






Communication

Replicative Study in Performance-Related Genes of Brazilian Elite Soccer Players Highlights Genetic Differences from African Ancestry and Similarities between Professional and U20 Youth Athletes

Tane Kanope ¹, Caleb G. M. Santos ^{2,3,*} , Feliciano Marinho ⁴, Gustavo Monnerat ⁵, Mario Campos-Junior ⁶, Ana Carolina P. da Fonseca ^{6,7} , Verônica M. Zembrzuski ⁶ , Miller de Assis ¹ , Michael W. Pfaffl ² 
and Eduardo Pimenta ¹

- ¹ UFMG Soccer Science Center, School of Physical Education, Physiotherapy and Occupational Therapy, Federal University of Minas Gerais, Belo Horizonte 31250810, Brazil; tane.ufrj@gmail.com (T.K.); empimenta@uol.com.br (E.P.)
 - ² Department of Animal Physiology and Immunology, School of Life Sciences, Technical University of Munich, Weihenstephaner Berg 3, 85354 Freising, Germany
 - ³ Teaching and Research Division, Brazilian Army Institute of Biology, Rio de Janeiro 20911270, Brazil
 - ⁴ Instituto Hermes Pardini SA, Vespasiano 33200000, Brazil
 - ⁵ Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941902, Brazil
 - ⁶ Human Genetics Laboratory, Oswaldo Cruz Institute, Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro 21040360, Brazil
 - ⁷ Laboratory of Immunopharmacology, Oswaldo Cruz Institute, Rio de Janeiro 21045900, Brazil
- * Correspondence: calebguedes@gmail.com or guedes.caleb@eb.mil.br; Tel.: +55-21-38902135



Citation: Kanope, T.; Santos, C.G.M.; Marinho, F.; Monnerat, G.; Campos-Junior, M.; da Fonseca, A.C.P.; Zembrzuski, V.M.; de Assis, M.; Pfaffl, M.W.; Pimenta, E. Replicative Study in Performance-Related Genes of Brazilian Elite Soccer Players Highlights Genetic Differences from African Ancestry and Similarities between Professional and U20 Youth Athletes. *Genes* **2023**, *14*, 1446. <https://doi.org/10.3390/genes14071446>

Academic Editors: Juan Del Coso, Agnieszka Maciejewska-Skrendo, Paweł Cięszczyk, Alison V. September and Ildus I. Ahmetov

Received: 15 June 2023
Revised: 10 July 2023
Accepted: 12 July 2023
Published: 14 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Classically, genetic association studies have attempted to assess genetic polymorphisms related to human physiology and physical performance. However, the heterogeneity of some findings drives the research to replicate, validate, and confirmation as essential aspects for ensuring their applicability in sports sciences. Genetic distance matrix and molecular variance analyses may offer an alternative approach to comparing athletes' genomes with those from public databases. Thus, we performed a complete sequencing of 44 genomes from male Brazilian first-division soccer players under 20 years of age (U20_BFDSC). The performance-related SNP genotypes were obtained from players and from the "1000 Genomes" database (European, African, American, East Asian, and South Asian). Surprisingly, U20_BFDSC performance-related genotypes had significantly larger F_{ST} levels ($p < 0.00001$) than African populations, although studies using ancestry markers have shown an important similarity between Brazilian and African populations (12–24%). U20_BFDSC were genetically similar to professional athletes, showing the intense genetic selection pressure likely to occur before this maturation stage. Our study highlighted that performance-related genes might undergo selective pressure due to physical performance and environmental, cognitive, and sociocultural factors. This replicative study suggests that molecular variance and Wright's statistics can yield novel conclusions in exercise science.

Keywords: population genetics; genetic distance matrix; SNP panel; soccer; athletic performance

1. Introduction

Physical performance is a multifactorial trait to which genetic data can potentially be applied for precision medicine approaches [1]. Classically, genetic association studies (GASs) are used to better understand the effects of genetic variability. Hundreds of previous studies have attempted to assess genetic polymorphisms and link their allele frequencies to human physiology and performance phenotypes. Based on GAS polymorphisms, there are lists of genes commonly associated with sports performance (Table 1) which support

additional deeper investigations using animal models, human tissues, and cells to validate the candidate genes.

Thus, important single-nucleotide polymorphisms (SNPs), like the non-sense rs17602729 in the *ACTN3* gene, have had their mechanisms described and related to sports performance. In this case, the absence of α -actinin-3 releases more calcineurin favoring endurance adaptation [2]. Additionally, the total or partial lack of adenosine monophosphate deaminase-1 due to *AMPD1* rs1760272, a key enzyme for purine metabolism, was previously related to ATP recovery and fatigue during exercise, a core valence for soccer players. In addition, peroxisome proliferator-activated receptor (PPAR)-related genes and their SNPs likely exert a complex, modulatory influence on fatty acid and carbohydrate oxidation pathways. However, their specific mechanisms remain to be understood [3]. Moreover, SNPs like *FTO* rs9939609 appeared in more than one GAS related to fitness and body fat but without real validation [4].

Alleles that enhance essential traits for specific physical activities likely experience selective pressures in elite athletes without consistently affecting unrelated genes [5]. However, the heterogeneity of some findings, even at the beginning of the genomic era, drives the research for replication, validation, and confirmation of the findings as essential aspects for ensuring the quality and applicability of their results in sports or medical sciences. Replicating positive findings favors the validity of genotype–phenotype associations and will avoid biases. For this reason, it is essential to observe various aspects using established analysis methods, genetic variants, the definition of phenotype, and ethnic/admixed groups, as well as independent but similar data sets [6].

Table 1. Summarized information about single-nucleotide polymorphism (SNP) genetic panel.

Groups of SNPs	Previous Reports
<i>ACTN3</i> rs17602729	Most studied polymorphism for sports science related to strength/power performance [2,7]
<i>AMPD1</i> rs1760272	Nonsense; related to ATP recovery and fatigue [8]
<i>PPARGC1A</i> , <i>MCT1</i> , <i>CYP1A2</i> rs8192678G, rs10494, rs76255134	Metabolic pathways and energy expenditure [9,10]
<i>COL5A1</i> , <i>MMP3</i> rs12722, rs591058	Stability and health of muscle structure [11,12]
<i>FTO</i> rs9939609	Physical fitness in several whole-genome GASs [13]
<i>16 CHRM2</i> , <i>15 CHRM2</i> rs8191992, rs324640	Heart rate recovery and electrical conduction [14]

ACTN3 (α -actinin-3), *AMPD1* (adenosine monophosphate deaminase 1), *PPARGC1A* (PPARG coactivator 1 alpha), *MCT1* (solute carrier family 16 member 1), *CYP1A2* (cytochrome P450 family 1 subfamily A member 2), *COL5A1* (collagen type V alpha 1 chain), *MMP3* (matrix metalloproteinase 3), *FTO* (α -ketoglutarate-dependent dioxygenase), *16 CHRM2* (cholinergic receptor muscarinic 2—intron 6), *15 CHRM2* (cholinergic receptor muscarinic 2—intron 5).

The recent results in 25 male Brazilian soccer players [15] showed an alternative use of genetic panels and GASs, approaching a concurrent contribution of SNPs based on genetic matrix distance and ethnic aspects between athletes or public genomic databases, like the “1000 Genomes” database. Notably, matrix distance-based studies using proteomic and metabolomic data in a pairwise or multidimensional analysis have shown biomarker patterns by clustering groups based on moments of acute exercise [16–18], performance [19–21], and sports disciplines or physiological variables [22]. However, as a population genetics method, human research on this subject has been mainly applied to the forensic field. Thus, this short communication aims to investigate the replicability of a “genetic distance approach” based on SNPs and Wright statistics for genetic distances [23], as an alternative to traditional GASs, in an unrelated, new, younger, and larger population of soccer athletes, to determine if the U20 athletes are genetically related to professional ones and other ethnicities available in 1000 Genomes database.

2. Materials and Methods

We selected an independent study group of soccer players for replication. We performed a complete sequencing of 44 genomes from male Brazilian first-division soccer players under 20 years of age (U20_BFDSC) to compare them with a group of 25 probands (BFDSC). DNA was obtained from buccal cells using a MagMAX™ DNA Multi-Sample kit (ThermoFisher, Waltham, MA, USA). Genomic libraries were constructed from 1 µg of DNA using a TruSeq DNA PCR Free kit (Illumina®, San Diego, CA, USA). The libraries were sequenced on the NovaSeq6000 platform (Illumina®, San Diego, CA, USA) with an average depth of 44×. Read mapping and genome assembly were performed based on the GRCh38 version of the human genome using DRAGEN Germline App v3.7.5 available in the BaseSpace cloud file deposit (Illumina, USA). Variant calling was performed using the DRAGEN tool™. The SNP chosen genotypes were obtained from VCF files and checked manually from BAM files using the Integrative Genomics Viewer (IGV) tool (Broad Institute, University of California, USA) [24]. Hardy–Weinberg equilibrium was tested using the patch “genetics” of R software v3.02 (Vienna, Austria).

Genotypes from European, African, American, East Asian, and South Asian populations were obtained from the 1000 Genomes database using the Ensembl home browser (www.ensembl.org, accessed on 1 October 2022). The genetic distances based on molecular variance (F_{ST}) were calculated using Arlequin software v3.5 (Bern, Switzerland) [25]. The principal component analysis (PCA) was performed using Past3 software (Oslo, Norway), and the phylogenetic tree was constructed using Interactive Tree Of Life (iTOL) v6 [26] (available at <https://itol.embl.de>, accessed on 21 January 2023). This study was approved by the Local Research Ethical Committee (76189817.0.0000.5235 and 69253417.1.0000.5149) according to the ethical standards of the Helsinki Declaration. All participants signed a written consent form to participate.

3. Results

Following the genetic distance matrix based on population genotypes [27], we analyzed an additional larger cohort of Brazilian first-division soccer players under 20 years of age ($n = 44$, U20_BFDSC) to compare the new genetic data with another cohort of professional players ($n = 25$, BFDSC) and 2503 genomes from the 1000 Genomes database (European = 502, African = 661, American = 347, East Asian = 504, South Asian = 489). Unlike traditional genetic association studies that study SNPs individually, we used a sports performance-related SNP panel in a genetic distance approach based on Wright’s statistics [23].

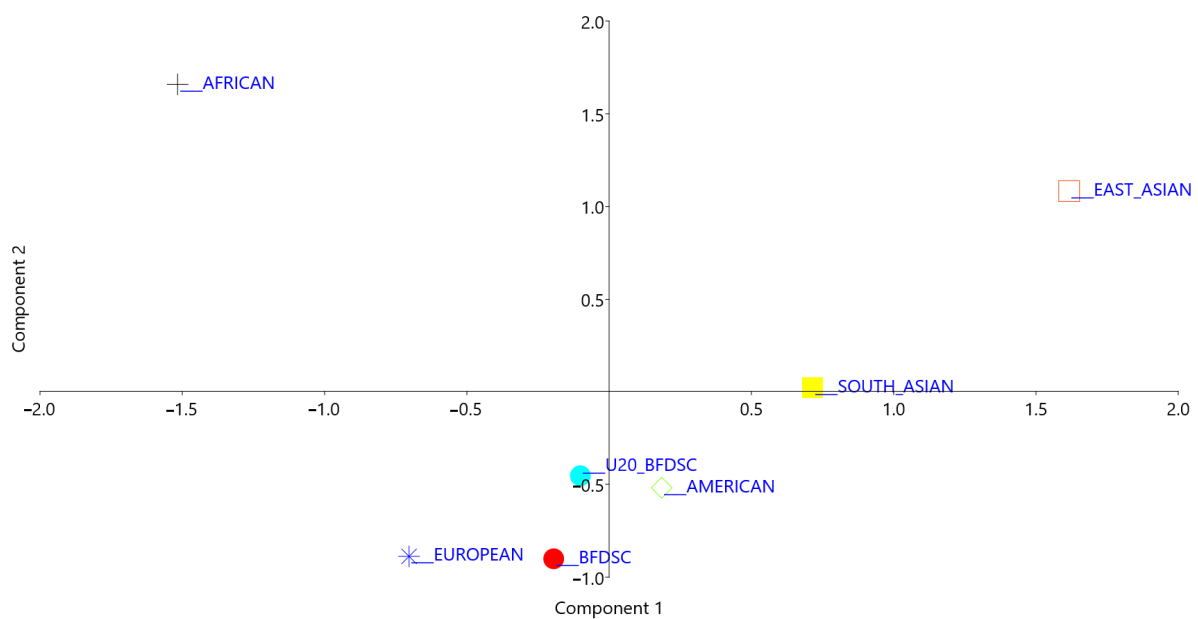
After testing for replicability, our results corroborated the previous minor cohort of 25 participants [1]. Although ancestry studies based on genome-wide arrays have shown an important similarity between south-eastern Brazilian and European, African, and Native American populations, the panel of 10 specific performance-related SNPs used in this study revealed very low F_{ST} values between U20_BFDSC and BFDSC, European, and American populations (Table 2). Larger F_{ST} and highly significant values were observed when distance matrices based on the U20_BFDSC genotypes were compared with those of East or South Asian and African populations ($p < 0.00001$) (Table 2). F_{ST} values between 0.05 and 0.15 were classified as moderate, and above 0.15 as high genetic distance.

In the multidimensional PCA analysis, the U20_BFDSC group clustered closely and in the same quadrant as the European and the previous BFDSC group population, near and in different quadrants from the American population, but far from the African or East Asian populations (Figure 1). The PCA was performed using the eigenvalue strategy for variance analysis. The first (58.4%) and the second (32.5%) components accounted for around 91% of the total variance, confirming the significance of the F_{ST} values and the statistical robustness of the model.

Table 2. F_{ST} values for comparisons between populations from 1000 Genomes database and U20_BFDSC.

Groups Compared to U20_BFDSC	F_{ST}
BFDSC	0.00100 n.s.
AFRICAN	0.10125 ***
AMERICAN	0.03097 ***
EAST_ASIAN	0.13803 ***
EUROPEAN	0.01534 ***
SOUTH_ASIAN	0.07152 ***

n.s. not significant; *** $p < 0.00001$. Abbreviations: BFDSC = Brazilian professional first-division soccer club, U20_BFDSC = Brazilian first-division soccer club players under 20 years of age, F_{ST} = genetic molecular variances. The molecular genetic variance between populations ($F_{ST} < 0.05$ —low, $F_{ST} = 0.05$ – 0.15 —moderate, $F_{ST} > 0.15$ —high).

**Figure 1.** Principal component analysis regarding the distance matrix of molecular variance (F_{ST}) values of continental populations from the 1000 Genomes database compared with U20_BFDSC and BFDSC.

No statistically significant genetic molecular variances (F_{ST}) were observed between the U20_BFDSC and BFDSC groups, and they remained very close on the phylogenetic tree (Figure S1), supporting replication of the results. In addition, we cannot ignore the possibility that linkage disequilibrium of non-genotyped SNPs may contribute to the clustering results obtained through the ten chosen markers.

4. Discussion

GASs can provide valuable insights, but achieving statistical confidence requires large sample sizes. To evaluate GAS results, a significance level of 5×10^{-8} based on classical Bonferroni's correction is typically used [28]. However, population genetics studies using Wright's statistics demonstrated significant group clustering with lower sampling. For example, in non-human models, matrices of genetic distances based on genotypes could separate genetically close groups containing as many as 17, 26, or fewer individuals by employing appropriate sets of SNPs [27,28]. Theoretical studies based on simulations indicate that even small sample sizes can produce accurate and unbiased estimates of F_{ST} when using enough and related bi-allelic markers like SNPs [29].

The Brazilian population is highly diverse and admixed, with a significant ancestral influence from Caucasians (61–77%), Africans (12–24%), Native Americans (10–15%), and

Asians (10%) [30,31]. Surprisingly, previous cohorts [15] have shown lower African-related ancestry in sports-related genes, prompting a replicative investigation in a larger, independent cohort. Additionally, recent findings based on total genotype scores described that the genetic profile of elite youth soccer players depends on their maturity status, based on a comparison of ages around 11 and 17 years old [32,33]. Although Wright's statistics can distinguish slight genetic differences, the U20_BFDSC group was not different from the BFDSC group. It is reasonable to hypothesize that the U20 category may resemble professional athletes, as most of the intense genetic selection pressure likely occurs before this maturation stage.

The clustering ability of molecular variance may be used even in populations with different physical phenotypes to evaluate performance. Although Wright's statistics-based studies have a powerful tool for clustering, they are sometimes less informative because they show the global molecular variance and do not point to which specific gene plays a vital role in that phenotype. It is possible to cluster genetically related subjects or distinguish them from others, but further studies are necessary to validate a potential variant related to the phenotype. In this case, good science is mandatory in selecting homogeneous populations to identify the genetic trait. Furthermore, clustering related subjects from the same ethnicity and similar phenotypes in a multifactorial trait could be harder to distinguish and may demand more genes and DNA sequencing costs.

The study of performance-related variants could allow for the design of a supposedly optimal profile. However, selecting an athlete based on molecular variance closer to some overrepresented clusters raises ethical concerns. As a multifactorial trait, soccer performance depends on genetics and environmental aspects. Without significant validation, there is no evidence to support a genetics-based selection of subjects, which poses a real risk of data misinterpretation that could compromise the integrity of the sport.

5. Conclusions

Finally, our replication results highlighted the possibility that, in male elite soccer players, performance-related genes might undergo selective pressure due to physical performance and environmental, cognitive, and sociocultural factors, justifying the slightly different results concerning the African or Asian population in studies with markers of ancestry that are not related to sports performance. In addition, the U20 athletes may resemble professional athletes, although even younger players could show genetic differences between them. The possibility of using Wright's statistics in larger samples, based on variability data from athletes' whole genomes (consortia), could generate conclusions never seen before in exercise science.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes14071446/s1>, Figure S1. Phylogenetic tree (UPGMA model) regarding genetic distance distribution from genotypes of the single nucleotide polymorphisms selected from U20_BFDSC, BFDSC, and continental populations from the "1000 Genomes database". U20 indicates under 20 years; BFDSC indicates Brazilian first-division soccer club.

Author Contributions: Conceptualization, T.K., C.G.M.S., M.d.A., G.M. and E.P.; methodology, C.G.M.S.; validation, M.C.-J., A.C.P.d.F., V.M.Z. and F.M.; formal analysis, C.G.M.S.; investigation, T.K. and M.d.A.; resources, T.K.; data curation, M.W.P.; writing—original draft preparation, C.G.M.S.; writing—review and editing, C.G.M.S., M.C.-J., A.C.P.d.F., V.M.Z., F.M., G.M. and M.W.P.; visualization, C.G.M.S.; supervision, E.P.; project administration, E.P.; funding acquisition, T.K., M.d.A. and E.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES)—14/2021 and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ)—E-26/204.117/2021.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Brazilian Research Ethical Committee (76189817.0.0000.5235-approved on 14 December 2017 and 69253417.1.0000.5149-approved on 2 February 2018.).

Informed Consent Statement: All participants signed a written consent form to participate.

Data Availability Statement: The datasets are fully available on request.

Acknowledgments: The authors would like to thank the “Instituto Hermes Pardini” for analysis and technical support and the Brazilian Army Institute of Biology for performing the functional genetics analysis.

Conflicts of Interest: All authors have read and approved this manuscript, which contains original material only and has not been published elsewhere in whole or in part. There are no conflicts of interest, including those involving financial interests.

References

1. Monnerat-Cahli, G.; Paulúcio, D.; Moura Neto, R.; Silva, R.; Pompeu, F.; Budowle, B.; Santos, C. Letter to the Editor: A Genetic-Based Algorithm for Personalized Resistance Training. *Biol. Sport* **2017**, *1*, 27–29. [[CrossRef](#)] [[PubMed](#)]
2. Seto, J.T.; Quinlan, K.G.R.; Lek, M.; Zheng, X.F.; Garton, F.; MacArthur, D.G.; Hogarth, M.W.; Houweling, P.J.; Gregorevic, P.; Turner, N. ACTN3 Genotype Influences Muscle Performance through the Regulation of Calcineurin Signaling. *J. Clin. Investig.* **2013**, *123*, 4255–4263. [[CrossRef](#)] [[PubMed](#)]
3. Hong, F.; Pan, S.; Guo, Y.; Xu, P.; Zhai, Y. PPARs as Nuclear Receptors for Nutrient and Energy Metabolism. *Molecules* **2019**, *24*, 2545. [[CrossRef](#)]
4. Winkler, T.W.; Justice, A.E.; Graff, M.; Barata, L.; Feitosa, M.F.; Chu, S.; Czajkowski, J.; Esko, T.; Fall, T.; Kilpeläinen, T.O.; et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet.* **2015**, *11*, e1005378. [[CrossRef](#)] [[PubMed](#)]
5. Santos, C.G.M.; Pimentel-Coelho, P.M.; Budowle, B.; de Moura-Neto, R.S.; Dornelas-Ribeiro, M.; Pompeu, F.a.M.S.; Silva, R. The Heritable Path of Human Physical Performance: From Single Polymorphisms to the “next Generation”. *Scand. J. Med. Sci. Sports* **2016**, *26*, 600–612. [[CrossRef](#)]
6. Uffelmann, E.; Huang, Q.Q.; Munung, N.S.; de Vries, J.; Okada, Y.; Martin, A.R.; Martin, H.C.; Lappalainen, T.; Posthuma, D. Genome-Wide Association Studies. *Nat. Rev. Methods Primers* **2021**, *1*, 59. [[CrossRef](#)]
7. MacArthur, D.G.; North, K.N. ACTN3: A Genetic Influence on Muscle Function and Athletic Performance. *Exerc. Sport Sci. Rev.* **2007**, *35*, 30–34. [[CrossRef](#)]
8. Hafen, P.S.; Law, A.S.; Matias, C.; Miller, S.G.; Brault, J.J. Skeletal Muscle Contraction Kinetics and AMPK Responses Are Modulated by the Adenine Nucleotide Degrading Enzyme AMPD1. *J. Appl. Physiol.* **2022**, *133*, 1055–1066. [[CrossRef](#)]
9. Gineviciene, V.; Jakaitiene, A.; Aksenov, M.; Aksenova, A.; Druzhevskaya, A.; Astratenkova, I.; Egorova, E.; Gabdrakhmanova, L.; Tubelis, L.; Kucinskas, V.; et al. Association Analysis of ACE, ACTN3 and PPARGC1A Gene Polymorphisms in Two Cohorts of European Strength and Power Athletes. *Biol. Sport* **2016**, *33*, 199–206. [[CrossRef](#)]
10. Kikuchi, N.; Fuku, N.; Matsumoto, R.; Matsumoto, S.; Murakami, H.; Miyachi, M.; Nakazato, K. The Association Between MCT1 T1470A Polymorphism and Power-Oriented Athletic Performance. *Int. J. Sports Med.* **2016**, *38*, 76–80. [[CrossRef](#)]
11. Heffernan, S.M.; Kilduff, L.P.; Erskine, R.M.; Day, S.H.; Stebbings, G.K.; Cook, C.J.; Raleigh, S.M.; Bennett, M.A.; Wang, G.; Collins, M.; et al. COL5A1 Gene Variants Previously Associated with Reduced Soft Tissue Injury Risk Are Associated with Elite Athlete Status in Rugby. *BMC Genom.* **2017**, *18*, 29–37. [[CrossRef](#)]
12. Larruskain, J.; Celorrio, D.; Barrio, I.; Odriozola, A.; Gil, S.M.; Fernandez-Lopez, J.R.; Nozal, R.; Ortuzar, I.; Lekue, J.A.; Aznar, J.M. Genetic Variants and Hamstring Injury in Soccer: An Association and Validation Study. *Med. Sci. Sports Exerc.* **2018**, *50*, 361–368. [[CrossRef](#)] [[PubMed](#)]
13. Gill, R.; Stratigopoulos, G.; Lee, J.H.; Leibel, R.L. Functional Genomic Characterization of the FTO Locus in African Americans. *Physiol. Genom.* **2019**, *51*, 517–528. [[CrossRef](#)] [[PubMed](#)]
14. van de Vegte, Y.J.; Tegegne, B.S.; Verweij, N.; Snieder, H.; van der Harst, P. Genetics and the Heart Rate Response to Exercise. *Cell. Mol. Life Sci.* **2019**, *76*, 2391–2409. [[CrossRef](#)]
15. Monnerat, G.; Maior, A.S.; Tannure, M.; Back, L.K.F.C.; Santos, C.G.M. Single-Nucleotide-Polymorphism-Panel Population-Genetics Approach Based on the 1000 Genomes Database and Elite Soccer Players. *Int. J. Sports Physiol. Perform.* **2019**, *14*, 711–717. [[CrossRef](#)]
16. San-Millán, I.; Stefanoni, D.; Martinez, J.L.; Hansen, K.C.; D’Alessandro, A.; Nemkov, T. Metabolomics of Endurance Capacity in World Tour Professional Cyclists. *Front. Physiol.* **2020**, *11*, 578. [[CrossRef](#)] [[PubMed](#)]
17. Khoramipour, K.; Gaeini, A.A.; Shirzad, E.; Gilany, K.; Chashniam, S.; Sandbakk, Ø. Metabolic Load Comparison between the Quarters of a Game in Elite Male Basketball Players Using Sport Metabolomics. *Eur. J. Sport Sci.* **2021**, *21*, 1022–1034. [[CrossRef](#)] [[PubMed](#)]
18. Schraner, D.; Kastenmüller, G.; Schönfelder, M.; Römisch-Margl, W.; Wackerhage, H. Metabolite Concentration Changes in Humans After a Bout of Exercise: A Systematic Review of Exercise Metabolomics Studies. *Sports Med. Open* **2020**, *6*, 11. [[CrossRef](#)] [[PubMed](#)]
19. Khoramipour, K.; Gaeini, A.A.; Shirzad, E.; Gilany, K.; Chamari, K.; Sandbakk, Ø. Using Metabolomics to Differentiate Player Positions in Elite Male Basketball Games: A Pilot Study. *Front. Mol. Biosci.* **2021**, *8*, 639786. [[CrossRef](#)] [[PubMed](#)]

20. Monnerat, G.; Sánchez, C.A.R.; Santos, C.G.M.; Paulucio, D.; Velasque, R.; Evaristo, G.P.C.; Evaristo, J.A.M.; Nogueira, F.C.S.; Domont, G.B.; Serrato, M.; et al. Different Signatures of High Cardiorespiratory Capacity Revealed with Metabolomic Profiling in Elite Athletes. *Int. J. Sports Physiol. Perform.* **2020**, *15*, 1156–1167. [[CrossRef](#)]
21. Cao, B.; Liu, S.; Yang, L.; Chi, A. Changes of Differential Urinary Metabolites after High-Intensive Training in Teenage Football Players. *Biomed. Res. Int.* **2020**, *2020*, 2073803. [[CrossRef](#)] [[PubMed](#)]
22. Al-Khelaifi, F.; Diboun, I.; Donati, F.; Botrè, F.; Alsayrafi, M.; Georgakopoulos, C.; Suhre, K.; Yousri, N.A.; Elrayess, M.A. A Pilot Study Comparing the Metabolic Profiles of Elite-Level Athletes from Different Sporting Disciplines. *Sports Med. Open* **2018**, *4*, 2. [[CrossRef](#)] [[PubMed](#)]
23. Hellwege, J.N.; Keaton, J.M.; Giri, A.; Gao, X.; Velez Edwards, D.R.; Edwards, T.L. Population Stratification in Genetic Association Studies. *Curr. Protoc. Hum. Genet.* **2017**, *95*, 1.22.1–1.22.23. [[CrossRef](#)]
24. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant Review with the Integrative Genomics Viewer. *Cancer Res.* **2017**, *77*, e31–e34. [[CrossRef](#)]
25. Excoffier, L.; Lischer, H.E.L. Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Mol. Ecol. Resour.* **2010**, *10*, 564–567. [[CrossRef](#)] [[PubMed](#)]
26. Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)]
27. Cornetti, L.; Tschirren, B. Combining Genome-wide Association Study and F_{ST}-based Approaches to Identify Targets of Borrelia-mediated Selection in Natural Rodent Hosts. *Mol. Ecol.* **2020**, *29*, 1386–1397. [[CrossRef](#)] [[PubMed](#)]
28. Yang, S.; Li, X.; Li, K.; Tang, Z. A Genome-Wide Scan for Signatures of Selection in Chinese Indigenous and Commercial Pig Breeds. *BMC Genet.* **2014**, *15*, 7. [[CrossRef](#)]
29. Willing, E.-M.; Dreyer, C.; van Oosterhout, C. Estimates of Genetic Differentiation Measured by F_{ST} Do Not Necessarily Require Large Sample Sizes When Using Many SNP Markers. *PLoS ONE* **2012**, *7*, e42649. [[CrossRef](#)]
30. Barbosa, F.B.; Cagnin, N.F.; Simioni, M.; Farias, A.A.; Torres, F.R.; Molck, M.C.; Araujo, T.K.; Gil-Da-Silva-Lopes, V.L.; Donadi, E.A.; Simões, A.L. Ancestry Informative Marker Panel to Estimate Population Stratification Using Genome-Wide Human Array. *Ann. Hum. Genet.* **2017**, *81*, 225–233. [[CrossRef](#)]
31. Andrade, R.B.; Amador, M.A.T.; Cavalcante, G.C.; Leitão, L.P.C.; Fernandes, M.R.; Modesto, A.A.C.; Moreira, F.C.; Khayat, A.S.; Assumpção, P.P.; Ribeiro-Dos-Santos, A.; et al. Estimating Asian Contribution to the Brazilian Population: A New Application of a Validated Set of 61 Ancestry Informative Markers. *G3 Genes Genomes Genet.* **2018**, *8*, 3577–3582. [[CrossRef](#)] [[PubMed](#)]
32. Murtagh, C.F.; Brownlee, T.E.; Rienzi, E.; Roquero, S.; Moreno, S.; Huertas, G.; Lugiuratto, G.; Baumert, P.; Turner, D.C.; Lee, D.; et al. The Genetic Profile of Elite Youth Soccer Players and Its Association with Power and Speed Depends on Maturity Status. *PLoS ONE* **2020**, *15*, e0234458. [[CrossRef](#)] [[PubMed](#)]
33. Hall, E.C.R.; Baumert, P.; Larruskain, J.; Gil, S.M.; Lekue, J.A.; Rienzi, E.; Moreno, S.; Tannure, M.; Murtagh, C.F.; Ade, J.D.; et al. The Genetic Association with Injury Risk in Male Academy Soccer Players Depends on Maturity Status. *Scand. J. Med. Sci. Sports* **2022**, *32*, 338–350. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.