Original Investigation

Are GWAS-identified SNPs associated with sprint athletic status? A replication study with three different cohorts

Running head: Gene polymorphisms for top-level sprinters

João Paulo L. F. Guilherme[1], Ekaterina A. Semenova[2,3], Hirofumi Zempo[4], Gabriel L. Martins[1], Antonio H. Lancha Junior[1], Eri Miyamoto-Mikami[5], Hiroshi Kumagai[5,6], Takuro Tobina[7], Keisuke Shiose[8], Ryo Kakigi[9], Takamasa Tsuzuki[10], Noriko Ichinoseki-Sekine[5,11], Hiroyuki Kobayashi[12], Hisashi Naito[5], Oleg V. Borisov[2,13], Elena S. Kostryukova[2], Nikolay A. Kulemin[2], Andrey K. Larin[2], Edward V. Generozov[2], Noriyuki Fuku[5], Ildus I. Ahmetov[2,14,15,16]

[1] Laboratory of Applied Nutrition and Metabolism, School of Physical Education and Sport, University of São Paulo, São Paulo, Brazil; [2] Department of Molecular Biology and Genetics, Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia; [3] Department of Biochemistry, Kazan Federal University, Kazan, Russia; [4] Faculty of Health and Nutrition, Tokyo Seiei College, Tokyo, Japan; [5] Graduate School of Health and Sports Science, Juntendo University, Chiba, Japan; [6] Research Fellow of Japanese Society for the Promotion of Science, Tokyo, Japan; [7] Faculty of Nursing and Nutrition, University of Nagasaki, Nagasaki, Japan; [8] Faculty of Education, University of Miyazaki, Miyazaki, Japan; [9] Faculty of Medicine, Juntendo University, Tokyo, Japan; [10] Faculty of Pharmacy, Meijo University, Aichi, Japan; [11] Faculty of Liberal Arts, The Open University of Japan, Chiba, Japan; [12] Department of General Medicine, Mito Medical Center, Tsukuba University Hospital, Ibaraki, Japan; [13] Institute for Genomic Statistics and Bioinformatics, University Hospital Bonn, Bonn, Germany; [14] Laboratory of Molecular Genetics, Kazan State Medical University, Kazan, Russia; [15] Sports Genetics Laboratory, St Petersburg Research Institute of Physical Culture, St Petersburg, Russia; [16] Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom.

Corresponding author: João Paulo Limongi França Guilherme, School of Physical Education and Sport – University of Sao Paulo, Professor Mello Moraes Avenue 65, São Paulo 05508-030, SP, Brazil. Phone: +55 11 30913096. E-mail: jplfguilherme@hotmail.com

Abstract word count: 250 words

Text-only word count: 3,943 words

Number of figures and tables: 1 figure and 3 tables

Abstract

***Purpose*:** This study was aimed to replicate previously GWAS-identified sprint-related polymorphisms in three different cohorts of top-level sprinters and to further validate obtained results in functional studies. ***Methods*:** A total of 240 Japanese, 290 Russians and 593 Brazilians were evaluated in a case–control approach. Of these, 267 were top-level sprint/power athletes. In addition, the relationship between selected polymorphisms and muscle fiber composition was evaluated in 211 Japanese and 287 Finnish individuals. ***Results*:** The G-allele of the rs3213537 polymorphism was overrepresented in Japanese (OR: 2.07, $P = 0.024$) and Russian (OR: 1.93, $P = 0.027$) sprinters compared to endurance athletes and associated with increased proportion of fast-twitch muscle fibers in Japanese ($P = 0.02$) and Finnish ($P = 0.041$) individuals. Meta-analysis of data from the four cohorts confirmed that the presence of the G/G genotype rather than G/A+A/A genotypes increased the odds ratio of being a sprinter compared to controls (OR: 1.54, $P = 0.005$), endurance athletes (OR: 1.79, $P = 0.001$) or controls + endurance athletes (OR: 1.61, $P = 0.001$). Furthermore, male sprinters with the G/G genotype were found to have significantly faster personal times in the 100-m dash than those with G/A+A/A genotypes (10.50 ± 0.26 vs. 10.76 ± 0.31, $P = 0.014$). ***Conclusion*:** The rs3213537 polymorphism found in the *CPNE5* gene was identified as a highly replicable variant associated with sprinting ability and increased proportion of fast-twitch muscle fibers, in which the homozygous genotype for the major allele (i.e., the G/G genotype) is preferable for performance.

Keywords: athletes; copine-V; genetics; sprint performance; synaptic plasticity

Introduction

The sprint ability is a core capacity that underlies performance in many individual sports as well as team sports. Naturally, pure sprint athletes (e.g., 100-m runners) perform better on physiological and mechanical variables of sprint performance [1]. A velocity-oriented force–velocity profile is a major contributing factor for a better sprint performance [2]. The maximal sprint velocity and mean power produced over the event distance strongly influence performance [3]. During a sprint task, power output demand increases exponentially with velocity and the best sprinters accelerate over a longer distance than their lower performing counterparts [4].

Sports performance is the combined result of numerous intrinsic and extrinsic factors, that is, the interaction between genetic factors and the environmental stimulus. Although training and other environmental stimulus are critical to performance achievement, individual performance thresholds can be determined by our genetic make-up. Twin studies have reported moderate to high heritability estimates for maximum movement speed as well as for other sprint and power phenotypes [5, 6] and so it has been proposed that elite sprint performance strongly depends on genetic characteristics.

Like other sports phenotypes, the sprint ability is a complex and polygenic phenomenon guided by the interaction of multiple genes and most likely gene variants. There are several polymorphisms that have been associated with elite power and sprint athletic status [7]. In particular, some of them were also associated with faster sprint times [8, 9]; however, many of the polymorphisms suggested as favorable to sprinters were evaluated using case–control approaches that have not yet been replicated in subsequent studies or independent samples [7]. Replication studies are of paramount importance to better evaluate and characterize performance-relevant polymorphisms. The same association in independent samples indicates a greater relevance between the polymorphism and the target phenotype.

Recently, Pickering et al. [10] first performed a genome-wide association study (GWAS) to identify sprint-related genetic variants. These authors exposed a set of new polymorphisms associated with short-distance sprints in youth football players, some of which were replicated in an independent cohort of Polish women. The replication of these findings in top-level athletes of different ethnicities would be interesting, since only one cohort of Russian athletes validated the most associated polymorphisms.

Therefore, the purpose of this study was to replicate GWAS-identified sprint-related polymorphisms in three different cohorts of top-level sprinters. A secondary purpose of this study was to evaluate the relationship between these polymorphisms and the proportion of fast-twitch muscle fibers in two different cohorts. First, the selected polymorphisms were evaluated for sprinter athletic status and proportion of fast-twitch muscle fibers in a Japanese cohort. Subsequently, the most consistent polymorphism was evaluated for sprinter athletic status in two other cohorts from Russia and Brazil, and evaluated for proportion of fast-twitch muscle fibers in Finnish individuals. Since the target phenotype is the sprint ability, sprinters were compared to non-athletes (controls) or endurance athletes (the metabolic demands required to perform sprint or endurance events are opposites of each other).

Methods

Table 1 shows the polymorphisms selected for use in this study. All of them are single nucleotide polymorphisms (SNPs) and were selected based on a previous study [10] and according to the following criteria: biallelic polymorphisms located on autosomal chromosomes, two replications in the initial study and minor allele frequency > 1% in the Japanese population. Although the rs12688220 and rs8064257 polymorphisms also showed two replications in the initial study, they were not included because they did not meet the inclusion criteria.

All cohorts included in this study had their procedures conducted according to the Declaration of Helsinki ethical principles for research involving human subjects. The Japanese studies were approved by the ethics committee of the Juntendo University and Fukuoka University. The Finnish study was approved by the coordinating ethics committee of the Hospital District of Helsinki and Uusimaa (this data was used with permission; Database of Genotypes and Phenotypes (dbGaP) Study Accession: phs000867.v1.p1). The Russian study was approved by the ethics committee of the Federal Research and Clinical Center of Physical-chemical Medicine of the Federal Medical and Biological Agency of Russia. The Brazilian study was approved by the ethics committee of the School of Physical Education and Sport, University of Sao Paulo, São Paulo, Brazil. A written informed consent was obtained from each participant.

The Japanese Cohort

The Japanese study involved 114 athletes (91 males and 23 females), of which 54 were sprint/power athletes (100-400 m runners, jumpers and throwers; mean age ± SD: 28 ± 7 years) and 60 endurance runners (800 m to marathon; mean age ± SD: 24 ± 3 years). All of these athletes were international-level competitors. The control group comprised 126 healthy Japanese individuals.

Total DNA was isolated from saliva or venous blood using the Oragene・DNA Collection Kit (DNA Genotek, Ontario, Canada) or the QIAamp DNA blood Maxi Kit (QIAGEN, Hilden, Germany), respectively. The total DNA content was measured using the NanoDrop 8000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Subsequently, DNA samples were adjusted to a concentration of 50 ng/μL with Tris-EDTA buffer and stored at 4°C. Total DNA samples were genotyped using the HumanOmniExpress Beadchip (Illumina, San Diego, CA, USA) to genotype > 700,000 SNPs, according to the manufacturer's instructions. Genotype calls were performed with Illumina GenomeStudio software and PLINK was used for quality control checks and association analyses.

Evaluation of skeletal muscle fiber types

First, a cohort of 203 Japanese healthy individuals (98 men and 105 women, with age range 20-79 years) performed muscle biopsy was used for the association study between sprint-related polymorphisms and muscle fiber composition. Muscle samples were obtained from the belly of the vastus lateralis and myosin heavy chain (MHC) isoforms were determined by performing glycerol SDS-PAGE, as previously described [11]. These individuals had their DNA samples isolated from venous blood and the polymorphisms were genotyped using the Japonica SNP array [12].

Second, muscle fiber composition in 287 Finnish individuals (167 men, age 59.5 ± 8.1 years; 120 women, age 60.7 ± 7.4 years) from the FUSION study was estimated based on the expression of the myosin heavy chain 1 (*MYH1*), myosin heavy chain 2 (*MYH2*), myosin heavy chain 7 (*MYH7*), Ca^{2+} ATPase A1 and Ca^{2+} ATPase A2 genes, as previously described [13]. Muscle samples were obtained from the vastus lateralis using a conchotome, under local anesthesia with 20 $mg \cdot ml^{-1}$ lidocaine hydrochloride without epinephrine [14]. DNA samples were extracted from the blood and the polymorphisms were genotyped using the HumanOmni2.5-4v1_H BeadChip array (Illumina, San Diego, CA, USA).

The Russian Cohort

The Russian study involved 173 athletes (99 males and 74 females; mean age ± SD: 31.3 ± 7.5 years), of which 70 were elite sprinters (100-400 m runners, 500-1000 m speed skaters, 50 m swimmers) and 103 elite endurance athletes (biathletes, rowers, cross-country skiers, 3-10 km runners, 800-1500 m swimmers and triathletes). All of these athletes were international-

level competitors, of which 30 (13 sprinters and 17 endurance athletes) were highly elite athletes (i.e., prize winners in international competitions). The control group comprised 117 healthy unrelated citizens (66 males and 51 females, mean age ± SD: 47.9 ± 4.8 years), without any competitive sport experience. This Russian cohort is independent of the one previously published [10].

Molecular genetic analysis was performed with DNA samples obtained from leukocytes (venous blood). Four millilitres of venous blood was collected in tubes containing EDTA (Vacuette EDTA tubes; Greiner Bio-One, Kremsmünster, Austria). Blood samples were transported to the laboratory at 4°C, and DNA was extracted on the same day. DNA extraction and purification were performed using a commercial kit according to the manufacturer's instructions (Technoclon, Moscow, Russia), which included chemical lysis, selective DNA binding on silica spin columns and ethanol washing. Extracted DNA quality was assessed by agarose gel electrophoresis. The genotyping process was performed using HumanOmni1-Quad BeadChips or HumanOmniExpress BeadChips (Illumina, San Diego, CA, USA) to genotype > 900,000 SNPs. The assay required 200 ng of DNA sample as input with a concentration of at least 50 ng/µl. Exact concentrations of DNA in each sample were measured using a Qubit Fluorometer (Invitrogen, Waltham, MA, USA). All further procedures were performed according to the instructions of the Infinium High-Density Assay.

The Brazilian Cohort

The Brazilian study involved 305 athletes (200 males and 105 females; mean age ± SD: 25.4 ± 6.9 years), of which 143 were elite sprinters (100-400 m runners, 50-200 m swimmers, canoeing and cycling) and 162 endurance athletes (rowers, > 1.5 km runners, 400-1500 m swimmers and triathletes). While 36% of these athletes were nationally prominent competitors, 64% were international-level competitors. The control group comprised 288 healthy Brazilian individuals (187 males and 101 females, mean age ± SD: 29.6 ± 8.1 years), without any competitive sport experience.

Genomic DNA of the Brazilian participants was isolated from buccal epithelial cells obtained from mouthwashes with a 0.9% saline solution prepared with DNA- and DNAse-free water as previously described [8]. Briefly, the DNA samples were extracted using chloroform, precipitated using ethanol and resuspended with 1× Tris-EDTA buffer. DNA quantification and quality assessment were performed using the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The A260/A280 ratio was used to evaluate the quality of the sample, which values between 1.7 and 2.1 were considered acceptable The genotyping process was performed using a pre-designed specific TaqMan® SNP Genotyping Assay (Applied Biosystems, Foster city, CA, USA), according to the manufacturer's instructions and using the Rotor-Gene Q PCR cycler (Qiagen, Hilden, Germany). A scatter plot showing the endpoint fluorescence signals (i.e., an increase in VIC or FAM fluorescent signal) was used to discriminate the genotypes. The transcript alleles were used similarly to that previously used.

Association with sprint performance

To further investigate the influence of sprint-related polymorphisms on sprint performance, a sample of 37 top-level 100-m runners (28 Brazilians and 9 Russians) had their personal best sprint running time in the 100-m dash at official events compared between the genotypes of the selected polymorphism. Athlete’s personal records were acquired using the International Association of Athletics Federations (IAAF) database, available online at https://www.worldathletics.org/athletes. Only athletes with performance data available on the IAAF database were included in the study.

Statistical analysis

First of all, the Chi-square test (χ^2) was used to test for the presence of the Hardy-Weinberg equilibrium (HWE) in each control group. A departure from HWE was observed when $\chi^2 > 3.84$ (i.e., $P > 0.05$). Thereafter, the frequencies of genotypes or alleles were compared between sprinters and ethnically-matched controls or ethnically-matched endurance athletes using the χ^2 test or Fisher's exact test when appropriate. Differences in the proportion of muscle fiber types between groups with different genotypes were analyzed using unpaired *t*-test and one-way ANOVA. The unpaired *t*-test was also used to evaluate the influence of the selected polymorphism on 100-m sprint performance. The significance level was established at $P < 0.05$.

For the pooled analysis of the Japanese, Russian and Brazilian cohorts, meta-analysis was conducted using the Review Manager (RevMan) computer program version 5.3 (*Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014*). Data from a previously published Russian case-control study [10] were also used in the meta-analysis, that is, data from 99 Russian highly elite athletes (i.e., Winners of World Championships, World Cups or Olympic Games; 43 sprinters and 56 endurance athletes) and 173 controls were also included in the meta-analysis. The DerSimonian and Laird random-effects model was used to calculate weighted odds ratio (OR) and its 95% confidence interval (95% CI). The test of overall effect was assessed using the Z score with the significance level established at $P < 0.05$. Heterogeneity between studies was assessed using the standard χ^2 test (Cochran Q test) and the I^2 statistic.

Results

Case–control association study

Of the three polymorphisms evaluated in the Japanese case-control study, only the rs3213537 of the copine 5 (*CPNE5*) gene was found to be significant. The G-allele of the rs3213537 polymorphism was overrepresented in Japanese sprinters compared to endurance athletes (80.6 vs. 66.7%; $P = 0.024$) and associated with increased proportion of glycolytic fast-twitch (IIx) muscle fibers in Japanese male controls (G/G (n = 69) 24.7 ± 9.4%, G/A (n = 22) 22.1 ± 6.9%, A/A (n = 7) 16.7 ± 7.0%; $P = 0.02$). A strong trend towards an increase in the proportion of type IIx fibers was observed in Japanese males even when the analysis was adjusted for age ($P = 0.061$). In addition, the G-allele of the rs3213537 polymorphism was also was associated with increased proportion of fast-twitch muscle fibers in 287 Finnish individuals adjusted for sex and age (G/G (n = 189) 55.8 ± 14.9%, G/A (n = 91) 54.1 ± 14.2%, A/A (n = 7) 43.1 ± 16.6%; $P = 0.041$). Based on these associations, the rs3213537 polymorphism was selected for replication in the Russian and Brazilian cohorts. Given the low frequency of the homozygous genotype for the minor allele (i.e., the A/A genotype), the rs3213537 polymorphism was analyzed only under the dominant model (G/G vs. G/A+A/A). Of note, the transcript alleles instead of the genomic alleles were used to facilitate the link between this study and previous data (discovery stage) [10], that is, the G/A alleles represent the genomic C/T alleles.

Table 2 shows the genotype distribution and allele frequency of the rs3213537 polymorphism in the three cohorts evaluated. Similar to that observed in the Japanese cohort, the G-allele was overrepresented in Russian sprinters compared to endurance athletes (86.4 vs. 76.7%; $P = 0.027$) or controls (86.4 vs. 77.8%; $P = 0.042$). Indeed, the G/G genotype was overrepresented in Russian sprinters compared to endurance athletes (74.3 vs. 56.3%; $P = 0.017$). In the Brazilian cohort, there was a difference of ≈ 5% of G/G genotype carriers between sprinters and the two other groups (controls and endurance athletes), but this

difference was not statistically significant. However, the direction of effect observed in the Brazilian cohort was the same as in the Japanese and Russian cohorts (OR > 1.2).

Meta-analysis

Meta-analysis showed that, in the pooled data of the Japanese, two Russian (including data from the previous study [10]) and Brazilian cohorts, the G-allele frequency was significantly higher in sprinters compared with controls (P = 0.004), endurance athletes (P = 0.002) or controls + endurance athletes (P = 0.002), as shown in Table 3. Indeed, presence of the G/G genotype rather than G/A+A/A genotypes increased the chance of being a top-level sprinter compared to controls (OR: 1.49, 95% CI: 1.10–2.01; P = 0.005), endurance athletes (OR: 1.79, 95% CI: 1.26–2.55; P = 0.001) or controls + endurance athletes (OR: 1.58, 95% CI: 1.19–2.10; P = 0.002). There was no evidence of heterogeneity between studies.

Sprint performance

Figure 1 shows the comparison of the personal best times in 100-m performance between male sprinters with the G/G genotype (n = 26) and male sprinters with the G/A+A/A genotypes (n = 11). Male sprinters with the G/G genotype have been found to have significantly faster personal times (10.50 ± 0.26 s vs. 10.76 ± 0.31 s, P = 0.014).

Discussion

This study aimed to replicate potential sprint-related polymorphisms recently identified by a GWAS in three independent cohorts of top-level sprinters, as well as to evaluate their relationship with the proportion of fast-twitch muscle fibers. The main finding of this investigation involving 1,875 subjects was that the G-allele of the rs3213537 polymorphism was more frequent in sprinters and associated with increased proportion of fast-twitch muscle fibers in Japanese and Finnish individuals and the 100-m sprint performance in Brazilian and Russian sprinters, particularly the homozygotes (i.e., carriers of the G/G genotype). Meta-analysis of 310 sprinters compared with 694 non-athletes and 381 endurance athletes showed that carriers of the G/G genotype were ≈ 1.6 times more likely to be a sprinter.

The rs3213537 is an intronic polymorphism found in the *CPNE5* gene located at the 6p21.2 region of the chromosome 6. Copines are a family of calcium-dependent, membrane-binding proteins that are evolutionary conserved from protozoans to humans [15]. Present in all major mammalian organs, copines may play fundamental roles in eukaryotic cell processes [16]. Copine proteins contain two N-terminal C2 domains that involve residues important for calcium and phospholipid binding and a C-terminal A domain that may be involved in protein–protein interactions [15]. This well characterized structure, especially the C2 domains, suggests their involvement in processes of signal transduction or membrane trafficking, which occurs in a calcium-dependent manner [16]. However, their biological roles have not yet been fully defined.

There are at least eight different human copine proteins, which were referred to using roman numerals. Some of them (copine-I, -II and -III) are ubiquitously expressed, while the others have a more restricted expression profile [17]. As an example, copine-VI is a cytosolic protein strongly expressed in hippocampal excitatory neurons that has been shown to affect the structural plasticity of the dendritic spine in response to presynaptic activity [18]. Synaptic calcium signals lead to copine-VI translocation from the cytosol to the postsynaptic spine membranes, where they can serves as a calcium sensor that links neuronal activity to the subsequent long-term changes in synaptic structure by altering actin cytoskeleton morphology [19]. It was shown that copine-VI is responsible for the recruitment and local activation of the Rac family small GTPase 1 (Rac1) protein, which, in turn, activates the Rac1-PAK-LIMK1-Cofilin pathway and cause actin re-arrangement in favor of the long-lasting, stable

strengthening of excitatory synapses [19]. The molecular events underlying copine-V (encoded by *CPNE5* gene) are less understood, however, there may be some resemblance to other copine proteins, such as copine-VI, as they are structurally highly similar. Nonetheless, they can be expressed in different brain regions or tissues and interact with different proteins.

Based on animal research, copine-V has been shown to play a key role in the development of the central nervous system as it is highly expressed during the embryonic brain development [20]. Its expression decreases dramatically in the adult brain, remaining expressed in some non-neural tissues such as the heart, lung and muscles [21]. Nevertheless, although its expression may be low in the cortex and almost undetectable in the cerebellum of the adult brain, *CPNE5* is moderately expressed in the striatum of adult mice that have learned a complex motor task [22]. Alterations in neuronal ensemble activity and synaptic plasticity of the striatum are highly relevant for efficient human motor actions because it is the foundation for long-term motor learning or motor memory [22, 23].

There is evidence supporting that a lack of motor memory may be detrimental to power and sprint performance [24]. Individuals with superior working memory are able to perform faster and more accurate in motor tasks due to a better neural efficiency [25]. Although with training, both neural activity and performance can be improved. Repetitive activation of the same neuronal circuit induces the clustering of new spines in postsynaptic membranes, favoring motor performance as it strengthens the dynamics of synaptic transmission [26]. Thus, the most effective neural communication favors sprint performance. There are synaptic inputs at the central and peripheral levels, directly influencing the rapid activation of muscles [27]. The ability of the neuromuscular system to increase contractile activity when muscle activation is intended to be performed as quickly as possible, referred to as Rate of Force Development (RFD), is considered vital for athletes requiring high-speed motor actions such as sprinters. Cross-sectional studies have shown that top-level sprint/power athletes are characterized by a markedly greater RFD [28]. Moreover, athletes with a higher RFD demonstrated faster sprint times [29]. Additional contributions may also occur due to differences in muscle fiber type composition—the RFD is faster in type II fibers [27].

Whether *CPNE5* rs3213537 mutant carriers have impaired motor memory or muscle recruitment ability remains to be established, but the homozygous genotype for the major allele (i.e., the G/G genotype) was associated with fast-twitch muscle fibers and faster times in the 100-m event, which is considered the standard measure of the sprint ability of human bipedal locomotion [2]. Based on its role in the central nervous system, *CPNE5* polymorphisms were previously associated with alcohol dependence and obesity [20]. In particular, the mutant allele of the *CPNE5* rs3213537 polymorphism was strongly associated with alcohol abuse [20], which adversely impacts athletic performance in a number of different ways, including mood instability and sensory-motor system dysfunction [30].

As mentioned earlier, the *CPNE5* rs3213537 is a gene variant occurring within an intron (genomic position and change: g.36748144C>T based on the Genome Reference Consortium Human Build 38). Introns harbour polymorphisms that can influence the expression of the genes that host them and modulate the genotype–phenotype relationship. Thus, this polymorphism may modulate *CPNE5* expression and its calcium-modulated signal transduction. Interestingly, the interaction between copines and membranes occurs at concentrations of calcium that are likely to occur in the cytosol of stimulated cells but not in resting cells [16], and therefore, calcium-regulated phenotypes may be affected by mutations in the *CPNE5* gene. Of particular interest, during neuromuscular junction formation, muscle fibers are intrinsically pre-specialized by clustering postsynaptic proteins, whereas the proper patterning of postsynaptic protein clusters in the center of developing muscle fibers and the subsequent innervation by the motor nerve critically depend on calcium signals [31].

Collectively, we speculate that the G/G genotype may be involved in synaptic plasticity and muscle fiber specificity in a way that favors sprint performance.

The present study has some limitations. Our muscle fiber composition study included only non-athlete individuals of a wide age range. However, if the polymorphism is associated with increased proportion of fast-twitch muscle fibers in untrained individuals, these individuals (carriers of the associated variant) are expected to respond better to sprint training. Power training, like that used by sprinters, seems to conserve the pre-training number of fast-twitch fibers while increasing their fiber cross-sectional area, particularly type IIx fibers [32], favouring a higher RFD [33]. Type IIx fibers have the highest muscle fiber conduction velocity [34] and are considered key determinants of the RFD, especially in power-trained individuals [33]. Power output in type IIx fibers was 2-fold higher than type IIa fibers and 14-fold greater than type I fibers [35]. In line with this, the G-allele of the rs3213537 polymorphism was previously associated with the 10-m performance in a cohort of untrained Polish women [10], as well as associated with the 100-m performance in elite athletes. Although our case-control study included metabolically similar athletes, the performance association study evaluated only runners. Additional studies evaluating other sprint-oriented disciplines will be interesting, given that there may be differences between sports disciplines.

Practical Applications

The GWAS represents a promising and productive way to study sports-related phenotypes by providing a number of new candidate polymorphisms—that need to be evaluated in independent cohorts of different ethnicities and using different methodological approaches to better assess the relationship between the polymorphisms and traits of interest. In this regard, collaborative efforts involving well characterized athlete cohorts of different ethnic backgrounds will be of critical importance for further progress. In the present study, based on data from different cohorts, it is plausible to assume that the rs3213537 polymorphism (G/G genotype) may be part of a favorable genetic profile for sprinters. Notwithstanding, it is important to emphasize that sports phenotypes are complex and polygenic phenomena and should therefore be interpreted with caution.

Conclusion

The G/G genotype of the *CPNE5* gene rs3213537 polymorphism was associated with sprint athletic status and performance. While the G-allele was associated with the proportion of fast-twitch muscle fibers in Japanese and Finnish individuals, the G/G genotype was associated with faster personal times in the 100-m sprint performance among elite athletes from Brazil and Russia. It is worth mentioning that a complex network of genes contributes to sports performance, and the *CPNE5* rs3213537 is just one of several variants that can make-up the genetic profile of the elite athlete.

Acknowledgments

The FUSION study was conducted by the FUSION GoT2D/T2D-GENES Investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The data from the FUSION study reported here were supplied by the University of Eastern Finland, University of Michigan, and Broad Institute. This manuscript was not prepared in collaboration with Investigators of the FUSION study and does not necessarily reflect the opinions or views of the FUSION study, or the NIDDK. The Russian study was supported in part by grant from the Russian Science Foundation (Grant No. 17-15-01436: “Comprehensive analysis of the contribution of genetic, epigenetic and environmental factors in the individual variability of the composition of human muscle fibers”). The Brazilian study was supported by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (São Paulo

Research Foundation), Brazil (FAPESP; Grant No. 2012/22516-6) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil (CNPq; Grant No. 301213/2015-1). The results of the current study do not constitute endorsement of the product by the authors or the journal.

References

1. Haugen TA, Breitschadel F, Seiler S. Sprint mechanical variables in elite athletes: Are force-velocity profiles sport specific or individual? *PLoS One.* 2019;14(7):e0215551.
2. Morin JB, Bourdin M, Edouard P, Peyrot N, Samozino P, Lacour JR. Mechanical determinants of 100-m sprint running performance. *Eur J Appl Physiol.* 2012;112(11):3921-3930.
3. Slawinski J, Termoz N, Rabita G, et al. How 100-m event analyses improve our understanding of world-class men's and women's sprint performance. *Scand J Med Sci Sports.* 2017;27(1):45-54.
4. Haugen T, Seiler S, Sandbakk O, Tonnessen E. The Training and Development of Elite Sprint Performance: an Integration of Scientific and Best Practice Literature. *Sports Med Open.* 2019;5(1):44.
5. Missitzi J, Geladas N, Klissouras V. Heritability in neuromuscular coordination: implications for motor control strategies. *Med Sci Sports Exerc.* 2004;36(2):233-240.
6. Missitzi J, Geladas N, Klissouras V. Genetic variation of maximal velocity and EMG activity. *Int J Sports Med.* 2008;29(3):177-181.
7. Maciejewska-Skrendo A, Sawczuk M, Cięszczyk P, Ahmetov I. Genes and Power Athlete Status. In: Barh D, Ahmetov I, ed. *Sports, Exercise, and Nutritional Genomics: Current Status and Future Directions*. Academic Press; 2019:41-72.
8. Guilherme JPLF, Silva MS, Bertuzzi R, Lancha Junior AH. The AGTR2 rs11091046 (A>C) Polymorphism and Power Athletic Status in Top-Level Brazilian Athletes. *J Sports Sci.* 2018;36(20):2327-2332.
9. Papadimitriou ID, Lucia A, Pitsiladis YP, et al. ACTN3 R577X and ACE I/D gene variants influence performance in elite sprinters: a multi-cohort study. *BMC Genomics.* 2016;17:285.
10. Pickering C, Suraci B, Semenova EA, et al. A Genome-Wide Association Study of Sprint Performance in Elite Youth Football Players. *J Strength Cond Res.* 2019;33(9):2344-2351.
11. Kumagai H, Tobina T, Ichinoseki-Sekine N, et al. Role of selected polymorphisms in determining muscle fiber composition in Japanese men and women. *J Appl Physiol (1985).* 2018;124(5):1377-1384.
12. Kawai Y, Mimori T, Kojima K, et al. Japonica array: improved genotype imputation by designing a population-specific SNP array with 1070 Japanese individuals. *J Hum Genet.* 2015;60(10):581-587.
13. Taylor DL, Jackson AU, Narisu N, et al. Integrative analysis of gene expression, DNA methylation, physiological traits, and genetic variation in human skeletal muscle. *Proc Natl Acad Sci U S A.* 2019;116(22):10883-10888.
14. Scott LJ, Erdos MR, Huyghe JR, et al. The genetic regulatory signature of type 2 diabetes in human skeletal muscle. *Nat Commun.* 2016;7:11764.
15. Creutz CE, Tomsig JL, Snyder SL, et al. The copines, a novel class of C2 domain-containing, calcium-dependent, phospholipid-binding proteins conserved from Paramecium to humans. *J Biol Chem.* 1998;273(3):1393-1402.
16. Tomsig JL, Creutz CE. Biochemical characterization of copine: a ubiquitous Ca2+-dependent, phospholipid-binding protein. *Biochemistry.* 2000;39(51):16163-16175.

17. Cowland JB, Carter D, Bjerregaard MD, Johnsen AH, Borregaard N, Lollike K. Tissue expression of copines and isolation of copines I and III from the cytosol of human neutrophils. *J Leukoc Biol.* 2003;74(3):379-388.
18. Burk K, Ramachandran B, Ahmed S, et al. Regulation of Dendritic Spine Morphology in Hippocampal Neurons by Copine-6. *Cereb Cortex.* 2018;28(4):1087-1104.
19. Reinhard JR, Kriz A, Galic M, et al. The calcium sensor Copine-6 regulates spine structural plasticity and learning and memory. *Nat Commun.* 2016;7:11613.
20. Wang KS, Zuo L, Pan Y, Xie C, Luo X. Genetic variants in the CPNE5 gene are associated with alcohol dependence and obesity in Caucasian populations. *J Psychiatr Res.* 2015;71:1-7.
21. Ding X, Jin Y, Wu Y, et al. Localization and cellular distribution of CPNE5 in embryonic mouse brain. *Brain Res.* 2008;1224:20-28.
22. D'Amours G, Bureau G, Boily MJ, Cyr M. Differential gene expression profiling in the mouse brain during motor skill learning: focus on the striatum structure. *Behav Brain Res.* 2011;221(1):108-117.
23. Wachter T, Rohrich S, Frank A, et al. Motor skill learning depends on protein synthesis in the dorsal striatum after training. *Exp Brain Res.* 2010;200(3-4):319-323.
24. Nanci Maria F, Eric D, Mario B, Emmanuel VP. Comparison of peak muscle power between Brazilian and French girls. *Am J Hum Biol.* 2002;14(3):364-371.
25. Behmer Jr. LP, Fournier LR. Working memory modulates neural efficiency over motor components during a novel action planning task: an EEG study. *Behav Brain Res.* 2014;260:1-7.
26. Fu M, Yu X, Lu J, Zuo Y. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. *Nature.* 2012;483(7387):92-95.
27. Maffiuletti NA, Aagaard P, Blazevich AJ, Folland J, Tillin N, Duchateau J. Rate of force development: physiological and methodological considerations. *Eur J Appl Physiol.* 2016;116(6):1091-1116.
28. Tillin NA, Jimenez-Reyes P, Pain MT, Folland JP. Neuromuscular performance of explosive power athletes versus untrained individuals. *Med Sci Sports Exerc.* 2010;42(4):781-790.
29. Tillin NA, Pain MT, Folland J. Explosive force production during isometric squats correlates with athletic performance in rugby union players. *J Sports Sci.* 2013;31(1):66-76.
30. McDuff DR, Baron D. Substance use in athletics: a sports psychiatry perspective. *Clin Sports Med.* 2005;24(4):885-897, ix-x.
31. Kaplan MM, Sultana N, Benedetti A, et al. Calcium Influx and Release Cooperatively Regulate AChR Patterning and Motor Axon Outgrowth during Neuromuscular Junction Formation. *Cell Rep.* 2018;23(13):3891-3904.
32. Terzis G, Spengos K, Methenitis S, Aagaard P, Karandreas N, Bogdanis G. Early phase interference between low-intensity running and power training in moderately trained females. *Eur J Appl Physiol.* 2016;116(5):1063-1073.
33. Methenitis S, Spengos K, Zaras N, et al. Fiber Type Composition and Rate of Force Development in Endurance- and Resistance-Trained Individuals. *J Strength Cond Res.* 2019;33(9):2388-2397.
34. Methenitis S, Karandreas N, Spengos K, Zaras N, Stasinaki AN, Terzis G. Muscle Fiber Conduction Velocity, Muscle Fiber Composition, and Power Performance. *Med Sci Sports Exerc.* 2016;48(9):1761-1771.
35. Trappe S, Luden N, Minchev K, Raue U, Jemiolo B, Trappe TA. Skeletal muscle signature of a champion sprint runner. *J Appl Physiol (1985).* 2015;118(12):1460-1466.

Figure caption

Figure 1 Association between the rs3213537 polymorphism and the best 100-m personal time in Brazilian and Russian male sprinters. The dashed line represents the qualifying time for the Tokyo 2020 Olympic Games (10.05 s).

Table 1 Description of the polymorphisms evaluated in the present study

Polymorphism	Location (position†)	Consequence	REF/ALT	Sprint Allele	Previous association
rs3213537	Chromosome 6 (36748144)	Intron variant	G/A	G (major allele)	SFP, SPW, AS
rs1929877	Chromosome 9 (78799771)	Intergenic variant	A/G	G (minor allele)	SFP, AS, MF
rs17347590	Chromosome 20 (48525214)	Intergenic variant	C/A	C (major allele)	SFP, SPW, MF

Legend: REF, Reference allele; ALT, Alternate allele; SFP, Associated with sprint performance in young British football players; SPW, Associated with sprint performance in healthy young Polish women; AS, Associated with sprint/power athlete status; MF, Associated with proportion of fast-twitch muscle fibers in Russian physically active subjects. †Genomic position based on GRCh38 (Genome Reference Consortium Human Build 38).

524 **Table 2** Genotype distribution and allele frequency of the rs3213537 polymorphism in the
525 Japanese, Russian and Brazilian cohorts

		Genotypes (%)				Comparisons: *P*-value (Effect Direction)	
Group	*n*	G/G	G/A	A/A	G allele	G/G vs. G/A+A/A	Alleles (G vs. A)
Japanese sprint/power athletes	54	63.0	35.2	1.9	80.6	1.000	1.000
Japanese endurance athletes	60	45.0	43.3	11.7	66.7	0.062 (OR: 2.08)	0.024 (OR: 2.07)
Japanese controls	116	57.8	35.3	6.9	75.4	0.615 (OR: 1.24)	0.334 (OR: 1.35)
Russian sprinters	70	74.3	24.3	1.4	86.4	1.000	1.000
Russian endurance athletes	103	56.3	40.8	2.9	76.7	0.017 (OR: 2.24)	0.027 (OR: 1.93)
Russian controls	117	62.4	30.8	6.8	77.8	0.110 (OR: 1.74)	0.042 (OR: 1.82)
Brazilian sprinters	143	73.4	23.8	2.8	85.3	1.000	1.000
Brazilian endurance athletes	162	68.5	28.4	3.1	82.7	0.378 (OR: 1.27)	0.439 (OR: 1.21)
Brazilian controls	288	67.7	29.5	2.8	82.5	0.266 (OR: 1.32)	0.331 (OR: 1.23)

526 Underlined values indicate an association trend ($0.05 < P < 0.07$), and double underlined values
527 indicate nominal associations ($P < 0.05$). Legend: OR, Odds Ratio.
528

Table 3 Meta-analysis of the association between the rs3213537 polymorphism and sprinter athlete status

Comparison	Model	OR (95% CI)	Heterogeneity	Test for overall effect
Sprint/Power athletes *vs*. Controls	G/G vs. G/A+A/A	1.49 (1.10–2.01)	χ^2 = 1.94 (*P* = 0.58); I^2 = 0%	Z = 2.57 (*P* = 0.01)
	Alleles (G vs. A)	1.47 (1.13–1.92)	χ^2 = 2.91 (*P* = 0.41); I^2 = 0%	Z = 2.85 (*P* = 0.004)
Sprint/Power athletes *vs*. Endurance athletes	G/G vs. G/A+A/A	1.79 (1.26–2.55)	χ^2 = 3.30 (*P* = 0.35); I^2 = 9%	Z = 3.25 (*P* = 0.001)
	Alleles (G vs. A)	1.70 (1.22–2.35)	χ^2 = 3.70 (*P* = 0.30); I^2 = 19%	Z = 3.16 (*P* = 0.002)
Sprint/Power athletes *vs*. Controls + Endurance athletes	G/G vs. G/A+A/A	1.58 (1.19–2.10)	χ^2 = 2.48 (*P* = 0.48); I^2 = 0%	Z = 3.13 (*P* = 0.002)
	Alleles (G vs. A)	1.55 (1.18–2.03)	χ^2 = 3.35 (*P* = 0.34); I^2 = 11%	Z = 3.16 (*P* = 0.002)

Comparisons are expressed as Odds Ratio (OR) and 95% Confidence Interval (95% CI). Heterogeneity between studies was assessed using the Cochran Q test (χ^2) and the I^2 statistic.