

Title: Mitochondrial biogenesis related endurance genotype score and sports performance in athletes

Authors: Nir Eynon^a, Jonatan R Ruiz^b, Yoav Meckel^a, María Morán^{c*}, Alejandro Lucia^{d*}

*Equal contribution

Affiliations:

^aGenetics and Molecular Biology Laboratory, Life Sciences Division, The Zinman College of Physical Education and Sport Sciences at Wingate Institute, Israel.

^bDepartment of Biosciences and Nutrition at NOVUM, Unit for Preventive Nutrition, Karolinska Institute, Huddinge, Sweden

^cCentro de Investigación Hospital 12 de Octubre and CIBERER, Spain

^dUniversidad Europea de Madrid, Spain

Corresponding Author: Mr. Nir Eynon. Genetics and Molecular Biology Laboratory, Life Sciences Division, The Zinman College of Physical Education and Sport Sciences at the Wingate Institute, 42902, Israel. Phones: Office + (972) 9 8639368; Home (972) 4 6306334. Fax: + (972) 9 8639365. E-Mail: eynon@wincol.ac.il

Key Words: Genetics, mitochondria, endurance athletes, sprinters

Abstract

We determined the probability of individuals having the ‘optimal’ mitochondrial biogenesis related endurance polygenic profile, and compared the endurance polygenic profile of Israeli (Caucasian) endurance athletes (n=74), power athletes (n=81), and non-athletes (n=240). We computed a mitochondrial biogenesis related ‘endurance genotype score’ (EGS, scoring from 0 to 100) from the accumulated combination of six polymorphisms in the *PPARGCIA-NRF-TFAM* pathway. Some of the variant alleles of the polymorphisms studied were so infrequent, that the probability of possessing an ‘optimal’ EGS (=100) was 0% in the entire study population. However, the EGS was significantly higher ($P < 0.001$) in endurance athletes (38.9 ± 17.1) compared with controls (30.6 ± 12.4) or power athletes (29.0 ± 11.2). In summary, although the probability of an individual possessing a theoretically ‘optimal’ genetic background for endurance sports is very low, in general endurance athletes have a polygenic profile that is more suitable for mitochondrial biogenesis.

1. Introduction

Besides decreasing mortality and disability in late life (Chakravarty et al. 2008), one of the numerous health benefits of regular aerobic (endurance) exercise is improved skeletal muscle capacity for oxygen consumption, which is a direct result of higher mitochondrial content (Hood 2009). However, the underlying mechanisms which explain exercise-induced mitochondrial biogenesis remain to be clearly elucidated. This is a question of broad interest as mitochondrial biogenesis is critical for the normal function of cells, and mitochondria are both the target and source of radicals in cells, as well as one of the most important organelles for understanding the ageing process (Vina et al. 2009).

Mitochondrial synthesis is stimulated by the peroxisome proliferator-activated receptor γ coactivator 1 α (PPARGC1A)-nuclear respiratory factor (NRF)- mitochondrial transcription Factor A (TFAM) pathway. Briefly, the peroxisome proliferator-activated receptor delta (PPARD) induces promotion of PPARGC1A (Berger and Moller 2002), which is the first stimulator of mitochondrial biogenesis. The NRF1 and NRF2 are intermediate transcription factors that stimulate the synthesis of TFAM, and the latter is the final effector activating the replication of mitochondrial DNA molecules (Garesse and Vallejo 2001; Kanki et al. 2004; Larsson et al. 1998). Growing interest has focused on elucidating the factors that can affect the PPARGC1A-NRF-TFAM pathway. Physical inactivity and ageing affect this pathway, leading to reduced muscle aerobic capacity, increased tendency for mitochondrially mediated apoptosis, and muscle sarcopenia (Hood 2009). Regular exercise, however, can ameliorate this metabolic dysfunction, improve endurance, and help maintain muscle mass (Vina et al. 2009) while reducing disability and mortality in late life (Chakravarty et al. 2002).

Genetic variants in the *PPARGC1A*-*NRF-TFAM* pathway genes, per se (or in combination with exercise) can also have an effect. For instance, the functional C294T (rs2016520) polymorphism in the *PPARD* gene and the Gly482Ser (rs8192678) polymorphism in the *PPARGC1A* gene may play a key role in mRNA and/or protein activity. The C allele of the *PPARD* C294T polymorphism (located in exon 4) is associated with higher transcriptional activity of the *PPARD* promoter, by inducing a binding site for the Sp-1 transcription factor (Skogsberg et al. 2003). The minor *PPARGC1A* Ser482 allele is associated with lower muscle *PPARGC1A* mRNA (Ling et al. 2004) and with lower endurance exercise capacity (Lucia et al. 2005). The G allele of the A/G variant in intron 3 of the β_1 subunit of the *NRF2* gene (rs7181866) is associated with endurance athletic status (Eynon et al. 2009d) and with the maximal oxygen uptake response to endurance exercise training (He et al. 2008). However, and despite the essential role of TFAM in transcription and replication of mammalian mitochondrial DNA [in fact the *TFAM* knockout mouse model exhibits depletion of mtDNA and abolished oxidative phosphorylation (Larsson et al. 1998)], we recently found no association between three polymorphisms in the *TFAM* gene (rs1937, rs2306604 and rs1049432) and aerobic exercise capacity at the pre-training state or in response to endurance training (He et al. 2007).

To better understand how genetic factors modulate mitochondrial biogenesis, elite endurance athletes represent a somewhat unique model of study. These individuals have undergone extreme physiological adaptations (e.g., in muscle oxidative capacity, as a result of increased mitochondrial content) that are likely to be the consequence of years of training as well as of the interaction between exercise training and a favorable polygenic profile (i.e., a combination of ‘endurance-oriented’ genetic variants), which is yet to be identified (Ruiz et al. 2009).

In the present study we computed the theoretically optimal endurance polygenic profile associated with mitochondrial biogenesis in a Caucasian population from Israel. We selected six polymorphisms in the *PPARGCIA-NRF-TFAM* pathway genes that are known to be associated with endurance exercise performance capacity: (i) *NRF2* A/C (rs12594956) (Eynon et al. 2009a); (ii) *NRF2* A/G (rs7181866) (Eynon et al. 2009d; He et al. 2008); (iii) *NRF2* C/T (rs8031031) (Eynon et al. 2009a); (iv) peroxisome proliferator-activated receptor alpha (*PPARA*) intron 7 G/C (rs4253778) (Ahmetov et al. 2006; Eynon et al. 2009b); (v) *PPARD* C294T (Eynon et al. 2009c; Skogsberg et al. 2003); and (vi) *PPARGCIA* Gly482Ser (Eynon et al. 2009c; Lucia et al. 2005).

Our purpose was twofold. First, we determined the probability that individuals might possess the ‘optimal’ mitochondrial biogenesis related endurance polygenic profile. Second, we compared the endurance polygenic profile in Israeli (Caucasian) endurance athletes, power athletes, and healthy non-athletes (controls). We also investigated whether the ‘elite’ athletes (international level) had a more favorable endurance polygenic profile than the ‘sub-elite athletes’ (national level). We hypothesized that the mitochondrial biogenesis related endurance polygenic profile would be better in endurance athletes than in the other study participants, and that elite endurance athletes would also have a better profile than their national level peers.

2. Material and Methods

The study was approved by the Helsinki Committee of the Hillel-Yaffe Medical Center, Hadera, Israel. A written informed consent was obtained from each participant. The study conformed to the standards set by the latest revision of the Declaration of Helsinki.

2.1. Participants

One hundred and fifty-five track and field athletes (119 men and 36 women, age= 35.9 ± 12.2 yrs) volunteered to participate in the study. Athletes were included in the study sample only if they had participated in national/international track and field championships. Athletes were divided into two main groups: (i) an endurance-type group that included 74 (60 men) long distance runners whose main events were the 10,000m run and the marathon, and (ii) a power type group that included 81 (59 men) sprinters whose main event was the 100-200m dash (Table 1).

According to their individual best performance, athletes within each group were further divided into two subgroups: elite-level, i.e. those who had represented Israel in world track and field championships or in the Olympic Games [28 men (13 endurance), and 18 women (7 endurance)]; and national-level [91 men (47 endurance) and 18 women (7 endurance)] (Table 1).

The control group included 240 non-athletic healthy individuals (170 men and 70 women, age= 26 ± 3 yrs) who were randomly selected from the Israeli population. Controls were not engaged in physical activity on a regular basis. All subjects, athletes and non-athletes, were Israeli Caucasians for ≥ 3 generations, with an equivalent ratio of Jews coming from Arab countries (non-Ashkenazi) and Jews coming from Europe (Ashkenazi) in each group (2:1 ratio, respectively).

2.2. Genotyping

Genomic DNA was extracted from peripheral EDTA treated anti-coagulated blood using a standard protocol. Genotype analyses were performed as explained below in the Genetics and Molecular Biology Laboratory of the Zinman College of Physical Education and Sport Sciences at the Wingate Institute, Netanya, Israel. To ensure proper internal control, for each genotype analysis we used positive and negative controls from different DNA aliquots which were previously genotyped by the same method according to recent recommendations for replicating genotype-phenotype association studies (Chanock et al. 2007). For all polymorphisms, we used the polymerase chain (PCR) reaction and the resulting restriction fragment length polymorphism (RFLP) analysis was scored by two experienced and independent investigators who were blind to subject data.

Information on the primers, PCR annealing temperature, restriction enzyme, and fragments obtained for each allele, respectively, for all studied polymorphisms is shown in Table 2.

2.3. Probability of having an ‘optimal’ polygenic profile for mitochondrial biogenesis in the Caucasian (Israeli) population

We calculated the probability of any given Caucasian (Israeli) individual possessing every ‘optimal’ genotype for one up to all six polymorphisms by using the typical frequency of ‘optimal’ genotypes observed in an Israeli (Caucasian descent for ≥ 3 generations) population (Table 3). The studied polymorphisms were ranked in alphabetical succession based on their official symbols. Based on the observed typical

frequencies of the ‘optimal’ genotypes, we produced a scale indicating how the probability of possessing an 'optimal' genetic profile for mitochondrial biogenesis decreases as the number of polymorphisms included in the profile increases.

2.4. Mitochondrial biogenesis related endurance genotype score determination

We computed the combined influence of all six polymorphisms studied following the model used elsewhere (Ruiz et al. 2009; Ruiz et al. 2010; Williams and Folland 2008).

First, we scored each genotype within each polymorphism (Table 3). We assumed an additive model (equaling 0, 1 or 2), that is, on the basis of the number of alleles associated with higher potential for mitochondrial biogenesis that were carried by each subject for each polymorphism. Thus, we assigned a genotype score (GS) of 2, 1 and 0 to each individual genotype theoretically associated to highest, medium or lowest potential for mitochondrial biogenesis, respectively.

Second, we summed the GS of each single genotype ($GS_{NRF2\ A/C} + GS_{NRF2\ A/G} + GS_{NRF2\ C/T} + GS_{PPARA\ intron\ 7\ G/C} + GS_{PPARD\ T294C} + GS_{PPARGCIA\ G482S}$), which allowed us to construct an ‘endurance GS’ (EGS).

Third, the EGS was transformed to a 0-100 scale for easier interpretation, as follows: $EGS = (100/12) \times (GS_{NRF2\ A/C} + GS_{NRF2\ A/G} + GS_{NRF2\ C/T} + GS_{PPARA\ intron\ 7\ G/C} + GS_{PPARD\ T294C} + GS_{PPARGCIA\ G482S})$ where 12 is the result of multiplying 6 (number of studied polymorphisms) by 2, which is the score given to the ‘optimal’ or preferable genotype. An EGS of 100 represents an ‘optimal’ polygenic profile for mitochondrial

biogenesis -that is, that all GS are 2. In contrast, an EGS of 0 represents the ‘worst’ possible profile for mitochondrial biogenesis, that is, all GS are 0.

We created a data set of 50,000 hypothetical Israeli individuals, each with a randomly generated polygenetic profile (for all six polymorphisms) based on the frequency of each genotype for the Israeli population observed in our laboratory (Table 4). Finally, we examined the distribution of EGSs within this virtual population.

2.5. Statistical analysis

We calculated the mean the EGS obtained in the three study groups. We compared the EGS of endurance, power athletes and non-athletes (controls) with one-way analysis of variance (ANOVA), and used Tukey *post hoc* test for between-group comparisons. We also performed ANOVA to compare the EGS between elite- and national-level athletes within each group of endurance and power athletes. Furthermore, we performed one-way analysis of covariance to examine the EGS difference between groups after adjusting for sex and age (entered in the model as covariates). We evaluated the ability of the EGS to correctly classify potential endurance athletes from non-athletes (0=non-athlete, 1=endurance athlete) by receiver operating characteristic (ROC) curves (Zweig and Campbell 1993). We calculated the area under the ROC curve (AUC) and 95% confidence intervals (95%CI). Finally, we used binary logistic regression to study the relationship between EGS and endurance athletic status.

Finally, we investigated whether all six polymorphisms are actually better than any lesser combination. We conducted stepwise multivariate logistic regression, where genotypes were analysed as a recessive trait, and were entered as independent variables. The phenotype (athlete group) was entered as dependent variable, where the control group was coded as 0 (reference group), the endurance group was coded as 1 and the power group was coded as 2. The goodness of fit of the model was evaluated using Hosmer-Lemeshow statistic and its accuracy was assessed by calculating the area under the Receiver Operating Characteristic (ROC) curve (AUC) with 95% confidence intervals (CI). All statistical analyses were performed with the PASW (v. 18.0 for WINDOWS, Chicago).

3. Results

Table 3 shows the genotype frequencies for the three study groups. The typical frequency of ‘optimal’ genotype for each polymorphism in Israeli (Caucasian) population is shown in Table 4. The probability of a given Israeli (Caucasian) individual possessing an ‘optimal’ polygenic profile for endurance athletic status was 43% when considering just one polymorphism (the AA genotype for the *NRF2* rs12594956 polymorphism) and it was reduced to ~0.9% and to 0.044 when adding a second (rs7181866) and third (rs8031031) polymorphism, respectively. The probability of an Israeli individual for possessing an ‘optimal’ polygenic profile was 0%.

The mean EGS was significantly higher in endurance athletes compared with controls and power athletes (both $P < 0.001$) whereas the EGS was similar ($P = 0.964$) between the latter two groups (Figure 1). The EGS tended to be higher in elite-level endurance athletes compared with endurance athletes of national-level (44.6 ± 21.7 vs. 36.9 ± 14.6 respectively, $P = 0.084$), whereas there were no differences between elite and national power athletes (27.9 ± 13.9 vs. 29.5 ± 9.7 respectively, $P = 0.536$). The results did not change when the analyses were adjusted for sex and age (data not shown). We observed that three (4.1%) endurance athletes (all of elite level) had 5 optimal endurance genotypes, versus 0 subjects in the other two groups (Table 4). To note is that these three athletes with 5 optimal genotypes were among the very best competitors in the elite endurance group, i.e. top 1 in 10,000m (28 min 37 s) and top 2 (2 h 14 min 52 s) and 4 in marathon (2 h 15 min 45 s). It is also noteworthy that none of the endurance athletes had an ‘optimal’ endurance genotype profile.

The ROC analysis showed a significant discriminating accuracy of EGS in identifying an endurance athlete (AUC=0.634; 95%CI: 0.561-0.707; $P<0.001$; sensitivity: 0.716, specificity: 0.454) (Figure 2). The corresponding EGS value at this point was 27.16. Logistic regression analysis showed that subjects with EGS above 27.16 had an increased odds ratio (OR) of being an endurance athlete when compared to those with an EGS below this value (OR: 2.10; 95%CI: 1.193-3.697; $P=0.010$).

The frequency distributions of EGSs derived from 240 non-athletes (controls) and that obtained from 74 Israeli endurance athletes, and 81 Israeli power athletes are depicted in Figure 3.

The results of the stepwise multivariate logistic regression for the comparison of controls vs. endurance athletes indicated that two polymorphisms were significantly associated with endurance performance: *NRF2* A/G (OR, 2.044, 95%CI: 1.189-3.514) and *PPARGC1A* Gly482Ser (OR, 6.952, 95%CI: 2.213-21.844). The multivariate model combining genotypic data significantly predicted endurance performance (model $\chi^2=6.682$, d.f.=1, probability value 0.01), and variance explained by the function was 8.3% (Nagelkerke's pseudo- R^2). The ROC analysis showed a significant discriminating accuracy of the model, with an AUC=0.63 (95% CI: 0.547-0.699). When comparing endurance vs. power athletes, the model included three polymorphisms significantly associated with endurance performance: *NRF2* A/C (OR, 2.222, 95%CI: 1.129-4.374), *NRF2* A/G (OR, 7.946, 95%CI: 0.952-66.300), and *PPARA* intron 7 G/C (OR, 7.010, 95%CI: 0.913-60.457). The multivariate model combining genotypic data significantly predicted endurance performance (model $\chi^2=4.522$, d.f.=1, probability value 0.03), and variance

explained by the function was 15.8% (Nagelkerke's pseudo- R^2). The ROC analysis showed a significant discriminating accuracy of the model, with an AUC=0.66 (95% CI: 0.575-0.747).

4. Discussion

This is a novel study on the mitochondrial biogenesis related endurance polygenic profile of competitive athletes of the same ethnic origin (Caucasian, Israelis). One major finding was that the probability for the occurrence of individuals with the 'optimal' polygenic profile for mitochondrial biogenesis (EGS=100) is very low, that is, 0% in the Israeli population, including Olympic-level athletes. In fact, the EGS distribution was shifted to the left (mean EGS < 50) in all three groups (Figure 3), i.e., far from the 'optimal' score. Though three controls and one power athlete had four optimal genotypes (vs. n=0 in the elite endurance athlete group), the only three study participants with five optimal endurance genotypes were elite endurance athletes; in addition, the mean EGS was significantly higher in endurance athletes than in the other two groups, and also tended to be higher in elite-level endurance athletes compared with their peers of lower (national) level. When we investigated the question of whether all six polymorphisms are actually better than any lesser combination, we found that when comparing the endurance vs. the control group, only two polymorphisms were associated with endurance performance: *NRF2* A/G and *PPARGC1A* Gly482Ser. When comparing the endurance vs. the power group, there were three polymorphisms associated with endurance performance: *NRF2* A/C, *NRF2* A/G, and *PPARA* intron 7 G/C.

Taken together our data suggest that, although the probability of there being humans with a theoretically ‘optimal’ genetic background for endurance sports is very low, the competitive performance status of endurance athletes is most likely not only the result of stringent training regimens. Such performance would seem to arise from the interaction between environmental factors (i.e., years of intense training stimuli) and an overall ‘favourable’ (though not necessarily ‘optimal’) genetic endowment. This is in agreement with recent findings with other Caucasian (Spanish) athletes (Ruiz et al. 2009; Ruiz et al. 2010; Santiago et al. 2009). On the other hand, to note is that those polymorphisms for which we found greater differences between endurance athletes and controls (i.e., greater frequency of individuals with a GS of 2) were the three polymorphisms in *NRF2*, and *PPARGC1A* Gly482Ser (Table 4). Although the reason for this finding is not apparent given the key role of each of the four genes we studied in mitochondrial biogenesis, the *PPARGC1A* gene could play the most important influence in human aerobic capacity since this gene is a coactivator of the subset of *OXPHOS* genes that control mitochondrial biogenesis, glucose and lipid transportation and oxidation, and skeletal muscle fiber-type formation (Terada et al. 2002; Tunstall et al 2002). The fact that exercise training increases *PPARGC1A* mRNA levels, and that transgenic over-expression of *PPARGC1A* mRNA corresponds with an increased resistance of contracting muscle to fatigue (Lin et al. 2002), indicate that *PPARGC1A* and exercise are part of a co-regulatory feedback loop (Lucia et al. 2005).

Identifying genetic profiles associated with exercise endurance capacity and with mitochondrial biogenesis (involving the *PPARGC1A*-*NRF*-*TFAM* pathway) is of interest not only from a sports performance point of view, but also from a broader health perspective, especially in our ageing, sedentary western societies. Indeed, mitochondrial biogenesis is critical for the normal function of cells (Vina et al. 2009) and the

PPARGC1A-NRF-TFAM pathway is impaired by chronic physical inactivity. This can result in reduced muscle aerobic capacity (which is in itself an important health indicator), increased tendency for mitochondrially-mediated apoptosis, and, ultimately, accelerated sarcopenia (Viña et al. 2009). Further, research on animal models shows that impaired mitochondrial function (due to decreased amounts of transcription factors required for mitochondrial biogenesis) might be the link explaining the association between low fitness levels and increased cardiovascular/metabolic disease risk (Wisloff et al. 2005). In contrast, regular exercise can ameliorate muscle aerobic capacity, attenuate sarcopenia and reduce disability and mortality risk over the entire lifespan (Chakravarty et al. 2008; Vina et al. 2009). More research, including the analysis of other polymorphisms, is thus necessary to identify genetic variants that help explain individual variability in the magnitude of beneficial exercise-induced mitochondrial adaptations.

We also observed that the mean EGS was significantly higher in endurance than in power athletes, whereas the latter had an EGS similar to that of the general population. In addition, the ROC analysis showed a significant discriminating accuracy of the EGS model in identifying an endurance athlete (i.e., vs. a power athlete), and logistic regression analysis showed that subjects with EGS above 27.16 had ~2 times more possibilities of being a competitive endurance athlete than those with lower scores. Thus, an important finding of our EGS model is that, from a genetic point of view, a distinction can be made between athletes excelling in endurance or 'aerobic' sports vs. those more suited for power/sprint events. It seems unlikely to find an individual with a polygenic profile suitable to excel in both power and endurance sport events, as is also supported from recent data (Ruiz et al. 2010). Such genotype distinction (endurance vs. power) can be attributed to the fact that the phenotype

traits that determine performance in both types of events are probably different. For endurance sports, oxidative metabolism adaptations and thus mitochondrial biogenesis are limiting factors whereas other muscle phenotypes (including the ability to produce fast, powerful muscle contractions through the rapid energetic supply of anaerobic pathways) are more important for power athletes. In fact, there seems to exist a “trade-off”, achieved through human evolution by balancing natural selection, between sprint and endurance phenotypic traits; as such, an individual would be inherently predisposed toward performance in either sprint/power or endurance events (Garland et al. 1990).

On the other hand, our study is not without limitations. The major drawback comes from the fact that we did not assess phenotypes that are indicators, at least indirectly, of muscle mitochondrial function, notably maximum oxygen uptake (VO_{2max}). This variable reflects maximal rates of oxygen utilization (and maximal sustainable rates of oxidative ATP production) in skeletal muscle (Levine et al. 2008); it is also a strong marker of population-based fitness and disease/mortality risk (Blair et al. 1996; Myers et al. 2002). Further research is thus necessary to determine the association between the EGS and VO_{2max} in athletic and non-athletic cohorts. This approach would help improve the knowledge on the heritability of muscle oxidative potential.

In summary, we studied six candidate polymorphisms associated with mitochondrial biogenesis (*NRF2* A/C, rs12594956; *NRF2* A/G, *NRF2* C/T, *PPARA* intron 7 G/C, *PPARD* C294T, and *PPARGCIA* Gly482Ser) and computed a polygenic profile that seems to distinguish elite endurance athletes from both power athletes and the non-athletic population. It is noteworthy that the possibility of a given Israeli (Caucasian) individual possessing a theoretically ‘optimal’ mitochondrial biogenesis related endurance polygenic profile for endurance sports performance

and for the polymorphisms we studied is almost non-existent. However, endurance athletes are in general favoured by a polygenic profile that is more suitable for mitochondrial biogenesis.

Acknowledgements

This study was partially funded by the Swedish Council for Working Life and Social Research (FAS) and Fondo de Investigaciones Sanitarias (FIS, # ref. # PS09/00194).

References

- Ahmetov, II, Mozhayskaya IA, Flavell DM, Astratenkova IV, Komkova AI, Lyubaeva EV, Tarakin PP, Shenkman BS, Vdovina AB, Netroba AI, Popov DV, Vinogradova OL, Montgomery HE, Rogozkin VA 2006. PPARalpha gene variation and physical performance in Russian athletes. *Eur J Appl Physiol* 97: 103-108.
- Berger J, Moller DE 2002. The mechanisms of action of PPARs. *Annu Rev Med* 53: 409-435.
- Blair SN, Kampert JB, Kohl HW, 3rd Barlow CE, Macera CA, Paffenbarger RS, Jr, Gibbons LW 1996. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *JAMA* 276: 205–210.
- Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF, Jr., Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS 2007. Replicating genotype-phenotype associations. *Nature* 447: 655-660.
- Chakravarty EF, Hubert HB, Lingala VB, Fries JF 2008. Reduced disability and mortality among aging runners: a 21-year longitudinal study. *Arch Intern Med* 168: 1638-46.
- Eynon N, Alves AJ, Sagiv M, Yamin C, Sagiv M, Meckel Y 2009a. Interaction between SNPs in the NRF2 gene and elite endurance performance. *Physiol Genomics*

- Eynon N, Meckel Y, Sagiv M, Yamin C, Amir R, Sagiv M, Goldhammer E, Duarte JA, Oliveira J 2009b. Do PPARGC1A and PPARalpha polymorphisms influence sprint or endurance phenotypes? *Scand J Med Sci Sports* 20:e145-150.
- Eynon N, Oliveira J, Meckel Y, Sagiv M, Yamin C, Sagiv M, Amir R, Duarte JA 2009c. The guanine nucleotide binding protein beta polypeptide 3 gene C825T polymorphism is associated with elite endurance athletes. *Exp Physiol* 94: 344-349.
- Eynon N, Sagiv M, Meckel Y, Duarte JA, Alves AJ, Yamin C, Sagiv M, Goldhammer E, Oliveira J 2009d. NRF2 intron 3 A/G polymorphism is associated with endurance athletes' status. *J Appl Physiol* 107: 76-79.
- Garesse R, Vallejo CG 2001. Animal mitochondrial biogenesis and function: a regulatory cross-talk between two genomes. *Gene* 263: 1-16.
- Garland TJ, Bernnet AF, Daniels CB 1990. Heritability of locomotor performance and its correlates in a natural population. *Experientia* 46: 530-533.
- He Z, Hu Y, Feng L, Bao D, Xi Y, Wen L, Lucia A 2007. Relationship between TFAM gene polymorphisms and endurance capacity in response to training. *Int J Sports Med* 28: 1059-1064.
- He Z, Hu Y, Feng L, Li Y, Liu G, Xi Y, Wen L, Lucia A 2008. NRF-1 genotypes and endurance exercise capacity in young Chinese men. *Br J Sports Med* 42: 361-366.
- Hood DA 2009. Mechanisms of exercise-induced mitochondrial biogenesis in skeletal muscle. *Appl Physiol Nutr Metab* 34: 465-472.

Larsson NG, Wang J, Wilhelmsson H, Oldfors A, Rustin P, Lewandoski M, Barsh GS, Clayton DA. Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet* 1998; 18: 231-236

Kanki T, Ohgaki K, Gaspari M, Gustafsson CM, Fukuooh A, Sasaki N, Hamasaki N, Kang D 2004. Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA. *Mol Cell Biol* 24: 9823-9834.

Larsson NG, Wang J, Wilhelmsson H, Oldfors A, Rustin P, Lewandoski M, Barsh GS, Clayton DA 1998. Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet* 18: 231-236.

Levine BD 2008. VO₂max: what do we know, and what do we still need to know? *J Physiol* 586: 25-34.

Lin J, Wu H, Tarr PT, Zhang CY, Wu Z, Boss O, Michael LF, Puigserver P, Isotani E, Olson EN, Lowell BB, Bassel-Duby R and Spiegelman BM 2002. Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres. *Nature* 418: 797-801.

Ling C, Poulsen P, Carlsson E, Ridderstrale M, Almgren P, Wojtaszewski J, Beck-Nielsen H, Groop L, Vaag A 2004. Multiple environmental and genetic factors influence skeletal muscle PGC-1alpha and PGC-1beta gene expression in twins. *J Clin Invest* 114: 1518-1526.

Lucia A, Gomez-Gallego F, Barroso I, Rabadan M, Bandres F, San Juan AF, Chicharro JL, Ekelund U, Brage S, Earnest CP, Wareham NJ, Franks PW 2005. PPARGC1A genotype (Gly482Ser) predicts exceptional endurance capacity in European men. *J Appl Physiol* 99: 344-348.

Myers J, Prakash M, Froelicher V, Do, D, Partington, S, Atwood, JE 2002. Exercise capacity and mortality among men referred for exercise testing. *N Engl J Med* 4: 793-801.

Ruiz JR, Arteta D, Buxens A, Artieda M, Gomez-Gallego F, Santiago C, Yvert T, Moran M, Lucia A 2010. Can we identify a power-oriented polygenic profile? *J Appl Physiol* 108: 561-566.

Ruiz JR, Gomez-Gallego F, Santiago C, Gonzalez-Freire M, Verde Z, Foster C, Lucia A 2009. Is there an optimum endurance polygenic profile? *J Physiol* 587: 1527-1534.

Santiago C, Ruiz JR, Muniesa CA, Gonzalez-Freire M, Gomez-Gallego F, Lucia A 2009. Does the polygenic profile determine the potential for becoming a world-class athlete? Insights from the sport of rowing. *Scand J Med Sci Sports*

Skogsberg J, Kannisto K, Cassel TN, Hamsten A, Eriksson P, Ehrenborg E 2003. Evidence that peroxisome proliferator-activated receptor delta influences cholesterol metabolism in men. *Arterioscler Thromb Vasc Biol* 23: 637-643.

Terada S and Tabata I 2004. Effects of acute bouts of running and swimming exercise on PGC-1alpha protein expression in rat epitrochlearis and soleus muscle. *Am J Physiol Endocrinol Metab* 286: E208-216.

Tunstall RJ, Mehan KA, Wadley GD, Collier GR, Bonen A, Hargreaves M and Cameron-Smith D 2002. Exercise training increases lipid metabolism gene expression in human skeletal muscle. *Am J Physiol Endocrinol Metab* 283: E66-72.

Vina J, Gomez-Cabrera MC, Borrás C, Froio T, Sanchis-Gomar F, Martinez-Bello VE, Pallardo FV 2009. Mitochondrial biogenesis in exercise and in ageing. *Adv Drug Deliv Rev* 61: 1369-1374.

Williams AG, Folland JP 2008. Similarity of polygenic profiles limits the potential for elite human physical performance. *J Physiol* 586: 113-121.

Wisloff U, Najjar SM, Ellingsen O, Haram PM, Swoap S, Al-Share Q, Fernstrom M, Rezaei K, Lee SJ, Koch LG, Britton SL 2005. Cardiovascular risk factors emerge after artificial selection for low aerobic capacity. *Science* 307: 418-420.

Zweig MH, Campbell G 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 39: 561-577.

Table 1. Main characteristics of the three study groups. Data are mean±SD

	Endurance athletes (n=74)*		Sprinters (n=81)**		Controls (n=240)	
	Male (n=60)	Female (n=14)	Male (n=59)	Female (n=22)	Male (n=170)	Female (n=70)
Elite-level (n=46)	2h 19 min 57 s ± 2 min (n=14)	2 h 44 min 20 s ± 3 min (n=6)	10.43 ±0.15 s (n=15)	11.80±0.1 s (n=11)		
National-level (n=109)	2 h 44 min 6 s ± 25 min (n=46)	3 h 5min 20 s ± 35 min (n=8)	10.85± 0.26 (n=44)	12.22±0.34 s (n=11)		

*Best marathon time.

**Best 100 m sprint time.

Table 2. Information on genotyping methods for each polymorphism.

Gene, polymorphism	Reference ID	Primers 5' →3'	Annealing temperature	Restriction enzyme	Obtained fragment
<i>PPARGC1A</i> Gly482Ser	rs8192678	F - 5' TAAAGATGTCTCCTCTGATT 3' R- 5' GGAGACACATTGAACAATGAATAGGATTG 3'	50°	<i>HPAII</i>	Ser482 allele → 378 bp Gly482 allele →209 and 169 bp
<i>PPARA</i> intron 7 G/C	rs4253778	F- 5' ACAATCACTCCTTAAATATGGTGG '3 R- 5' AAGTAGGGACAGACAGGACCAGTA '3	59°	<i>Taq I</i>	C allele → 266 bp G allele →216 and 50 bp
<i>PPARD</i> T294C	rs2016520	R- 5' GAAGGAGCAGGAGCAGAAGA 3' R- 5' CAGTCATAGCTCTGGCATCG 3'	59°	<i>Bsl I</i>	T allele → 187 bp C allele →141 and 46 bp
<i>NRF2</i> A/C	rs12594956	F- 5' TAAAATGAATAAAGGTGGGGGT '3 R- 5' TAAGAGTGGAAGGGTGGAGAA '3	53°	<i>mfe I</i>	C allele → 407 bp A allele →277 and 130 bp
<i>NRF2</i> A/G	rs7181866	F- 5' AGTTTAGTGTCTCCCAGTGT 3' R 5' CTTAGTTTTCTTGTATCCGT 3'	50°	<i>Rsa I</i>	G allele → 483 bp A allele →284 and 199 bp
<i>NRF2</i> C/T	rs8031031	F- 5' CTAAAATGTGAGGGAAGGAAGA '3 R- 5' ATAGAGAGATAGGACTAAGGAC '3	57°	<i>Rsa I</i>	C allele → 208 bp T allele →158 and 50 bp

For all polymorphisms, the general cycle of PCR consisted of denaturation at 94°C for 5 min, 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min, extension at 72°C for 1 min, and a final extension step of 10 min at 72°C. All restriction enzymes were obtained from New England Biolabs (Beverly, MA, USA).

Table 3. Studied polymorphisms and genotype frequencies in the Caucasian (Israeli) population, and in Caucasian (Israeli) endurance and power athletes.

Symbol	Gene	Polymorphism	Genotypes (2= 'optimal' endurance genotype)	Frequency in controls (%)	Frequency in endurance elite athletes (%)	Frequency in power elite athletes (%)
<i>NRF2</i>	nuclear respiratory factor 2	A/C (rs12594956)	0=CC, 1=AC, 2=AA	11, 46, 43	5, 37, 58	18, 47, 35
		A/G (rs7181866)	0=AA, 2=AG and GG	98, 2	88, 12	98, 2
		C/T (rs8031031)	0=CC, 2=CT and TT	96, 4	89, 11	99, 1
<i>PPARA</i>	Peroxisome proliferator-activated receptor alpha	Intron 7 G/C (rs4253778)	0=CC, 1=CG, 2=GG	68, 28, 4	62, 28, 10	72, 27, 1
<i>PPARD</i>	Peroxisome proliferator-activated receptor delta	T294C (rs2016520)	0=TT, 1=CT, 2=C/C	39, 49, 12	39, 50, 11	47, 39, 14
<i>PPARGCIA</i>	Peroxisome proliferator-activated receptor-gamma, coactivator 1, alpha	Gly(G)482Ser(S) (rs8192678)	0=SS, 1=GS, 2=GG	18, 49, 33	0, 50, 50	13, 44, 43

Table 4. Probability of possessing an ‘optimal’ genetic endurance profile by number of polymorphisms in Israeli population.

Polymorphisms influencing endurance performance	New gene included at each stage	Typical frequency (%) of ‘optimal’ genotype in Israeli population	Probability of possessing an ‘optimal’ profile	
			% chance	Approximate odds ratio
1	<i>NRF2</i> A/C (rs12594956)	43	43.000	1 : 2
2	<i>NRF2</i> A/G (rs7181866)	2	0.900	1 : 111
3	<i>NRF2</i> C/T (rs8031031)	4	0.044	1 : 2273
4	<i>PPARA</i> (rs4253778)	4	0.044	1 : 2273
5	<i>PPARD</i> (rs2016520)	12	0	0
6	<i>PPARGCIA</i> (rs8192678)	33	0	0

Data obtained from a data set of 50 000 hypothetical Israeli individuals, each with a randomly generated genetic profile (for all 6 polymorphisms) based on the typical frequency of each genotype. See text for gene abbreviations. Polymorphisms are sorted in alphabetical order.

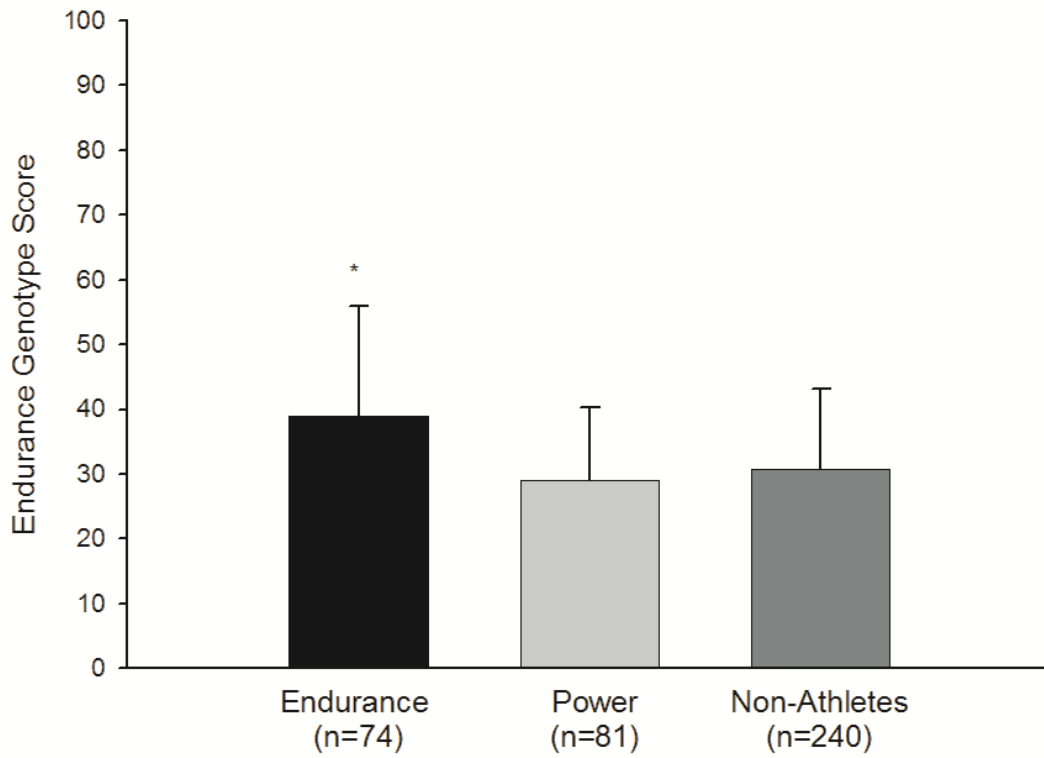


Figure 1. Endurance Genotype Score in endurance athletes, power athletes, and non-athletes (controls). * $P < 0.001$ for endurance vs. power and control. $P = 0.964$ for power vs. control.

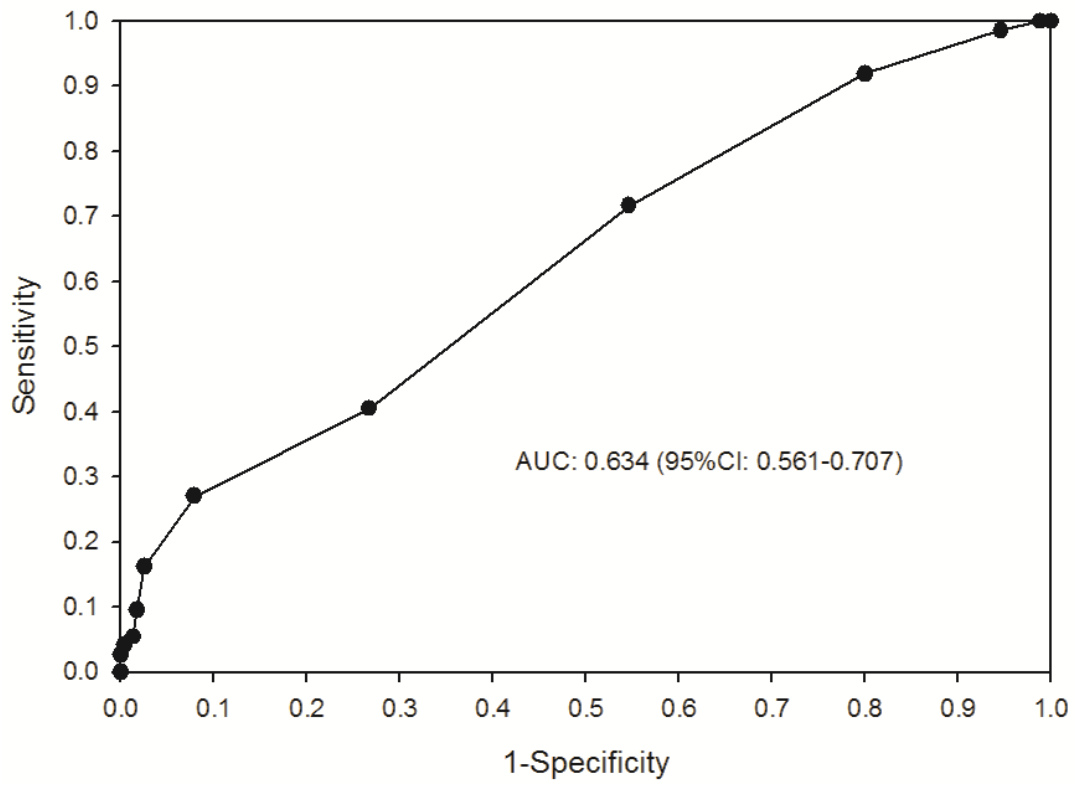


Figure 2. Receiver operating characteristic curve (ROC) summarizing the ability of endurance genotype score to classify potential endurance athletes from non-athletes (controls). AUC indicates the area under the curve (95% confidence intervals).

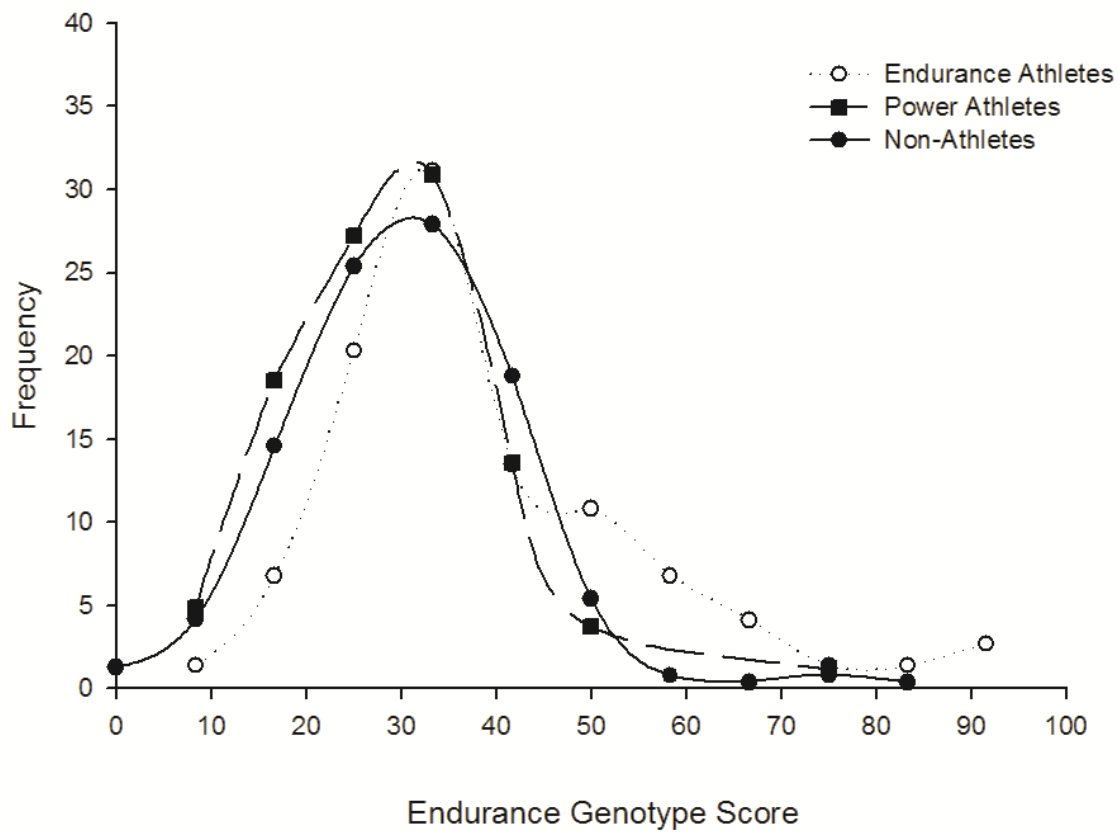


Figure 3. Frequency distribution of endurance genotype scores derived from 240 Israeli (Caucasian) individuals, 74 Israeli endurance athletes, and 81 Israeli power athletes.

The kurtosis statistic (mean±SE) in the non-athletic (control) Israeli sample was: 1.860±0.313; in the 74 Israeli endurance athletes: 1.541±0.552; and in the Israeli power athletes: 2.413±0.529.

