- Trabajo realizado por la Biblioteca Digital de la Universidad CEU-San Pablo

- Me comprometo a utilizar esta copia privada sin finalidad lucrativa, para fines de investigación y docencia, de acuerdo con el art. 37 de la M.T.R.L.P.I. (Modificación del Texto Refundido de la Ley de Propiedad Intelectual del 7 julio del 2006)

Medicina Sportiva
Med Sport 13 (4): 251-255, 2009
DOI: 10.2478/v10036-009-0039-9

AQUAPORIN-1 GENE DNA VARIATION PREDICTS PERFORMANCE IN HISPANIC MARATHON RUNNERS

Juan L. Martínez[1(B,E,G)], Aleshka Carrión[5(E,F)], Maria E. Florián[5(E,F)], José A. Martín[2(B,G)], Juan R. López-Taylor[3(B,G)], Thomas D. Fahey[4(E)],, Miguel A. Rivera[5(A,C,D,E)]

[1]Sport Medicine and Applied Sciences Unit, Sinaloense Sports Institute, Sinaloa, México
[2]Dept of Physiotherapy, Faculty of Medicine, University San Pablo CEU, Madrid, Spain
[3]Physical Activity and Applied Sport Sciences Institute, University of Guadalajara, Jalisco, México
[4]Dept of Kinesiology, California State University, Chico, CA USA
[5]Department of Physical Medicine, Rehabilitation and Sports Medicine, School of Medicine, University of Puerto Rico, San Juan, Puerto Rico

Abstract

Introduction and aim of the study: We examined the association between a DNA sequence variant in the aquaporin-1 (AQP1) gene and performance in Hispanic marathon runners.

Materials and methods: DNA samples (N = 784) were obtained from apparently healthy, biologically unrelated, ethnically matched Hispanic marathon runners who participated in international level marathons in Spain and Mexico. Cases (n = 396; men= 225; women=171) finished in the top 3rd percentile for their age and gender, while Controls (n = 388; men=221 women=167) finished in the lowest 3rd percentile. The AQP1 gene DNA sequence variant was detected by the tetra-primer ARMS-PCR procedure which targeted an area of 984 base pairs in the 3' untranslated region that included the polymorphic site (C - G).

Results: The three expected AQP1 genotypes were observed in both groups (Cases: CC = 15%. GC = 42%, GG = 43%; and Controls: CC = 10%, GC = 39%, GG = 51%). These genotype frequencies were in Hardy-Weinberg equilibrium (X^2, $P \geq 0.05$) and significantly ($X^2 = 6.94$, $P=0.03$) different between the Cases and Controls. The Allelic frequencies (Cases; C = 0.36, G = 0.64 and Controls; C = 0.30, G = 0.70) were significantly ($X^2 = 7.55$, $P=0.005$) different between groups. The Odds ratio = 1.35; 95% confidence interval = 1.08 – 1.67, suggested that the C allele was more prevalent in Cases than Controls.

Conclusions: This study showed an association between a DNA sequence variant in the aquaporin-1 gene or in its surrounding region and performance in Hispanic marathon runners.

***Key words**: genetics, chromosome 7, athletes, aerobic capacity, running, endurance*

Introduction

The present study is part of a larger investigation aimed at identifying genes contributing to endurance performance in Hispanics. The possible contribution of DNA sequence variation in genes encoding for the water channel family (AQPs) might explain the mechanism for rapid movement of water needed during endurance exercise. AQPs are a family of small transmembrane proteins expressed in many cell types involved in fluid transport (1). The AQPs family consists of known 13 members (1). The Aquaporin-1 (AQP1) is the best known and most studied of this family. The AQP1 gene is located on chromosome 7, region p14, it extends 17 kilobase pairs and contains 4 exons and 3 introns (2). This gene displays more than 150 DNA sequence variations (3). AQP1 gene encodes for a protein responsible **for** transporting large amounts of water across cell membranes.

The AQP1 polypeptide is a complex, sophisticated and regulated 28-kDa protein (2) with multiple functional capacities (4). It facilitates the rapid movement of water across cell membrane and functions as a cation channel (5). This protein exists as a tetramer with each subunit containing its own functionally independent pore (6-8). The water permeability and the ion conductance do not occur through the same pathway (5), as the individual subunit pores transports water and the central pore acts as a selectively ion channel that allows the flow of ions and water (4). This central pore is outlined by hydrophobic residues that exclude small chemicals species such as H^+ and NH_3 (9). The regulation of cation permeability of AQP1 through a cGMP-dependent signaling cascade has significant importance to the control of secretion and absorption of fluid, and regulation of cell volume in many tissues that highly express this channel (4). AQP1 has been identified in various tissues (10, 11), including red blood cells, endothelial cells, as well as smooth, skeletal and cardiac muscle

During osmotic stress, such as occurs during intense exercise (10, 11), AQP1 facilitates the transfer of water from the blood into the muscle (12), provides

osmotic protection, and promotes water reabsorption. In humans, AQP1 null individuals led normal lives and were entirely unaware of any physical limitations (6). However, they were unable to maintain fluid homeostasis when exposed to subacute or chronic fluid overload (6). This might reduce the amount of oxygen that could enter the blood and decrease endurance performance. The association between AQP1 genotype and endurance performance phenotype in humans has not been yet investigated. The present study investigated the allelic and genotypic frequency distribution of an AQP1 gene polymorphism and explored its association with performance in healthy Hispanic marathon runners.

Methods

Subjects

This investigation was a case-control study. The subjects were ethnically matched, biologically unrelated volunteers, who were recruited during the activities of the following long distance running events conducted between 2004 and 2008: Marathon of Madrid, Madrid, Spain; Gran Maratón Pacifico, Mazatlán, México; Guadalajara International Marathon, Guadalajara, México; Marathon of Culiacán, Culiacán, México. Gender specific values for the physical characteristics for the cases and controls are presented in Table 1.

Table 1. *Characteristics of 380 biologically unrelated marathon runners in the Hispanic Gene Study*

Group	Gender	Age (years)	Height (cm)	Weight (kg)
Cases (N=396)	Men (n=225)	33 ± 12	177 ± 9	60 ± 8
	Women (n=171)	29 ± 10	159 ± 9	50 ± 12
Controls (N=388)	Men (n=221)	38 ± 15	172 ± 5	64 ± 8
	Women (n=167)	32 ± 11	159 ± 10	50 ± 14

Values are mean ± standard deviation

The recruitment process included an interview to determine ethnic and geographic area of origin, place of residence, history of endurance events participation, cardiovascular disease, respiratory disease, arthritis activity, history of metabolic disease, as well as use of related medications. Any subject with known history of cardiovascular, respiratory, arthritis, metabolic disease or on medication was excluded from this report. Subjects were excluded from the study if they reported any incident that negatively affected their performance during the race (e.g., significant injury). The subjects' classification as cases or controls was based on the official results of the competitions.

Cases. (n = 396; men= 225; women=171) were adult Hispanic marathon runners who finished in the top 3rd percentile of their respective age and gender group in at least one of the above identified long distance running events. Controls (n = 388; men=221 women=167) were adult Hispanic marathon runners who finished in the lowest 3rd percentile of their respective age and gender. In all instances, the study received the approval by the Medical Commissions from each of the competitions. Written informed consent was obtained from all subjects.

DNA Isolation

Genomic DNA was purified from whole blood by means of the Promega Wizard Genomic DNA Purification Kit (Promega Corporation, Madison, WI). Briefly, 20 milliliters of venous peripheral blood was collected from each subject, in tubes containing EDTA. The first step in the purification procedure required the lysis of the red blood cells as well as the white blood cells and their nuclei. An RNase digestion step followed. The cellular proteins were removed by a salt-precipitation step leaving the high molecular weight genomic DNA in solution. Next, the genomic DNA was concentrated and desalted by isopropanol precipitation. Finally, the extracted DNA was rehydrated in 800 µl of Promega DNA Rehydration Solution (10mM Tris-HCl (pH 7.4) and 1mM EDTA (pH 8.0)) and stored at -80°C (Forma Scientific Freezer 916, Marietta, Ohio). The concentration and quality, as evaluated by the absorbance ratio at 280 vs. 260 nm, of the DNA was determined by spectrophotometry.

Genotype determination

Genotypes were determined by the tetra-primer Amplification Refractory Mutation System Polymerase Chain Reaction (ARMS-PCR) procedure (13), which targeted an area of 657 base pairs in the 3' untranslated region that included the polymorphic (G/C) site (restriction site #1049305). The present single nucleotide polymorphisms was selected due to its known high Heterozygosity index (H >0.42) which predicted sufficient numeric variability in the expected genotypes as to allow for adequate statistical comparisons. The polymerase chain reaction (PCR) was performed in a GeneAmp PCR system 9800 (Applied Biosystems GeneAmp 9800, Foster City, CA). Primers were as follows: Forward inner primer (C allele): TGGAATCGTCCCTATATCAGGGCGTC; Reverse inner primer (G allele): TGCCACTTTGCAGAAGGAGGTCAGTC; Forward outer primer (5' - 3'): ACCTGCATGGTCAAGCCTCTTATGGG; and Reverse outer primer (5' - 3'): TCTCTGCTTTGTAGCCTGTCTGCTCTGC. Total volume of the PCR

reaction mixture was 25-μL containing 5 – 10 ng of genomic DNA and Qiagen Taq master mix (Qiagen Inc, Ventura, CA). The amplification protocol was: 1) one cycle of denaturation at 95°C for 5 min; 2) 30 cycles of denaturation at 95°C for 60 s, annealing at 65°C for 60 s, and extension at 72°C for 30 s; and 3) one final 10 min elongation cycle at 72°C. Preventive contamination measures were taken by the inclusion of PCR reaction mixture without DNA (negative control) in every run of amplification. The resulting fragments were separated by horizontal electrophoresis on 1.5% agarose gels. Each gel was run for 60 min at 150 mA while refrigerated at 10°C, stained with ethidium bromide, and photographed under UV transmitted lights *using a Bio-Rad Gel-Doc 2000 system.*. A 100 bp ladder was used as length marker to estimate the PCR product size. The product size for the C allele was 233bp while that for the G allele was 476bp.

After all the genotyping work was completed a systematic random sample of subjects DNA aliquots was performed. The selected DNA samples were submitted to a new round of PCR and genotyping as described above. In case of a mismatch between the first and second genotyping, the sample was resubmitted to a PCR and genotyping cycle. All the above described work (DNA extraction, storage and management, as well as PCR and genotyping was performed at the facilities of the Exercise Physiology and Genetics Laboratory of the Department of Physical Medicine, Rehabilitation and Sports Medicine at the School of Medicine of the University of Puerto Rico, San Juan.

Statistical analysis

The Pearson Chi-square test was used to examine differences in allele frequencies between genders, cases vs. controls and to determine whether the genotype prevalence distribution differed from that expected under the Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium describes a state in which genotype frequencies are constant from generation to generation and in which genotype frequencies are a product of allele frequencies. This condition prevails in situations in which there are an absence of mutations, migrations, non-random mating and environmental factors favoring particular genotypes.

The polymorphism information content (PIC) and the heterozygosity index (H) were calculated according to procedures described elsewhere (14, 15). The PIC (14) value is an indicator of the probability that a marker locus is informative. It is estimated from the frequency of heterozygotes that are determined by the number of marker alleles and their respective frequencies. H (15) is a probability measure of the likelihood that a randomly selected subject is heterozygous for any two alleles at a given gene locus. Odds ratios with 95% confidence interval as well as the Chi-square tests were performed using the Epi Info 3.5.1 software package (16). *P* values ≤0.05 were statistically significant.

Results

Genotype frequencies

Table 2 shows that the three expected AQP-1 genotypes were observed in the Cases and Controls. These observed genotype frequencies were in agreement with those expected under Hardy-Weinberg equilibrium (X^2, $P \geq 0.05$) and were not significantly (X^2, $P \geq 0.05$) different between genders. Given that there were similar genotypic (not shown) frequency distributions in men and women, for both Cases and Controls, the data for both genders were pooled. Chi square test on the pooled data revealed a significant ($X^2 = 6.94$, df=2. $P=0.03$) difference in the genotype prevalence distribution between the Cases and Controls.

Allelic frequencies

n both genders, within Cases and Controls, the C allele was the less frequently observed (Table 2). These observed allelic frequency distributions within the Cases as well as Controls revealed no gender differences (not shown; X^2, $P \geq 0.05$). Chi square test on the gender pooled allelic frequency distribution *(Cases; C = 0.36, G = 0.64 and Controls; C = 0.30, G = 0.70)* revealed significant differences ($X^2=7.55$, df=1, $P=0.005$) between Cases and Controls. Figure 1 illustrates the prevalence of the AQP-1 C-allele carrier

Table 2. *Absolute and relative (%) allele and genotype frequencies for the AQP1 gene DNA (C – G) sequence variant in the 3' untranslated region for 380 biologically unrelated marathon runners of the Hispanic Gene Study*

	Genotypes*			Alleles**				
Groups	CC (%)	CG (%)	GG (%)	C (%)	G (%)	Odds Ratio	95% Confidence Interval	P Value
Cases§ (n=396)	59 (0.15)	167 (0.42)	170 (0.43)	285 (0.36)	506 (0.64)	1.35	1.08 – 1.67	0.005
Controls‡ (n=388)	39 (0.10)	151 (0.39)	198 (0.51)	229 (0.30)	547 (0.70)			

* $X^2 = 6.94$, df=2. P=0.03; **$X^2= 7.55$ df=1, P=0.005; §Hardy-Weinberg equilibrium = 0.09; ‡ Hardy-Weinberg equilibrium = 0.21

status in Cases and Controls. Fifty- seven percent (57%) of the Cases were carriers of the C-allele versus 49% for the Controls. *The calculated Odds ratio (OR= 1.35) and its 95% confidence interval (OR* = 1.08 – 1.67), suggested that the C allele was more likely (*P*=0.005) prevalent in the Cases than in the Controls.

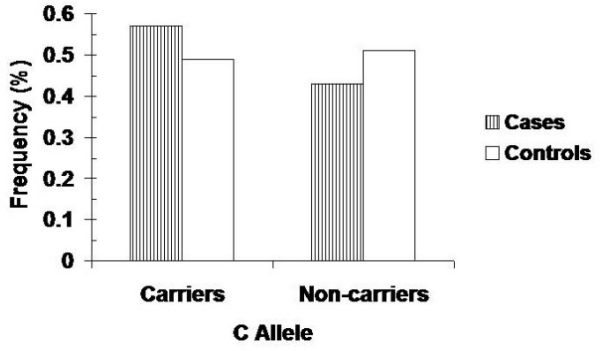

Figure 1. *Prevalence of genotype frequencies by C allele carrier status, X^2=4.0, df=1, P=0.03*

Degree of polymorphism

The calculated measures of the degree of polymorphism at the AQP-1 polymorphic site were: Cases; H=0.46; PIC=0.35 and Controls; H= 0.42; PIC= 0.33.

Discussion

The present study determined the allelic frequencies and genotype prevalence for a SNP (G → C) in the 3' untranslated region of the AQP1 gene and examined the extent of association of this polymorphism with endurance performance status in Hispanic marathon runners. Recent evidence associates the 3' untranslated region (3' UTR) of messenger RNA (mRNA) with the regulation of gene expression (17). The 3' UTR controls the nuclear export, subcellular targeting, and rates of translation and degradation of DNA (17). Genes controlled by the sequence of the 3' UTR are generally regulatory proteins and their irregular expression may have serious effects on humans (18).

The body mass of humans is approximately 70% water (19). Regulation of water flow across cell membranes is essential for maintaining an appropriate fluid balance within the cell (19), which in turn is a critical factor in health status as well as running performance. AQP1 channels allow the flux of water at a rate of approximately 10^9-10^{10} molecules per second per channel, with an activation energy as low as the one associated with the self-diffusion rate in bulk water (20-22). The water selectively of AQP1 is caused by the size exclusion effect (9, 22). The narrow or mid-region at the channel imposes a strict restriction so that there is only one water molecule at a time allowed to pass through (22). When coming close to the mid-region, the water molecule undergoes a rotation leading to the formation of hydrogen bonds between its oxygen and specific amino groups (23). This rotation and its effect is essential for protons exclusion (24). The monomeric pores of AQP1 have strong hydrogen bonds between the protein and water molecules inside the channel, this makes the monomers less gas permeable than the central pore (19). Recent evidence suggests that the hydrophobic central pore of AQP1 is permeable to both CO_2 and O_2 (19).

Endurance exercise has been defined as the ability to perform cardiovascular exercise, while endurance performance is the capacity of sustaining repeated muscle contraction over a long period of time (25). Such capacity is fulfilled by cellular (mitochondrial) utilization of oxygen, an adequate fuel provision, as well as an appropriate body fluid balance (25). An adequate body fluid balance may improve blood flow through the smallest vessels which will enhance oxygen delivery to working muscle mass (25). Inadequate body fluid balance has a detrimental effect on exercise capacity that would impair endurance performance (26).

In conclusion, these data support an association between a DNA sequence variant in the aquaporin-1 gene or in its surrounding region and marathon performance in Hispanic runners. The findings also suggest that inter-individual variability in marathon performance could be partly explained by molecular mechanisms, such as DNA sequence variations, in a thermoregulatory mediator. The allelic frequencies and genotype prevalence for the AQP1 C/G SNP reported here is novel information. As previously stated this study is the first attempt to examine the association between of the present AQP1 C/G SNP with endurance performance level in humans. Therefore, there are no basis for comparisons. The present data on Hispanics provide the means for further research and replication of findings. It should be of interest to corroborate if the C allele, indeed provide some advantage regarding marathon running performance.

Acknowledgements

Funding received from CRC-RCMI 1P20RR11126 and RCMI #G12RR03051.

References

1. Verkman AS. More than just water channels: Unexpected cellular roles of aquaporins. *J Cell Sci* 2005; 118: 3225-32.
2. Moon C, Preston GM, Griffin CA, et al. The human aquaporin-chip gene. *J Biol Chem* 1993; 268: 15772-8.
3. National Center for Biotechnology Information, U.S. National Library of Medicine [Internet]. Bethesda (MD): Entrez Single Nucleotide Polymorphism; [cited 2006 May 23]. Available from: http://preview.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=1049305.
4. Yool AJ, Weinstein AM. New roles for old holes: Ion channel function in aquaporin-1. *News Physiol Sci* 2002; 17: 68-72.

5. Saparov SM, Kozono D, Rothe U, et al. Water and ion permeation of aquaporin-1 in planar lipid bilayers. *J Biol Chem* 2001; 276: 31515-20.
6. Agre P, King LS, Masato Y, et al. Aquaporin water channels - from atomic structure to clinical medicine. *J Physiol* 2002; 542: 3-16.
7. Shi L-b, Skach WR, Verkman AS. Functional independence of monomeric chip28 water channels revealed by expression of wild-type mutant heterodimers. *J Biol Chem* 1994; 269: 10417-22.
8. Zhu F, Tajkhorshid E, Schulten K. Theory and simulation of water permeation in aquaporin-1. *Biophys J* 2004; 86: 50-7.
9. Ren G, Reddy VS, Cheng A, et al. Visualization of a water-selective pore by electron crystallography in vitreous ice. *Proc Natl Acad Sci USA* 2001; 98: 1398-403.
10. Au CG, Cooper ST, Lo HP, et al. Expression of aquaporin 1 in human cardiac and skeletal muscle. *J Mol Cell Cardiol* 2004; 36: 655-62.
11. Butler TL, Au CG, Yang B, et al. Cardiac aquaporin expression in humans, rats, and mice. *Am J Physiol Heart Circ Physiol* 2006; 291: H705-H13.
12. Frigeri A, Nicchia GP, Balena R, et al. Aquaporins in skeletal muscle: Reassessment of the functional role of aquaporin-4. *FASEB J* 2004; 18: 905-7.
13. Ye S, Dhillon S, Ke X, et al. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucl Acids Res* 2001; 29: 1-8.
14. Botstein D, White RL, Skolnick M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980; 32: 314–31.
15. Ott J. *Analysis of human genetic linkage*. 3rd Ed. Baltimore, MD:The John Hopkins University Press, 1991.
16. Harbage B, Dean AG. Distribution of epi info software: An evaluation using the internet. *Am J Prev Med* 1999; 16: 314-7.
17. Conne B, Stutz A, Vassalli J-D. The 3' untranslated region of messenger rna: A molecular 'hotspot' for pathology? *Nature Med* 2000; 6: 637-41
18. Grzybowska EA, Wilczynska A, Siedlecki JA. Regulatory functions of 3'utrs *Biochem Biophys Res Commun* 2001; 288: 291-5.
19. Wang Y, Tajkhorshid E. Molecular mechanisms of conduction and selectivity in aquaporin water channels. *J Nutr* 2007; 137: 1509S-15S.
20. de Groot BL, Frigato T, Helms V, et al. The mechanism of proton exclusion in the aquaporin-1 water channel. *J Mol Biol* 2003; 333: 279-93.
21. Heymann JB, Engel A. Aquaporins: Phylogeny, structure, and physiology of water channels. *News Physiol Sci* 1999; 14: 187-93.
22. Kong Y, Ma J. Dynamic mechanisms of the membrane water channel aquaporin-1 (aqp1). *Proc Natl Acad Sci USA* 2001; 98: 14345-9.
23. de Groot BL, Grubmuller H. Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and glpf. *Science* 2001; 294: 2353-7.
24. Murata K, Mitsuoka K, Hirai T, et al. Structural determinants of water permeation through aquaporin-1. *Nature* 2000; 407: 599-605.
25. Brooks G, Fahey T, Baldwin K. *Exercise physiology. Human bioenergetics and its applications* 4th. New York (NY): McGraw Hill, 2005.
26. von Duvillard SP, Braun WA, Markofski M, et al. Fluids and hydration in prolonged endurance performance. *Nutrition* 2004; 20: 651-6.

Received:
Accepted: December 15, 2009
Published:

Address for correspondence:
Miguel A. Rivera
Department of Physical Medicine, Rehabilitation & Sports Medicine
Main Building Office A204
School of Medicine, University of Puerto Rico
San Juan, P.R. 00936
787-751-9625 (Phone)
787-754-1471 (Fax)
email: miguel.rivera10@upr.edu; miguelrivera@prtc.net

Juan L. Martínez: jl_mtz12@hotmail.com
Aleshka Carrion: aleshkacarrion@gmail.com
Maria E Florian: mariaflorian@gmail.com
José A. Martín: jamurria@ceu.es
Juan R. López-Taylor: taylor@cucs.udg.mx
Thomas D. Fahey: discusdoc@aol.com