2001;535(2):313–22.
18. Radom-Aizik S, Zaldivar F, Haddad F, Cooper DM. Impact of brief exercise on peripheral blood NK cell gene and microRNA expression in young adults. J Appl Physiol. 2013;114(5):628–36.
19. Kim H-A, Kim M-C, Kim N-Y, Ryu D-Y, Lee H-S, Kim Y. Integrated analysis of microRNA and mRNA expressions in peripheral blood leukocytes of Warmblood horses before and after exercise. J Vet Sci. 2018;19(1):99–106.
20. Basso C, Corrado D, Marcus FI, Nava A, Thiene G. Arrhythmogenic right ventricular cardiomyopathy. Lancet. 2009;373(9671):1289–300.
21. Freel KM, Morrison LR, Thompson H, Else RW. Arrhythmogenic right ventricular cardiomyopathy as a cause of unexpected cardiac death in two horses. Vet Rec Case Rep. 2013;1(1):ec3000.

22. Sawant Abhishek C, Bhonsale A, te Riele Anneline SJM, Tichnell C, Murray B, Russell Stuart D, Tandri H, Tedford Ryan J, Judge Daniel P, Calkins H, et al. Exercise has a Disproportionate Role in the Pathogenesis of Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy in Patients Without Desmosomal Mutations. J Am Heart Assoc. 3(6):e001471.
23. Kusakabe T, Motoki K, Hori K. Mode of interactions of human Aldolase Isozymes with cytoskeletons. Arch Biochem Biophys. 1997;344(1):184–93.
24. Harris SJ, Winzor DJ. Enzyme kinetic evidence of active-site involvement in the interaction between aldolase and muscle myofibrils. Biochim Biophys Acta Protein Struct Mol Enzymol. 1987;911(1):121–6.
25. Arnold H, Pette D. Binding of glycolytic enzymes to structure proteins of the muscle. Eur J Biochem. 1968;6(2):163–71.
26. Tochio T, Tanaka H, Nakata S, Hosoya H. Fructose-1,6-bisphosphate aldolase a is involved in HaCaT cell migration by inducing lamellipodia formation. J Dermatol Sci. 2010;58(2):123–9.
27. Hu H, Juvekar A, Lyssiotis Costas A, Lien Evan C, Albeck John G, Oh D, Varma G, Hung Yin P, Ullas S, Lauring J, et al. Phosphoinositide 3-kinase regulates glycolysis through mobilization of Aldolase from the actin cytoskeleton. Cell. 2016;164(3):433–46.
28. Carr D, Knull H. Aldolase-tubulin interactions: removal of tubulin C terminals impairs interactions. Biochem Biophys Res Commun. 1993;195(1):289–93.
29. Walsh JL, Knull HR. Heteromerous interactions among glycolytic enzymes and of glycolytic enzymes with F-actin: effects of poly (ethylene glycol). Biochim Biophys Acta Protein Struct Mol Enzymol. 1988;952:83–91.
30. Clarke FM, Masters CJ. On the association of glycolytic enzymes with structural proteins of skeletal muscle. Biochim Biophys Acta. 1975;381(1):37–46.
31. Begue G, Raue U, Jemiolo B, Trappe S. DNA methylation assessment from human slow- and fast-twitch skeletal muscle fibers. J Appl Physiol (Bethesda, Md : 1985). 2017;122(4):952–67.
32. Trenerry MK, Carey KA, Ward AC, Cameron-Smith D. STAT3 signaling is activated in human skeletal muscle following acute resistance exercise. J Appl Physiol. 2007;102(4):1483–9.
33. Muslin AJ. Akt2: a critical regulator of Cardiomyocyte survival and metabolism. Pediatr Cardiol. 2011;32(3):317–22.
34. Volland C, Bremer S, Hellenkamp K, Hartmann N, Dybkova N, Khadjeh S, Kutschenko A, Liebetanz D, Wagner S, Unsöld B, et al. Enhanced cardiac TBC1D10C expression lowers heart rate and enhances exercise capacity and survival. Sci Rep. 2016;6(1):33853.
35. Ivanov SV, Ward JM, Tessarollo L, McAreavey D, Sachdev V, Fananapazir L, Banks MK, Morris N, Djurickovic D, Devor-Henneman DE, et al. Cerebellar Ataxia, seizures, premature death, and cardiac abnormalities in mice with targeted disruption of the Cacna2d2 gene. Am J Pathol. 2004;165(3):1007–18.
36. Kanazawa Y, Nagano M, Koinuma S, Sujino M, Minami Y, Sugiyo S, Takeda I, Shigeyoshi Y. Basement membrane recovery process in rat soleus muscle after exercise-induced muscle injury. Connect Tissue Res. 2020:1–12.
37. Mayosi Bongani M, Fish M, Shaboodien G, Mastantuono E, Kraus S, Wieland T, Kotta M-C, Chin A, Laing N, Ntusi Ntobeko BA, et al. Identification of cadherin 2 (CDH2) mutations in Arrhythmogenic right ventricular cardiomyopathy. Circ Cardiovasc Genet. 2017;10(2):e001605.
38. Bottini N, Vang T, Cucca F, Mustelin T. Role of PTPN22 in type 1 diabetes and other autoimmune diseases. Semin Immunol. 2006;18(4):207–13.
39. Tarlinton D, Light A, Metcalf D, Harvey RP, Robb L. Architectural defects in the spleens of Nkx2-3-deficient mice are intrinsic and associated with defects in both B cell maturation and T cell-dependent immune responses. J Immunol. 2003;170(8):4002.
40. Alpay F, Zare Y, Kamalludin MH, Huang X, Shi X, Shook GE, Collins MT, Kirkpatrick BW. Genome-wide association study of susceptibility to infection by Mycobacterium avium subspecies paratuberculosis in Holstein cattle. PLoS One. 2014;9(12):e111704.
41. Marino R, Capoferri R, Panelli S, Minozzi G, Strozzi F, Trevisi E, Snel GGM, Ajmone-Marsan P, Williams JL. Johne's disease in cattle: an in vitro model to study early response to infection of Mycobacterium avium subsp. paratuberculosis using RNA-seq. Mol Immunol. 2017;91:259–71.
42. Neibergs HL, Settles ML, Whitlock RH, Taylor JF. GSEA-SNP identifies genes associated with Johne's disease in cattle. Mamm Genome. 2010;21(7):419–25.
43. Guarini AR, Lourenco DAL, Brito LF, Sargolzaei M, Baes CF, Miglior F, Misztal I, Schenkel FS. Genetics and genomics of reproductive disorders in Canadian Holstein cattle. J Dairy Sci. 2019;102(2):1341–53.
44. Paylor R, Hirotsune S, Gambello MJ, Yuva-Paylor L, Crawley JN, Wynshaw-Boris A. Impaired learning and motor behavior in heterozygous Pafah1b1 (Lis1) mutant mice. Learn Mem. 1999;6(5):521.
45. Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmi J, Kääriäinen H, Kere J. The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. PLoS Genet. 2005;1(4):e50.
46. Kim J, Williams FJ, Dreger DL, Plassais J, Davis BW, Parker HG, Ostrander EA. Genetic selection of athletic success in sport-hunting dogs. Proc Natl Acad Sci U S A. 2018;115(30):E7212–e7221.
47. Koide T, Banno M, Aleksic B, Yamashita S, Kikuchi T, Kohmura K, Adachi Y, Kawano N, Kushima I, Nakamura Y, et al. Common variants in MAGI2 gene are associated with increased risk for cognitive impairment in schizophrenic patients. PLoS One. 2012;7(5):e36836.
48. Szumlinski KK, Lominac KD, Kleschen MJ, Oleson EB, Dehoff MH, Schwartz MK, Seeberg PH, Worley PF, Kalivas PW. Behavioral and neurochemical phenotyping of Homer1 mutant mice: possible relevance to schizophrenia. Genes Brain Behav. 2005;4(5):273–88.
49. Deng X, Shibata H, Takeuchi N, Rachi S, Sakai M, Ninomiya H, Iwata N, Ozaki N, Fukumaki Y. Association study of polymorphisms in the glutamate transporter genes SLC1A1, SLC1A3, and SLC1A6 with schizophrenia. Am J Med Genet B Neuropsychiatr Genet. 2007;144B(3):271–8.
50. Morar B, Dragović M, Waters FAV, Chandler D, Kalaydjieva L, Jablensky A. Neuregulin 3 (NRG3) as a susceptibility gene in a schizophrenia subtype with florid delusions and relatively spared cognition. Mol Psychiatry. 2011;16(8):860–6.
51. Münster-Wandowski A, Heilmann H, Bolduan F, Trimbuch T, Yanagawa Y, Vida I. Distinct Localization of SNAP47 Protein in GABAergic and Glutamatergic Neurons in the Mouse and the Rat Hippocampus. Front Neuroanat. 2017;11(56).
52. Terracciano A, Esko T, Sutin AR, de Moor MHM, Meirelles O, Zhu G, Tanaka T, Giegling I, Nutile T, Realo A, et al. Meta-analysis of genome-wide association studies identifies common variants in CTNNA2 associated with excitement-seeking. Transl Psychiatry. 2011;1(10):e49.
53. Martinez AF, Abe Y, Hong S, Molyneux K, Yarnell D, Löhr H, Driever W, Acosta MT, Arcos-Burgos M, Muenke M. An Ultraconserved brain-specific enhancer within ADGRL3 (LPHN3) underpins attention-deficit/hyperactivity disorder susceptibility. Biol Psychiatry. 2016;80(12):943–54.
54. Plani-Lam JH-C, Chow T-C, Siu K-L, Chau WH, Ng M-HJ, Bao S, Ng CT, Sham P, Shum DK-Y, Ingley E, et al. PTPN21 exerts pro-neuronal survival and neuritic elongation via ErbB4/NRG3 signaling. Int J Biochem Cell Biol. 2015;61:53–62.
55. Chen J, Lee G, Fanous AH, Zhao Z, Jia P, O'Neill A, Walsh D, Kendler KS, Chen X. The international schizophrenia C: two non-synonymous markers in PTPN21, identified by genome-wide association study data-mining and replication, are associated with schizophrenia. Schizophr Res. 2011;131(1):43–51.
56. Guo Y, Qiu H, Xiao S, Wu Z, Yang M, Yang J, Ren J, Huang L. A genome-wide association study identifies genomic loci associated with backfat thickness, carcass weight, and body weight in two commercial pig populations. J Appl Genet. 2017;58(4):499–508.
57. Sanchez M-P, Tribout T, Iannuccelli N, Bouffaud M, Servin B, Tenghe A, Dehais P, Muller N, Del Schneider MP, Mercat M-J, et al. A genome-wide association study of production traits in a commercial population of large white pigs: evidence of haplotypes affecting meat quality. Genet Sel Evol. 2014;46(1):12.
58. Hartati H, Utsunomiya YT, Sonstegard TS, Garcia JF, Jakaria J, Muladno M. Evidence of Bos javanicus x Bos indicus hybridization and major QTLs for birth weight in Indonesian Peranakan Ongole cattle. BMC Genet. 2015;16(1):75.
59. G. T. Pereira A, Utsunomiya YT, Milanesi M, RBP T, Carmo AS, HHR N, Carvalheiro R, Ajmone-Marsan P, Sonstegard TS, Sölkner J, et al. Pleiotropic Genes Affecting Carcass Traits in *Bos indicus* (Nellore) Cattle Are Modulators of Growth. PLoS One. 2016;11(7):e0158165.
60. Taye M, Lee W, Jeon S, Yoon J, Dessie T, Hanotte O, Mwai OA, Kemp S, Cho S, Oh SJ, et al. Exploring evidence of positive selection signatures in cattle breeds selected for different traits. Mamm Genome. 2017;28(11):528–41.
61. Lam S, Miglior F, Fonseca P, Seymour D, Asselstine V, Brito L, Schenkel F, Cánovas A. Identification of variants associated with divergent feed efficiency groups using multiple RNA-sequencing datasets from dairy and beef cattle; 2018.
62. Chang T, Xia J, Xu L, Wang X, Zhu B, Zhang L, Gao X, Chen Y, Li J, Gao H. A genome-wide association study suggests several novel candidate genes for

carcass traits in Chinese Simmental beef cattle. Anim Genet. 2018;49(4):312–6.
63. Jiang Z, Michal JJ, Chen J, Daniels TF, Kunej T, Garcia MD, Gaskins CT, Busboom JR, Alexander LJ, Wright RW Jr, et al. Discovery of novel genetic networks associated with 19 economically important traits in beef cattle. Int J Biol Sci. 2009;5(6):528–42.
64. Gholami M, Erbe M, Gärke C, Preisinger R, Weigend A, Weigend S, Simianer H. Population genomic analyses based on 1 million SNPs in commercial egg layers. PLoS One. 2014;9(4):e94509.
65. Messad F, Louveau I, Koffi B, Gilbert H, Gondret F. Investigation of muscle transcriptomes using gradient boosting machine learning identifies molecular predictors of feed efficiency in growing pigs. BMC Genomics. 2019;20(1):659.
66. Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW, et al. Signals of recent positive selection in a worldwide sample of human populations. Genome Res. 2009;19(5):826–37.
67. Low WY, Tearle R, Bickhart DM, Rosen BD, Kingan SB, Swale T, Thibaud-Nissen F, Murphy TD, Young R, Lefevre L, et al. Chromosome-level assembly of the water buffalo genome surpasses human and goat genomes in sequence contiguity. Nat Commun. 2019;10(1):260.
68. Li H, Durbin R. Fast and accurate short read alignment with burrows–wheeler transform. Bioinformatics. 2009;25(14):1754–60.
69. Nekrutenko A, Taylor J. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. Nat Rev Genet. 2012;13(9):667–72.
70. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8.
71. Patterson N, Price AL, Reich D. Population structure and Eigenanalysis. PLoS Genet. 2006;2(12):e190.
72. Shaun P, Benjamin N, Kathe TB, Lori T, Ferreira MAR, David B, Julian M, Pamela S, Bakker PIW, De DMJ. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007.
73. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population-structure. Evolution. 1984;38(6):1358–70.
74. Rubin C-J, Megens H-J, Martinez Barrio A, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, et al. Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci U S A. 2012;109(48):19529–36.
75. Szpiech ZA, Hernandez RD. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol Biol Evol. 2014;31(10):2824–7.
76. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81(5):1084–97.
77. Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. Recent and ongoing selection in the human genome. Nat Rev Genet. 2007;8(11):857–68.
78. Zeng K, Shi S, Wu CI. Compound tests for the detection of hitchhiking under positive selection. Mol Biol Evol. 2007;24(8):1898–908.
79. Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, et al. Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet. 2019;51(5):885–95.
80. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li C-Y, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39(suppl_2):W316–22.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC