Genotype scores in energy and iron-metabolising genes are higher in elite endurance athletes than in non-athlete controls

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Applied Physiology, Nutrition, and Metabolism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>apnm-2020-0174.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Varillas Delgado, David; Universidad Francisco de Vitoria, Faculty of Medicine, Research Unit Telleria Orriols, Juan Jose; University of Valladolid, Faculty of Medicine Monge Martin, Diana; Universidad Francisco de Vitoria, Faculty of Medicine Del Coso, Juan; Rey Juan Carlos University, Centre for Sport Studies</td>
</tr>
<tr>
<td>Novelty bullets: points that summarize the key findings in the work:</td>
<td>Genetic profile in energy generation and iron-metabolising genes in elite endurance athletes is different than non-athlete´s. There is an implication of an &quot;optimal&quot; genetic profile in the selected genes favouring endurance sporting performance.</td>
</tr>
<tr>
<td>Keyword:</td>
<td>physical endurance, sports performance, sport, genetic profile, AMPD1 protein, human HFE protein, human PPARGC1A protein</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue? :</td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>
i) Title

Genotype scores in energy and iron-metabolising genes are higher in elite endurance athletes than in non-athlete controls

ii) Authors

David Varillas Delgado¹, Juan José Tellería Orriols², Diana Monge Martín³, Juan Del Coso⁴.

iii) Corresponding author

David Varillas Delgado david.varillas@ufv.es

Universidad Francisco de Vitoria, Faculty of Medicine, Pozuelo de Alarcón, Madrid, Spain.

Elite and high-performance athletes research group.

Phone: +34 917091400 ext. 1965

iv) Affiliations

¹ Universidad Francisco de Vitoria, Faculty of Medicine, Research Unit, Pozuelo de Alarcón, Madrid, Spain. david.varillas@ufv.es ORCID 0000-0001-5026-2701

² University of Valladolid, Valladolid, Spain telleria@med.uva.es ORCID 0000-0003-1923-8345

³ Universidad Francisco de Vitoria, Faculty of Medicine, Pozuelo de Alarcón, Madrid, Spain d.monge@ufv.es ORCID 0000-0002-3593-1820

⁴ Rey Juan Carlos University, Centre for Sport Studies, Fuenlabrada, Madrid, Spain. juan.delcoso@urjc.es ORCID 0000-0002-5785-984X
**Background:** Information about the association of energy and iron-metabolising genes with endurance performance is scarce. The objective of this investigation was to compare the frequencies of polymorphic variations of genes involved in energy generation and iron metabolism in elite endurance athletes vs. non-athlete controls.

**Methods:** Genotype frequencies in 123 male elite endurance athletes (75 professional road cyclists and 48 elite endurance runners) and 122 male non-athlete participants were compared by assessing four genetic polymorphisms: \(\text{AMPD1} \) c.34C/T (rs17602729), \(\text{PPARGC1A} \) c.1444G/A (rs8192678) \(\text{HFE}_{H63D} \) c.187C/G (rs1799945) and \(\text{HFE}_{C282Y} \) c.845G/A (rs1800562). A weighted genotype score (w-TGS: from 0 to 100 arbitrary units; a.u.) was calculated by assigning a corresponding weight to each polymorphism.

**Results:** In the non-athlete population, the mean w-TGS value was lower (39.962±14.654 a.u.) than in the group of elite endurance athletes (53.344±17.053 a.u.). The binary logistic regression analysis showed that participants with a w-TGS>38.975 a.u had an odds ratio of 1.481 (95%CI: 1.244-1.762; p<0.001) for achieving elite athlete status. **Conclusions:** The genotypic distribution of polymorphic variations involved in energy generation and iron metabolism was different in elite endurance athletes vs. controls. Thus, an optimal genetic profile in these genes might contribute to physical endurance in athlete status.

**Keywords:** physical endurance; sports performance, sport; genetic profile; AMPD1 protein; human HFE protein; human PPARGC1A protein.
Novelty

1. Genetic profile in energy generation and iron-metabolising genes in elite endurance athletes is different than non-athlete’s.

2. There is an implication of an "optimal" genetic profile in the selected genes favouring endurance sporting performance.
1 Introduction

Endurance performance is related to a complex phenotype, influenced by a myriad of intrinsic and extrinsic factors (Lundby et al., 2017). Among the intrinsic factors, the likelihood of becoming an endurance athlete is influenced by the athlete’s skeletal muscle fibre composition, maximal cardiac output and oxygen uptake (VO$_{2\text{max}}$) during exercise, metabolic efficiency and total haemoglobin mass (Joyner & Coyle, 2008).

Interestingly, most of these traits are strongly influenced by genetics while some of them can be positively modified with endurance exercise training. Thus, the determinism of endurance athlete status is explained by the optimal combination of genetic predisposition and adequate physical conditioning (Eynon et al., 2013).

It has been shown that at least 120 genetic markers are linked to elite athlete status (Ahmetov & Fedotovskaya, 2015) and almost all chromosomes contain at least one gene associated with sport performance. From these genetic markers of performance, more than 70 are associated with endurance-type sports activities although only about a dozen genes have shown positive associations with elite athlete status in three or more studies (Ahmetov et al., 2016; Ahmetov & Fedotovskaya, 2015; Ahmetov et al., 2009; Varillas Delgado et al., 2019). Interestingly, most of the genetic variants associated with endurance performance codified proteins related to cellular metabolism (Ahmetov et al., 2009; Varillas Delgado et al., 2019) and muscle and cardiovascular function (Ahmetov & Fedotovskaya, 2015).

Studies on genetic variants which influence elite endurance performance have shown that several genes associated with metabolic efficiency might entail an improvement in endurance capacity through decreased oxidative stress (Al-Khelaifi et al., 2018; Fikenzer et al., 2018; Lee et al., 2017; Petibois et al., 2002). Nevertheless, the whole
metabolic genetic profile related to an elite endurance athlete’s status is not completely understood. The Adenosine Monophosphate Deaminase isoform 1 (AMPD1) is an important regulator of energy metabolism in the muscle fibre that shifts the equilibrium of the myokine reactions towards ATP production by converting AMP into inosine monophosphate (IMP) (Fedotovskaya et al., 2013; Gineviciene et al., 2014; Maciejewska-Skrendo et al., 2019). Previous investigations have found that carrying the T allele in one polymorphism in the AMPD1 gene (c.34C/T; rs17602729) might reduce the likelihood of being an elite endurance athlete (Cieszczyk et al., 2011; Gronek et al., 2018) because it might be associated with a reduced VO$_{2\text{max}}$ and lower response to endurance training (Thomaes et al., 2011). Moreover, the peroxisome proliferator activated receptor $\gamma$ coactivator 1$\alpha$ (PGC1$\alpha$) is a transcriptional coactivator of the peroxisome proliferator-activated receptor (PPAR) family, which regulates the expression of several genes associated with substrate oxidation, mitochondrial biogenesis and muscle fibre conversion (Peplonska et al., 2017). PGC1$\alpha$ is encoded by the PPARGC1A gene, and recent meta-analyses have shown that endurance athletes had a higher frequency of the Gly/Gly genotype in one common polymorphism (rs8192678) of the PPARGC1A gene, suggesting that this polymorphism might facilitate endurance performance (Chen et al., 2019; Petr et al., 2019; Tharabenjasin et al., 2019). Finally, genetics play a significant role in interindividual differences in serum iron parameters. The homeostatic iron regulator protein (HFE), codified by the HFE gene, regulates iron reabsorption (Grealy et al., 2015; Janssen & Swinkels, 2009; Ruiz et al., 2009). Individuals with C/G or GG genotypes in the c.187C/G variant (rs1799945) of this gene possessed higher circulating iron concentrations which ultimately produce a higher haemoglobin concentration (Barbara et al., 2016). A recent investigation has found that
the frequencies of the \textit{HFE} CG/GG genotypes were higher in endurance athletes and were associated with greater VO$_{2\text{max}}$ in men athletes (Semenova et al., 2020).

There is a consensus in the scientific community about the importance of the combined influence of several genetic variants, rather than the existence of one “endurance gene”, for excelling in endurance performance. The complex interaction of genetic variants (Pickering et al., 2019) might help to explain individual variations in human endurance performance and thus, the possession of an optimal polygenic profile seems necessary to succeed in endurance sports (Guth & Roth, 2013; Moran & Pitsiladis, 2017; Sarzynski et al., 2017). A previous investigation that calculated a potentially ‘perfect’ polygenic score in endurance athletes, by accounting the number of favourable alleles in seven candidate genes (including \textit{AMPD1}, \textit{PPARGC1A} and \textit{HFE}), found that elite endurance athletes had a higher polygenic score than the control population (Ruiz et al., 2009). This outcome highlights the necessity of having several favourable alleles in candidate genes for achieving elite athlete status, at least in endurance exercise.

Thus, the scientific information that interrelates the influence of the \textit{AMPD1}, \textit{PPARGC1A} and \textit{HFE} genes on endurance performance is scarce, and further confirmation is needed to clearly depict the requirement of possessing several favourable alleles in candidate genes to achieve elite athlete status in endurance exercise disciplines. The main objective of this study was to compare the frequencies of polymorphic variations of genes involved in energy generation and iron metabolism in elite endurance athletes vs. non-athlete controls.

2 Materials and methods

2.1 Study population
The study involved 123 elite endurance athletes (75 professional road cyclists and 48 elite endurance runners) and 122 men non-athlete participants (sedentary controls). All participants in the group of endurance athletes and sedentary controls were male. An analysis of the influence of liver-metabolising genes on elite athlete status has been published elsewhere with this same sample (Varillas Delgado et al., 2019). All the elite endurance athletes (25.8±4.2 years, range = 18-42 years) had tested negative for doping substances in controls made by the World Anti-doping Agency. The elite runners had a validated high level and elite sports records in endurance competitions: five athletes ran the marathon in less than 2h 10 min, 12 athletes ran the half-marathon in less than 1h 03 min and the remaining 31 athletes participated in competitions of 10000 m and 5000 m recording times below 30 min and 14 min respectively. Some of the athletes achieved finalist positions in the marathon and the 10000 m in the European Championships, with gold and silver medals in the European Cross-Country Championship. The professional cyclists had participated in the Union Cycliste Internationale (UCI) World-Tour events, including Grand Tours, classic cycle races, other one-day races or stage races (often in all of them). Ten of the cyclists reached one of the top five positions in the Grand Tours: Tour de France, Giro d’Italia and Vuelta a España. Both runners and cyclists were men, due to the small number of high-level women athletes in Spain who met the inclusion criteria. The sample of non-athlete controls was composed of healthy men matched by age to the athletes (27.9±5.1 years, range = 19-42 years); they were non-smokers and did not suffer from chronic or acute illnesses at the time of sampling. Informed consent was obtained from all the participants in the study. The study protocol was approved by the Committee of Institutional Ethics (University of Valladolid) and complied with the Declaration of Helsinki for Human Research of 1974 (last modified in 2000). Participants' rights and confidentiality were protected during the whole...
experiment, and the genetic information was used only for the purposes included in this investigation.

2.2 Genotypes

2.2.1 Target genes

In this case-control investigation, the following functional single nucleotide polymorphisms (SNPs) were genotyped:

- c.34C/T polymorphism (p.Gln12X) of *AMPD1* gene (location: 1p13) contributing to the appearance of a premature stop codon, which leads to some related metabolic muscle diseases due to the AMPD activity deficiency (Feng et al., 2017; Fischer et al., 2005). Lack of the muscle-specific isoform of AMPD can cause a metabolic myopathy, with exercise-induced muscle symptoms such as early fatigue, cramps and/or myalgia (Gross, 1997).

- c.1444G/A polymorphism (p.Gly482Ser) of *PPARGC1A* gene (location: 4p15.2) is a transcriptional coactivator of many different transcription factors and nuclear receptors. It can act through direct interaction with a transcription factor, control energy expenditure and regulate fat oxidation as well as non-oxidative glucose metabolism (Maciejewska-Skrendo et al., 2019). It is responsible for the induction of reactive oxygen species (ROS). Because ROS have been implicated as contributors to both the onset and the progression of insulin resistance, this gene might play a role in the development of type 2 diabetes mellitus (T2DM) and obesity (Baar, 2004).

- c.187C/G polymorphism (p.His63Asp) of *HFE* gene (HFE<sub>H63D</sub>) (location: 6p21.3) causes a heterogenic metabolic syndrome which is due to the unchecked transfer of iron into the bloodstream and its toxic effects on parenchymatous organs (Barbara et al.,
2016), inducing liver iron overload, are related to the risk of hepatocellular carcinoma in otherwise predisposed patients (Ropero et al., 2007) and a risk factor for nephropathy in type 2 diabetic patients (Moczulski et al., 2001).

- c.845G/A polymorphism (p.Cys282Tyr) of \textit{HFE} gene (HFE\textsubscript{C282Y}) (location: 6p21.3), causes an excessively increased absorption of dietary iron and affects the normal activity of another protein, hepcidin, a negative regulator of iron homeostasis (Katsarou et al., 2019), causing liver cirrhosis and severe liver disease (Grosse et al., 2018; Juzenas et al., 2016) and is related to various tumour types; colorectal (Chen et al., 2013) and breast (Liu et al., 2013).

2.2.2 Deoxyribonucleic acid (DNA) extraction and genotyping

- Nucleic acid purification

Genomic DNA was obtained from ethylenediaminetetraacetic acid (EDTA) anticoagulated blood samples according to standard phenol-chloroform procedures, followed by precipitation with ethanol.

- Genotyping

\textit{AMPD1}, \textit{PPARGC1A} and \textit{HFE} genotyping were carried out by direct polymerase chain reaction (PCR) amplification and subsequent agarose gel electrophoresis in 2\% agarose gel, followed by specific restriction fragment analysis, as previously described (Anderson et al., 2000; Steffensen et al., 1998; Su et al., 2008). All PCR reactions were carried out in 20\mu l of the total volume, with DNA concentrations between 125-250\mu gr.

\textit{AMPD1} genotyping

The \textit{AMPD1} (c.34C/T) p.Gln12X was genotyped by PCR using an Eppendorf thermal cycler, using the forward primer 5’-CTTCATACAGCTGAAGAGACA-3’ and the
reverse primer 5’-GAATCCAGAAAAGCCATGAGC-3’. The PCR mixture and thermal-time profile were denatured at 94 °C for 5 min. The amplification step consisted of 30 cycles of 1 min at 94 °C, 1 min at 56 °C and 1 min at 72 °C, with a final extension of 5 min at 72 °C. After restriction enzyme digestion by NspI (ThermoFisher Scientific, USA), the restriction products were separated by electrophoresis on a 2% agarose gel.

**PPARGC1A genotyping**

The **PPARGC1A** (c.1444G>A) p.Gly482Ser was genotyped by PCR using an Eppendorf thermal cycler, using the forward primer 5’-CAAGTCCTCCAGTCCTCAC-3´ and the reverse primer 5’-GGGGTCTTTGAGAAAATAAGG-3’. The PCR mixture and thermal-time profile were denatured at 94 °C for 5 min. The amplification step consisted of 38 cycles of 45 s at 95 °C, 45 s at 60 °C and 45 s at 72 °C, with a final extension of 10 min at 72 °C. After digestion by MspI (ThermoFisher Scientific, USA), electrophoresis was carried out with separation of the restriction fragments in a 2% agarose gel.

**HFE genotyping**

In the HFE gene we studied two polymorphisms: (c.187C/G) p.His63Asp (**HFE**H63D) and (c.845G/A) p.Cys282Tyr (**HFE**C282Y). For the His63Asp polymorphism, the forward primer 5´-ACATGGTTAAGGCCTGTTGC-3´ and reverse primer 5´-GCCACATCTGGCTTGAAATT-3´ were used, and for p.Cys282Tyr polymorphism (HFEC282Y) forward primer 5´-CAATGGGGATGGGACCTACC-3´ and reverse primer 5´-GCTCTCATCAGTCACATACCCCAG-3´. The PCR mixture and thermal-cycle profile were first denatured at 94 °C for 3 min. The amplification step consisted of 40 cycles of 30 s at 94 °C, 30 s at 60 °C (for **HFE**H63D) and 30 s at 64 °C (for **HFE**C282Y).
and 30 s at 72 °C, with a final extension of 8 min at 72 °C. After restriction enzyme
digestion by BclI (ThermoFisher Scientific, USA) for \textit{HFE}_{H63D} and MspI
(ThermoFisher Scientific, USA) for \textit{HFE}_{C282Y}, the restriction fragments were separated
by electrophoresis on a 2% agarose gel.

2.3 \textit{Polygenic potential for endurance performance in the Spanish population}

The combined influence of the four polymorphisms studied was calculated using a
weighted total genotype score (w-TGS). Initially, genotypes from each SNP were coded
according to the number of alleles with potential benefits for endurance performance
(Table 1; (Ruiz et al., 2009; Semenova et al., 2020)). For this codification, we used an
additive model (Williams & Folland, 2008) as follow: a score of 2 was assigned to the
"optimal" or preferable endurance genotype (i.e., homozygosity for the allele previously
associated to endurance performance), a score of 1 was assigned to heterozygote
genotype, while a score of 0 was assigned to the less optimal genotype. Afterwards,
these scores were weighted by using β-coefficients for each SNP (Table 1), based on the
assumption that each SNP of interest have independent effects and contribute in an
additive manner on endurance performance. To calculate the β-coefficient of each SNP,
a multivariable regression analysis was conducted to assess the partial contribution of
each SNP to the status of elite endurance athlete (coded as 1) or to control (coded as 0).
The relative contribution of each SNP in relation to the status of elite endurance athlete
was calculated as follows:

\[
\text{SNP partial contribution} = \left( \frac{\beta\text{-coefficient for SNP}}{\sum \text{of all } \beta\text{-coefficient}} \right)
\]

The score within each SNP (i.e., 2, 1 and 0) was then weighted by its partial
contribution and a weighted genotype score was obtained for each SNP (w-GS).
Afterwards, all w-GS were summed to obtain a unique w-TGS for each participant.
(theoretical range: 0–8 a.u.). Lastly, this value was transformed to a 0-100 a.u. scale to improve the comparison with previous investigations with a different number of SNP investigated (Ruiz et al., 2009; Varillas Delgado et al., 2019) using the following formulae:

\[ w-TGS = \left( \frac{(w-GS_{AMDP1} + w-GS_{PPARGC1A} + w-GS_{HFE_{H63D}} + w-GS_{HFE_{C282Y}})}{8} \right) \times 100 \]

With this approach, a w-TGS of 100 (a.u.) represents the "perfect" polygenic profile for endurance performance and a w-TGS of 0 a.u. would be the "worst" possible profile for endurance performance.

2.4 Statistical analysis

Compliance of Hardy-Weinberg Equilibrium (HWE) in each SNP was tested using \( \chi^2 \) tests. The statistical average and kurtosis were calculated using the Statistical Package for the Social Sciences (SPSS), v.21.0 for Windows (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp). The probability of having an "optimal" endurance genotype for one to four polymorphisms between elite endurance athletes and non-athletes was calculated using the \( \chi^2 \) test with fixed \( \alpha \) error of 0.05. The genotypic frequencies of the polymorphisms in AMDP1, PPARGC1A, HFE_{H63D} and HFE_{C282Y} variants were compared between elite endurance athletes and non-athletes, using a \( \chi^2 \) test with fixed \( \alpha \) error of 0.05. The ability of w-TGS to correctly distinguish potential elite endurance athletes from non-athletes (0 = non-athlete, 1 = elite) was assessed using receiver operating characteristic (ROC) curves (Zweig & Campbell, 1993). With that purpose, the area under the ROC curve (AUC) was calculated with confidence intervals of 95% (95%CI). Finally, a binary logistic...
3 Results

3.1 Single SNP analysis

All the SNPs analysed met the HWE. In the AMPD1 variant, the group of endurance athletes showed a higher frequency in the “optimal” genotype (C/C 79.67%) when compared to the non-athlete group (C/C 66.39%; p=0.019). For PPARGC1A, the “optimal” genotype in elite endurance athletes (G/G 62.61%) was higher than in the non-athlete population (G/G 53.29%; p=0.011). In the HFEH63D, the distribution of the genotypes was different in elite endurance athletes and non-athletes (p<0.001). Specifically, a higher frequency in the “optimal” genotype was found in athletes (G/G 6.51%) vs. non-athletes (G/G 0.00%; Table 2). However, there were no between-group differences in the genotype frequencies of the HFEC282Y gene (p=0.986; Table 2). In any case, there was no statistically significant differences in the genotypic distribution between elite endurance cyclists and elite endurance runners in any gene (data not shown).

3.2 Weighted-total genotype score

The mean value of the w-TGS was lower in the control population (39.962±14.654 a.u., statistical kurtosis: -0.672±0.435) than in the group of elite endurance athletes (53.344±17.053 a.u., statistical kurtosis: -0.234±0.433; p<0.001). The distributions of frequency of individuals according to their w-TGS is represented in Figure 1. The w-TGS distribution of elite endurance athletes was shifted right with respect to the distribution of non-athletes (p=0.001). ROC analysis showed significant discriminatory
accuracy of w-TGSs in the identification of elite endurance athletes (AUC=0.721; 95%CI: 0.658-0.785; p<0.001) with a sensitivity of 0.837 and a specificity of 0.574 (Figure 2). The corresponding w-TGS value at this point was 38.975 a.u. Binary logistic regression analysis showed that participants with a w-TGS higher than 38.975 a.u. had an odds ratio (OR) of 1.481 (95%CI: 1.244-1.762; p<0.001) of being elite endurance athletes, compared to those with a w-TGS below this cut-off value.

4 Discussion

Previous research has been satisfactory in finding links between potential genetic markers associated with enhanced physiological functioning and elite endurance performance (Ahmetov & Fedotovskaya, 2015; Ruiz et al., 2009; Varillas Delgado et al., 2019). Interestingly, most of the genes previously associated with endurance performance codify proteins related to cellular metabolism and muscle and cardiovascular function. However, the information about the association of energy and iron-metabolising genes with elite endurance athlete status is unknown. This investigation represents the first attempt, using a polygenic model, to determine whether polymorphic variations in energy and iron-metabolising genes had a joint effect on the probability of becoming an elite endurance athlete. The main outcome of this investigation is that there is a significant ‘favourability’ in the genetic profile studied for elite endurance athletes versus non-athletes, which is represented in the single comparisons of the distribution of three out of the four genes studied (Table 2). However, the addition of all the genes investigated, estimated by the total genotype score, was even clearer to determine the polygenic influence of these genes on the endurance athlete status. Thus, these results suggest that there is an endurance-specific polygenic profile in energy metabolism and iron modulation variants that is more suitable for human endurance exercise performance.
Although outstanding endurance exercise performance in sports such as cycling and running might be facilitated by an optimal polygenic profile in numerous key genes, the current analysis indicates that the influence of the genes investigated here is strong enough to differentiate elite athletes from non-athletes (Ahmetov et al., 2009). Perhaps, the clear differentiation between the group of elite endurance athletes and the control group in the genotypic distribution of the genes under investigation, and in the w-TGS, even with this low number of genes, is due to the high-performance status of endurance athletes. Elite athletes with a pure endurance-oriented phenotype and world-class performance, like the ones studied here, are seldom gathered together in genotype:phenotype association studies, and the majority of studies in the field have typically focused on endurance-related phenotype traits (Grealy et al., 2015; Yvert et al., 2016). Of note, the results of the current investigation should not be translated to other forms of exercise and sports because the “optimum” genotype profile probably does differ between endurance- and more power-oriented or intermittent sports (Al-Khelaifi et al., 2018).

The current analysis shows a higher w-TGS in elite endurance athletes than in non-athlete controls. The cut-off value of 38.975 a.u. in the 0-100 w-TGS scale was effective to discriminate the likelihood of being an endurance athlete with respect to non-athletes. However, these results also suggest the unlikely nature of finding an individual with a polygenic profile equivalent to 100 w-TGS, even elite endurance athletes and in a w-TGS made with only four genes. Interestingly, the Ruiz et al. study (2009) also found a difference in the w-TGS profile of Spanish elite endurance athletes (runners and cyclists) and non-athlete controls when investigating 7 different genes associated with performance. The current investigation is innovative because it confirms a more favourable w-TGS in elite endurance athletes, even when using a lower number of
genes, while it might be more accurate because eliminates some SNPs that have been
discarded as influential for endurance performance (Del Coso et al., 2019). In any case,
the outcomes of this investigation confirms some of the findings by Ruiz et al. study
(Ruiz et al., 2009) and clearly depict that elite endurance performance might be
obtained without a w-TGS score close to 100 a.u.

While previous investigations have found that the AMPD1 C allele may help athletes to
attain elite status in sprint/power-based sports (Gineviciene et al., 2014; Thomaes et al.,
2011), the current investigation suggests that this allele might also benefit endurance
performance. Interestingly, 79.7% of the elite endurance athletes were homozygous for
the AMPD1 C allele. Although it has been found that heterozygosity in this
polymorphism does not impede outstanding endurance performance (Rubio et al.,
2005), the results of the current analysis suggest that C/C homozygosity in the AMPD1
gene might be the optimal genotype to excel in endurance sport (Grealy et al., 2015;
Lucia et al., 2009; Rubio et al., 2005; Ruiz et al., 2009).

The distribution of the PPARGC1A genotype was different between elite athletes and
the non-athlete population. However, a high proportion of non-athletes contained the
optimal G/G genotype for this gene (Table 2). The importance of the G/G genotype in
the PPARGC1A gene has been previously found when comparing samples of elite
Turkish and Brazilian athletes with control populations (Guilherme et al., 2018; Tural et
al., 2014) and its influence on performance has been associated with the induction of
enhanced mitochondrial biogenesis associated with endurance training (Baar, 2004).
However, the importance of this gene is not exclusive to endurance sports because a
higher than expected proportion of the G/G genotypes was also present in strength
based sports (Guilherme et al., 2018; Peplonska et al., 2017). The current analysis is
innovative because it interrelates the optimal PPARGC1A genotype with other genes associated with metabolism. This outcome suggests that, although the sole presence of the G/G genotype does not guarantee outstanding endurance performance, it might favour this phenotype in the presence of other optimal genetic profiles of genes key for performance.

The tendency of endurance athletes to develop iron deficiency can trigger anaemia over time. For this reason, special care usually should be taken to avoid the mechanisms that cause this deficiency in elite athletes (Burden et al., 2015; Coates et al., 2017; Nikolaidis et al., 2018). The mutation of HFE H63D is associated with a higher capacity for iron absorption without causing hemochromatosis. However, the polymorphism HFE C282Y is more related to hemochromatosis (Chicharro et al., 2004; Zoller & Vogel, 2004). In a study by Chicharro et al., (Chicharro et al., 2004) carried out with Spanish elite athletes, the frequency of G/G homozygotes for the HFE H63D variant was 3.7% in athletes and 3.1% in non-athletes, these frequencies being similar to our research (6.5% and 0.0% for athletes and non-athletes, respectively, Table 2). Low frequencies were also present in the G/A heterozygosity for HFE C282Y polymorphism with 3.1% in athletes and 4.5% in controls (Chicharro et al., 2004), comparable to the 7.3% and 7.4% found in our study. These results indicate that the proportion of elite endurance athletes with optimal genotype profiles for the HFE gene is low while heterozygosity in either HFE H63D or HFE C282Y polymorphic variants is more present in elite athletes than in the control population (Hermine et al., 2015). Accordingly, although the likelihood of having an optimal profile in the two polymorphisms of the HFE gene is minimal even in elite endurance athletes, heterozygosity might confer an intermediate phenotype in terms of iron absorption that might favour endurance performance.
The current analysis presents some limitations that should be discussed to adequately understand the scope of the investigation. The relatively small sample of endurance athletes precludes us from drawing definite conclusions. Yet, due to the limited nature of the population under investigation, we believe this limitation is justifiable as there are hardly better endurance specialists in Spain. Numerous genetic variants that have not been included in the model are likely to appear in the foreseeable future, which can also explain individual variations in the potential for attaining elite endurance athletic status. In addition, this study has only focused on genetic data while it does not contain information that associates the genotype-phenotype in these athletes, which will need to be completed in subsequent research. Future research is also necessary in elite endurance women, as the influence of some polymorphisms might differ between sexes in Spanish Caucasian elite athletes.

Conclusions

The genotypic distribution of polymorphic variations involved in energy generation and iron metabolism was different in elite Spanish endurance athletes vs. controls. These results confirm the beneficial influence of an optimal genetic profile to obtain elite athlete status and widen the importance of genetics to become an elite endurance athlete.

Acknowledgements

The authors appreciate the participation of all the athletes for the development of the study, the support of the Spanish High Council of Sports (CSD) in the execution of this study; and thank Dr. Angel Martín Pastor for his invaluable help in collecting these elite endurance athletes for the study.
DECLARATIONS

• Ethics approval and consent to participate

Informed consent was obtained from all the participants in the study. The study protocol was approved by the Committee of Institutional Ethics (University of Valladolid) and complied with the Declaration of Helsinki for Human Research of 1974 (last modified in 2000).

• Consent for publication

Not applicable

• Availability of data and material

All data generated or analysed during this study are included in this published article (and its supplementary information files).

• Competing interests

David Varillas Delgado, Juan del Coso, Diana Monge Martín and Juan José Tellería Orriols declare that they have no competing interests.

• Funding

The present study has been funded by the Spanish High Council of Sports (CSD), through the project "Study and validation of genetic polymorphisms that predict a better performance in endurance sports" (15/UPB10/08).

• Authors' contributions
DVD carried out the genetic study and recruitment of participants, as well as the statistical study, forming part of his doctoral thesis.

JJTO helped to search for the genes involved, collaborated in writing and advise for its edition and in the improvement of the methodological aspects.

DMM collaborated in writing of paper.

JDC, as the senior author, reviewed the work, collaborated in writing, and advise for its edition and in the improvement of the methodological aspects.
References


La mutación H63D del gen HFE se asocia con un riesgo aumentado de carcinoma hepatocelular.


1 **TABLES**

2 **Table 1.** Studied polymorphisms, score assigned to each genotype for the calculation of the total genotype score, and genotype frequencies in the Spanish population obtained from a public data base.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene</th>
<th>Polymorphism</th>
<th>Genotype score</th>
<th>β-coefficient</th>
<th>Weighted genotype score</th>
<th>Iberian population (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMPD1</td>
<td>Adenosine monophosphate deaminase 1</td>
<td>c.34C&gt;T (p.Gln12X)</td>
<td>0=TT 1=CT 2=CC</td>
<td>0.675</td>
<td>0=TT 1.0=CT 2.0=CC</td>
<td>0 28 72</td>
</tr>
<tr>
<td>PPARGC1A</td>
<td>Peroxisome Proliferator Activated Receptor γ Coactivator α</td>
<td>c.1444G&gt;A (p.Gly482Ser)</td>
<td>0=AA 1=GA 2=GG</td>
<td>0.383</td>
<td>0=AA 0.6=GA 1.2=GG</td>
<td>12 53 35</td>
</tr>
<tr>
<td>HFE&lt;sub&gt;H63D&lt;/sub&gt;</td>
<td>Hemochromatosis variant H63D</td>
<td>c.187C&gt;G (p.His63Asp)</td>
<td>0=CC 1=GC 2=GG</td>
<td>1.425</td>
<td>0=CC 2.2=GC 4.4=GG</td>
<td>58 34 8</td>
</tr>
<tr>
<td>HFE&lt;sub&gt;C282Y&lt;/sub&gt;</td>
<td>Hemochromatosis variant C282Y</td>
<td>c.845G&gt;A (p.Cys282Tyr)</td>
<td>0=GG 1=GA 2=AA</td>
<td>0.149</td>
<td>0=GG 0.2=GA 0.4=AA</td>
<td>92 8 0</td>
</tr>
</tbody>
</table>

4 Data for Spanish population have been obtained in [www.ensembl.org](http://www.ensembl.org).
Table 2. Distribution of elite endurance athletes and non-athletes in the polymorphisms studied.

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Elite endurance athletes (n=123)</th>
<th>Non-athletes (n=122)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AMPD1</strong>&lt;br&gt;rs17602729</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/T</td>
<td>0 (0.00%)</td>
<td>0 (0.00%)</td>
<td>0.019</td>
</tr>
<tr>
<td>C/T</td>
<td>25 (20.33%)</td>
<td>41 (33.61%)</td>
<td></td>
</tr>
<tr>
<td>C/C</td>
<td>98 (79.67%)</td>
<td>81 (66.39%)</td>
<td></td>
</tr>
<tr>
<td><strong>PPARGC1A</strong>&lt;br&gt;rs8192678</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0 (0.00%)</td>
<td>8 (6.55%)</td>
<td>0.011</td>
</tr>
<tr>
<td>G/A</td>
<td>46 (37.39%)</td>
<td>49 (40.16%)</td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>77 (62.61%)</td>
<td>65 (53.29%)</td>
<td></td>
</tr>
<tr>
<td><strong>HFE</strong>&lt;br&gt;H63D&lt;br&gt;rs1799945</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/C</td>
<td>47 (38.21%)</td>
<td>88 (72.13%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>G/C</td>
<td>68 (55.28%)</td>
<td>34 (27.87%)</td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>8 (6.51%)</td>
<td>0 (0.00%)</td>
<td></td>
</tr>
<tr>
<td><strong>HFE</strong>&lt;br&gt;C282Y&lt;br&gt;rs1800562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>114 (92.68%)</td>
<td>113 (92.62%)</td>
<td>0.986</td>
</tr>
<tr>
<td>G/A</td>
<td>9 (7.32%)</td>
<td>9 (7.38%)</td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0 (0.00%)</td>
<td>0 (0.00%)</td>
<td></td>
</tr>
</tbody>
</table>
FIGURE CAPTIONS

Figure 1. Distribution of individuals according to their weighted total genotype score in elite endurance athletes and in non-athlete control population.

(*) The distribution is different from the distribution of non-athletes at $p<0.001$.

Figure 2. ROC curve summarizing the ability of the weighted total genotype score to distinguish potential elite endurance athletes from non-athletes.
Figure 1. Distribution of individuals according to their weighted total genotype score in elite endurance athletes and in non-athlete control population.

(*) The distribution is different from the distribution of non-athletes at p<0.001.
Figure 2. ROC curve summarizing the ability of the weighted total genotype score to distinguish potential elite endurance athletes from non-athletes.

AUC 0.721 (0.658 - 0.785)