GENETIC POLYMORPHISMS RELATED TO VO_{2MAX} ADAPTATION ARE ASSOCIATED WITH ELITE RUGBY UNION STATUS AND COMPETITIVE MARATHON PERFORMANCE

Elliott C R Hall[1], Sandro S de Almeida[2], Shane M Heffernan[3], Sarah J Lockey[4], Adam J Herbert[5], Peter Callus[1], Stephen H Day[6], Charles R Pedlar[7,8], Courtney Kipps[8], Malcolm Collins[9], Yannis P Pitsiladis[10], Mark A Bennett[3], Liam P Kilduff[3], Georgina K Stebbings[1], Robert M Erskine[8,11] & Alun G Williams[1,8]

[1]*Sports Genomics Laboratory, Department of Sport and Exercise Sciences, Manchester Metropolitan University, Manchester, UK;*
[2]*Hospital Israelita Albert Einstein, São Paulo, Brazil;*
[3]*Applied Sports, Technology, Exercise and Medicine Research Centre (A-STEM), College of Engineering, Swansea University, Swansea, UK;*
[4]*Faculty of Health, Education, Medicine and Social Care, Anglia Ruskin University, Chelmsford, UK;*
[5]*School of Health Sciences, Birmingham City University, Birmingham, UK;*
[6]*School of Medicine and Clinical Practice, University of Wolverhampton, Wolverhampton, UK;*
[7]*Faculty of Sport, Health and Applied Science, St Mary's University, Twickenham, UK;*
[8]*Institute of Sport, Exercise and Health, University College London, London, UK;*
[9]*Division of Exercise Sciences and Sports Medicine, Department of Human Biology, University of Cape Town (UCT), Cape Town, South Africa;*
[10] *Collaborating Centre of Sports Medicine, University of Brighton, Eastbourne, UK, UK;*
[11]*School of Sport and Exercise Science, Liverpool John Moores University, Liverpool, UK;*

Corresponding author details:
Dr Elliott C R Hall, PhD
Department of Sport and Exercise Sciences,
Manchester Metropolitan University,
Manchester, M1 5GD, United Kingdom
Email: elliott.hall@mmu.ac.uk

Key words:
Sports; genomics; athletic; heritability; cardiorespiratory

Abstract

Purpose: Genetic polymorphisms have been associated with the adaptation to training in maximal oxygen uptake ($\dot{V}O_2$max[please standardize throughout the manuscript to show rate]). However, the genotype distribution of selected polymorphisms in athletic cohorts is unknown, with their influence on performance characteristics also undetermined. This study investigated whether the genotype distributions of three polymorphisms previously associated with $\dot{V}O_2$max training adaptation are associated with elite athlete status and performance characteristics in runners and rugby athletes, where for both sports whom aerobic metabolism is important.

Methods: Genomic DNA was collected from 732 men, including 165 long-distance runners, 212 elite rugby union athletes and 355 non-athletes. Genotype and allele frequencies of *PRDM1* rs10499043 C/T, *GRIN3A* rs1535628 G/A and *KCNH8* rs4973706 T/C were compared between athletes and non-athletes. Personal best marathon times in runners, as well as in-game performance variables and playing position of rugby athletes, were analysed according to genotype.

Results: Runners with *PRDM1* T-alleles recorded faster marathon times than CC homozygotes (02:27:55 ± 00:07:32 h *vs.* 02:31:03 ± 00:08:24 h, $p = 0.023$). Elite rugby athletes had 1.57 times greater odds of possessing the *KCNH8* TT genotype than non-athletes (65.5% *vs.* 54.7%, $\chi^2 = 6.494$, $p = 0.013$). No other associations were identified.

Conclusions: This study is the first to demonstrate that polymorphisms previously associated with $\dot{V}O_{2max}$ training adaptations in non-athletes are also associated with marathon performance (*PRDM1*) and elite rugby union status (*KCNH8*). These genotypes and alleles previously associated with superior endurance training adaptation appear to be advantageous for superior performance in long-distance running and achieving elite rugby union status.

Introduction

Exercise-related phenotypes are determined by the interaction of genetics and the environment [1]. For many phenotypes, individual differences remain when environmental factors are controlled, highlighting the important contribution of heritable factors [2]. The discovery of genes and common genetic variants that are associated with quantifiable phenotypes can, therefore, help to elucidate the mechanisms that contribute to such individual differences.

Cardiorespiratory fitness was positively associated with health outcomes and can be improved by regular aerobic activity [3]. The maximal rate of O_2 uptake ($\dot{V}O_{2max}$) describes the maximal amount of O_2 per unit of time that can be delivered to peripheral organs, such as skeletal muscle, and is the standard measurement of cardiorespiratory fitness [4]. Findings from the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study estimate that the heritability of $\dot{V}O_{2max}$ in the untrained state is approximately 50% [5]. A subsequent report involving the same cohort estimated the heritability of the adaptation of $\dot{V}O_{2max}$ following a 20-week endurance training program to be 47% [2]. These data suggest that not only are some individuals predisposed to superior cardiovascular fitness in the absence of exercise stimuli, but that the magnitude to which an individual can adapt to aerobic exercise training is also genetically influenced.

The benefit of cardiorespiratory fitness to athletic performance is reflected by superior $\dot{V}O_{2max}$ amongst athletes compared to non-athletes and may be explained, in part, by the deliberate exposure of athletes to prolonged exercise training [4]. It is also possible that individuals with genetic variants that predispose them to better training adaptation are more likely to reach the elite level, because they can improve their baseline cardiorespiratory fitness to a greater magnitude than those with less favourable genetics. The association of specific polymorphisms with athlete status, through the overrepresentation of a particular genotype

compared to the general population, supports the notion that genetic variation can enhance an individual's chances of becoming an elite athlete [6]. Indeed, there are specific genotypes that are more common amongst elite endurance athletes [7] and elite athletes from team sports such as rugby [8,9] and soccer [10] than the general population. Nonetheless, the polygenic nature of physiological traits means that the discovery of additional variants remains key to understanding the genetic contribution to athletic performance.

After determining the heritability of $\dot{V}O_{2max}$ training adaptations, Bouchard and colleagues performed a Genome Wide Association Study (GWAS) to identify genomic loci associated with the variance in training adaptation [11]. Twenty-one single nucleotide polymorphisms (SNPs) were individually associated with the magnitude of $\dot{V}O_{2max}$ improvement, and in combination explained 48.6% of the variance in adaptation between individuals. The three SNPs contributing the most to inter-individual differences in $\dot{V}O_{2max}$ training adaptation were PR/SET domain 1 (*PRDM1*) rs10499043 C/T (7.0%), glutamate ionotropic receptor NMDA type subunit 3A (*GRIN3A*) rs1535628 G/A (5.2%), and potassium voltage-gated channel subfamily H member 8 (*KCNH8*) rs4973706 T/C (4.5%). However, to our knowledge, no study has sought to replicate these associations or investigate whether the distribution of those genotypes associated with $\dot{V}O_{2max}$ training adaptation differs between the general population and groups where enhanced training adaptations may be advantageous, such as elite athletes. In addition, team sports such as rugby union include different playing positions with variable match demands [12,13] and differences in aerobic performance between these positions [14] . This suggests that some athletes may have an inherited benefit of an enhanced capacity for cardiorespiratory adaptation. Furthermore, the relationship between estimated $\dot{V}O_{2max}$ and the effects of fatigue on tackling technique in rugby league [15] suggests cardiorespiratory fitness could be an important contributor to match outcomes. However, it is not known whether in-game performance variables are associated with genetic variability.

Thus, the purpose of the present study was to determine whether three SNPs previously related to $\dot{V}O_{2max}$ training adaptation are associated with elite athlete status amongst long-distance runners and rugby union athletes, and whether genotypes of these SNPs are associated with long-distance running and elite rugby union performance. We hypothesised that the alleles associated with greater training adaptations of $\dot{V}O_{2max}$ (*PRDM1* T-allele, *GRIN3A* A-allele and *KCNH8* T-allele) would (i) be overrepresented in athletes compared to the general population, (ii) be associated with superior performance amongst long-distance runners and favourable in-game performance in rugby union athletes, and (iii) differ in frequency according to the playing position of elite rugby union athletes.

Materials and Methods

Participants

This study recruited a total of 732 Caucasian male participants including 212 rugby athletes), 165 long-distance runners and 355 healthy non-athletes. Rugby athletes all competed in rugby union and included 73.1% British, 14.2% South African and 10.4% Irish, with other nationalities each contributing 0.5%. All rugby athletes were considered elite having competed regularly (>5 matches) since 1995 in the highest professional league in the UK, Ireland or South Africa, and were recruited as part of the RugbyGene Project (described in detail by Heffernan and colleagues [16]). Of these athletes, 53.8% had competed at international level, with 99.1% of those international players representing a "High Performance Union" (Regulation 16, www.worldrugby.org). Long-distance runners were primarily recruited from the London Marathon Expo between 2012 and 2014, in addition to national/regional athletic clubs and organisations in the UK. Runners included 91.5% British and 1.2% Polish, with other nationalities each contributing 0.5%. Inclusion criteria for runners was a personal best (PB) marathon time of $\leq$ 3 hours verified using official online records (www.thepowerof10.info). Non-athlete participants were 355 healthy, unrelated recreationally active males recruited

through mail-outs, posters and word of mouth. Due to assay availability, *KCNH8* rs4973706 genotype data was only available for 362 participants, including 139 rugby athletes and 223 non-athletes. Participant characteristics are described in Table 1. Manchester Metropolitan University Ethics Committee approved the study and written informed consent was obtained from each participant.

Table 1. Characteristics of participants analysed for *PRDM1*, *GRIN3A* and sub-sample for *KCNH8*.

Group	***n***	**Age** (y)	**Height** (m)	**Mass** (kg)	**BMI** (kg/m^2)
Non-athletes					
PRDM1 & GRIN3A	355	27 (15)	1.79 (0.07)	78.0 (11.4)*	24.5 (3.5)*
KCNH8	223	23 (7)	1.79 (0.06)	77.6 (11.8)	24.2 (3.3)
Rugby Union					
PRDM1 & GRIN3A	221	28 (7)	1.86 (0.07)**	102.4 (11.4)**	29.7 (3.1)**
KCNH8	147	26 (5)	1.86 (0.07)***	102.8 (12.3)***	29.7 (3.0)***
Runners					
PRDM1 & *GRIN3A*	165	36 (9)	1.76 (0.06)	66.9 (6.8)	21.0 (2.0)

Data are mean (standard deviation)
* greater than runners ($p < 0.0001$)
** greater than non-athletes and runners ($p < 0.0001$)
*** greater than non-athletes ($p < 0.0001$)

Procedures

Sample collection

Blood (~68% of samples), buccal swab (~23%) or saliva (~9%) samples were obtained via the following protocols. Blood was drawn from a superficial forearm vein into an EDTA tube and stored in sterile tubes at -20°C until processed. Saliva samples were collected into Oragene DNA OG-500 collection tubes (DNA Genotek, Ottawa, Ontario, Canada) according to the manufacturer's protocol and stored at room temperature until processed. Sterile buccal swabs (Omni swab; Whatman, Springfield, Mill, UK) were rubbed against the buccal mucosa of the cheek for ~30 s. Tips were ejected into sterile tubes and stored at -20°C until processed.

DNA isolation

DNA isolation was performed using a QIAamp DNA Blood Mini kit and standard spin column protocol according to the manufacturer's instructions (Qiagen, West Sussex, UK). Briefly, 200 µL of whole blood/saliva, or one buccal swab, was lysed and incubated, the DNA washed, and the eluate containing isolated DNA stored at 4°C.

Genotyping

Samples were genotyped for the *PRDM1* (rs10499043 C/T), *GRIN3A* (rs1535628 G/A) and *KCNH8* (rs4973706 T/C) SNPs by combining 5 µL Genotyping Master Mix (Applied Biosystems, Paisley, UK), 4.3 µL H_2O, 0.5 µL assay mix (Applied Biosystems), and 0.2 µL of purified DNA (~9 ng), for samples derived from blood and saliva. For DNA derived from buccal swabs, 5 µL Genotyping Master Mix was combined with 3.5 µL H_2O, 0.5 µL assay mix, and 1 µL DNA solution (~9 ng DNA). Either a Chromo4 (Bio-Rad, Hertfordshire, UK) or a StepOnePlus real-time system (Applied Biosystems) was used. Briefly, denaturation began at 95°C for 10 min, with 40 cycles of incubation at 92°C for 15 s before annealing and extension at 60°C for 1 min. Initial genotyping analysis was performed with Opticon Monitor software version 3.1 (Bio-Rad) or StepOnePlus software version 2.3 (Applied Biosystems). All samples were analysed in duplicate and were in 100% agreement.

Rugby union positional groups

To further assess genotype and allele frequencies in rugby union, athletes were allocated to subgroups: forwards (props, hookers, locks, flankers, number eights) and backs (scrum halves, fly halves, centres, wingers, full backs). Due to diverse physiological demands within rugby union, athletes were further divided into positional groups based on similarities in their movement patterns [12] as front five (props, hookers, locks), back row (flankers, number eights),

half backs (scrum halves, fly halves), centres, and back three (wings and full backs). The rugby athletes' playing positions are shown in Table 2.

Table 2. Distribution of rugby athletes according to playing position. Data are number of athletes (% of all athletes)

	Analysis of *PRDM1* and *GRIN3A*	Analysis of *KCNH8*
	n = 212	*n* = 139
Forwards *vs.* Backs		
Forwards	114 (53.8)	79 (56.8)
Backs	98 (46.2)	60 (43.2)
Positional sub-groups		
Front Five	66 (31.1)	47 (33.8)
Back Row	50 (23.6)	34 (25.5)
Half Backs	41 (19.3)	22 (15.8)
Centres	27 (12.7)	18 (12.9)
Back Three	28 (13.2)	18 (12.9)

Rugby union in-game performance variables

In-game performance data for 112 of the 212 rugby athletes was obtained from Opta Sports (London, UK) for all matches during eight seasons (2012-13 to 2019-20) of rugby union competition in the highest professional competitive leagues in England (Premiership) and Wales/Ireland/Scotland/Italy/South Africa (Celtic/PRO12/PRO14). Athletes were included for analysis where performance data was available for a minimum of 320 competitive minutes, with an equivalent of 39.9 ± 27.0 80-min matches per player. The analysed variables were: number of carries per 80 min; metres gained in possession per 80 min; number of penalties conceded per 80 min; number of successful tackles per 80 min; percentage of successful tackles during all matches.

Data analysis

Statistical analyses were conducted using SPSS for Windows version 25.0 (IBM Statistics, Chicago, Illinois). Genotype distributions and allele frequencies of athletes and non-athletes were compared by χ^2 goodness-of-fit test. Genotype distribution was analysed using additive (AA *vs.* Aa *vs.* aa) and recessive (AA *vs.* Aa+aa) models due to low minor allele frequencies. Odds ratios (OR) were calculated where genotype distribution differed between groups. Genotype distribution and allele frequencies according to rugby playing position were compared using the χ^2 test of independence. The relationships between long-distance runners' PB marathon time and *PRDM1* and *GRIN3A* genotypes were analysed in a recessive model only (due to low minor allele frequency) by independent samples t-test. The relationship between rugby union forwards and backs with performance variables was analysed by independent samples t-test. The association between *PRDM1* (n = 112), *GRIN3A* (n = 112) and *KCNH8* (n = 95) genotype and rugby union in-game performance variables were analysed in a recessive model only (due to low minor allele frequency) by one-way ANCOVA, with first rugby union subgroups (forwards and backs), then positional groups (front five, back row, half backs, centres and back three), as covariates. P values < 0.05 were considered statistically significant. All data are presented as mean (standard deviation).

Results

Genotype distribution

Genotype distributions across all groups for each SNP are described in Table 3. Although not statistically significant (p = 0.054, OR = 1.44), 26.7% of long-distance runners carried the *PRDM1* T-allele (CT/TT) compared to 20.6% of non-athletes. *KCNH8* TT genotype was overrepresented in rugby athletes compared to non-athletes (65.5% *vs.* 54.7%, χ^2 = 6.494, p = 0.013, OR = 1.57, Fig. 1). There were no other differences in genotype frequencies between groups ($p \geq 0.148$).

Table 3. Genotype distribution for *PRDM1* and *GRIN3A* according to athlete group. Data are number of individuals (%)

		Group				
SNP	**Genotype**	**Non-athlete**	**Rugby**	**Runner**	**Total**	**MAF**
		n = 355	*n* = 212	*n* = 165	*n* = 732	*n* = 732
PRDM1	CC	282 (79.4)	163 (76.9)	121 (73.3)	566 (76.6)	0.12
rs10499043	CT/TT	73 (20.6)	49 (23.1)	44 (26.7)	166 (23.4)	
GRIN3A	GG	293 (82.5)	175 (82.6)	136 (82.4)	604 (82.5)	0.09
rs1535628	GA/AA	62 (17.5)	37 (17.4)	29 (17.6)	128 (17.5)	

MAF, minor allele frequency

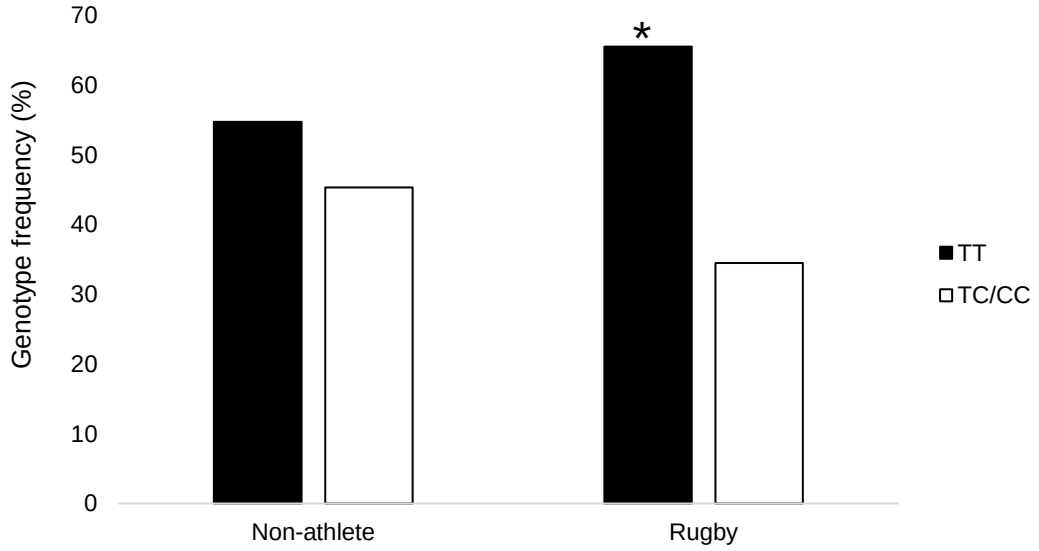

Fig 1. *KCNH8* rs4973706 genotype distribution in non-athletes and rugby athletes. * greater than non-athletes (p = 0.013)

Runner PB marathon times

Runners with the *PRDM1* T-allele (CT/TT) had faster PB marathon times than those with the CC genotype (02:27:55 ± 00:07:32 h *vs.* 02:31:03 ± 00:08:24 h, p = 0.023; Fig. 2). There was no association between PB marathon time and *GRIN3A* genotype.

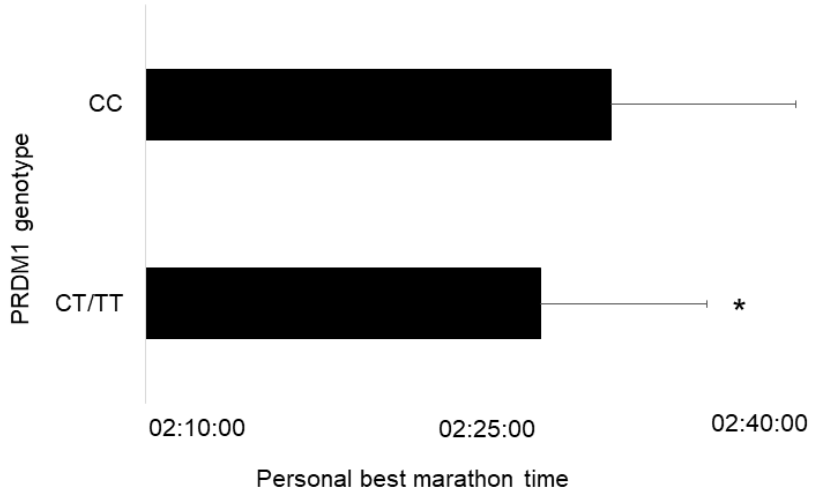

Fig 2. Runners' PB marathon time according to *PRDM1* rs10499043 genotype. * lower time than CC (p = 0.023).

Rugby union positional groups

No differences in genotype distribution were observed between forwards and backs for *PRDM1* (T-allele carriers 22.0% *vs.* 24.5 respectively, p = 0.744), *GRIN3A* (A-allele carriers 18.4% *vs.* 16.3%, p = 0.720) and *KCNH8* (C-allele carriers 29.1% *vs.* 41.7%, p = 0.150). Similarly, no differences in *PRDM1*, *GRIN3A* or *KCNH8* genotype distribution were observed according to rugby athletes' playing position ($p \geq 0.228$).

Rugby union in-game performance variables

There was no association of genotype with in-game performance variables adjusted for playing position ($p \geq 0.131$). Regardless of genotype, backs carried the ball forward for a greater distance per 80 min than forwards (32.2 ± 15.0 m *vs* 12.5 ± 11.0 m, $p < 0.0005$). Compared to backs, forwards completed more successful tackles per 80 min (9.8 ± 2.1 *vs* 5.9 ± 2.3, $p < 0.0005$), had a higher percentage of successful tackles (89.7 ± 4.14 *vs* 82.9 ± 6.7, $p < 0.0005$) and conceded more penalties per 80 min (1.0 ± 0.5 *vs* 0.4 ± 0.2, $p < 0.0005$). Performance data not presented.

Discussion

The aim of this study was to determine whether three SNPs previously linked to $\dot{V}O_{2max}$ training adaptations were associated with athlete status and performance characteristics in elite rugby athletes and long-distance runners. The main findings were that in runners, the *PRDM1* T-allele was associated with faster marathon running times and tended to be overrepresented compared to non-athletes, and that elite rugby athletes had 1.57 greater odds of possessing the *KCNH8* TT genotype than non-athletes. These findings confirm our primary hypothesis, that the alleles and genotypes associated with athlete status and athletic performance in the present study are in concordance with those previously associated with greater $\dot{V}O_{2max}$ improvement [11]. In contrast, the *GRIN3A* SNP was not associated with any of the variables investigated in this study, suggesting it does not affect elite status or performance in runners or rugby athletes, whilst there was no relationship between any SNP and rugby union in-game performance variables.

Runners with the *PRDM1* CT/TT genotype recorded ~3 min (2.1%) faster personal best marathon times than CC homozygotes, suggesting that carrying at least one *PRDM1* T-allele is favourable to endurance running performance. The T-allele also tended to be more common amongst runners than non-athletes. The rs10499043 SNP is a C>T substitution located 287 kb from *PRDM1*, previously known as *BLIMP1*, which encodes a protein that represses β-interferon gene expression and may be involved in skeletal muscle fiber differentiation [17], although that has not been shown in human tissue. *PRDM1* may also be a target of epigenetic downregulation [18], though a functional link to cardiorespiratory fitness is unknown. Whilst $\dot{V}O_{2max}$ was not measured in this study, our finding that runners with the *PRDM1* T-allele recorded faster personal best marathon times than CC homozygotes (02:27:55 h *vs* 02:31:03 h) suggests that a genetic predisposition to achieve greater training-induced improvements in

$\dot{V}O_{2max}$ might contribute to superior running performance. Nevertheless, 73.3% of runners in this study, all of whom recorded marathon times below 03:00:00 h, did not carry the *PRDM1* T-allele. These data reaffirm the notion that while high-level marathon performance is dependent on several factors including major and obvious environmental ones like training volume, some of the variation in marathon performance at high levels of the sport could be genetically influenced [19,20]. When considered alongside the superior VO_{2max} improvements in TT homozygotes and the proportion of $\dot{V}O_{2max}$ improvement attributed to this SNP [11], our findings suggest further investigation of this SNP in human endurance performance is warranted. If replicated in independent populations, *in vitro* studies should seek to determine a functional link between *PRDM1* rs10499043 and relevant biology including aspects of muscle differentiation.

In a sub-sample of the study cohort, the *KCNH8* TT genotype was overrepresented in elite rugby athletes compared to non-athletes. The rs4973706 SNP is a T>C substitution located 268kb from *KCNH8,* which is principally expressed within the human nervous system [21]. *KCNH8* includes a potassium voltage-gated channel and is a member of the human Elk K^+ channel gene family [22], which has diverse functions including regulating heart rate, insulin secretion, neurotransmitter release and epithelial electrolyte transport [22]. Due to association of the TT genotype with greater $\dot{V}O_{2max}$ adaptation [11], and the importance of aerobic fitness to repeated-effort performance of rugby league athletes [23], we hypothesised an overrepresentation of the TT genotype in rugby union athletes compared to non-athletes. Indeed, cardiorespiratory fitness contributes to elite rugby union performance [13,24] and endurance training is fundamental to elite clubs' athlete preparation [25]. Consequently, heritable factors predisposing a greater magnitude of $\dot{V}O_{2max}$ improvement during training are likely to contribute to athletes' ability to reach the highest level of competition in rugby union. The association described in the present study is the first association of this SNP with elite rugby status, and whilst rugby

athletes had 1.57 times greater odds of have the TT genotype than non-athletes, ~36% of rugby athletes in this study lack the TT genotype, demonstrating that other factors including other genetic variants [8,9] contribute to elite rugby status. No studies have investigated the rs4973706 variant since the association with $\dot{V}O_{2max}$ improvement [11], so as far as we are aware, the functional role of this SNP is unknown. However, the exercise-induced rise in ATP-sensitive potassium channel expression, which promotes reduced cardiac energy consumption under escalating workloads as an adaptive response to exercise [26] suggests genetic variations in *KCNH8*, a gene related to potassium channel pathways, might influence the inter-individual capacity for cardiorespiratory adaptation. The findings described here and the previous association with $\dot{V}O_{2max}$ adaptation demonstrate the need for replication in larger athletic and non-athletic cohorts and for mechanistic studies of the rs4973706 variant in relation to cardiac function and $\dot{V}O_{2max}$.

The *GRIN3A* rs1535628 variant was not associated with athlete status, running performance, rugby playing position or rugby performance variables. That SNP lies 516 kb upstream of *GRIN3A*, which is widely expressed in neural cells and is involved in the development of synaptic elements [27]. Other *GRIN3A* polymorphisms are associated with conditions such as Kawasaki disease [28] and schizophrenia [29], yet the functional consequence of the rs1535628 SNP remains undescribed. Less than 1% of participants in the present study had the AA genotype previously associated with superior $\dot{V}O_{2max}$ adaptation [11], with a low minor allele frequency across all groups potentially limiting the power to detect associations. Furthermore, the present study investigated runners and rugby athletes, and it is possible that the *GRIN3A* rs1535628 SNP is only associated with cardiorespiratory fitness improvement of non-athletes when they first begin training, as investigated in HERITAGE and the subsequent GWAS [11]. While further studies are warranted to replicate the original association, the present

study suggests this SNP is unlikely to influence athlete status and performance in runners or rugby athletes.

No SNP was associated with rugby union playing position or in-game performance variables in the present study. We hypothesised that differences would exist because rugby athletes exhibit different movement patterns according to their playing position [12] and because of reported differences in aerobic field test performance between playing positions [14]. Previous associations of *ACTN3* [8] and *FTO* [30] genotypes with playing position in similar populations, where the functional consequences of both SNPs are better understood, permits logical speculation regarding each association. However, lack of association of *ACE* and *COL5A1* SNPs with playing position [8,9] demonstrates that although some SNPs may be advantageous to certain positions, others may be more broadly associated with superior athletic ability in rugby players. Indeed, no genotypes were associated with rugby union performance variables in the present study, indicating that some SNPs are advantageous to general rugby union ability, but do not contribute to the number of key actions performed by individual athletes. In light of our finding that several in-game performance variables differ between forwards and backs (as expected), future studies should seek to determine whether other SNPs - including those previously associated with playing position in rugby [8,30] - are associated with these or other performance variables relevant to that particular SNP. Despite the association of *KCNH8* with elite rugby status, the genotypes recorded in this study do not appear to differ between playing positions or relate to the frequency or success of specific playing actions.

In addition to novel findings, the present study also has limitations. Firstly, assessing $\dot{V}O_{2max}$ directly may have helped to determine whether the associations discovered in this study are linked to cardiorespiratory fitness, although from a practical perspective that is virtually impossible in large cohorts of high-level athletes. Secondly, only male Caucasians were investigated to control for the effects of sex and geographic ancestry. Accordingly, these

findings should be replicated in women and participants with different ancestry. The present study included athletes from long-distance running and rugby union, meaning the influence of these SNPs in other sports remains unknown. Finally, the lack of *KCNH8* genotype data for all participants, particularly in runners, highlights the need for further investigation of this SNP in relation to athletic status and performance.

The present study is the first to demonstrate associations of the *PRDM1* rs10499043 SNP with marathon running performance and the *KCNH8* rs4973706 SNP with elite athlete status in rugby union. The alleles associated with superior performance and elite athlete status in the present study are the same as those previously associated with greater $\dot{V}O_{2max}$ adaptation. This suggests that at least some SNPs, and thus physiological mechanisms that modulate the extent of training adaptations, are common to both untrained individuals and trained athletes.

Acknowledgements:

The authors thank all athletes and non-athletes for their time and willingness to participate. We also thank Hannah Dines for assistance during data collection and analysis.

References

1. Puthucheary Z, Skipworth JR, Rawal J, et al. Genetic influences in sport and physical performance. *Sports Med.* 2011;41(10):845-859.
2. 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 (1985).* 1999;87(3):1003-1008.
3. Erikssen G, Liestøl K, Bjørnholt J, et al. Changes in physical fitness and changes in mortality. *Lancet.* 1998;352(9130):759-762.

4. Jones AM, Carter H. The effect of endurance training on parameters of aerobic fitness. *Sports Med.* 2000;29(6):373-386.
5. Bouchard C, Daw EW, Rice T, et al. Familial resemblance for VO2max in the sedentary state: the HERITAGE family study. *Med Sci Sports Exerc.* 1998;30(2):252-258.
6. Ahmetov, II, Fedotovskaya ON. Current Progress in Sports Genomics. *Adv Clin Chem.* 2015;70:247-314.
7. Ahmetov I, Kulemin N, Popov D, et al. Genome-wide association study identifies three novel genetic markers associated with elite endurance performance. *Biol Sport.* 2015;32(1):3-9.
8. Heffernan SM, Kilduff LP, Erskine RM, et al. Association of ACTN3 R577X but not ACE I/D gene variants with elite rugby union player status and playing position. *Physiol Genomics.* 2016;48(3):196-201.
9. Heffernan SM, Kilduff LP, Erskine RM, et al. COL5A1 gene variants previously associated with reduced soft tissue injury risk are associated with elite athlete status in rugby. *BMC Genomics.* 2017a;18(Suppl 8):820.
10. Eynon N, Ruiz JR, Yvert T, et al. The C allele in NOS3 -786 T/C polymorphism is associated with elite soccer player's status. *Int J Sports Med.* 2012;33(7):521-524.
11. Bouchard C, Sarzynski MA, Rice TK, et al. Genomic predictors of the maximal O_2 uptake response to standardized exercise training programs. *J Appl Physiol (1985).* 2011;110(5):1160-1170.
12. Cahill N, Lamb K, Worsfold P, Headey R, Murray S. The movement characteristics of English Premiership rugby union players. *J Sports Sci.* 2013;31(3):229-237.
13. Brazier J, Antrobus M, Stebbings GK, et al. Anthropometric and physiological characteristics of elite male rugby athletes. *J Strength Cond Res.* 2020;34(6):1790-1801.

14. Quarrie KL, Handcock P, Toomey MJ, Waller AE. The New Zealand rugby injury and performance project. IV. Anthropometric and physical performance comparisons between positional categories of senior A rugby players. *Br J Sports Med.* 1996;30(1):53-56.
15. Gabbett TJ. Influence of fatigue on tackling technique in rugby league players. *J Strength Cond Res.* 2008;22(2):625-632.
16. Heffernan SM, Kilduff LP, Day SH, Pitsiladis YP, Williams AG. Genomics in rugby union: A review and future prospects. *Eur J Sport Sci.* 2015;15(6):460-468.
17. Beermann ML, Ardelt M, Girgenrath M, Miller JB. Prdm1 (Blimp-1) and the expression of fast and slow myosin heavy chain isoforms during avian myogenesis in vitro. *PLoS One.* 2010;5(4):e9951.
18. Nie K, Zhang T, Allawi H, et al. Epigenetic down-regulation of the tumor suppressor gene PRDM1/Blimp-1 in diffuse large B cell lymphomas: a potential role of the microRNA let-7. *Am J Pathol.* 2010;177(3):1470-1479.
19. Rivera MA, Fahey TD, López-Taylor JR, Martínez JL. The Association of Aquaporin-1 Gene with Marathon Running Performance Level: a Confirmatory Study Conducted in Male Hispanic Marathon Runners. *Sports Med Open.* 2020;6(1):16.
20. Stebbings GK, Williams AG, Herbert AJ, et al. TTN genotype is associated with fascicle length and marathon running performance. *Scand J Med Sci Sports.* 2018;28(2):400-406.
21. Zou A, Lin Z, Humble M, et al. Distribution and functional properties of human KCNH8 (Elk1) potassium channels. *Am J Physiol Cell Physiol.* 2003;285(6)C1356-1366.

22. Ellinghaus E, Ellinghaus D, Krusche P, et al. Genome-wide association analysis for chronic venous disease identifies EFEMP1 and KCNH8 as susceptibility loci. *Sci Rep.* 2017;7:45652.
23. Gabbett TJ, Stein JG, Kemp JG, Lorenzen C. Relationship between tests of physical qualities and physical match performance in elite rugby league players. *J Strength Cond Res.* 2013;27(6):1539-1545.
24. Smart DJ, Hopkins WG, Gill ND. Differences and changes in the physical characteristics of professional and amateur rugby union players. *J Strength Cond Res.* 2013;27(11):3033-3044.
25. Jones TW, Smith A, Macnaughton LS, French DN. Strength and Conditioning and Concurrent Training Practices in Elite Rugby Union. *J Strength Cond Res.* 2016;30(12):3354-3366.
26. Zingman LV, Zhu Z, Sierra A, et al. Exercise-induced expression of cardiac ATP-sensitive potassium channels promotes action potential shortening and energy conservation. *J Mol Cell Cardio.* 2011;51(1):72-81.
27. Das S, Sasaki YF, Rothe T, et al. Increased NMDA current and spine density in mice lacking the NMDA receptor subunit NR3A. *Nature.* 1998;393(6683):377-381.
28. Lin YJ, Chang JS, Liu X, et al. Association between GRIN3A gene polymorphism in Kawasaki disease and coronary artery aneurysms in Taiwanese children. *PLoS One.* 2013;8(11):e81384.
29. Takata A, Iwayama Y, Fukuo Y, et al. A population-specific uncommon variant in GRIN3A associated with schizophrenia. *Biol Psychiatry.* 2013;73(6):532-539.
30. Heffernan SM, Stebbings GK, Kilduff LP, et al. Fat mass and obesity associated (FTO) gene influences skeletal muscle phenotypes in non-resistance trained males and elite rugby playing position. *BMC Genet.* 2017b;18(1):4.