GENETIC POLYMORPHISMS RELATED TO VO₂MAX ADAPTATION ARE ASSOCIATED WITH ELITE RUGBY UNION STATUS AND COMPETITIVE MARATHON PERFORMANCE

Elliott C R Hall¹, Sandro S Almeida², Shane M Heffernan³, Sarah J Lockey⁴, Adam J Herbert⁵, Peter Callus¹, Stephen H Day⁶, Charles R Pedlar⁷,⁸, Courtney Kipps⁸, Malcolm Collins⁹, Yannis P Pitsiladis¹⁰, Mark A Bennett³, Liam P Kilduff³, Georgina K Stebbings¹, Robert M Erskine⁸,¹¹ & Alun G Williams¹,⁸

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

Preferred running head: VO₂max-related polymorphisms in rugby and runners

Abstract word count: 250 words

Text-only word count: 3,458 words

Number of figures and tables: 2 figures and 3 tables
Abstract

**Purpose:** Genetic polymorphisms have been associated with the adaptation to training in maximal oxygen uptake ($\dot{V}O_2\text{max}$). 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 $\dot{V}O_2\text{max}$ training adaptation are associated with elite athlete status and performance characteristics in runners and rugby athletes, competitors for 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 marathon times ~3 min faster than CC homozygotes (02:27:55 ± 00:07:32 h vs. 02:31:03 ± 00:08:24 h, $p = 0.023$). 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 $\dot{V}O_2\text{max}$ training adaptations in non-athletes are also associated with marathon performance (*PRDM1*) and elite rugby union status (*KCNH8*). The genotypes and alleles previously associated with superior endurance training adaptation appear to be advantageous in long-distance running and achieving elite status in rugby union.

**Key words:** genomics; exercise; heritability; endurance; polymorphism
Introduction

Exercise-related phenotypes are determined by the interaction of genetics and the environment. For many phenotypes, individual differences remain when environmental factors are controlled, highlighting the important contribution of heritable factors. 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 is positively associated with health outcomes and can be improved by regular aerobic activity. The maximal rate of \( \dot{V}O_2 \) uptake (\( \dot{V}O_2\max \)) 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. Findings from the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study estimate that the heritability of \( \dot{V}O_2\max \) in the untrained state is approximately 50%. A subsequent report involving the same cohort estimated the heritability of the adaptation of \( \dot{V}O_2\max \) following a 20-week endurance training program to be 47%. 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 \( \dot{V}O_2\max \) amongst athletes compared to non-athletes and may be explained, in part, by the deliberate exposure of athletes to prolonged exercise training. 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. Indeed, there are specific genotypes that are more common amongst elite endurance athletes and elite athletes from team sports such as rugby and soccer 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. Once new associations are discovered between specific variants and traits of interest, it is important that researchers can independently replicate those findings. Reproducible data reduces the risk of false positive results based on single studies, and subsequently helps identify genes and/or variants for further mechanistic investigation regarding their functional consequences on physiological processes.

After determining the heritability of $\dot{V}O_2$ max training adaptations, Bouchard and colleagues performed a Genome Wide Association Study (GWAS) to identify genomic loci associated with the variance in training adaptation. Twenty-one single nucleotide polymorphisms (SNPs) were individually associated with the magnitude of $\dot{V}O_2$ max 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 $\dot{V}O_2$ max 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 $\dot{V}O_2$ max 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 and differences in aerobic performance between these positions. Some athletes may, therefore, have an inherited benefit of an enhanced capacity for cardiorespiratory adaptation. Furthermore, the relationship between estimated $\dot{V}O_2$ max and
the effects of fatigue on tackling technique in rugby league\textsuperscript{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 $\dot{V}O_{2\text{max}}$ 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 $\dot{V}O_{2\text{max}}$ ($\text{PRDM1} \ T$ allele, $\text{GRIN3A} \ A$ allele and $\text{KCNH8} \ T$ 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.

**Methods**

**Subjects**

This study recruited 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\textsuperscript{16}). Of these athletes, 53.8\% had competed at international level, with 99.1\% of those 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%. The inclusion criterion 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. This study was conducted in accordance with the Declaration of Helsinki and all participants gave written informed consent. Ethical approval was granted by Manchester Metropolitan University.

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 \(-20^\circ C\) 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 \(-20^\circ C\) until processed.

DNA isolation

DNA isolation was performed using a QIAamp DNA Blood Mini kit and standard spin column protocol according to manufacturer 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 H2O, 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 H2O, 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 patterns12 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.

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 were available for a minimum of 320 competitive minutes, equivalent to 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.

Statistical analysis

Statistical analyses were conducted using SPSS for Windows version 25.0 (IBM Statistics, Chicago, Illinois). Genotype distributions and allele frequencies of athletes versus non-athletes, athlete sub-groups, and of athlete sub-groups versus non-athletes, were compared by χ² 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 χ² test of independence. The associations of PRDM1 and GRIN3A genotypes with long-distance runners’ PB marathon time were analysed in a recessive model only (due to low minor allele frequency) by independent samples t-test. Performance variables were compared between rugby union forwards and backs 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, after Bonferroni adjustment for multiple comparisons. All data are presented as mean ± standard deviation.

**Results**

*Hardy Weinberg Equilibrium (HWE) and Genotype distribution*

Genotype distributions across all groups for each SNP are described in Table 3. All were in HWE ($\chi^2 \leq 0.773, p \geq 0.379$). Although not statistically significant ($p = 0.056$, OR = 1.27 (95% confidence intervals (CI) 0.89-1.79)), 24.7% of all athletes carried the *PRDM1* T allele (CT/TT) compared to 20.6% of non-athletes. Similarly, although not statistically significant ($p = 0.054$, OR = 1.44 (95% CI 0.91-2.16), 26.7% of runners carried one or more T allele compared to 20.6% of non-athletes. *KCNH8* TT genotype was overrepresented in rugby athletes compared to non-athletes (65.5% vs. 54.7%, $\chi^2 = 6.494, p = 0.013$, OR = 1.57 (95% CI 1.01-2.43), Fig. 1). There were no other differences in genotype frequencies between groups ($p \geq 0.148$).

*Runner PB marathon times*

Runners with the *PRDM1* T allele (CT/TT) had ~3 min faster PB marathon times than those with the CC genotype (02:27:55 ± 00:07:31 h vs. 02:31:03 ± 00:08:24 h, $p = 0.023$; Fig. 2). There were four T allele carriers (CT/TT) amongst the 10 fastest runners, with a T allele frequency of 0.25 in those 10 compared to 0.13 in the remaining 155 runners. However, there was no overall association between PB marathon time and *GRIN3A* genotype.

*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 \textit{KCNH8} (C allele carriers 29.1% vs. 41.7%, $p = 0.150$). Similarly, no differences in \textit{PRDM1}, \textit{GRIN3A} or \textit{KCNH8} genotype distribution were observed according to rugby athletes’ playing position ($p \geq 0.228$).

\textit{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 \(\pm\) 15.0 m vs 12.5 \(\pm\) 11.0 m, $p < 0.0005$). Compared to backs, forwards completed more successful tackles per 80 min (9.8 \(\pm\) 2.1 vs 5.9 \(\pm\) 2.3, $p < 0.0005$), had a higher percentage of successful tackles (89.7 \(\pm\) 4.14 vs 82.9 \(\pm\) 6.7, $p < 0.0005$) and conceded more penalties per 80 min (1.0 \(\pm\) 0.5 vs 0.4 \(\pm\) 0.2, $p < 0.0005$). Performance data are not presented.

\textbf{Discussion}

The aim of this study was to determine whether three SNPs previously linked to \(\dot{V}O_2\text{max}\) 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 \textit{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 times greater odds of possessing the \textit{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 align with those previously associated with greater \(\dot{V}O_2\text{max}\) improvement.\textsuperscript{11} In contrast, the \textit{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, although that has not been shown in human tissue. PRDM1 may also be a target of epigenetic downregulation, though a functional link to cardiorespiratory fitness is unknown. Whilst \( \text{V}O_{2}\text{max} \) 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 \( \text{V}O_{2}\text{max} \) might contribute to superior running performance. Indeed, the PRDM1 T allele and CT/TT genotype were more frequent in the 10 fastest runners than in the remaining runners (25.0% vs 13.2%, and 40.0% vs 25.8%, respectively). 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. When considered alongside the superior \( \text{V}O_{2}\text{max} \) improvements in TT homozygotes and the proportion of \( \text{V}O_{2}\text{max} \) improvement attributed to this SNP, our findings suggest further investigation of this SNP in human endurance performance is warranted. If replicated in independent populations, \textit{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 VO2max 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 VO2max improvement during training likely 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 possessing the TT genotype than non-athletes, ~36% of rugby athletes in this study lack the TT genotype, demonstrating that other factors including other genetic variants8,9 contribute to elite rugby status. No studies have investigated the rs4973706 variant since the association with VO2max 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 exercise26 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 VO_{2\text{max}} adaptation demonstrate the need for replication in larger athletic and non-athletic cohorts and for mechanistic studies of the rs4973706 variant vis-à-vis cardiac function and VO_{2\text{max}}.

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 involved in the development of synaptic elements.\textsuperscript{27} Other GRIN3A polymorphisms are associated with conditions such as Kawasaki disease\textsuperscript{28} and schizophrenia,\textsuperscript{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 VO_{2\text{max}} adaptation,\textsuperscript{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.\textsuperscript{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\textsuperscript{12} and because of reported differences in aerobic field test performance between playing positions.\textsuperscript{14} Previous associations of ACTN3\textsuperscript{8} and FTO\textsuperscript{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\textsuperscript{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 or success 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\textsuperscript{8,30} - are associated with these or other performance variables relevant to that particular SNP. Despite the association of \textit{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.

**Practical Applications**

This study presents novel associations of the \textit{PRDM1} rs10499043 SNP with marathon performance and the \textit{KCNH8} rs4973706 SNP with the attainment of elite rugby union status, adding to a growing body of evidence surrounding the heritability of athletic traits and identifying polymorphisms that merit further examination. It is also important to note the limitations of this investigation. Firstly, assessing $\dot{V}O_{2\text{max}}$ 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, and the lack of \textit{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. Most importantly, our results should be replicated in independent cohorts and different contexts before the investigated SNPs should be used in commercial genetic testing.

Conclusions
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 VO2max 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. Further investigation is required to confirm whether the magnitude of cardiorespiratory training adaptation that occurs in elite runners and rugby athletes is genotype dependent.

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


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

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Age (y)</th>
<th>Height (m)</th>
<th>Mass (kg)</th>
<th>BMI (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-athletes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>PRDM1 &amp; GRIN3A</em></td>
<td>355</td>
<td>27 (15)</td>
<td>1.79 (0.07)</td>
<td>78.0 (11.4)*</td>
<td>24.5 (3.5)*</td>
</tr>
<tr>
<td><em>KCNH8</em></td>
<td>223</td>
<td>23 (7)</td>
<td>1.79 (0.06)</td>
<td>77.6 (11.8)</td>
<td>24.2 (3.3)</td>
</tr>
<tr>
<td>Rugby Union</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>PRDM1 &amp; GRIN3A</em></td>
<td>212</td>
<td>28 (7)</td>
<td>1.86 (0.07)**</td>
<td>102.4 (11.4)**</td>
<td>29.7 (3.1)**</td>
</tr>
<tr>
<td><em>KCNH8</em></td>
<td>139</td>
<td>26 (5)</td>
<td>1.86 (0.07)***</td>
<td>102.8 (12.3)***</td>
<td>29.7 (3.0)***</td>
</tr>
<tr>
<td>Runners</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>PRDM1 &amp; GRIN3A</em></td>
<td>165</td>
<td>36 (9)</td>
<td>1.76 (0.06)</td>
<td>66.9 (6.8)</td>
<td>21.0 (2.0)</td>
</tr>
</tbody>
</table>

31. Data are mean (standard deviation)
32. * greater than runners (*) < 0.0005)
33. ** greater than non-athletes and runners (** < 0.0005)
34. *** greater than non-athletes (** < 0.0005)
Table 2. Distribution of rugby athletes according to playing position. Data are number of athletes (% of all athletes)

<table>
<thead>
<tr>
<th></th>
<th>Analysis of PRDM1 and GRIN3A</th>
<th>Analysis of KCNH8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 212$</td>
<td>$n = 139$</td>
</tr>
<tr>
<td><strong>Forwards vs. Backs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forwards</td>
<td>114 (53.8)</td>
<td>79 (56.8)</td>
</tr>
<tr>
<td>Backs</td>
<td>98 (46.2)</td>
<td>60 (43.2)</td>
</tr>
<tr>
<td><strong>Positional sub-groups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Front Five</td>
<td>66 (31.1)</td>
<td>47 (33.8)</td>
</tr>
<tr>
<td>Back Row</td>
<td>50 (23.6)</td>
<td>34 (25.5)</td>
</tr>
<tr>
<td>Half Backs</td>
<td>41 (19.3)</td>
<td>22 (15.8)</td>
</tr>
<tr>
<td>Centres</td>
<td>27 (12.7)</td>
<td>18 (12.9)</td>
</tr>
<tr>
<td>Back Three</td>
<td>28 (13.2)</td>
<td>18 (12.9)</td>
</tr>
</tbody>
</table>
Table 3. Genotype distribution for *PRDM1* and *GRIN3A* according to athlete group. Data are number of individuals (%)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>Group</th>
<th>Non-athlete</th>
<th>All athletes</th>
<th>Rugby</th>
<th>Runner</th>
<th>Total</th>
<th>MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>n</em> = 355</td>
<td><em>n</em> = 377</td>
<td><em>n</em> = 212</td>
<td><em>n</em> = 165</td>
<td><em>n</em> = 732</td>
<td></td>
</tr>
<tr>
<td><em>PRDM1</em></td>
<td>CC</td>
<td></td>
<td>282 (79.4)</td>
<td>284 (75.3)</td>
<td>163 (76.9)</td>
<td>121 (73.3)</td>
<td>566 (76.6)</td>
<td>0.12</td>
</tr>
<tr>
<td>rs10499043</td>
<td>CT/TT</td>
<td></td>
<td>73 (20.6)</td>
<td>93 (24.7)</td>
<td>49 (23.1)</td>
<td>44 (26.7)</td>
<td>166 (23.4)</td>
<td></td>
</tr>
<tr>
<td><em>GRIN3A</em></td>
<td>GG</td>
<td></td>
<td>293 (82.5)</td>
<td>311 (82.5)</td>
<td>175 (82.6)</td>
<td>136 (82.4)</td>
<td>604 (82.5)</td>
<td>0.09</td>
</tr>
<tr>
<td>rs1535628</td>
<td>GA/AA</td>
<td></td>
<td>62 (17.5)</td>
<td>66 (17.5)</td>
<td>37 (17.4)</td>
<td>29 (17.6)</td>
<td>128 (17.5)</td>
<td></td>
</tr>
</tbody>
</table>

*MAF, minor allele frequency*
Figure 1. *KCNH8* rs4973706 genotype distribution in non-athletes and rugby athletes. * greater than non-athletes ($p = 0.013$)
Figure 2. Runners’ PB marathon time according to PRDM1 rs10499043 genotype. * faster time than CC ($p = 0.023$).