

# Genetics and sport performance: current challenges and directions to the future

CDD. 20.ed. 575.1  
796.092  
796.5

João Paulo Limongi França GUILHERME\*  
Aline Cristina Capparelli TRITTO\*  
Kathryn Nance NORTH\*\*  
Antonio Herbert LANCHÁ JUNIOR\*  
Guilherme Giannini ARTIOLI\*

\*Escola de Educação Física e Esporte, Universidade de São Paulo.

\*\*Institute for Neuroscience and Muscle Research, University of Sidney - Austrália.

## Abstract

In recent years there has been a great progress in molecular biology techniques, which has facilitated the researches on influence of genetics on human performance. There are specific regions of DNA that can vary between individuals. Such variations (i.e., polymorphisms) may, in part, explain why some individuals have differentiated responses to certain stimuli, including the responses to sports training. In a particular sport, the presence of specific polymorphisms may contribute to high levels of performance. Since 1998, several polymorphisms have been associated with athletic phenotypes; however the accumulation of information generated over these 15 years shows that the influence of genetics to sport is extremely complex. In this review, we will summarise the current status of the field, discussing the implications of available knowledge for the practice of professionals involved with the sport and suggesting future directions for research. We also discuss topics related to the importance of polygenic profile characterization of athletes, methods for the identification of new polymorphisms associated with physical performance, the use of genetic testing for predicting competitive success, and how crucial is the genetic profile for the success athletes in competition.

KEY WORDS: Genetic; Polymorphism; Athletes; Sports; Performance.

## Introduction

The determinants of human athletic performance have long been a challenging field of study in sport sciences. Sports performance is an enormously complex multifactorial phenomenon, and is determined by numerous intrinsic (e.g., genetics, motor behavior, physiological and psychological profile) and extrinsic factors (e.g., training, nutrition, development opportunities and overall health conditions) as well as by the interaction between them<sup>1</sup>. Although it is impossible to set a unique formula to make anyone becoming a successful athlete, it is widely accepted that any individual who is highly committed and dedicated to training is able to improve athletic performance. Likewise, to be a top-level athlete, several years of dedication to an organized and rigid training system is indeed a prerequisite, although not a guarantee of success. However, a few athletes

seem to be exceptionally gifted and demonstrate extraordinarily high performance levels even before taking part in training programs; some athletes demonstrate better responses to training than others, or may be able to consistently sustain high levels of performance over their competitive career<sup>2-3</sup>.

Despite the awareness of the genetic influences on competitive success, genetics of sports performance is a quite recent area of investigation. As a consequence, the currently available knowledge is largely incipient and some authors consider we are at infant stages of the area<sup>4</sup>. Hence, every effort aiming at improving our understanding on this phenomenon is of great importance.

The earlier studies on genetics of human performance were focused on estimating the heritability of different complex traits. Using approaches such as twin-studies

and familial aggregation studies, investigators were able to estimate the percentage contribution of genetic factors to muscle fibre type distribution and enzyme activities<sup>5-6</sup>, bone density and muscle strength<sup>7</sup>, aerobic<sup>8</sup> and anaerobic capacities<sup>9</sup>, among other performance-relevant variables. Although unquestionably relevant, these studies did not provide specific information on which particular genes and genetic variants would be involved in such genetic influences. The first polymorphism related to sport performance (i.e., angiotensin-conversion enzyme, ACE) was not identified until 1998<sup>10-11</sup>.

Over the past 15 years, the advances in biotechnology and molecular biology tools have facilitated a rapid increase in the identification of structural genetic variations capable of exerting some influence on the phenotypes related to athletic performance<sup>12</sup>. Association studies made possible the establishment of a human gene map for exercise<sup>13</sup> and constant updates reflect the rapid increase in the number of polymorphisms that entered in this map<sup>14-15</sup>. To date, more than 200 polymorphisms have been associated with some feature related to physical exercise<sup>15</sup>, and it is expected that this number will increase in the following years<sup>15</sup>. However, only about 20 of these >200 genetic variants were specifically observed in athletes<sup>4</sup>. Furthermore, most of these genes and variants have failed to confirm association in replication studies<sup>16</sup>, so that less than 10 genetic variants have been consistently associated

with sports performance<sup>17</sup>. Considering that the human genome has over 20,000 genes and that each gene may present an enormous diversity of common variants that could theoretically influence some performance-related phenotype, it is extremely likely that our currently knowledge represents only a small fraction of the genetic factors that influence sports performance. Hence, numerous new genetic variants are yet to be discovered, and we still barely understand how genes interact with each other and with environmental factors.

High-level sports performance is an extremely complex phenotype and genetic background is only one of its multiple contributory factors. It is likely that the contribution of heritability to a particular phenotype will largely depend on the specific sport discipline, among other factors. Even if only the genetic factors are considered, sports success remains an extremely complex phenomenon because it is a multigenic trait<sup>18</sup>.

Understanding the influence of genetics in sport is a critical step to unravel the determinants of sports excellence. However, the challenges are enormous and the experimental approaches to address this highly complex and multifaceted phenotype are somewhat limited. In this critical review, we discuss the potential and limitations of research methods on genetics of sports performance, the implications of this knowledge in “real-world” sport settings, as well as directions for future researches.

## Basic concepts

For the benefit of reader, the terms and concepts most important for understanding this discussion are briefly revised below.

Genome refers to the entire collection of genetic material hereditarily transmitted to the next generation that a given species possesses<sup>19</sup>. In the case of most eukaryotes, including humans, the genome is encoded in DNA sequences. These sequences are composed by four different nitrogen bases (Adenine, Thymine, Cytosine and Guanidine - A, T, C and G, respectively) that, by binding in sequence, make a single DNA strand (e.g., AACGGT is a sequence of nitrogen bases forming a single-stranded DNA). Each nitrogen base also binds to its complementary base (i.e., A binds to T and C binds to G) so that a single DNA strand is attached to a complementary single DNA strand, forming a double-stranded DNA. The

DNA molecules are predominantly found inside the nucleus, although a small portion is also found inside mitochondria. Inside the nucleus, the genome is organized in 23 different pairs of chromosomes.

The human genome has over 3 billion nitrogen base pairs. Of this total, only ~5% are encoding regions. A gene is considered a specific region of the genome whose DNA sequence encodes a biologically active product, which is, in most cases, a RNA molecule that ultimately results in a protein. Every individual has two copies of each gene, which are called alleles. In the coding sequences, a sequence of 3 nucleotides encodes for 1 specific amino acid, which will take place in the peptide chain. These 3-nucleotide sequences encoding for 1 amino acid are called codons.

Although all human beings share all the same genes, they display some slight structural variations in

their nitrogen base sequences. It is estimated that only 0.1% of the genome varies between individuals<sup>20</sup>. However, this minor portion of the genome explains the enormous phenotypic diversity that exists among humans. These variations may occur as: 1) changes in a single base pair (for example, in the sequence AACGGT the nucleotide G is swapped by a nucleotide A, so the variant sequence is AACAGT); 2) deletions of a single base pair (for example, in the sequence AACGGT, the nucleotide G is deleted, so the variant sequence is AACGT); 3) insertions of single base pairs (for example, in the sequence AACGGT, the nucleotide C is inserted, so the variant sequence is AACCGT); or 4) changes, deletions or insertions in two or more base pairs. The most common type of variation in the human genome consists in a change in a single base pair<sup>21</sup>. This type of variant can be a single nucleotide polymorphism (SNP) or a point mutation, depending on how prevalent it is in the population and on the impact it has on phenotype (see further discussion).

Despite being non-coding sequences, 95% of the genome do play fundamental physiological roles, especially by regulating the rate of genes expression (i.e., whether or not a gene will be transcribed into RNA to produce a protein, and how much RNA will be produced from its corresponding gene). Therefore, genetic variants may encompass both coding and non-coding regions of the genome<sup>22</sup>. Genetic variants that occur in coding regions may affect the sequence of amino acids in a protein and, depending on the type of variant, the protein structure can be slightly or severely affected. Obviously, the more a protein structure is affected, the greater its physiological impact tends to be. The variant protein may be less functional or even not functional at all. On the other hand, when genetic variants occur in non-coding regions of the genome, the protein structure is normally unaffected and its physiological impact tends to be less pronounced. In these cases, it is more likely that the rate of the gene expression is affected.

Some genetic variations are rare and some are common. When a variant appears in less than 1% of the population, it is considered a mutation; when its frequency in the population is greater than 1%, it is considered a polymorphism<sup>19</sup>. Normally, a mutation has a greater impact on physiological functions than

a polymorphism<sup>23</sup>. As a consequence, mutations tend to have some health impact whereas polymorphisms tend to account for normal phenotypic variation. However, there are cases of rare mutations that do not lead to a disease, as well as there are common polymorphisms that increase the likelihood of an individual to develop a disease. Whether these variants should be considered mutations or polymorphisms is still a matter under debate and are not under the scope of this review.

As discussed earlier, there are two copies of each gene in the genome<sup>a</sup>. One allele is found in a specific region of a specific chromosome whilst the other allele (which is exactly the same gene but not formed by exactly the same sequence) is found in the same region of the homologous chromosome. Considering a given variation in a given gene, an individual may have one or two copies of the most frequent variant (which is often referred as the “normal” copy) and/or one or two copies of the least frequent variant (which is often referred as the polymorphic or mutated copy). Therefore, for this specific variation, the genotype of an individual can be: 1) homozygous (two alleles of a “normal” copy of the gene); 2) heterozygous (one allele of a “normal” copy and the other of a polymorphic copy of the gene); or 3) homozygous (two alleles of a polymorphic or mutated copy of the gene).

Phenotypic traits are observable characteristics controlled by genes. Thus, a given genotype affects a given phenotype to some extent. Some traits are controlled by one single gene, and they are referred to as monogenic traits. Normally, it is relatively easy to establish a link between a genotype and a phenotype in monogenic traits since they obey the Mendelian logic of inheritance. On the other hand, polygenic traits are far more complex because they are influenced by several genes, as well as by multiple non-genetic environmental factors. Due to its multifactorial nature, it is normally difficult to establish a strong association between one single genetic variant and a complex phenotype, which often imposes a hurdle to the studies attempting to identify specific genes that influence a complex phenotype. This explains, at least in part, why the associations between genetic polymorphisms and athletic performance are normally weak and frequently not confirmed in replication studies.

## Genetic influences on quantitative traits and sports performance

Sports performance is an extremely complex phenotypic trait, which is in turn influenced, although not determined, by many other traits, such as muscle fibre type distribution, aerobic power and capacity, anaerobic power and capacity, and trainability of physical capacities<sup>24</sup>.

Most traits that are relevant to sports performance are quantitative, meaning that they are possible to be measured and quantified. Some examples of quantitative traits that are relevant to physical performance are: body composition, aerobic power and muscle strength. In some cases, the final outcome of sport performance can also be a quantitative trait. For examples, swimming distance times, running races, jumps, throws and all other sports in which final performance is quantifiable can be considered

quantitative traits. In other cases, however, sports performance “per se” is not a quantitative trait. This is the case of unpredictable sports, such as team sports, individual sports that depend on nature’s conditions (e.g., surfing and sailing) and individual sports that depend on opponents’ actions (e.g., combat sports). Theoretically, some performance-relevant quantitative traits are strongly influence by genetic factors, which is also the case of some “predictable sports” (FIGURE 1). On the other hand, other traits as well as “unpredictable sports” are less influenced by genetic factors (FIGURE 1) and, therefore, genotype-phenotype relationships are less likely to be established. This must be kept in mind when performing association studies, as latter discussed in this review.

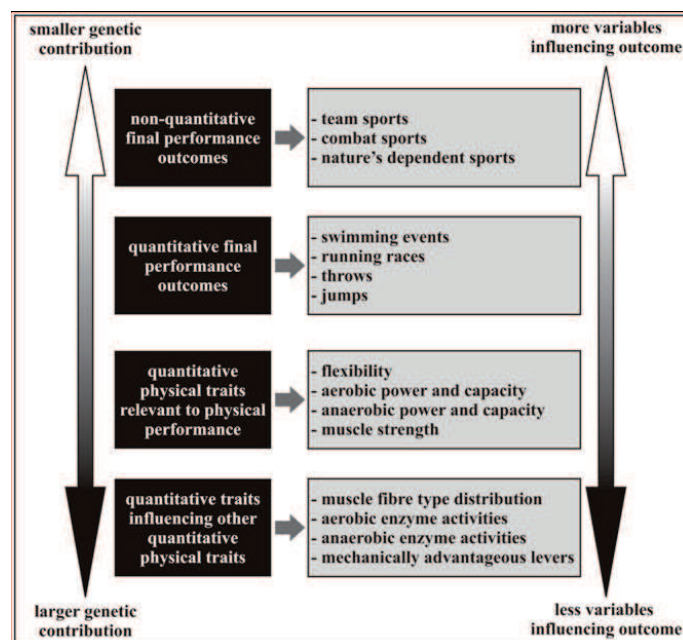


FIGURE 1 - Contribution of genetic factors to performance-relevant quantitative traits.

## Experimental approaches for studying genetics of sport performance

Over the past 15 years, there has been a great effort to identify, at the molecular level, variations in DNA sequence that may contribute to sports performance or to any trait that is relevant to sport performance,

so-called genotype-phenotype correlations. However, in complex polygenic multifactorial traits, genotype-phenotype correlations are often elusive and difficult to be clearly identified.

## Candidate genes association studies

One of the most frequent experimental approaches for assessing genotype-phenotype correlations is the genetic association. In genetic association studies, a candidate polymorphism is correlated with a performance-relevant trait. For example, the frequency of a candidate polymorphism is compared between two highly distinct populations: elite athletes and non-athletes. If the polymorphism is significantly more frequent in the athletic group, it is assumed that this polymorphism is associated with athletic status and contributes to elite performance. In general, a polymorphism is considered a candidate based on the physiological role of the gene and on how the difference in nucleotide sequence affects gene function and/or expression.

Association studies can be divided into three main categories: 1) case-control studies, which compare the frequency of genotypes in a cohort of controls (non-athletes) and a cohort of elite athletes; 2) cross-sectional studies, which compare selected physiological and/or performance data between different genotypes<sup>25</sup> and 3) longitudinal studies, in which responses to a given intervention (e.g., exercise training or diet) are compared between genotype

groups. All approaches are important to demonstrate the relevance of a genetic variant to performance<sup>26</sup>. However, polymorphisms emerging from association studies remain as “candidate” genetic variants until the association is replicated in other independent cohorts and a plausible biological explanation for the impact of the polymorphism is formulated<sup>25</sup>.

Case-control association studies are relatively cost-effective and easy to be performed, especially when a large number of DNA samples from top-level athletes and controls is readily available, which makes this approach interesting to initially screen potential candidate polymorphisms. In contrast, association studies do not provide cause-effect relationships, meaning that establishing an association between a candidate polymorphism and elite athletic status is not sufficient to accept a candidate polymorphism as valid. Thus, providing further evidence on the influence of that polymorphism on physical performance it is extremely important. This evidence can be produced using the cross-sectional and longitudinal prospective approaches detailed in the previous paragraph, as well as by determining a physiological role of the polymorphism on sports-related phenotypes. The critical steps for producing compelling evidence on the role of a genetic variant in enhancing sports performance are schematised in FIGURE 2.

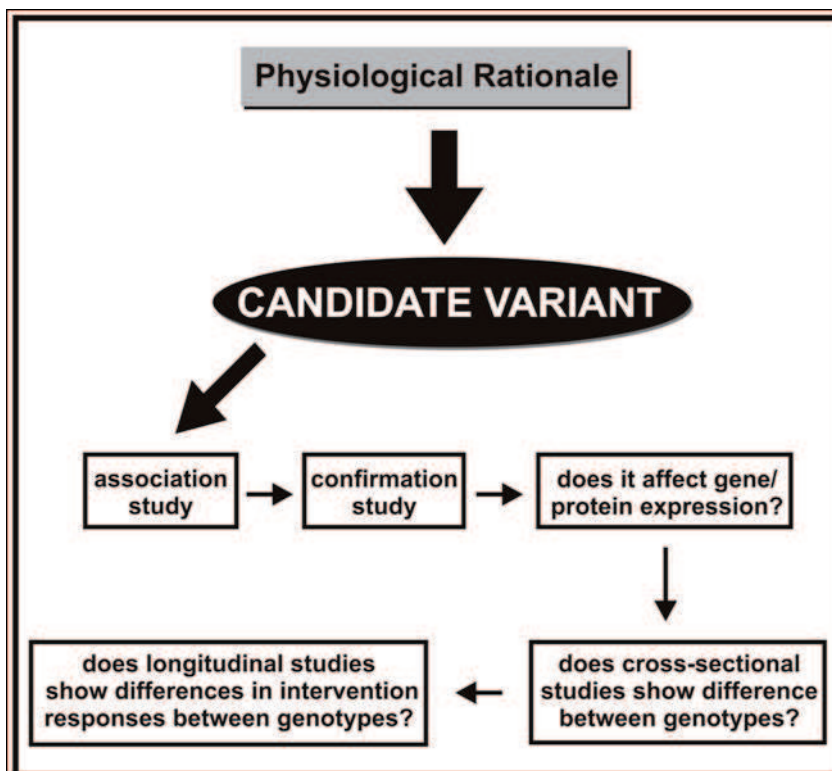


FIGURE 2 - Critical steps and experimental approaches necessary to validate a candidate gene as relevant to sports performance.

Although relatively cost-effective and straightforward, association studies often face some obstacles that may hamper researchers to securely draw conclusions from the data obtained. To circumvent these problems, some measures are recommended when performing association studies.

In some cases, a particular polymorphism correlates with athletic phenotypes in a group of individuals from a specific region and, later, the same association is observed in a different set of individuals from a distinct region<sup>27-28</sup>. This positive replication (i.e., a consistent association) strengthens the evidence for the influence of the polymorphism on phenotype. However, even if the replication occurs in more than one population, it does not mean that the same association will be found in every population of the world<sup>29</sup>. In fact, associations found in a study are frequently not replicated in subsequent studies<sup>30-33</sup>. Depending on the characteristics of the studies that showed the association (e.g., sample size, ethnic background and homogeneity of athletic cohort in terms of competitive level and sports disciplines), such inconsistencies may be interpreted as: 1) the polymorphism is not relevant to physical performance; 2) the polymorphism is relevant in a population with a specific ethnic background; 3) the polymorphism is relevant for some specific sports disciplines.

According to the ethnic background of the studies cohort, the frequency of each genotype can vary dramatically<sup>33-34</sup>. Therefore, all polymorphisms reported in the literature should be replicated in different populations<sup>4,16</sup>. It is possible that some polymorphisms are relevant to performance only in some specific regions or under some specific conditions, whereas other polymorphisms may have a more “universal” effect.

### Genome-wide studies

One major disadvantage of candidate-gene studies is that only one or a few (in the case of polygenic profile studies) genes can be assessed at a time. In view of this, new methods allowing the screening of the whole genome were developed.

Genome-wide linkage studies (GWLS) were the first approach to analyse genetic markers across the entire genome. This method identifies chromosomal regions that harbour genes affecting quantitative traits over generations (i.e., it identifies quantitative trait loci)<sup>24</sup>. GWLS have been used to discover QTL associated with a variety of diseases and other phenotypes<sup>25</sup>.

More recently, technological advances originated another technique, the so-called genome-wide association study (GWAS), capable of identifying genes, rather than genomic loci, associated with a phenotype<sup>35</sup>. Unlike GWLS, which requires familial data (the basic unit of observation is a pair of parents, usually brothers), GWAS studies analyse individual data<sup>24</sup>. This new approach is becoming increasingly popular in the search for variations that contribute to complex traits<sup>24</sup>.

While candidate gene studies are driven by the theoretical impact that a variant would have on physical performance, GWAS do not make any prior assumption regarding genes and variants involved with physical performance<sup>2</sup>. Due to our limited ability to select candidate polymorphisms for association with performance based only on available theory, the design of candidate-genes association studies will be always restricted to a certain degree<sup>24</sup>. In this sense, the fact that GWAS studies are “theory-free” and that polymorphism selection is based on observational data makes this approach more robust. This increases the chances of finding new and perhaps unexpected genes and variants affecting physical performance, which might open new venues for investigation.

Despite being considered more robust than candidate genes studies, GWAS are very expensive and, therefore, not widely used in sports sciences or not used in truly large athletic cohorts<sup>35</sup>. In 2008, the cost of a run containing two human genomes (30x coverage - Illumina sequencing machine - HiSeq 2000) was around US \$60.000<sup>36</sup>. Even though this value has been falling and in 2011, the price of a similar run was around US \$10.000<sup>36</sup>, meaning that a very large study with adequate statistical power would reach extremely high costs.

Because of the multivariate nature of GWAS, the p value must be times lower than usual to accept the correlation as significant (normally  $5 \times 10^{-8}$  rather than 0.05)<sup>2,37</sup>, which obviously require a very large sample size to achieve a desirable statistical power. By their very nature, cohorts of fairly homogenous top-elite athletes are small. Thus, reaching desirable statistical power in GWAS studies can be an enormous challenge in sports sciences. The study by PITSILADIS and WANG<sup>36</sup> describes a good example on how sample size can affect GWAS outcomes, reporting two studies designed to quantify the heritability of stature, a very stable and easy to be measured variable, using the GWAS approach. In one study, a set of ~30,000 individuals was analysed,

and it was found that less than 5% of the variation in the phenotype is explained by genetics. In contrast, the other study using a much larger set of individuals ( $n \cong 180.000$ ) showed that genetics explained 10% of the variance in height.

Likewise, the number of SNPs included in the analysis affects the GWAS statistical power. The greater the number of SNPs, the greater the number of comparisons carried out by statistical analysis, so the value accepted to consider statistical significance is reduced<sup>16</sup>. Another limitation of GWAS is that, despite screening for genetic variants along the entire genome, only SNPs are detected in this analysis, meaning that not all types of genetic variations will be captured<sup>4,24</sup>, such as copy number variants.

A massive amount of information is generated by one GWAS. However, as with any other association study, the results of GWAS can be vague or uninformative<sup>38</sup>. Within the same study, the following situations can be possible: 1) a significant association is found between a SNP and a phenotype, and the result is replicated in other cohorts (true-positive); 2) a significant association between the SNP and the phenotype is found, but replication studies fail to confirm it in other cohorts (false-positive); 3) GWAS was not sensitive enough to detect associations which were already confirmed in previous studies (lost-results)<sup>38</sup>. It has been shown that non-replicating results and inconsistency in the magnitude of results (heterogeneity) are common flaws in GWAS<sup>38</sup>.

In the last 5 years, several investigations were carried out using GWAS for complex phenotypes, mainly pathologies<sup>35</sup>. In 2011, BOUCHARD et al.<sup>39</sup> conducted the first study applying GWAS in a group of individuals (non-athletes) before and after a training period. They examined the association of SNPs with the  $VO_{2max}$  responses to physical training. The authors managed to perform the study in a quite large sample (> 1000 subjects) and to compare the results in different cohorts. The study successfully identified genetic variants associated with training responses, and then it was possible to construct a panel of 21 SNPs, which accounted for 49% of the variance in  $VO_{2max}$  responses ( $p < 0.05$ ). To date, there are no studies using the GWAS approach in a cohort of athletes. Although it is acceptable that access to large cohorts of truly elite athletes represents a tremendous barrier to high-quality genome-wide studies in this population, a collective effort from research groups around world is indeed necessary to undertake this extremely relevant type of study.

## The role of sample size

The number of participants is probably the most important limitation of genetic association studies<sup>26</sup>, especially when referring to elite athletes<sup>40</sup>. Because performance-relevant polymorphisms are normally found in low frequencies in a given population<sup>12</sup>, reduced sample sizes will probably return very small numbers of individuals presenting the rare genotype. Each polymorphism exhibits distinct frequencies in different populations, so the sample size necessary to detect an association can vary according to the polymorphism and to the population analysed<sup>17,41</sup>. Sample size is directly related to the statistical power of the study, and a reduced number of participants can hamper the drawing of firm conclusions<sup>4</sup>. This limitation becomes even more evident when multiple comparisons enter into statistical model (e.g., in GWAS and polygenic studies), which makes the analysis more rigid and lowers the set level of significance<sup>16</sup>.

Some authors claim that this limitation is justifiable since there is a very limited number of high-level athletes in most regions and countries. This makes the collection of a sufficiently large cohort almost impossible and it explains why studies with elite athletes usually assess a small number of individuals (i.e.,  $n > 100$ )<sup>42</sup>. Indeed, this is a very strong and truthful argument and, regrettably, no much can be done to enlarge athletic cohorts, especially in countries where competitive sport is less prominent. However, researchers in this area should endeavour to maximize the number of athletes in their cohorts. In fact, some authors advocate that the research groups around the world should create an international consortium with DNA samples from worldwide elite athletes<sup>4,16</sup>. This is probably the best way to circumvent this relevant limitation in sample size. However, researchers should be conscious to the fact that this procedure could result in heterogeneous groups regarding both athletic and ethnic backgrounds, which would add extra confounding variables to the analysis.

Alternatively, increasing the number of participants in the non-athletic (control) group can be an appealing manner to minimize problems with low sample sizes and underpowered studies. For example, by running a statistical simulation, we observe that a group of athletes ( $n = 100$ ; frequency of the rare allele = 30%) when compared to a control group ( $n = 100$ ; frequency of the rare allele = 20%), no significant differences are observed for the allele frequency between groups ( $\chi^2 = 2.16$ ,  $p = 0.14$ ). However, as shown in FIGURE 3, even if the number of athletes

remains unchanged ( $n = 100$ ; frequency of the rare allele = 30%), it is possible to detect a significant difference between groups by increasing the number of controls to 325 (frequency of the rare allele remains 20%) ( $\chi^2 = 3.85$ ,  $p = 0.0497$ ). As displayed in FIGURE 3, successive increases in the number of controls are paralleled by increases in statistical power.

However, at a certain point, further large increases in sample size result in merely slight increases in power. In view of this, researchers are advised to increase the number of controls to a maximum in their analysis. Possibly, a control sample size between 1000 and 1500 will yield a desirably good statistical power for the majority of the situations.

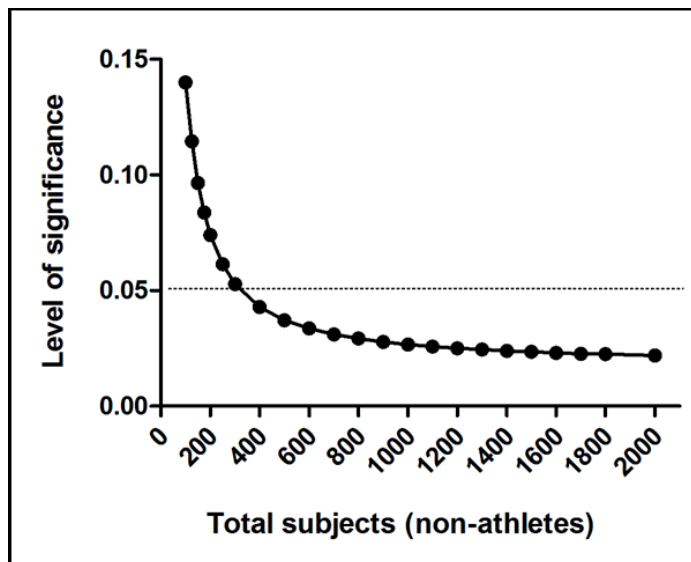


FIGURE 3 - Level of significance found considering a fixed sample of athletes and varying the amount of non-athletic subjects (controls).

### The importance of making comparisons between homogeneous groups

It is widely accepted that each sport discipline has its particular physiological, psychological, anthropometrical and biomechanical demands, which directly influence the characteristics that would most contribute to competitive success<sup>43-45</sup>. In fact, it has been considered that the polymorphisms influencing endurance sports are different from the polymorphisms influencing sprint sports, which display a greater demand for muscle strength and power<sup>46</sup>. Considering that some (if not all) of these characteristics are in part genetically determined, studies comparing genotypes between athletes should carefully select which sports disciplines will be included in a given group when categorizing athletes.

Most studies have been categorized athletes into two opposing groups, namely sprint/power-orientated sports vs. endurance-orientated sports<sup>31,47</sup>. Although this is an interesting approach, since it would group together sports under the influence of the same genes and polymorphisms, the inclusion of sports in the same group that are not really analogous can represent a major flaw in the analysis. For

example, football, 200-m swimming and powerlifting could be classified as power/sprint-orientated sports, even though the determinants of these disciplines are clearly different. Even nearly identical sprint/power-orientated sports could have quite different demands, such as 50-m and 400-m swimming. In view of this, it becomes evident that researchers should wisely include in the same group only athletes who compete in truly comparable disciplines.

The study by MUNIESA et al.<sup>48</sup> compared the distribution of 6 polymorphisms between endurance runners (> 5000 meters) and endurance cyclists, sports that are usually clustered on the same group due to their similar metabolic demand. Interestingly, the frequency of the I and D alleles (indel polymorphism in the ACE gene) was different between these groups. Despite the similarities regarding the metabolic demands, factors such as mechanical efficiency and movement economy may have influenced the differential role of the ACE polymorphism in these sports<sup>48</sup>.

Another matter of concern is how to categorize athletes according to competitive levels. Ideally, a polymorphism that exerts some influence on sports performance should be capable of differentiating not



only athletes from non-athletes, but also elite athletes from non-elite athletes. Thus, the criteria utilised to categorise athletes into competitive groups can be also confounding factors in association studies. Usually, “high-elite” athletes are those who participate in Olympic Games or World Championships; “elite” are those who participate in international-level competitions (e.g., continental championships), “sub-elite” participate at the national-level; and “non-elite” in state- and regional-level competitions. Although slightly different classifications have been used in literature<sup>40</sup>, this is an easy and straightforward manner to classify athletes according to competitive levels. However, researchers from different countries should be cautious in relation to how a particular sport is developed in that country. For example, judo is a very popular and developed sport in Brazil, whilst rugby is not. On the other hand, judo is not a popular and developed sport in New Zealand, whereas rugby is probably the most prominent sport. In this context, reaching national judo level in Brazil is probably more difficult than reaching international high-elite levels in judo in New Zealand. Likewise, a world-cup rugby player in Brazil would probably have less of the necessary characteristics to succeed in rugby than a national-level New Zealand athlete. Because of these national disparities, authors should ponder the most appropriate way to classify the athletes from their cohorts taking into consideration how the specific sport is developed.

### Single gene approach vs. polygenic profile

Most studies on genetics of sports performance have assessed the impact of only one variant on athletic performance<sup>46</sup>. Due to the multifactorial nature of sports performance, the effect of a single variant is most likely to be small<sup>49</sup>. According to FLUECK et al.<sup>50</sup>, one important drawback of the single gene approach in sports sciences is that the selected gene may only marginally contribute to performance, or it may not be the “bottleneck” of performance, since many other genes may compensate the altered function of the polymorphic gene. The regulation of every physiological system involved in exercise performance and training responses is dependent on a complex network of interconnected genes<sup>51</sup>. Thus, it is assumed that the contribution of genetics to physical performance relies on the combined action of different genes and, consequently, a number of different polymorphisms

accounts to the variation in sports excellence and training responses.

Despite their limitations, single gene studies are important to first identify polymorphisms associated with performance, especially when it is not possible to perform GWAS. However, because one single polymorphism would only account for a minor part of the total variation, a combination of polymorphisms (i.e., a polygenic profile) would provide a better model to explain how changes at the molecular level in DNA would affect the phenotype.

The concept of polygenic profile is now becoming more solid. It was recently shown that healthy individuals who possess more alleles associated with aerobic metabolism present better responses to aerobic training (according to  $VO_{2max}$  responses to training)<sup>39</sup>. Likewise, endurance athletes with a greater amount of alleles already associated with endurance performance have a greater chance of being successful in endurance-orientated sports<sup>52</sup>. Therefore, it is becoming increasingly clear that high-level athletes are individuals genetically distinguished, meaning that they probably present a combination of numerous polymorphic alleles associated with physical performance<sup>12</sup>.

In 2008, WILLIAMS and FOLLAND<sup>17</sup> proposed the calculation of the total score genotyping (TGS). To calculate TGS, the polymorphisms that have been previously associated with a target phenotype must be selected. Each allele receives a score, which ranges from 0 to 2, according to the existing genotype. Homozygous genotypes associated with the phenotype receive the score “2”, while heterozygotes receive the score “1”. Homozygous for the allele not associated with performance returns the score “zero”. After determining the scores for each polymorphism, these are summed and the result is expressed in a scale ranging from 0 to 100, so that the TGS is obtained. Using this score, it is possible to assess the balance between several selected different polymorphisms that are knowingly relevant to performance<sup>41</sup>.

From 2009 to the present day, the TGS has been used by some research groups and has proved to be a sensitive tool to differentiate endurance from strength/power athletes<sup>31,47,53-54</sup>. Additionally, the use of TGS has revealed that high-elite athletes have a polygenic profile significantly different from the general population, which probably makes athletes more favourable to sports success<sup>18</sup>.

The TGS assumes a dose-response effect, i.e., the more associated alleles an athlete has, the better his/her genetic profile for sports success. Hence, it is assumed that there is an additive effect

of polymorphisms<sup>4</sup>. Moreover, the sensibility of the TGS to differentiate individuals with different genetic predisposition to excel in a given type of sport seems to be dependent on the number of polymorphisms included in the calculation<sup>18</sup>.

The choice of which polymorphisms are included in the TGS calculation should be made in light of the characteristics of the studied population. For example, if a study aims to determine the polygenic profile of sprinters, polymorphisms associated with muscle size and strength, as well as anaerobic energy metabolism and other sprint-orientated phenotypes should be inserted in the formula<sup>41</sup>. Besides, the choice of the polymorphisms should include only those truly consistently associated with the phenotype. Otherwise, the TGS will lack power in differentiating athletes between non-athletes or sprinters endurance athletes<sup>17,41</sup>.

A potential limitation of the TGS is the fact that it considers all polymorphisms as influencing athletic performance to the same extent<sup>4</sup>. With the currently available knowledge, it is impossible to determine the exact weight that each polymorphism should have in TGS calculation, simply because it is not yet determined the weight of each polymorphism in the regulation of the different sports-related phenotypes<sup>55</sup>.

The number of genes that are modulated in response to physical training is relatively large. In contrast, the scientific literature shows a relatively limited number of polymorphisms associated with physical performance in athletes, i.e., about 20 genes<sup>17,41</sup>. The scores so far published have included 6-10 polymorphisms in the TGS calculation, which clearly represents only a small fraction of the genes that putatively affect sports-related phenotypes<sup>52,56</sup>. As a consequence, the optimal combination of polymorphisms that would result in an ideal TGS model is yet to be determined<sup>49</sup>. However, this does not preclude the TGS to be successfully used to distinguish endurance athletes from strength/power athletes and from non-athletes. Furthermore, the TGS will probably be updated and optimized in the upcoming years as the knowledge on single genes associated with performance will improve<sup>42</sup>. One important caveat, however, is that as the number of polymorphisms is included in the calculation increases, the likelihood of finding an individual with optimum genotypes decreases exponentially<sup>17,41</sup>, which could cause TGS to lose sensitivity. Thus, an ideal model of TGS will include only the genes that most contribute to that specific trait. The exact number and the most important genes for each phenotype are fundamental questions that researchers should address in the future.

## Identification of new candidate polymorphisms

To choose a new polymorphism that will be analysed in an association study, it is first necessary to choose a candidate gene to be explored. A convenient way to choose a candidate gene is to observe which genes are modulated during or after exercise<sup>57</sup>. If a particular gene is consistently modulated, it would be worthy to search for structural variations in this gene that could influence phenotypes related to physical performance<sup>58</sup>.

Recently, TIMMONS et al.<sup>57</sup> screened, through the use of microarray, the mRNA expression in "vastus lateralis" muscle of healthy subjects who underwent 20 weeks of aerobic training. Afterwards, the authors searched for polymorphism in the genes whose expression accounted for the variation in  $VO_{2max}$  responses. Six SNPs that may account for the variability in the cardiorespiratory response to aerobic training were identified.

Although useful, microarrays are somewhat expensive and time-consuming. The enormous amount of data generated might also represent

another obstacle to the use of this technique. In order to identify new candidate polymorphisms influencing athletic performance, alternative methods other than microarrays can be employed. A considerable amount of information on mRNA expression and exercise is available in literature and potential genes are waiting to be explored<sup>59</sup>.

After choosing one candidate gene to be explored, an investigator should examine all genetic variations described for the gene, which can easily be done with online public databases (e.g., NCBI and UCSC)<sup>4</sup>. For the vast majority of the genes, several structural variations will be found, including single nucleotide, *indel* and copy number polymorphisms. The selection of the best candidates should consider previously available data. Nonetheless, multiple polymorphisms for each gene should be tested in order to verify which one presents the best correlation with the phenotype<sup>25</sup>.

There is a natural tendency in sports research to study genes involved in motor activities. However, more recent studies indicate that genes involved

in psychological characteristics may also influence athletic performance<sup>60-61</sup>. Some prime examples are the genes *5HTT*, *BDNF* and *UCP2*<sup>62</sup> and future

studies should start focusing on the psychological aspects implicated in sports performance when selecting new candidate genes.

## Strategies to identify physiological roles of a polymorphism

As previously discussed, unravelling the underlying physiological mechanisms by which a genetic variant affect performance is crucial to definitively associate that variant to performance. However, this is often a very laborious and elusive task. Nevertheless, a few experimental approaches might be of great use for this purpose.

Knowing whether the structural DNA variant affects or not gene and protein expression at the tissue level is probably the most fundamental question. If the polymorphism severely affects protein structure (e.g., a non-sense or a frameshift polymorphism), it is usually easier to observe its impact on physiology and on phenotype. In fact, the best-characterized polymorphism impacting physical performance is the R577X in the *ACTN3* gene, which is a non-sense polymorphism leading to the synthesis of a non-functional protein<sup>63</sup>. This is because the absence of the alpha-actinin-3 protein in humans carrying both polymorphic alleles can be mimicked in knockout animals, providing a very interesting model to evaluate how the lack of the protein impacts muscle structure, muscle metabolism and performance<sup>64</sup>. However, most polymorphisms are SNPs, indel or copy number variants, and they may occur at non-coding genomic regions. Consequently, their impact on protein structure is less evident and the creation of a good animal model is often unfeasible. In these cases, human studies are indeed necessary.

The knowledge of the tissues and cells where the polymorphic gene is expressed should dictate the best approach to examine how the polymorphism affects gene and protein expression. For example, considering a variant in a gene that expresses exclusively in skeletal muscle, then muscle biopsies should be collected from a group of individuals with all genotypes. These

samples could be submitted to a variety of analysis, including gene expression (e.g., qPCR), protein expression (e.g., western blotting), and morphological inspection (e.g., light or electron microscopy). Further analysis should also be carried out depending on the role of the gene. Exemplifying, if the gene encodes for an enzyme, then comparing enzyme activity between genotypes is probably a very reasonable approach. Obviously, in some cases human tissue collection is not an option, especially if the gene is expressed in heart, bone, kidney, liver brain or any other organ where a biopsy is too invasive and unjustifiable. This would imply the necessity for alternative and more indirect approaches. These scenarios illustrate quite well how difficult establishing a physiological link between a polymorphism and a phenotype can be.

In cases that an animal model can be generated so the genetic variation in humans is properly mimicked, acceptable sample sizes for these animal studies are substantially low, since all other sources of variation are controlled. In contrast, when human studies are performed and groups of individual with different genotypes are compared, sample size becomes critical. Because studies with humans, especially the cross-sectional ones, normally have poor control of major intervenient variables (e.g., diet, exercise background, genetic background, use of medications and development conditions), it is imperative that all genotypes have a large enough number of participants in order to minimize the chances of groups being different due to confounding factors rather than to genotype itself. In this case, the importance of large samples is not only about statistical power, but to conveniently decrease the likelihood of assuming that genotype groups are different when the cause of the difference is not the genotype "per se".

## Genetic interactions

A still unexplored, yet interesting and promising area of research in sports genetics is how genes interact with other genes to modulate exercise

responses or to modulate a phenotype that is strongly related to performance<sup>4</sup>. It is well recognised that genes do not act in isolation. Rather, there are

complex interactions among many genes modulating phenotypes<sup>42</sup>. Some polymorphisms might not have a meaningful impact on athletic performance alone, but the presence of other polymorphisms, through gene-gene interactions, may enlarge its impact upon phenotype, so that not a presence of a single polymorphism but, instead, the combination of some specific polymorphisms may be potentially beneficial to some performance-related phenotypes<sup>4</sup>.

In addition to the interactions between genes, gene-environment interactions must be taken into consideration in future research, especially in light of

the multifactorial nature of sports performance. After certain environmental stimuli, some polymorphisms may have greater or lesser influence on the responses to those stimuli<sup>26</sup>. The replication of the results in different cohorts, under different environmental stimuli, is an indirect way to assess this interaction. The researchers' primary efforts have been focused on the discovery of novel polymorphisms with the capability to influence physical fitness components. Subsequently establishing a panorama for one or more polymorphisms, possible gene-gene and gene-environment interactions should be investigated.

## Rare variants

It is well accepted that the influence of common genetic variants on performance is slight, but the combination of several polymorphisms probably is more influential<sup>16</sup>. On the other hand, it is highly plausible that rare genetic variants or mutations may eventually confer a very significant advantage to performance. Rare variants, therefore, are probably a very important aspect of the genetic variability that underlies sports excellence. However, due to the extreme difficulty to identify and characterize these variants, this is still poorly understood<sup>15</sup>.

In most instances, mutations lead to loss of function, disease or disabilities, which adversely affects sports-related phenotypes<sup>24</sup>. A good example is the mutation found in the *PYGM* gene that results in a deficit in carbohydrate metabolism (i.e., McArdle's disease) and, as a consequence, the patient presents intolerance to physical exercise. Many other examples of mutations that would cause incompatibility with athletic phenotypes could be

provided. However, not all mutations are detrimental to physical performance. The notorious case of a mutation resulting in the absence of the myostatin protein illustrates how a rare variant can represent an extreme advantage to some performance-related phenotypes without causing any harm to health<sup>65</sup>. It must be noted, however, that it is currently unknown whether this specific mutation in the myostatin gene is in fact favourable to physical performance.

Because each rare variant will be carried by no more than a few individuals, identifying these rare variants and correlate them to a specific sports-related phenotype is not an easy task. Nonetheless, efforts should be made in order to identify new rare variants and broaden our knowledge about the genetics of the sports excellence. Maybe performing a genome-wide scan of individuals presenting abnormally high phenotypes such as muscle mass, strength, flexibility and aerobic capacity and extraordinarily excellent performance in sports is a reasonable approach to search for new rare variants.

## The use of genetic markers to detect sports talent

The molecular mechanisms influencing athletic performance take place and are regulated by genes; these are gradually being revealed. Some authors argue that testing for the presence of key genetic variants in youth can be used as a way to select potential athletes<sup>61</sup>. The early identification of potential elite athletes could, in theory, optimize training plans and competitive stimuli during growth and development, therefore increasing the chances of reaching the peak of physical performance<sup>61</sup>. This process of identifying talent by

means of polymorphisms (genetic testing) could, in principle, be revolutionary to the field of sport<sup>25</sup>.

The publication of YANG et al.<sup>66</sup> describing the influence of the polymorphism R577X in *ACTN3* gene on athletic performance had a major impact on the scientific community and in general population. Some authors<sup>67</sup> have stated that the association of *ACTN3* with the physical performance is sufficient for using this polymorphism as a tool to select potential athletes. In 2004, just one year after

the publication of Yang and colleague's paper, an Australian company and after, in 2008, an American company began to develop and sell, at a cost of US \$170, claiming that the genetic test would determine to which sport the person would have a higher "vocation"<sup>68</sup>. Currently, there are at least 7 companies that sell genetic tests, now with a broader range of polymorphisms being offered<sup>61</sup>.

Despite the strong appeal of these genetic tests, the information provided by them is useless to anyone who seeks the competitive career in any sport. As discussed throughout this paper, sports performance is multifactorial. Genetics comprises only one of many contributing factors. Moreover, the genes and variants that may have a positive effect on performance are numerous, all of them under the regulation of an extremely complex network of other genes and variants, among other factors. Therefore, the presence or the absence of a few polymorphisms will never have such a predictive power. The literature presents many good examples of very successful athletes who, nevertheless, presented a genetic profile that could be considered "unfavourable". Regarding the *ACTN3*, which is one of the best-characterized polymorphisms<sup>69</sup>, there are athletes achieving great results in strength activities that do not display the RR genotype<sup>70</sup>.

The Council of Europe Bioethics Convention and the Genetic Information Non-Discrimination Act in the US consider the use of the genetic testing to predict performance an unethical action and they are currently evaluating whether they should be banned<sup>71</sup>. Besides the absence of scientific validity and the lack of useful information, three other important ethical issues must be highlighted in relation to information misuse: 1) the autonomy of the individual to not to know their genetic information, since some polymorphisms associated physical

performance may also be associated with diseases; 2) invasion of privacy, as a personal information can be passed to other people and used in a discriminative way; 3) professional misconduct and diminished opportunities, as coaches and trainers may not be willing to invest their time and efforts in an individual without "genetic predisposition" to sports<sup>61</sup>.

With the intention to find gifted individuals at early stages, the major target of genetic testing would be children and youths. Some authors point that fanatic relatives could take their children to train very hard, if they believe that the children have more changes based on results of genetic testing<sup>72</sup>. Knowingly, excessive training loads at young ages could represent a health harm, and result in overtraining, burnout and sport abandonment<sup>73</sup>. It is important to note that, as highlighted by HAWLEY<sup>74</sup>, the athlete who knows your genetic information may feel discouraged to keep practicing, especially if the information is negative for his/her expectation.

A potential use of genetic testing is related to the individualized prescription of exercise<sup>26</sup>. Several researchers argue that genetics should be used as an aid for improving exercise prescription and optimizing training responses in non-athletes<sup>75-80</sup>. Although some authors have already published training recommendations according to athletes' genotype<sup>67</sup>, it must be noted that there is no currently available information to support any genotype-directed training for athletes. In fact, there is no study addressing how genotypes influence training responses in athletes. Although in principle the proposal of genetic tests is revolutionary, at the moment, we still have no evidence supporting the creation of any model with acceptable predictive value. At present, the use of genetic tests has no advantage over the traditional methods for talent identification already used by trainers<sup>25</sup>.

## Conclusion

The future of genetic studies involving athletes is promising. In recent years, many polymorphisms have been associated with athletic phenotypes, but definitive confirmation of association and the underlying physiological mechanisms are proven difficult tasks. The challenges to progress in this novel area are enormous, but a variety of experimental approaches can be used to unravel part of the mystery. Researchers and the general population should be

conscious about the implications of the misuse of the genetic information. While some people may claim that genetic information could be used to detect talent and to drive athletic development, it must be noted that there is no scientific evidences for the predictive value of genetic in sports. The most appropriate statement at the moment is that genetics is only one out of many contributing factors to the athletic performance, and sometimes it may play only

secondary roles. It will be a long way until we know exactly what is the role of genetics for each sport and which are, at the molecular level, the variants accounting for this and how they work.

## Note

a. In fact, some genes are present in more than one copy per haploid genome (considering only the 23 non-redundant chromosomes), which means that there are more than two copies of these genes per diploid genome (considering the entire 46 chromosomes).

## References

1. Kiss MAPDM, Böhme MTS, Mansoldo AC, Degaki E, Regazzini M. Performance and sports talent. *Rev Paul Educ Fís.* 2004;19(n.Esp):89-100.
2. Gibson WT. Key concepts in human genetics: understanding the complex phenotype. *Med Sport Sci.* 2009;54:1-10.
3. Tucker R, Collins M. What makes champions? A review of the relative contribution of genes and training to sporting success. *Br J Sports Med.* 2012;46:555-61.
4. Eynon N, Ruiz JR, Oliveira J, Duarte JA, Birk R, Lucia A. Genes and elite athletes: a roadmap for future research. *J Physiol.* 2011;589:3063-70.
5. Komi PV, Viitasalo JH, Havu M, Thorstensson A, Sjodin B, Karlsson J. Skeletal muscle fibres and muscle enzyme activities in monozygous and dizygous twins of both sexes. *Acta Physiol Scand.* 1977;100:385-92.
6. Bouchard C, Simoneau JA, Lortie G, Boulay MR, Marcotte M, Thibault MC. Genetic effects in human skeletal muscle fiber type distribution and enzyme activities. *Can J Physiol Pharmacol.* 1986;64:1245-51.
7. Arden NK, Spector TD. Genetic influences on muscle strength, lean body mass, and bone mineral density: a twin study. *J Bone Miner Res.* 1997;12:2076-81.
8. Bouchard C, An P, Rice T, et al. Familial aggregation of VO(2max) response to exercise training: results from the HERITAGE Family Study. *J Appl Physiol.* 1999;87:1003-8.
9. Calvo M, Rodas G, Vallejo M, et al. Heritability of explosive power and anaerobic capacity in humans. *Eur J Appl Physiol.* 2002;86:218-25.
10. Gayagay G, Yu B, Hambly B, et al. Elite endurance athletes and the ACE I allele-the role of genes in athletic performance. *Hum Genet.* 1998;103:48-50.
11. Montgomery HE, Marshall R, Hemingway H, et al. Human gene for physical performance. *Nature.* 1998;393:221-2.
12. Ostrander EA, Huson HJ, Ostrander GK. Genetics of athletic performance. *Annu Rev Genomics Hum Genet.* 2009;10:407-29.
13. Bray MS, Hagberg JM, Perusse L, et al. The human gene map for performance and health-related fitness phenotypes: the 2006-2007 update. *Med Sci Sports Exerc.* 2009;41:35-73.
14. Hagberg JM, Rankinen T, Loos RJ, et al. Advances in exercise, fitness, and performance genomics in 2010. *Med Sci Sports Exerc.* 2011;43:743-52.
15. Roth SM, Rankinen T, Hagberg JM, et al. Advances in exercise, fitness, and performance genomics in 2011. *Med Sci Sports Exerc.* 2012;44:809-17.
16. Bouchard C. Overcoming barriers to progress in exercise genomics. *Exerc Sport Sci Rev.* 2011;39:212-7.
17. Williams AG, Folland JP. Similarity of polygenic profiles limits the potential for elite human physical performance. *J Physiol.* 2008;586:113-21.
18. Lucia A, Moran M, Zihong H, Ruiz JR. Elite athletes: are the genes the champions? *Int J Sports Physiol Perform.* 2010;5:98-102.
19. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Molecular biology of the cell.* New York: Garland Science; 2008.
20. IHGSC. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature.* 2004;431:931-45.
21. Salisbury BA, Pungliya M, Choi JY, Jiang R, Sun XJ, Stephens JC. SNP and haplotype variation in the human genome. *Mutat Res.* 2003;526:53-61.

22. Mills RE, Pittard WS, Mullaney JM, et al. Natural genetic variation caused by small insertions and deletions in the human genome. *Genome Res.* 2011;21:830-9.
23. Gonzaga-Jauregui C, Lupski JR, Gibbs RA. Human genome sequencing in health and disease. *Annu Rev Med.* 2012;63:35-61.
24. Bouchard C, Rankinen T, Timmons JA. Genomics and genetics in the biology of adaptation to exercise. *Compr Physiol.* 2011;1:1603-48.
25. MacArthur DG, North KN. Genes and human elite athletic performance. *Hum Genet.* 2005;116:331-9.
26. Rankinen T, Bouchard C. Gene-exercise interactions. *Prog Mol Biol Transl Sci.* 2012;108:447-60.
27. Lucia A, Gomez-Gallego F, Barroso I, et al. PPARGC1A genotype (Gly482Ser) predicts exceptional endurance capacity in European men. *J Appl Physiol.* 2005;99:344-8.
28. Maciejewska A, Sawczuk M, Cieszczyk P, Mozhayskaya IA, Ahmetov II. The PPARGC1A gene Gly482Ser in Polish and Russian athletes. *J Sports Sci.* 2012;30:101-13.
29. He Z, Hu Y, Feng L, et al. Is there an association between PPARGC1A genotypes and endurance capacity in Chinese men? *Scand J Med Sci Sports.* 2008;18:195-204.
30. Eynon N, Ruiz JR, Meckel Y, et al. Is the -174 C/G polymorphism of the IL6 gene associated with elite power performance? A replication study with two different Caucasian cohorts. *Exp Psychol.* 2010;96:156-62.
31. Ruiz JR, Eynon N, Meckel Y, et al. GNB3 C825T Polymorphism and elite athletic status: a replication study with two ethnic groups. *Int J Sports Med.* 2010;32:151-3.
32. He ZH, Hu Y, Li YC, et al. Are calcineurin genes associated with athletic status? A function, replication study. *Med Sci Sports Exerc.* 2011;43:1433-40.
33. Puthuchery Z, Skipworth JR, Rawal J, Loosemore M, Van Someren K, Montgomery HE. The ACE gene and human performance: 12 years on. *Sports Med.* 2011;41:433-48.
34. Ash GI, Scott RA, Deason M, et al. No association between ACE gene variation and endurance athlete status in Ethiopians. *Med Sci Sports Exerc.* 2011;43:590-7.
35. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet.* 2012;90:7-24.
36. Pitsiladis Y, Wang G. Necessary advances in exercise genomics and likely pitfalls. *J Appl Physiol.* 2011;110:1150-1.
37. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet.* 2008;9:356-69.
38. Ioannidis JP. Non-replication and inconsistency in the genome-wide association setting. *Hum Hered.* 2007;64:203-13.
39. Bouchard C, Sarzynski MA, Rice TK, et al. Genomic predictors of the maximal O<sub>2</sub> uptake response to standardized exercise training programs. *J Appl Physiol.* 2011;110:1160-70.
40. Puthuchery Z, Skipworth JR, Rawal J, Loosemore M, Van Someren K, Montgomery HE. Genetic influences in sport and physical performance. *Sports Med.* 2011;41:845-59.
41. Hughes DC, Day SH, Ahmetov II, Williams AG. Genetics of muscle strength and power: polygenic profile similarity limits skeletal muscle performance. *J Sports Sci.* 2011;29:1425-48.
42. Ruiz JR, Arteta D, Buxens A, et al. Can we identify a power-oriented polygenic profile? *J Appl Physiol.* 2010;108:561-6.
43. Artioli GG, Gualano B, Franchini E, Batista RN, Polacow VO, Lancha Junior AH. Physiological, performance, and nutritional profile of the Brazilian Olympic Wushu (kung-fu) team. *J Strength Cond Res.* 2009;23:20-5.
44. Franchini E, Del Vecchio FB, Matsushigue KA, Artioli GG. Physiological profiles of elite judo athletes. *Sports Med.* 2011;41:147-66.
45. Alves CR, Pasqua L, Artioli GG, et al. Anthropometric, physiological, performance, and nutritional profile of the Brazil National Canoe Polo Team. *J Sports Sci.* 2012;30:305-11.
46. Buxens A, Ruiz JR, Arteta D, et al. Can we predict top-level sports performance in power vs endurance events? A genetic approach. *Scand J Med Sci Sports.* 2011;21:570-9.
47. Ruiz JR, Gomez-Gallego F, Santiago C, et al. Is there an optimum endurance polygenic profile? *J Physiol.* 2009;587:1527-34.
48. Muniesa CA, Gonzalez-Freire M, Santiago C, et al. World-class performance in lightweight rowing: is it genetically influenced? A comparison with cyclists, runners and non-athletes. *Br J Sports Med.* 2010;44:898-901.
49. Gonzalez-Freire M, Santiago C, Verde Z, et al. Unique among unique. Is it genetically determined? *Br J Sports Med.* 2009;43:307-9.
50. Flueck M, Vaughan D, Westerblad H. Linking genes with exercise: where is the cut-off? *Eur J Appl Physiol.* 2010;110:1095-8.
51. Hoppeler H, Baum O, Lurman G, Mueller M. Molecular mechanisms of muscle plasticity with exercise. *Compr Physiol.* 2011;1:1383-412.

52. Aahmetov II, Williams AG, Popov DV, et al. The combined impact of metabolic gene polymorphisms on elite endurance athlete status and related phenotypes. *Hum Genet.* 2009;126:751-61.
53. Santiago C, Ruiz JR, Muniesa CA, Gonzalez-Freire M, Gomez-Gallego F, Lucia A. Does the polygenic profile determine the potential for becoming a world-class athlete? Insights from the sport of rowing. *Scand J Med Sci Sports.* 2010;20:e188-94.
54. Eynon N, Ruiz JR, Meckel Y, Moran M, Lucia A. Mitochondrial biogenesis related endurance genotype score and sports performance in athletes. *Mitochondrion.* 2011;11:64-9.
55. Dias RG. Genetics, human physical performance and gene doping: common sense versus the scientific reality. *Rev Bras Med Esporte.* 2011;17:62-70.
56. Gomez-Gallego F, Santiago C, Gonzalez-Freire M, et al. Endurance performance: genes or gene combinations? *Int J Sports Med.* 2009;30:66-72.
57. Timmons JA, Knudsen S, Rankinen T, et al. Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. *J Appl Physiol.* 2010;108:1487-96.
58. Keller P, Vollaard NB, Gustafsson T, et al. A transcriptional map of the impact of endurance exercise training on skeletal muscle phenotype. *J Appl Physiol.* 2010;110:46-59.
59. Coffey VG, Hawley JA. The molecular bases of training adaptation. *Sports Med.* 2007;37:737-63.
60. Pearson, H. Physiology: freaks of nature? *Nature.* 2006;444:1000-1.
61. Roth SM. Critical overview of applications of genetic testing in sport talent identification. *Recent Pat DNA Gene Seq.* 2012;6:247-55.
62. Lippi G, Longo UG, Maffulli N. Genetics and sports. *Br Med Bull.* 2009;93:27-47.
63. Yang N, Garton F, North K. alpha-actinin-3 and performance. *Med Sport Sci.* 2009;54:88-101.
64. MacArthur DG, Seto JT, Chan S, et al. An Actn3 knockout mouse provides mechanistic insights into the association between alpha-actinin-3 deficiency and human athletic performance. *Hum Mol Genet.* 2008;17:1076-86.
65. Schuelke M, Wagner KR, Stolz LE, et al. Myostatin mutation associated with gross muscle hypertrophy in a child. *N Engl J Med.* 2004;350:2682-8.
66. Yang N, MacArthur DG, Gulbin JB, et al. ACTN3 genotype is associated with human elite athletic performance. *Am J Hum Genet.* 2003;73:627-31.
67. Kambouris M, Ntalouka F, Ziogas G, Maffulli N. Predictive genomics DNA profiling for athletic performance. *Recent Pat DNA Gene Seq.* 2012;6:229-39.
68. Brooks MA, Tarini BA. Genetic testing and youth sports. *JAMA.* 2011;305:1033-4.
69. Pasqua LA, Artioli GG, Pires FO, Bertuzzi R. ACTN3 gene and sports performance: a candidate gene to success in short and long duration events. *Rev Bras Cineantropom Desempenho Hum.* 2011;13:477-83.
70. Lucia A, Oliván J, Gomez-Gallego F, Santiago C, Montil M, Foster C. Citius and longius (faster and longer) with no alpha-actinin-3 in skeletal muscles? *Br J Sports Med.* 2007;41:616-7.
71. McNamee MJ, Muller A, Van Hilvoorde I, Holm S. Genetic testing and sports medicine ethics. *Sports Med.* 2009;39:339-44.
72. Collier R. Testing the ethics of genetic testing in sports. *CMAJ.* 2012;184:E45-6.
73. Brenner JS. Overuse injuries, overtraining, and burnout in child and adolescent athletes. *Pediatrics.* 2007;119:1242-5.
74. Hawley JA. Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1253.
75. Lightfoot JT. Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1249.
76. Pescatello LS. Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1247.
77. Roth SM. Perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1243-5.
78. Thomis M. Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1251.
79. Timmons JA. Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1250.
80. Wagner PD. Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *J Appl Physiol.* 2008;104:1246.



ENDEREÇO

João Paulo Limongi França Guilherme  
Escola de Educação Física e Esporte - USP  
Av. Prof. Mello Moraes, 65  
05508-030 - São Paulo - SP - BRASIL  
e-mail: jplfguilherme@usp.br

Recebido para publicação: 15/01/2013

Aceito: 11/02/2013