



LJMU Research Online

Guilherme, JPLF, Semenova, EA, Zempo, H, Martins, GL, Lancha Junior, AH, Miyamoto-Mikami, E, Kumagai, H, Tobina, T, Shiose, K, Kakigi, R, Tsuzuki, T, Ichinoseki-Sekine, N, Kobayashi, H, Naito, H, Borisov, OV, Kostryukova, ES, Kulemin, NA, Larin, AK, Generozov, EV, Fuku, N and Ahmetov, II

Are Genome-Wide Association Study Identified Single-Nucleotide Polymorphisms Associated With Sprint Athletic Status? A Replication Study With 3 Different Cohorts

<http://researchonline.ljmu.ac.uk/id/eprint/14698/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Guilherme, JPLF, Semenova, EA, Zempo, H, Martins, GL, Lancha Junior, AH, Miyamoto-Mikami, E, Kumagai, H, Tobina, T, Shiose, K, Kakigi, R, Tsuzuki, T, Ichinoseki-Sekine, N, Kobayashi, H, Naito, H, Borisov, OV, Kostryukova, ES, Kulemin, NA, Larin, AK, Generozov, EV, Fuku, N and Ahmetov, II (2021) Are

LJMU has developed [LJMU Research Online](http://researchonline.ljmu.ac.uk/) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

<http://researchonline.ljmu.ac.uk/>

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

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. Guilherme¹, Ekaterina A. Semenova^{2,3}, Hirofumi Zempo⁴, Gabriel L. Martins¹, Antonio H. Lancha Junior¹, Eri Miyamoto-Mikami⁵, Hiroshi Kumagai^{5,6}, Takuro Tobina⁷, Keisuke Shiose⁸, Ryo Kakigi⁹, Takamasa Tsuzuki¹⁰, Noriko Ichinoseki-Sekine^{5,11}, Hiroyuki Kobayashi¹², Hisashi Naito⁵, Oleg V. Borisov^{2,13}, Elena S. Kostryukova², Nikolay A. Kulemin², Andrey K. Larin², Edward V. Generozov², Noriyuki Fuku⁵, 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

40 **Abstract**

41 **Purpose:** This study was aimed to replicate previously GWAS-identified sprint-related
42 polymorphisms in three different cohorts of top-level sprinters and to further validate obtained
43 results in functional studies. **Methods:** A total of 240 Japanese, 290 Russians and 593
44 Brazilians were evaluated in a case–control approach. Of these, 267 were top-level
45 sprint/power athletes. In addition, the relationship between selected polymorphisms and
46 muscle fiber composition was evaluated in 211 Japanese and 287 Finnish individuals.
47 **Results:** The G-allele of the rs3213537 polymorphism was overrepresented in Japanese (OR:
48 2.07, $P = 0.024$) and Russian (OR: 1.93, $P = 0.027$) sprinters compared to endurance athletes
49 and associated with increased proportion of fast-twitch muscle fibers in Japanese ($P = 0.02$)
50 and Finnish ($P = 0.041$) individuals. Meta-analysis of data from the four cohorts confirmed
51 that the presence of the G/G genotype rather than G/A+A/A genotypes increased the odds
52 ratio of being a sprinter compared to controls (OR: 1.54, $P = 0.005$), endurance athletes (OR:
53 1.79, $P = 0.001$) or controls + endurance athletes (OR: 1.61, $P = 0.001$). Furthermore, male
54 sprinters with the G/G genotype were found to have significantly faster personal times in the
55 100-m dash than those with G/A+A/A genotypes (10.50 ± 0.26 vs. 10.76 ± 0.31 , $P = 0.014$).
56 **Conclusion:** The rs3213537 polymorphism found in the *CPNE5* gene was identified as a
57 highly replicable variant associated with sprinting ability and increased proportion of fast-
58 twitch muscle fibers, in which the homozygous genotype for the major allele (i.e., the G/G
59 genotype) is preferable for performance.

60

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

62

63 Introduction

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

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

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

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

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

104

105 Methods

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

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

124

125 **The Japanese Cohort**

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

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

140

141 *Evaluation of skeletal muscle fiber types*

142 First, a cohort of 203 Japanese healthy individuals (98 men and 105 women, with age range
143 20-79 years) performed muscle biopsy was used for the association study between sprint-
144 related polymorphisms and muscle fiber composition. Muscle samples were obtained from the
145 belly of the vastus lateralis and myosin heavy chain (MHC) isoforms were determined by
146 performing glycerol SDS-PAGE, as previously described ¹¹. These individuals had their DNA
147 samples isolated from venous blood and the polymorphisms were genotyped using the
148 Japonica SNP array ¹².

149 Second, muscle fiber composition in 287 Finnish individuals (167 men, age 59.5 ± 8.1
150 years; 120 women, age 60.7 ± 7.4 years) from the FUSION study was estimated based on the
151 expression of the myosin heavy chain 1 (*MYH1*), myosin heavy chain 2 (*MYH2*), myosin
152 heavy chain 7 (*MYH7*), Ca²⁺ ATPase A1 and Ca²⁺ ATPase A2 genes, as previously described
153 ¹³. Muscle samples were obtained from the vastus lateralis using a conchotome, under local
154 anesthesia with 20 mg·ml⁻¹ lidocaine hydrochloride without epinephrine ¹⁴. DNA samples
155 were extracted from the blood and the polymorphisms were genotyped using the
156 HumanOmni2.5-4v1_H BeadChip array (Illumina, San Diego, CA, USA).

157

158 **The Russian Cohort**

159 The Russian study involved 173 athletes (99 males and 74 females; mean age \pm SD: $31.3 \pm$
160 7.5 years), of which 70 were elite sprinters (100-400 m runners, 500-1000 m speed skaters, 50
161 m swimmers) and 103 elite endurance athletes (biathletes, rowers, cross-country skiers, 3-10
162 km runners, 800-1500 m swimmers and triathletes). All of these athletes were international-

163 level competitors, of which 30 (13 sprinters and 17 endurance athletes) were highly elite
164 athletes (i.e., prize winners in international competitions). The control group comprised 117
165 healthy unrelated citizens (66 males and 51 females, mean age \pm SD: 47.9 ± 4.8 years),
166 without any competitive sport experience. This Russian cohort is independent of the one
167 previously published¹⁰.

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

181

182 **The Brazilian Cohort**

183 The Brazilian study involved 305 athletes (200 males and 105 females; mean age \pm SD: 25.4
184 ± 6.9 years), of which 143 were elite sprinters (100-400 m runners, 50-200 m swimmers,
185 canoeing and cycling) and 162 endurance athletes (rowers, > 1.5 km runners, 400-1500 m
186 swimmers and triathletes). While 36% of these athletes were nationally prominent
187 competitors, 64% were international-level competitors. The control group comprised 288
188 healthy Brazilian individuals (187 males and 101 females, mean age \pm SD: 29.6 ± 8.1 years),
189 without any competitive sport experience.

190 Genomic DNA of the Brazilian participants was isolated from buccal epithelial cells
191 obtained from mouthwashes with a 0.9% saline solution prepared with DNA- and DNase-
192 free water as previously described⁸. Briefly, the DNA samples were extracted using
193 chloroform, precipitated using ethanol and resuspended with 1 \times Tris-EDTA buffer. DNA
194 quantification and quality assessment were performed using the NanoDrop 2000
195 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The A260/A280 ratio
196 was used to evaluate the quality of the sample, which values between 1.7 and 2.1 were
197 considered acceptable. The genotyping process was performed using a pre-designed specific
198 TaqMan[®] SNP Genotyping Assay (Applied Biosystems, Foster city, CA, USA), according to
199 the manufacturer's instructions and using the Rotor-Gene Q PCR cycler (Qiagen, Hilden,
200 Germany). A scatter plot showing the endpoint fluorescence signals (i.e., an increase in VIC
201 or FAM fluorescent signal) was used to discriminate the genotypes. The transcript alleles
202 were used similarly to that previously used.

203

204 **Association with sprint performance**

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

212

213 **Statistical analysis**

214 First of all, the Chi-square test (χ^2) was used to test for the presence of the Hardy-Weinberg
 215 equilibrium (HWE) in each control group. A departure from HWE was observed when $\chi^2 >$
 216 3.84 (i.e., $P > 0.05$). Thereafter, the frequencies of genotypes or alleles were compared
 217 between sprinters and ethnically-matched controls or ethnically-matched endurance athletes
 218 using the χ^2 test or Fisher's exact test when appropriate. Differences in the proportion of
 219 muscle fiber types between groups with different genotypes were analyzed using unpaired t -
 220 test and one-way ANOVA. The unpaired t -test was also used to evaluate the influence of the
 221 selected polymorphism on 100-m sprint performance. The significance level was established
 222 at $P < 0.05$.

223 For the pooled analysis of the Japanese, Russian and Brazilian cohorts, meta-analysis
 224 was conducted using the Review Manager (RevMan) computer program version 5.3
 225 (*Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014*). Data from
 226 a previously published Russian case-control study¹⁰ were also used in the meta-analysis, that
 227 is, data from 99 Russian highly elite athletes (i.e., Winners of World Championships, World
 228 Cups or Olympic Games; 43 sprinters and 56 endurance athletes) and 173 controls were also
 229 included in the meta-analysis. The DerSimonian and Laird random-effects model was used to
 230 calculate weighted odds ratio (OR) and its 95% confidence interval (95% CI). The test of
 231 overall effect was assessed using the Z score with the significance level established at $P <$
 232 0.05. Heterogeneity between studies was assessed using the standard χ^2 test (Cochran Q test)
 233 and the I^2 statistic.

234

235 **Results**

236 **Case-control association study**

237 Of the three polymorphisms evaluated in the Japanese case-control study, only the rs3213537
 238 of the copine 5 (*CPNE5*) gene was found to be significant. The G-allele of the rs3213537
 239 polymorphism was overrepresented in Japanese sprinters compared to endurance athletes
 240 (80.6 vs. 66.7%; $P = 0.024$) and associated with increased proportion of glycolytic fast-twitch
 241 (IIx) muscle fibers in Japanese male controls (G/G (n = 69) $24.7 \pm 9.4\%$, G/A (n = 22) $22.1 \pm$
 242 6.9% , A/A (n = 7) $16.7 \pm 7.0\%$; $P = 0.02$). A strong trend towards an increase in the
 243 proportion of type IIx fibers was observed in Japanese males even when the analysis was
 244 adjusted for age ($P = 0.061$). In addition, the G-allele of the rs3213537 polymorphism was
 245 also associated with increased proportion of fast-twitch muscle fibers in 287 Finnish
 246 individuals adjusted for sex and age (G/G (n = 189) $55.8 \pm 14.9\%$, G/A (n = 91) $54.1 \pm$
 247 14.2% , A/A (n = 7) $43.1 \pm 16.6\%$; $P = 0.041$). Based on these associations, the rs3213537
 248 polymorphism was selected for replication in the Russian and Brazilian cohorts. Given the
 249 low frequency of the homozygous genotype for the minor allele (i.e., the A/A genotype), the
 250 rs3213537 polymorphism was analyzed only under the dominant model (G/G vs. G/A+A/A).
 251 Of note, the transcript alleles instead of the genomic alleles were used to facilitate the link
 252 between this study and previous data (discovery stage)¹⁰, that is, the G/A alleles represent the
 253 genomic C/T alleles.

254 Table 2 shows the genotype distribution and allele frequency of the rs3213537
 255 polymorphism in the three cohorts evaluated. Similar to that observed in the Japanese cohort,
 256 the G-allele was overrepresented in Russian sprinters compared to endurance athletes (86.4
 257 vs. 76.7%; $P = 0.027$) or controls (86.4 vs. 77.8%; $P = 0.042$). Indeed, the G/G genotype was
 258 overrepresented in Russian sprinters compared to endurance athletes (74.3 vs. 56.3%; $P =$
 259 0.017). In the Brazilian cohort, there was a difference of $\approx 5\%$ of G/G genotype carriers
 260 between sprinters and the two other groups (controls and endurance athletes), but this

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

263

264 **Meta-analysis**

265 Meta-analysis showed that, in the pooled data of the Japanese, two Russian (including data
266 from the previous study ¹⁰) and Brazilian cohorts, the G-allele frequency was significantly
267 higher in sprinters compared with controls ($P = 0.004$), endurance athletes ($P = 0.002$) or
268 controls + endurance athletes ($P = 0.002$), as shown in Table 3. Indeed, presence of the G/G
269 genotype rather than G/A+A/A genotypes increased the chance of being a top-level sprinter
270 compared to controls (OR: 1.49, 95% CI: 1.10–2.01; $P = 0.005$), endurance athletes (OR:
271 1.79, 95% CI: 1.26–2.55; $P = 0.001$) or controls + endurance athletes (OR: 1.58, 95% CI:
272 1.19–2.10; $P = 0.002$). There was no evidence of heterogeneity between studies.

273

274 **Sprint performance**

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

279

280 **Discussion**

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

290 The rs3213537 is an intronic polymorphism found in the *CPNE5* gene located at the
291 6p21.2 region of the chromosome 6. Copines are a family of calcium-dependent, membrane-
292 binding proteins that are evolutionary conserved from protozoans to humans ¹⁵. Present in all
293 major mammalian organs, copines may play fundamental roles in eukaryotic cell processes ¹⁶.
294 Copine proteins contain two N-terminal C2 domains that involve residues important for
295 calcium and phospholipid binding and a C-terminal A domain that may be involved in
296 protein–protein interactions ¹⁵. This well characterized structure, especially the C2 domains,
297 suggests their involvement in processes of signal transduction or membrane trafficking, which
298 occurs in a calcium-dependent manner ¹⁶. However, their biological roles have not yet been
299 fully defined.

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

311 strengthening of excitatory synapses¹⁹. The molecular events underlying copine-V (encoded
312 by *CPNE5* gene) are less understood, however, there may be some resemblance to other
313 copine proteins, such as copine-VI, as they are structurally highly similar. Nonetheless, they
314 can be expressed in different brain regions or tissues and interact with different proteins.

315 Based on animal research, copine-V has been shown to play a key role in the
316 development of the central nervous system as it is highly expressed during the embryonic
317 brain development²⁰. Its expression decreases dramatically in the adult brain, remaining
318 expressed in some non-neural tissues such as the heart, lung and muscles²¹. Nevertheless,
319 although its expression may be low in the cortex and almost undetectable in the cerebellum of
320 the adult brain, *CPNE5* is moderately expressed in the striatum of adult mice that have
321 learned a complex motor task²². Alterations in neuronal ensemble activity and synaptic
322 plasticity of the striatum are highly relevant for efficient human motor actions because it is
323 the foundation for long-term motor learning or motor memory^{22, 23}.

324 There is evidence supporting that a lack of motor memory may be detrimental to
325 power and sprint performance²⁴. Individuals with superior working memory are able to
326 perform faster and more accurate in motor tasks due to a better neural efficiency²⁵. Although
327 with training, both neural activity and performance can be improved. Repetitive activation of
328 the same neuronal circuit induces the clustering of new spines in postsynaptic membranes,
329 favoring motor performance as it strengthens the dynamics of synaptic transmission²⁶. Thus,
330 the most effective neural communication favors sprint performance. There are synaptic inputs
331 at the central and peripheral levels, directly influencing the rapid activation of muscles²⁷. The
332 ability of the neuromuscular system to increase contractile activity when muscle activation is
333 intended to be performed as quickly as possible, referred to as Rate of Force Development
334 (RFD), is considered vital for athletes requiring high-speed motor actions such as sprinters.
335 Cross-sectional studies have shown that top-level sprint/power athletes are characterized by a
336 markedly greater RFD²⁸. Moreover, athletes with a higher RFD demonstrated faster sprint
337 times²⁹. Additional contributions may also occur due to differences in muscle fiber type
338 composition—the RFD is faster in type II fibers²⁷.

339 Whether *CPNE5* rs3213537 mutant carriers have impaired motor memory or muscle
340 recruitment ability remains to be established, but the homozygous genotype for the major
341 allele (i.e., the G/G genotype) was associated with fast-twitch muscle fibers and faster times
342 in the 100-m event, which is considered the standard measure of the sprint ability of human
343 bipedal locomotion². Based on its role in the central nervous system, *CPNE5* polymorphisms
344 were previously associated with alcohol dependence and obesity²⁰. In particular, the mutant
345 allele of the *CPNE5* rs3213537 polymorphism was strongly associated with alcohol abuse²⁰,
346 which adversely impacts athletic performance in a number of different ways, including mood
347 instability and sensory-motor system dysfunction³⁰.

348 As mentioned earlier, the *CPNE5* rs3213537 is a gene variant occurring within an
349 intron (genomic position and change: g.36748144C>T based on the Genome Reference
350 Consortium Human Build 38). Introns harbour polymorphisms that can influence the
351 expression of the genes that host them and modulate the genotype–phenotype relationship.
352 Thus, this polymorphism may modulate *CPNE5* expression and its calcium-modulated signal
353 transduction. Interestingly, the interaction between copines and membranes occurs at
354 concentrations of calcium that are likely to occur in the cytosol of stimulated cells but not in
355 resting cells¹⁶, and therefore, calcium-regulated phenotypes may be affected by mutations in
356 the *CPNE5* gene. Of particular interest, during neuromuscular junction formation, muscle
357 fibers are intrinsically pre-specialized by clustering postsynaptic proteins, whereas the proper
358 patterning of postsynaptic protein clusters in the center of developing muscle fibers and the
359 subsequent innervation by the motor nerve critically depend on calcium signals³¹.

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

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

377 **Practical Applications**

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

388 **Conclusion**

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

396 **Acknowledgments**

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

410 Research Foundation), Brazil (FAPESP; Grant No. 2012/22516-6) and the Conselho Nacional
 411 de Desenvolvimento Científico e Tecnológico, Brazil (CNPq; Grant No. 301213/2015-1). The
 412 results of the current study do not constitute endorsement of the product by the authors or the
 413 journal.
 414

415 References

- 416 1. Haugen TA, Breitschadel F, Seiler S. Sprint mechanical variables in elite athletes: Are
 417 force-velocity profiles sport specific or individual? *PLoS One*. 2019;14(7):e0215551.
- 418 2. Morin JB, Bourdin M, Edouard P, Peyrot N, Samozino P, Lacour JR. Mechanical
 419 determinants of 100-m sprint running performance. *Eur J Appl Physiol*.
 420 2012;112(11):3921-3930.
- 421 3. Slawinski J, Termoz N, Rabita G, et al. How 100-m event analyses improve our
 422 understanding of world-class men's and women's sprint performance. *Scand J Med Sci*
 423 *Sports*. 2017;27(1):45-54.
- 424 4. Haugen T, Seiler S, Sandbakk O, Tonnessen E. The Training and Development of
 425 Elite Sprint Performance: an Integration of Scientific and Best Practice Literature.
 426 *Sports Med Open*. 2019;5(1):44.
- 427 5. Missitzi J, Geladas N, Klissouras V. Heritability in neuromuscular coordination:
 428 implications for motor control strategies. *Med Sci Sports Exerc*. 2004;36(2):233-240.
- 429 6. Missitzi J, Geladas N, Klissouras V. Genetic variation of maximal velocity and EMG
 430 activity. *Int J Sports Med*. 2008;29(3):177-181.
- 431 7. Maciejewska-Skrendo A, Sawczuk M, Ciężczyk P, Ahmetov I. Genes and Power
 432 Athlete Status. In: Barh D, Ahmetov I, ed. *Sports, Exercise, and Nutritional*
 433 *Genomics: Current Status and Future Directions*. Academic Press; 2019:41-72.
- 434 8. Guilherme JPLF, Silva MS, Bertuzzi R, Lancha Junior AH. The AGTR2 rs11091046
 435 (A>C) Polymorphism and Power Athletic Status in Top-Level Brazilian Athletes. *J*
 436 *Sports Sci*. 2018;36(20):2327-2332.
- 437 9. Papadimitriou ID, Lucia A, Pitsiladis YP, et al. ACTN3 R577X and ACE I/D gene
 438 variants influence performance in elite sprinters: a multi-cohort study. *BMC*
 439 *Genomics*. 2016;17:285.
- 440 10. Pickering C, Suraci B, Semenova EA, et al. A Genome-Wide Association Study of
 441 Sprint Performance in Elite Youth Football Players. *J Strength Cond Res*.
 442 2019;33(9):2344-2351.
- 443 11. Kumagai H, Tobina T, Ichinoseki-Sekine N, et al. Role of selected polymorphisms in
 444 determining muscle fiber composition in Japanese men and women. *J Appl Physiol*
 445 (1985). 2018;124(5):1377-1384.
- 446 12. Kawai Y, Mimori T, Kojima K, et al. Japonica array: improved genotype imputation
 447 by designing a population-specific SNP array with 1070 Japanese individuals. *J Hum*
 448 *Genet*. 2015;60(10):581-587.
- 449 13. Taylor DL, Jackson AU, Narisu N, et al. Integrative analysis of gene expression, DNA
 450 methylation, physiological traits, and genetic variation in human skeletal muscle. *Proc*
 451 *Natl Acad Sci U S A*. 2019;116(22):10883-10888.
- 452 14. Scott LJ, Erdos MR, Huyghe JR, et al. The genetic regulatory signature of type 2
 453 diabetes in human skeletal muscle. *Nat Commun*. 2016;7:11764.
- 454 15. Creutz CE, Tomsig JL, Snyder SL, et al. The copines, a novel class of C2 domain-
 455 containing, calcium-dependent, phospholipid-binding proteins conserved from
 456 Paramecium to humans. *J Biol Chem*. 1998;273(3):1393-1402.
- 457 16. Tomsig JL, Creutz CE. Biochemical characterization of copine: a ubiquitous Ca²⁺-
 458 dependent, phospholipid-binding protein. *Biochemistry*. 2000;39(51):16163-16175.

- 459 17. Cowland JB, Carter D, Bjerregaard MD, Johnsen AH, Borregaard N, Lollike K.
460 Tissue expression of copines and isolation of copines I and III from the cytosol of
461 human neutrophils. *J Leukoc Biol.* 2003;74(3):379-388.
- 462 18. Burk K, Ramachandran B, Ahmed S, et al. Regulation of Dendritic Spine Morphology
463 in Hippocampal Neurons by Copine-6. *Cereb Cortex.* 2018;28(4):1087-1104.
- 464 19. Reinhard JR, Kriz A, Galic M, et al. The calcium sensor Copine-6 regulates spine
465 structural plasticity and learning and memory. *Nat Commun.* 2016;7:11613.
- 466 20. Wang KS, Zuo L, Pan Y, Xie C, Luo X. Genetic variants in the CPNE5 gene are
467 associated with alcohol dependence and obesity in Caucasian populations. *J Psychiatr*
468 *Res.* 2015;71:1-7.
- 469 21. Ding X, Jin Y, Wu Y, et al. Localization and cellular distribution of CPNE5 in
470 embryonic mouse brain. *Brain Res.* 2008;1224:20-28.
- 471 22. D'Amours G, Bureau G, Boily MJ, Cyr M. Differential gene expression profiling in
472 the mouse brain during motor skill learning: focus on the striatum structure. *Behav*
473 *Brain Res.* 2011;221(1):108-117.
- 474 23. Wachter T, Rohrich S, Frank A, et al. Motor skill learning depends on protein
475 synthesis in the dorsal striatum after training. *Exp Brain Res.* 2010;200(3-4):319-323.
- 476 24. Nanci Maria F, Eric D, Mario B, Emmanuel VP. Comparison of peak muscle power
477 between Brazilian and French girls. *Am J Hum Biol.* 2002;14(3):364-371.
- 478 25. Behmer Jr. LP, Fournier LR. Working memory modulates neural efficiency over
479 motor components during a novel action planning task: an EEG study. *Behav Brain*
480 *Res.* 2014;260:1-7.
- 481 26. Fu M, Yu X, Lu J, Zuo Y. Repetitive motor learning induces coordinated formation of
482 clustered dendritic spines in vivo. *Nature.* 2012;483(7387):92-95.
- 483 27. Maffiuletti NA, Aagaard P, Blazevich AJ, Folland J, Tillin N, Duchateau J. Rate of
484 force development: physiological and methodological considerations. *Eur J Appl*
485 *Physiol.* 2016;116(6):1091-1116.
- 486 28. Tillin NA, Jimenez-Reyes P, Pain MT, Folland JP. Neuromuscular performance of
487 explosive power athletes versus untrained individuals. *Med Sci Sports Exerc.*
488 2010;42(4):781-790.
- 489 29. Tillin NA, Pain MT, Folland J. Explosive force production during isometric squats
490 correlates with athletic performance in rugby union players. *J Sports Sci.*
491 2013;31(1):66-76.
- 492 30. McDuff DR, Baron D. Substance use in athletics: a sports psychiatry perspective. *Clin*
493 *Sports Med.* 2005;24(4):885-897, ix-x.
- 494 31. Kaplan MM, Sultana N, Benedetti A, et al. Calcium Influx and Release Cooperatively
495 Regulate AChR Patterning and Motor Axon Outgrowth during Neuromuscular
496 Junction Formation. *Cell Rep.* 2018;23(13):3891-3904.
- 497 32. Terzis G, Spengos K, Methenitis S, Aagaard P, Karandreas N, Bogdanis G. Early
498 phase interference between low-intensity running and power training in moderately
499 trained females. *Eur J Appl Physiol.* 2016;116(5):1063-1073.
- 500 33. Methenitis S, Spengos K, Zaras N, et al. Fiber Type Composition and Rate of Force
501 Development in Endurance- and Resistance-Trained Individuals. *J Strength Cond Res.*
502 2019;33(9):2388-2397.
- 503 34. Methenitis S, Karandreas N, Spengos K, Zaras N, Stasinaki AN, Terzis G. Muscle
504 Fiber Conduction Velocity, Muscle Fiber Composition, and Power Performance. *Med*
505 *Sci Sports Exerc.* 2016;48(9):1761-1771.
- 506 35. Trappe S, Luden N, Minchev K, Raue U, Jemiolo B, Trappe TA. Skeletal muscle
507 signature of a champion sprint runner. *J Appl Physiol (1985).* 2015;118(12):1460-
508 1466.

509 **Figure caption**

510

511

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

515

516

517 **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

518 Legend: REF, Reference allele; ALT, Alternate allele; SFP, Associated with sprint performance in young
519 British football players; SPW, Associated with sprint performance in healthy young Polish women; AS,
520 Associated with sprint/power athlete status; MF, Associated with proportion of fast-twitch muscle fibers in
521 Russian physically active subjects. [†]Genomic position based on GRCh38 (Genome Reference Consortium
522 Human Build 38).
523

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

Group	<i>n</i>	Genotypes (%)			G allele	Comparisons: <i>P</i> -value (Effect Direction)	
		G/G	G/A	A/A		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	<u>0.062 (OR: 2.08)</u>	<u>0.024 (OR: 2.07)</u>
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	<u>0.017 (OR: 2.24)</u>	<u>0.027 (OR: 1.93)</u>
Russian controls	117	62.4	30.8	6.8	77.8	0.110 (OR: 1.74)	<u>0.042 (OR: 1.82)</u>
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

529 **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)	$\chi^2 = 1.94$ ($P = 0.58$); $I^2 = 0\%$	$Z = 2.57$ ($P = 0.01$)
	Alleles (G vs. A)	1.47 (1.13–1.92)	$\chi^2 = 2.91$ ($P = 0.41$); $I^2 = 0\%$	$Z = 2.85$ ($P = 0.004$)
Sprint/Power athletes vs. Endurance athletes	G/G vs. G/A+A/A	1.79 (1.26–2.55)	$\chi^2 = 3.30$ ($P = 0.35$); $I^2 = 9\%$	$Z = 3.25$ ($P = 0.001$)
	Alleles (G vs. A)	1.70 (1.22–2.35)	$\chi^2 = 3.70$ ($P = 0.30$); $I^2 = 19\%$	$Z = 3.16$ ($P = 0.002$)
Sprint/Power athletes vs. Controls + Endurance athletes	G/G vs. G/A+A/A	1.58 (1.19–2.10)	$\chi^2 = 2.48$ ($P = 0.48$); $I^2 = 0\%$	$Z = 3.13$ ($P = 0.002$)
	Alleles (G vs. A)	1.55 (1.18–2.03)	$\chi^2 = 3.35$ ($P = 0.34$); $I^2 = 11\%$	$Z = 3.16$ ($P = 0.002$)

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

532

533