The rs12594956 polymorphism in the *NRF-2* gene is associated with top-level Spanish athlete's performance status

Abstract

Objectives: To determine the association between the nuclear respiratory factor 2 (*NRF-2)* polymorphisms and elite athletic performance. **Design:** We compared the genotype and allele frequencies of the *NRF-2* A/C (rs12594956), *NRF-2* A/G (rs7181866), and *NRF*-2 C/T (rs8031031) polymorphisms between world-class endurance athletes (n=89), elite power-oriented athletes (n=38), and non-athletic controls (n=110) of the same Caucasian (Spanish) origin. **Methods:** Genomic DNA was extracted from peripheral EDTA-treated, anti-coagulated blood using a standard protocol. Genotyping was performed using polymerase chain reaction (PCR). **Results:** The frequency of the AA genotype of the *NRF-2* A/C (rs12594956) polymorphism was significantly higher in endurance athletes compared with power athletes ($P<0.01$) and controls ($P<0.01$) (48% versus 13% and 21%, respectively). The likelihood of having the AA (rs12594956) genotype was higher in elite endurance athletes compared with controls [odds ratio (OR): 3.536, 95% confidence interval (CI): 1.903-6.571] and elite power athletes (OR: 6.170, 95%CI: 2.206-17.253).

Conclusions: Our results suggest that the *NRF-2* A/C polymorphism might belong to a growing group of polymorphisms associated with endurance performance at the elite level. However, it is important to replicate these findings in other groups of elite athletes using larger sample sizes.

Key words: Genetics, transcription factors, exercise, polymorphism

1. Introduction

A key component of skeletal muscles is the mitochondrion, which provides the energy required for muscle contraction via oxidative phosphorylation, especially during endurance exercise. Not surprisingly, this type of exercise also stimulates mitochondrial biogenesis, which improves the ability of mitochondria to convert biochemical energy from nutrients into ATP.[1] In this regard, the nuclear respiratory factor (NRF-2) plays a key role in regulating mitochondrial biogenesis.[2] Recently researchers have investigated the possible contribution of mitochondrial-related genes, such as the nuclear respiratory factor 2 (*NRF-2*) gene, in attaining elite endurance status.[3-5]

The NRF-2 protein is a transcription factor that was discovered as the human homolog of the mouse GA-binding protein (GABP). A structural analysis of NRF-2 revealed a high degree of sequence identity with the mouse GABP subunits.[6, 7] The *NRF-2* gene has been linked to the transcriptional control of many genes involved in mitochondrial function and biogenesis suggestively through nucleo-mitochondrial interactions which enhance mitochondrial DNA (mtDNA) levels and the activity of oxidative phosphorylation.[2, 7, 8]

Several lines of evidence imply that the *NRF-2* gene, or its product (NRF-2), plays a functional role within skeletal muscles during exercise. The mRNA levels of nuclear respiratory factor 1 (*NRF-1*) and *NRF-2* are significantly induced as part of the adaptation of skeletal muscles to exercise training.[9, 10] Furthermore, a previous study suggested that the β_1-subunit of the *NRF2* gene, located on chromosome 15q21.2, might be linked with elevated maximal oxygen consumption (VO_{2max}) in response to a 20-week endurance training program.[11]

In a recent study with elite Israeli athletes, we observed an association between the *NRF2* A/G (rs7181866), *NRF-2* A/C (rs12594956) and *NRF-2* C/T (rs8031031) polymorphisms and elite endurance status. Endurance athletes presented a higher frequency of AG, AA

and CT genotypes (as well as higher frequency of G, A and T alleles) in the *NRF-2* A/G (rs7181866), *NRF2* A/C (rs12594956) and *NRF2* C/T (rs8031031) polymorphisms respectively, compared with ethnically-matched sprinters and non-athletic controls.[3, 4] Further comparisons between the sub-groups of elite and national-level endurance athletes revealed that the theoretically endurance-favorable genotypes, *NRF-2* AG (rs7181866), *NRF-2* AA (rs12594956) and *NRF-2* CT (rs8031031), were more frequent in the group of elite level athletes than in the national-level group.[3, 4] Support for an influential role of the aforementioned *NRF-2* polymorphisms on endurance phenotypes was also provided in a previous study in healthy Han Chinese men, in whom both VO_{2max} and running economy were associated with the aforementioned gene variants.[12]

Notwithstanding the above findings in Israeli athletes, it is important to replicate genotype:phenotype associations in the field of sports genetics as the ethnic/geographic background of the cohorts may influence the results.[13, 14] For instance, the association that our group found between elite power athletic status and the 174 G/C polymorphism of the interleiukin-6 (*IL6*) gene in a Caucasian (Spanish) cohort[15] was not corroborated in Israeli Caucasians.[16] The aim of the present study was therefore to compare the genotype and allele frequencies of the *NRF-2* A/C (rs12594956), *NRF-2* A/G (rs7181866) and *NRF-2* C/T (rs8031031) polymorphisms between elite endurance athletes, elite power-oriented athletes, and non-athletic controls of Spanish ancestry.

2. Methods

Written consent was obtained from each participant. The study was approved by the ethics committee of Universidad Europea de Madrid, Spain. The study followed the recommendations for replicating genotype-phenotype association studies.[17, 18] The study

population was all of the same Caucasian (Spanish) descent for ≥3 generations and was comprised of:

i. 89 unrelated male world-class endurance athletes aged 20-39 (19 endurance runners, 32 professional road cyclists, and 38 rowers). All of the endurance runners (mostly specialists in the 5,000 m, 10,000 m and marathon) had participated in at least one Olympiad, and some were Olympic finalists or European/World Champions; the cyclists were all Tour de France participants; all the rowers were world-class as they had won at least one bronze, silver or gold medal in the lightweight category in the World Championships held during 1997-2006. Their mean±SD maximal oxygen uptake (VO_{2max}) was 77.9±6.9 $ml \cdot kg^{-1} \cdot min^{-1}$ (range: 62-87) (runners), 74.5±6.9 $ml \cdot kg^{-1} \cdot min^{-1}$ (62-86) (cyclists), and 71.7±5.5 $ml \cdot kg^{-1} \cdot min^{-1}$ (58-87) (rowers).

ii. 38 unrelated elite male power athletes aged 20-33 years (jumpers, throwers and sprinters), including the best Spanish jumpers and sprinters in recent years. Thirteen of them were Olympians during the period 2000-2008. Their VO_{2max} averaged 60.1±5.0 $ml \cdot kg^{-1} \cdot min^{-1}$ (range: 50- 69).

iii. 110 unrelated healthy male non-athletic controls aged 19-32 years. All were students from the same university (*Universidad Europea de Madrid*, Spain). Inclusion and exclusion criteria for this group were to be free of any diagnosed cardiorespiratory disease and not to be engaged in competitive sports or in formal, supervised exercise training (i.e., performing less than 3 structured weekly sessions of strenuous exercise such as running, swimming, bicycling, and weight lifting). Their VO_{2max} averaged 45.6±2.8 $ml \cdot kg^{-1} \cdot min^{-1}$ (range: 45-60).

The VO_2max values of the athletes was obtained using a breath-by-breath system (Oxycon Pro System, Jaeger, Wuerzburg, Germany) during laboratory treadmill gas-analysis, cycle-ergometer or rower-ergometer tests performed until volitional exhaustion.

The VO_{2max} of the controls was estimated from the time to complete a 2,000-m running test. The tests were performed on a 400-m outdoor track under similar environmental conditions (temperature, ~ 23-24º C; relative humidity, 45-55%; barometric pressure, ~ 720 mmHg). None of the VO_2max values were determined specifically for the present study (i.e., we retrieved them from our database).

Genomic DNA was extracted from peripheral EDTA-treated, anti-coagulated blood using a standard protocol. Genotyping was performed using polymerase chain reaction (PCR). The reaction and the resulting restriction fragment length polymorphism (RFLP) analysis were scored by two experienced and independent investigators who were blind to the participants' data. This method was verified using direct sequencing analysis. Information on the primers, PCR annealing temperature, restriction enzyme, and fragments obtained for each allele, respectively, for *NRF-2* A/C (rs12594956) and *NRF-2* C/T (rs8031031), and *NRF-2* A/G (rs7181866) polymorphisms is shown in Table 1.

Genotyping of *NRF-2* A/C (rs12594956), C/T (rs8031031), and A/G were performed with polymerase chain reaction (PCR). PCR for the *NRF-2* A/C was performed by denaturation at 94°C for 5 min, 34 cycles of denaturation at 94°C for 1 min, annealing at 53°C for 1 min, and extension at 72°C for 1 min, and a final extension step of 10 min at 72°C. The amplified fragment subsequently underwent digestion by *Mfe**I*** (New England Biolabs, Beverly, MA) in a condition recommended by the supplier. The digested products were then electrophoresed in a 2% agarose gel. PCR for the *NRF-2* C/T (rs8031031) was performed by denaturation at 94°C for 5 min, 34 cycles of denaturation at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 72°C for 1 min, and a final extension step of 10 min at 72°C. The amplified fragment subsequently underwent digestion by *Rsa**I*** (New England Biolabs). PCR for the *NRF-2* A/G was performed by 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. The amplified fragment subsequently underwent digestion by *Rsa1* (New England Biolabs, Beverly, MA) in a condition recommended by the supplier. The digested products were then electrophoresed in a 2.5% agarose gel.

To ensure proper internal control, for each genotype analysis we used positive and negative controls from different DNA aliquots that were previously genotyped with the same method, according to recent recommendations for replicating genotype-phenotype association studies.[18] The restriction fragment length polymorphism (RFLP) results were scored by two experienced and independent investigators who were blind to the participants' data.

We assessed deviations of genotype distributions from the Hardy-Weinberg equilibrium (HWE, using the chi-squared ($\chi 2$) test) in controls only (not in cases); indeed, in genetic associations studies (as the present one) that follow a case:control design (instead of a single-cohort design), deviation from the HWE should only be tested in controls because they are supposedly representative of the general population. [17] We used the $\chi 2$, Yates corrected $\chi 2$ test, or Fischer exact test to compare the genotype and allele frequency of the NRF-2 A/C (rs12594956), NRF-2 A/G (rs7181866) and NRF2 C/T (rs8031031) polymorphisms in the three study groups. We conducted binary logistic regression analysis to determine the association between alleles and sports performance. All analyses were corrected (by genotype) for multiple comparisons (i.e. 0.05/3, $P \leq 0.016$). All statistical analyses were performed using the PASW (v. 18.0 for WINDOWS, Chicago).

3. Results

Genotype distribution of the rs7181866 and rs8031031 polymorphisms were in HWE in the control group (P>0.01), yet that of the rs12594956 polymorphism was not (P<0.05). The results on genotype and allele frequencies of the *NRF2* A/G (rs7181866), *NRF2* A/C (rs12594956) and *NRF2* C/T (rs8031031) polymorphisms are shown in Table 2 and Figure 1, respectively. The frequency of the AA genotype for the *NRF-2* A/C (rs12594956) polymorphism was significantly higher in endurance athletes compared with power athletes (P<0.01) and controls (P<0.01) (48% versus 13% and 21%, respectively). No other significant between-group difference was found for this polymorphism. The likelihood of having the AA (rs12594956) genotype was higher in elite endurance athletes compared with controls [odds ratio (OR): 3.5, 95% confidence interval (CI): 1.9-6.6] and elite power athletes (OR: 6.2, 95%CI: 2.2-17.2).

On the other hand, we found no significant between-group differences in the genotype or allele distributions of the *NRF-2* A/G (rs7181866) polymorphism. Finally, the frequency of the CT genotype and T allele of the *NRF-2* C/T (rs8031031) polymorphism were significantly higher in endurance athletes compared with the control group (P=0.017 and 0.019, respectively). No other between-group difference was found for this polymorphism.

4. Discussion

The rationale for performing the current study was based on the putative role of the *NRF-2* gene in the process of mitochondrial transcription and regulation, and its potential influence on endurance performance. While keeping in mind the limitation that stems from the low sample size of our cohorts, our main finding was the higher frequency of the AA genotype of the *NRF-2* A/C polymorphism (rs12594956) among the endurance

athletes compared with both a sample of power athletes and a sample of healthy, non-athletic controls. Therefore, it can be assumed that harboring the *NRF-2* AA polymorphism is associated with success in endurance-type sports. This assumption was strengthened by the fact that we recruited world-class elite-level athletes who were the best caliber in their competition event. However, more data are needed to corroborate the association between the *NFR-2* gene and long-term endurance athletic status, and most importantly, the trainability of endurance phenotype traits.

The NRF-2 protein is involved in the control of basic cellular processes, such as cell cycle progression[19], protein synthesis, and mitochondrial biogenesis.[20] Being a master coordinator of the expression of all cytochrome C oxidase (COX) subunits, NRF-2 regulates the mechanism that senses upstream energy signals, and possibly controls oxygen consumption in the cells.[8] Recently, it was demonstrated that NRF-2 induces many of the human proteins active in mitochondrial DNA transcription and replication, such as: transcription termination factor (mTERF), the RNA polymerase POLRMT, the B subunit of the DNA polymerase γ, and the DNA helicase TWINKLE.[21] These findings strengthen the growing evidence regarding the key role of NRF-2 in determining an endurance-favorable phenotype. There are in fact preliminary data supporting that the *NRF-2* A/C (rs12594956) polymorphism, in particular, belongs to a growing group of genetic polymorphisms that affect elite endurance performance. We previously showed that the *NRF-2* AA genotype was strongly associated with endurance athletic status in elite Israeli athletes.[4] In a following study, using a genotype score model, which determines the probability of an individual having the 'optimal' mitochondrial-biogenesis-related endurance polygenic profile, we compared the endurance polygenic profile of Israeli (Caucasian) endurance athletes, power athletes, and non-athletes.[5] The results indicated that the probability of a given Israeli (Caucasian) individual possessing

an 'optimal' endurance athletic polygenic profile was marginally dependent upon possessing the AA genotype for the *NRF-2* rs12594956 polymorphism.

We believe that the results of our study are overall valid, as all of the following criteria were met[17]: phenotypes were accurately assessed, participants were ethnically-matched, genetic assessment was accurate, reliable and unbiased, genotype distributions were in HWE (except for the rs12594956 polymorphism–see below), we adjusted our statistical analyses for multiple comparisons, and our results are in line with previous findings.[4] However, it also must be kept in mind that the genotype distribution of the *NRF-2* A/C (rs12594956) polymorphism was not in HWE in the control group, which limits, at least partly, the external validity of our findings. Reasons for deviation from HWE in a given population are the following: genetic drift, migration (i.e. gene flow), mutation, selection, and non-random mating. It is difficult to determine, without speculating too much, which of the aforementioned conditions occurred in our control group, given its limited size. On the other hand, the low sample size of our population samples does also limit the 'external validity' (and therefore generalizability) of our results.[17] Given the unique phenotype of elite athletes, we believe that the small sample size of our athletic cohorts is justifiable. Indeed, we have gathered almost all elite Spanish (world-class) athletes with a 'pure' power phenotype (weightlifters, sprinters or throwers) and a high proportion of the best athletes in the country with an endurance phenotype (runners, cyclists, rowers).

5. Conclusion

In conclusion, our main finding suggests that there is an association between the *NRF-2* A/C polymorphism and elite endurance athletics status. Although the limited sample size of our cohort and its single ethnic/geographic origin does limit the external validity of our findings, this particular polymorphism might belong to a growing group of

polymorphisms associated with endurance performance at the elite level. Further replication studies involving other cohorts of elite athletes, as well as functional studies, are needed however needed to clarify just how this polymorphism contributes to elite endurance performance and trainability.

Practical implications

- Via a simple blood test talent identification programs will be able to identify those who carry the *NRF-2* AA genotype, and who may have a greater chance of becoming an elite endurance athlete.
- This information will also assist to identify exactly which genes and gene pathways are involved in the process of becoming an elite athlete.

Acknowledgments

The present study was partially funded by the Spanish Ministry of Science and Innovation (RYC-2010-05957)

References

1. Hood DA. Mechanisms of exercise-induced mitochondrial biogenesis in skeletal muscle. *Appl Physiol Nutr Metab* 2009; 34(3):465-472.
2. Kelly DP, Scarpulla RC. Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. *Genes Dev* 2004; 18(4):357-368.
3. Eynon N, Sagiv M, Meckel Y, et al. NRF2 intron 3 A/G polymorphism is associated with endurance athletes' status. *J Appl Physiol* 2009; 107(1):76-79.
4. Eynon N, Alves AJ, Sagiv M, et al. Interaction between SNPs in the NRF2 gene and elite endurance performance. *Physiol Genomics* 2010; 41(1):78-81.
5. Eynon N, Ruiz JR, Meckel Y, et al. Mitochondrial biogenesis related endurance genotype score and sports performance in athletes. *Mitochondrion* 2011; 11(1):64-69.

6. Virbasius JV, Virbasius CA, Scarpulla RC. Identity of GABP with NRF-2, a multisubunit activator of cytochrome oxidase expression, reveals a cellular role for an ETS domain activator of viral promoters. *Genes Dev* 1993; 7(3):380-392.

7. Virbasius JV, Scarpulla RC. Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis. *Proc Natl Acad Sci U S A* 1994; 91(4):1309-1313.

8. Ongwijitwat S, Liang HL, Graboyes EM, et al. Nuclear respiratory factor 2 senses changing cellular energy demands and its silencing down-regulates cytochrome oxidase and other target gene mRNAs. *Gene* 2006; 374:39-49.

9. Baar K, Wende AR, Jones TE, et al. Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1. *Faseb J* 2002; 16(14):1879-1886.

10. Murakami T, Shimomura Y, Yoshimura A, et al. Induction of nuclear respiratory factor-1 expression by an acute bout of exercise in rat muscle. *Biochim Biophys Acta* 1998; 1381(1):113-122.

11. Bouchard C, Rankinen T, Chagnon YC, et al. Genomic scan for maximal oxygen uptake and its response to training in the HERITAGE Family Study. *J Appl Physiol* 2000; 88(2):551-559.

12. He Z, Hu Y, Feng L, et al. NRF2 genotype improves endurance capacity in response to training. *Int J Sports Med* 2007; 28(9):717-721.

13. Eynon N, Ruiz JR, Oliveira J, et al. Genes and elite athletes: a roadmap for future research. *J Physiol* 2011; 589(Pt 13):3063-3070.

14. Rankinen T, Roth SM, Bray MS, et al. Advances in exercise, fitness, and performance genomics. *Med Sci Sports Exerc* 2010; 42(5):835-846.

15. Ruiz JR, Buxens A, Artieda M, et al. The -174 G/C polymorphism of the IL6 gene is associated with elite power performance. *J Sci Med Sport* 2010; 13(5):549-553.

16. 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 Physiol* 2011; 96(2):156-162.

17. Attia J, Ioannidis JP, Thakkinstian A, et al. How to use an article about genetic association: B: Are the results of the study valid? *JAMA* 2009; 301(2):191-197.

18. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature* 2007; 447(7145):655-660.

19. Yang ZF, Mott S, Rosmarin AG. The Ets transcription factor GABP is required for cell-cycle progression. *Nature Cell Biol* 2007; 9(3):339-346.

20. Rosmarin AG, Resendes KK, Yang Z, et al. GA-binding protein transcription factor: a review of GABP as an integrator of intracellular signaling and protein-protein interactions. *Blood Cells Mol Dis* 2004; 32(1):143-154.

21. Bruni F, Polosa PL, Gadaleta MN, et al. Nuclear respiratory factor 2 induces the expression of many but not all human proteins acting in mitochondrial DNA transcription and replication. *J Biol Chem* 2010; 285(6):3939-3948.

Table 1. Information on genotyping methods for each polymorphism in the *NRF-2* gene.

polymorphism	Reference ID	Primers 5' →3'	Annealing temperature	Restriction enzyme	Obtained fragment
NRF2 A/C	rs12594956	F- 5' TAAAATGAATAAAGGTGGGGGT '3 R- 5' TAAGAGTGGAAGGGTGGAGAA '3	53°	*mfe I*	C allele → 407 bp A allele →277 and 130 bp
NRF2 A/G	rs7181866	F- 5' AGTTTAGTGTCTCCCAGTGT 3' R 5' CTTAGTTTTCTTGTATCCGT 3'	55°	*Rsa I*	G allele → 483 bp A allele →284 and 199 bp
NRF2 C/T	rs8031031	F- 5' CTAAAATGTGAGGGAAGGAAGA '3 R- 5' ATAGAGAGATAGGACTAAGGAC '3	57°	*Rsa I*	C allele → 208 bp T allele →158 and 50 bp

Table 2. Genotype distribution of the *NRF2* A/C, *NRF2* A/G and *NRF2* C/T polymorphisms in all groups. Data is presented as absolute and relative values (within parentheses).

	Athlete groups	**n**	**AA**	**AC**	**CC**	**χ^2, P value E vs. P**	**χ^2,P value E vs. C**	**χ^2, P value C vs. P**
rs12594956								
NRF-2 A/C	Endurance (E)	89	43 (48)	21 (24)	25 (28)	20, **<0.01**	18.1, **<0.01**	0.41, 0.52
	Power (P)	38	5 (13)	25 (66)	8 (21)			
	Controls (C)	110	23 (21)	66 (60)	21 (19)			
			AA	AG	GG			
rs7181866								
NRF-2 A/G	Endurance (E)	89	86 (97)	3 (3)	0 (0)	0.11, 0.74	1.78, 0.18	0.16, 0.69
	Power (P)	38	36 (95)	2 (5)	0 (0)			
	Controls (C)	110	100 (90)	10 (10)	0 (0)			
			CC	CT	TT			
rs8031031								
NRF-2 C/T	Endurance (E)	89	81 (90)	8 (10)	0 (0)	2.3, 0.13	5.68**, 0.017**	0.62, 0.42
	Power (P)	38	38 (100)	0 (0)	0 (0)			
	Controls (C)	110	109 (99)	1 (1)	0 (0)			

Figure Legends

Figure 1. Allele frequencies of the *NRF-2* A/C (rs12594956), *NRF-2* A/G (rs718186), and the *NRF-2* C/T (rs8031031) polymorphisms in Spanish (Caucasian, all males) elite endurance athletes, elite power athletes and controls.

Fig 1A:

Endurance vs. Power: χ^2=3.72, P=0.054
Endurance vs. Controls: χ^2=3.001, P=0.08
Power vs. Controls: χ^2=0.36, P=0.55

Fig 1B:

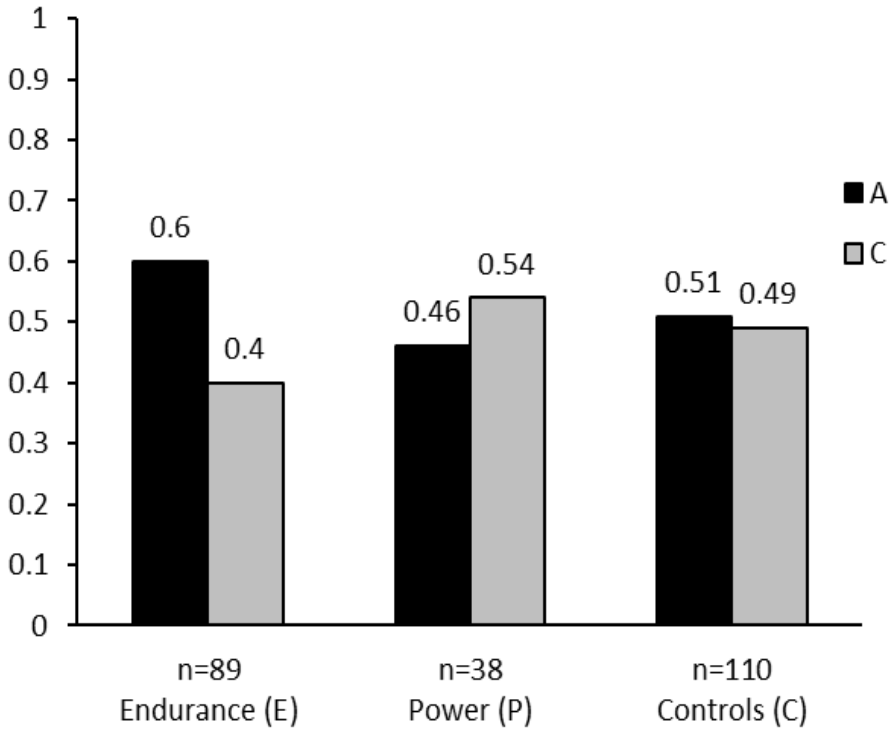
Endurance vs. Power: χ^2=0.25, P=0.62
Endurance vs. Controls: χ^2=1.72, P=0.19
Power vs. Controls: χ^2=0.15, P=0.7

Fig 1C :

Endurance vs. Power: χ^2=0.74, P=0.39
Endurance vs. Controls: $^{\chi 2}$=5.55, P=0.019
Power vs. Controls: χ^2=0.99, P=0.32

Figure 1A:

NRF-2 A/C (rs12594956)

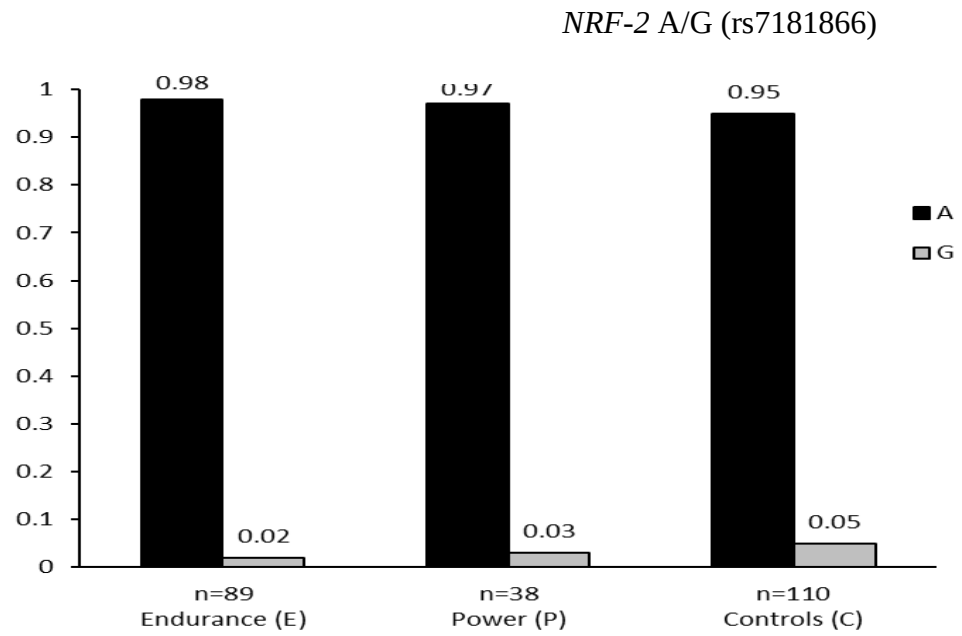


Endurance vs. Power: χ^2=3.72, *P*=0.054

Endurance vs. Controls: χ^2=3.001, *P*=0.08

Power vs. Controls: χ^2=0.36, *P*=0.5

Figure 1B:

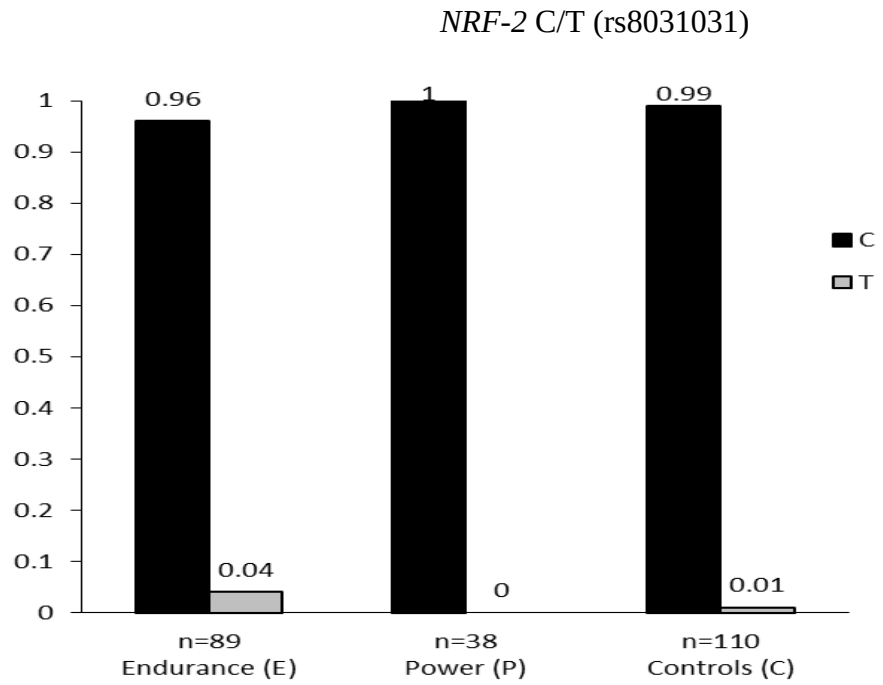


Endurance vs. Power: χ^2=0.25, *P*=0.62

Endurance vs. Controls: χ^2=1.72, *P*=0.19

Power vs. Controls: χ^2=0.15, *P*=0.7

Figure 1C:



Endurance vs. Power: χ^2=0.74, *P*=0.39

Endurance vs. Controls: χ^2=5.55, ***P*=0.019**

Power vs. Controls: χ^2=0.99, *P*=0.32