Genetic variation at RAB3GAP2 and its role in exercise-related adaptation and recovery

Kristoffer Ström  
Lund University

Nikolay Oskolkov  
Lund University

Tugce Karaderi  
University of Oxford

Sebastian Kalamajski  
Lund University

Ola Ekström  
Lund University

Eri Miyamoto-Mikami  
Juntendo University

Claes Ladenvall  
Lund University

Ellen Kakulidis  
Lund University

Steven Reid  
Lund University

Anna-Maria Dutius Andersson  
Lund University

Dmytro Kryvokhyzha  
Lund University

Enming Zhang  
Lund University

João Fadista  
Statens Serum Institut

Motoyuki Iemitsu  
Ritsumeikan University

Anubha Mahajan  
University of Oxford

Emma Ahlqvist  
Lund University  https://orcid.org/0000-0002-6513-2384

Rashmi Prasad  
Lund University  https://orcid.org/0000-0002-4400-6741

Kay Pruefer  
Max Planck Institute for Evolutionary Anthropology

María Sabater-Lleal  
Research Institute Hospital de la Santa Creu i Sant Pau - IIB Sant Pau  https://orcid.org/0000-0002-0128-379X

Nicholas Smith  
University of Washington

Abbas Abbas Dehghan  
Imperial College London

Lars Lind  
Uppsala University

Kerry McGawley  
Mid Sweden University

Andrew Morris  
University of Oxford

Mikko Lehtovirta  
Lund University
Article

Keywords: RAB3GAP2, muscle healing, exercise

DOI: https://doi.org/10.21203/rs.3.rs-701652/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Skeletal muscle fiber composition and capillary density influence physical performance and whole-body metabolic properties. ~45% of the variance in fiber type is heritable, which motivated us to perform a genome-wide association study of skeletal muscle histology from 656 Swedish men. Four independent variants were associated (p < 5x10^-8) with proportion of type IIX fibers or capillary-to-fiber ratio (C:F). The strongest signal localized to the rs115660502 variant, where the G-allele corresponded with increased C:F and reduced skeletal muscle expression of the proximal gene, RAB3 GTPase Activating Non-Catalytic Protein Subunit 2 (RAB3GAP2). The G-allele was less frequent in elite short-track sprinters and more frequent in endurance athletes than in matched non-athlete (population) controls; RAB3GAP2 expression was reduced by high-intensity intermittent training. RAB3GAP2 protein was not uniformly expressed in muscle tissue but localized to the endothelium and capillaries. Experimental reduction of RAB3GAP2 in human endothelial cells led to increased tube formation in vitro, to regulation of secreted factors promoting angiogenesis and T-cell activation, to reduced intracellular levels of von Willebrand factor (VWF) and, post-implantation, to increased endothelial cell density in vivo in mice. The amount of RAB3GAP2 in skeletal muscle was positively associated with exercise-induced release of VWF in vivo in humans. By regulating the release of protein factors (VWF, CD70, TNC, TNXB, MCP1, IGFBP3, COL1A1, TFPI2 and tPA), RAB3GAP2 influences fitness adaptation after exercise by improving muscle healing and promotion of capillary formation.

Introduction

One of the most pervasive threats to contemporary human health is the sedentary nature of our lifestyles. The beneficial impact of exercise on health is determined in part by the damage caused to skeletal muscle and the rate at which damaged muscle tissue repairs. Indeed, this cyclical process of damage and repair dictates fitness adaptation. These adaptations include skeletal muscle hypertrophy, improvements in substrate utilization, fiber type switching (from Type IIX to Type IIA) and changes leading to increased blood flow and perfusion such as increased capillary density.

Sustained aerobic exercise has many established health benefits including reductions in risk of cardiovascular disease. However, people vary in their response to exercise^1,2, especially those who are untrained, acute bouts of exhaustive exercise occasionally results in life-threatening events such as myocardial infarction, hemorrhagic stroke and provoked venous thromboembolism^3-6.

Genetics has proven powerful for the discovery of pathogenic pathways and processes across many diseases. Such knowledge has been exploited in drug development, where information about genetic perturbation can help narrow the search for drugable targets^7. Exercise is increasingly viewed as a therapeutic adjunct or alternative to drugs. As with drug development, it is plausible that identifying genetic perturbations influencing pathways through which exercise impacts human health might lead to discoveries that broaden therapeutic options, either by facilitating the development of exercise mimetics or by optimizing individualized exercise interventions.

The purpose of this study was to identify genetic variants associated with exercise-related muscle tissue phenotypes and elucidate the underlying molecular mechanisms that link exercise with health.

Results

GWAS of skeletal muscle fiber type distribution and capillary-to-fiber ratio

Skeletal muscle fiber type distribution and capillary density are highly heritable^8,9 and help determine energy metabolism and performance capacity^10-12. However, the specific genetic loci influencing fiber composition and capillary density are largely undefined. We began by undertaking genome-wide association studies (GWAS) for exercise-related traits, placing emphasis on phenotypes characterized using histology in human muscle biopsies. We identified three genome-wide significant (p<5 x 10^-8) signals for Type IIX fiber type (rs7776803, rs145631867 and rs149081100) and an additional signal for capillary-to-fiber ratio (C:F). The signals for C:F were derived from two variants within the same Chr 1 haplotype, refined to a single variant (rs115660502) after conditional analyses (Table 1, Figure 1a-d and Supplement figures S1-S3).

Skeletal muscle is highly vascularized, with capillaries comprising 2-3% of total muscle mass. C:F varies depending on fiber-type distribution and the intensity and frequency of exercise (Supplement table S1 – S3). In sedentary men and women, on average, ~4 capillaries surround Type I and IIa fibers and ~3 surround Type IIX fibers^13. In endurance trained athletes C:F is higher (~5-6 for Type I and IIa fibers and ~4 for Type IIX fibers)^13. The minor(major) G(A) allele at rs115660502 (associated with C:F) is frequent at 3-5% in European ancestry populations, but lower or absent in African and Asian populations (Supplement figure S4). We found that the same allele was roughly twice as frequent in elite endurance athletes relative to ethnicity-matched general population controls, and half the frequency in ethnicity-matched elite power sports athletes (see Materials and Methods, and Supplement figure S5). In summary, four independent variants were associated with either proportion of Type IIX fibers or C:F and the rs115660502 variant frequency differs between the general population and elite athletes.

The rs115660502 variant propagates its effect on C:F by influencing RAB3GAP2 expression

Having identified intronic or intergenic variants associated with key exercise-response phenotypes (fiber type and C:F), we examined if these variants propagate their effects through the expression of proximal (within a 1 Mb window) genes (Supplement table S4). The rs115660502 variant was associated with RAB3 GTPase activating protein subunit 2 gene (RAB3GAP2) expression (p<0.007). We found that both six-weeks of high-intensity intermittent training (HIIT) and presence of the G allele were associated with lower RAB3GAP2 mRNA expression in human skeletal muscle by ~10% (p = 0.002,
mRNA level was positively associated with change in circulating VWF concentration following the Wingate test (2128 ± 728 ng/ml [b ± SE], n = 32) (Figure 2c). The reduction of RAB3GAP2 expression following HIIT was replicated in a meta-analysis of 8 datasets with acute HIIT (−7%, \( P_{\text{meta}} = 2.9 \times 10^{-3}, n = 54 \)) and in 3 datasets with prolonged HIIT (−6%, \( P_{\text{meta}} = 1.5 \times 10^{-3}, n = 34 \)) in MetaMEX\(^\text{14} \). The RAB3GAP2 protein is widely expressed in human tissues, including skeletal muscle (Supplement Figure S6). Using immunohistochemistry, we localized the RAB3GAP2 protein to skeletal muscle capillaries and endothelium, but it was absent in muscle fibers and larger vessels (Figure 2d-i). This distinction in vascular localization was supported by the presence of RAB3GAP2 in human microvascular endothelial cells (HMEC) and absence in human umbilical vein endothelial cells (HUVEC) of macrovascular origin (Supplement Figure S7). rs115660502 is located close to the transcriptional start site of RAB3GAP2 (~20 kb upstream) and, in two ENCODE cell lines (HuH7 and Medullob)\(^\text{15} \), within a DNase hypersensitive region. Several transcription factors implicated in angiogenesis (e.g. STAT3, FOSL1, SIRT6, TAL1, and GATA2) bind in close proximity (500-1000 bp) to rs115660502 in the ENCODE database\(^\text{15} \). In summary, the G-allele and HIIT are both associated with lower RAB3GAP2 expression, and the protein is localized to the endothelium and capillaries.

**Reduced expression of RAB3GAP2 in human endothelial cells stimulates tube formation**

To test if RAB3GAP2 expression is causally related to capillary formation, RAB3GAP2 was knocked-down using CRISPR-Cas9 (mRNA by ~70 - 80% in HMECs), with corresponding protein reduction (Supplement figure S8). The level of knockdown was stable and sustained over 12 passages (Supplement figure S8).

We then undertook an *in vitro* tube formation assay using HMECs, where RAB3GAP2 knockdown led to a 2.8-fold increase in number of loops (i.e., ring structures) at 6 hours of differentiation and a 2.9-fold increase at 24 hours (\( \rho_{\text{Mann-Whitney}} = 8.0 \times 10^{-7} \) and \( \rho_{\text{Mann-Whitney}} = 7.1 \times 10^{-5} \) respectively, \( n = 11 \) and 13). The number of branching points increased by ~30% in RAB3GAP2 knocked-down cells compared with wildtype at 6 hours of differentiation (\( \rho_{\text{Mann-Whitney}} = 8.8 \times 10^{-9}, n = 11 \)) and increased by ~32% at 24 hours (\( \rho_{\text{Mann-Whitney}} = 7.6 \times 10^{-4}, n = 13 \)). The total tube length formed was ~22% longer in RAB3GAP2 knocked-down cells compared to wildtype at 6 hours of differentiation (\( \rho_{\text{Mann-Whitney}} = 8.0 \times 10^{-7}, n = 11 \)) and ~27% longer at 24 hours (\( \rho_{\text{Mann-Whitney}} = 2.4 \times 10^{-5}, n = 13 \)). Quantification and example images of tube formation in HMECs from the different conditions are shown in figure 3a-d.

In an *in vivo* angiogenesis plug assay, Matrigel containing HMECs with or without RAB3GAP2 knockdown were injected in the groin region of NGS mice. The resulting gel plugs were harvested after 7-days, sectioned, and stained using Masson’s Trichrome. The total HMEC cell area was ~50% larger in RAB3GAP2 knocked-down cells compared to wildtype (\( \rho_{\text{Mann-Whitney}} = 0.036, n = 3-5 \)). Quantification and example images of cell density of HMECs are shown in Figure 3e-f. In summary, experiments in human endothelial cells replicate the GWAS association with C:F, whereby knockout of RAB3GAP2 leads to a gain-of-function with increased tube formation *in vitro* and increased cell density *in vivo* in mice.

**RAB3GAP2 regulates the secretion of proteins promoting angiogenesis and T-cell activation**

RAB3GAP2 is known to regulate vesicle trafficking\(^\text{16} \). To identify secreted proteins regulated by RAB3GAP2 and driving the effects described above, we performed a RNA sequencing screen in HMECs with or without CRISPR-Cas9-mediated knockdown of RAB3GAP2 expression. 496 genes with lower and 993 genes with higher expression in RAB3GAP2 knocked down cells were identified at FDR 5% (Supplement table S5). In addition, 154 secreted proteins were screened in the medium of HMECs using 4 Olink panels with or without knockdown of RAB3GAP2 (Supplement table S6). Combining these results, RAB3GAP2 knockdown influenced both the mRNA expression and protein secretion of eight factors (Figure 4a-h). Four of these factors are known to stimulate angiogenesis, i.e., Tenascin C (TNC), Insulin like growth factor binding protein 3 (IGFBP3), Plasminogen activator tissue type or tPA (PLAT), C-C motif chemokine ligand 2 / MCP-1 (CCL2), all had higher concentration in RAB3GAP2 knocked down cells. Three of the eight factors are known to inhibit angiogenesis, i.e., Tenascin XB (TNXB), Collagen, type I, alpha 1 (COL1A1) and Tissue factor pathway inhibitor 2 (TFPI2), all had lower concentration in RAB3GAP2 knocked down cells. However, the most strongly regulated factor was the T-cell activating protein cluster of differentiation 70 (CD70), with a ~3-fold higher RNA expression in RAB3GAP2 knocked down cells (\( \rho_{\text{adj}} = 1.4 \times 10^{-10}, n = 8 \)) and a ~2.5-fold higher protein concentration in the medium from RAB3GAP2 knocked down cells compared to wildtype (\( \rho_{\text{Wilcoxon}} = 0.03, n = 6 \)). In summary, lower RAB3GAP2 levels lead to pro-angiogenic levels of proteins like TNC, IGFBP3, tPA and MCP-1 and possibly to T-cell activation by increased secretion of CD70.

**RAB3GAP2 regulates von Willebrand factor secretion**

RAB3GAP2 heterodimerizes with the catalytic subunit RAB3AP1 to inhibit RAB3 activity\(^\text{17} \). In endothelial cells, RAB3 (A and D isoforms) stimulate the release of WPB (Weibel-Palade bodies)\(^\text{18,19} \), vesicles that are highly enriched for von Willebrand factor (VWF) but also contain angiogenic factors\(^\text{20} \). Characterized as fluorescence intensity per cell, CRISPR-Cas9-mediated knockdown of RAB3GAP2 expression in HMECs suggested reduced intracellular VWF levels (\( n = 3, 20-24 \) cells per experiment, Figure 4i-j). This decrease was quantified in independent experiments using an ELISA assay to measure the intracellular VWF concentration. Intracellular VWF concentration was ~11% lower in RAB3GAP2 knocked-down cells compared with wildtype (\( \rho_{\text{Wilcoxon}} = 7.8 \times 10^{-3}, n = 8, \) Figure 4i), indicating increased VWF secretion.

Exercise acutely stimulates VWF secretion\(^\text{21} \), with VWF secretion being inversely related to physical fitness\(^\text{22} \). To test whether RAB3GAP2 acutely regulates exercise induced VWF secretion, we performed 30s Wingate tests in healthy male adults and examined basal expression of RAB3GAP2 in *vastus lateralis* muscle and circulating VWF before and after the exercise (Supplement table S7). Circulating VWF approximately doubled with exercise (\( \rho_{\text{Wilcoxon}} = 7.3 \times 10^{-12}, n = 39 \), Figure 4k). Moreover, exercise-induced VWF secretion was inversely associated with VO\(_{2}\text{max} \) (expressed in ml kg\(^{-1}\) min\(^{-1}\)) (~216 ± 77 ng/ml [b ± SE], \( p = 9.0 \times 10^{-5}, n = 38 \), adjusted for VWF before the Wingate test and peak power per kg body weight). Finally, we demonstrated that basal RAB3GAP2 mRNA level was positively associated with change in circulating VWF concentration following the Wingate test (2128 ± 728 ng/ml [b ± SE], \( p = 6.0 \times 10^{-5} \), n = 11) and ~4%, (\( \rho_{\text{Meta-FDR}} = 0.007, n = 69 \) vs 8) respectively (Figure 2a-b, Supplement table S4), which in turn was inversely correlated with C:F ratio (\( \rho_{\text{Meta-FDR}} = 0.38, p = 0.03, n = 32 \) (Figure 2c).
Given that the G allele at rs115660502 is associated with lower RAB3GAP2 expression and experimental lowering of RAB3GAP2 expression reduces intracellular VWF levels, we hypothesized that this allele is associated with higher circulating VWF levels in the non-exercised state, which we confirmed in a GWAS meta-analysis dataset \( n = 43,775 \) \( b(\pm SE) = 0.0132 \pm 0.0068 \) s.d. units per G allele, \( p = 0.05 \). In summary, the in vivo intervention study in humans supports RAB3GAP2 regulation of exercise-induced VWF secretion.

### Discussion

We undertook a sequence of integrated studies to discover, validate and mechanistically elucidate novel molecular features underlying skeletal muscle-mediated effects of exercise in health. We began by undertaking a GWAS of exercise-related skeletal muscle phenotypes and leveraged these findings to discover novel mechanisms linking exercise to health traits. Specifically, we show that a RAB3GAP2 variant (rs115660502) is associated with C:F ratio and that the frequency of the effect allele \( G \) varies between elite athletes and population controls. We uncovered that both the G allele and HIIT exercise training decrease RAB3GAP2 muscle expression, which, through either mechanism, regulates the secretion of VWF, T-cell activating and angiogenesis promoting proteins, resulting in enhanced muscle repair and capillary formation. Experimental reduction of RAB3GAP2 expression leads to increased angiogenesis both in vitro and in vivo. Finally, we show that the level of RAB3GAP2 in muscle modulates the exercise-induced secretion of VWF in vivo. Taken together, these results show that RAB3GAP2 plays a key role in muscle adaptation to exercise, both acutely by regulating VWF and CD70 release and chronically by mediating capillary formation.

RAB3GAP2 encodes a regulatory subunit that together with RAB2GAP1 (encoding the catalytic subunit) form the GTPase-activating complex with specificity for the RAB3 Rab-GTPase-activating proteins (RAB3A, RAB3B, RAB3C and RAB3D)\(^{16,17}\). The RAB3 protein family controls exocytosis of neurotransmitters and hormones\(^{16}\), and mutations in RAB3GAP2 cause Martsolf syndrome and Warburg Micro Syndrome 2, diseases characterized by delayed neurodevelopment and cognitive impairments\(^{25}\). To our knowledge, no previous studies have linked RAB3GAP2 to exercise-response phenotypes.

Exercise acutely stimulates VWF secretion\(^{21}\) with the quantity of VWF secreted being inversely related to physical fitness\(^{22}\), consistent with the broader known effects of exercise in cardiovascular risk. Our data help explain in part why an acute bout of strenuous exercise might transiently raise the propensity for thrombotic events (at least in high risk individuals), whereas regular exercise (i.e., repeated bouts of acute exercise that enhance aerobic fitness) lowers overall cardiovascular risk in the long-term in virtually all population groups.

Our findings suggest that in those carrying the rs115660502 G allele, skeletal muscle recovery may be accelerated; yet these same individuals may experience higher basal levels of VWF, an established risk factor for venous thrombosis\(^{26,27}\). A reduction of RAB3GAP2 activity, either genetically (via presence of the G allele) or by exercise, results in reduced intracellular VWF concentrations, presumably owing to increased basal secretion of WPBs. These conclusions are supported by the in vitro and in vivo experiments in HMECs, the higher circulating levels of resting VWF in carriers of the G allele and by the in vivo response to Wingate testing.

The secretion of VWF, CD70 and proteins promoting angiogenesis in response to exercise likely reflects the elevated risk of injury when intensive exercise ensues. This mechanism might have evolved to reduce risk of bleeding, minimize muscle microruptures and help repair muscle tissue after exercise, as well as to promote fitness adaptation through de novo capillary formation. This implies that the regulation of these factors is a rapid process that clearly outpaces de novo production of VWF and formation of new WPBs. Subsequently, the endothelium secretes all VWF present in its cells. However, the lower VWF concentration in G allele carriers and/or well-trained individuals suppresses this acute response (as observed with the Wingate test, both with respect to RAB3GAP2 in muscle and VO\(_{2\text{max}}\)), which may reduce risk of cardiovascular events associated with acute strenuous exercise. These observations may partially explain why exercise training reduces cardiovascular risk, whereas acute and intense exertion in untrained (or unhealthy/older) individuals can sometimes cause harm.

In summary, naturally occurring genetic perturbation of RAB3GAP2 promotes capillary formation and the release of VWF in a manner consistent with exercise. The higher frequency of the rs115660502 variant allele in highly trained athletes may be ergogenic, by enhancing exercise-induced capillary formation and muscle repair. However, this variant may also predispose risk of thrombotic events in sedentary individuals.

### Materials And Methods

#### Cohorts

GWAS (ULSAM, MM and MEI): we conducted a genome-wide association study (GWAS) in 656 men from three cohorts of Swedish ancestry. These cohorts were the Uppsala Longitudinal Study of Adult Men (ULSAM, n=482, Supplement table S8)\(^{29}\), the Malmö Men’s study (MM, n=128, Supplement table S9)\(^{30}\), and the Malmö Exercise Intervention study (MEI, n=46, Supplement table S10)\(^{31,32}\). The ULSAM cohort consists of men aged 71.0 ± 0.6 years [mean ± SD] and BMI 26.3 ± 3.4 kg/m\(^2\) [mean ± SD]. The MM cohort consists of men aged 65.9 ± 2.0 years [mean ± SD] and BMI = 26.4 ± 3.4 kg/m\(^2\) [mean ± SD], including 50 participants with diagnosed type 2 diabetes (T2D) at enrollment. The MEI cohort consists of men aged 37.7 ± 4.3 years [mean ± SD] and BMI = 28.0 ± 3.1 kg/m\(^2\) [mean ± SD], including 22 participants with and 24 without a first-degree family member with T2D.
WINGATE (MSAT): 39 healthy men (Age=36.5±8.3, BMI=24.1±2.5, VO$_{2\text{max}}$ = 52.0±8.1; [mean ± SD]) were enrolled in the study by advertising in social media and through local cycling clubs. Inclusion criteria as follows: 1) Male 2) Healthy, no medications 3) Age range between 20 and 55. Subjects were given both oral and written information about the experimental procedures before giving their written informed consent. Each participant went through three visits at different time points. All subjects completed all three visits. First visit involved a regular doctor's examination with blood samples and measuring anthropometric characteristics. Second visit consisted, after an overnight fast, of a Wingate test followed by muscle biopsy, VO$_{2\text{max}}$ was measured during the third and last visit. The study was approved by the local Ethics committee, Lund University (Dnr 2015/593). For determination of peak anaerobic power and VO$_{2\text{max}}$, subjects were instructed to perform only easy training during the 48 h prior each test. To determine peak anaerobic power, a 30-s all-out Wingate test$^{33}$ was conducted on a cycle ergometer (Monark Peak power). Before the test, a 5-minute low intensity ~150w warm-up, with instructions to perform a 5s high cadence drill each minute was performed. The test started with the subject pedaling as fast as possible. When a cadence of 120 rpm was reached, a braking resistance equivalent to 0.7 N×m×kg$^{-1}$ was applied to the freewheel and remained constant during the 30 seconds. Subjects were instructed to sit down throughout the test. Strong verbal encouragement was given throughout to ensure a maximal effort was provided. An incremental test to exhaustion was performed to determine VO$_{2\text{max}}$. The test started with 3 min of cycling at 3 W/kg (rounded down to nearest 10 W) and then increased by 35 W every 2 minutes until voluntary exhaustion or failure to maintain ≥60 rpm. Strong verbal encouragement was given throughout. VO$_{2}$ was measured using an Oxycon Pro (Jaeger GmbH, Germany) with a mixing chamber and a 30s sampling time. Gas sensors were calibrated according to instructions by the vendor before every test. Maximal oxygen uptake was determined as the mean of the two highest values attained during exercise from any 30-s period.

COHORTS OF ATHLETES: The four elite athlete cohorts were: 1) Swedish cross-country skiers (n = 15)$^{24}$. Fourteen junior cross-country skiers (8 males: age and VO$_{2\text{max}}$ were 18-20 years and 65.7 ± 2.6 [mean±SD] ml/kg/min, respectively; 6 females: age and VO$_{2\text{max}}$ were 18-20 years and 54.7 ± 4.7 [mean±SD] ml/kg/min, respectively) were recruited from two specialist Swedish ski schools. In addition, one senior male cross-country skier was recruited (age and VO$_{2\text{max}}$ were 38 years and 69.5 ml/kg/min). All participants competed at a national level, six were members of national junior development teams and one was a senior World championship medalist. All subjects provided written informed consent and the study protocol was approved by the Regional Ethical Review Board, Umeå University. 2) Spanish athletes from different disciplines (n = 141), including triathletes (n = 16)$^{35}$. All athletes were Spanish (Caucasian) men with solid expertise in international competitions. Blood/saliva samples were collected over 10 years in different places of Spain from athletes who were runners (mostly 5,000m and above), professional road cyclists, rowers, canoeists, and triathletes. Except for professional cyclists, inclusion criteria for this group were: having been finalist in ≥1 edition of Olympic Games or World/European championships (in either ‘absolute’ or under-23-year). Their age and VO$_{2\text{max}}$ were 19–32 years and 74.0±5.5 [mean±SD] ml/kg/min, respectively. Saliva samples were collected from control, disease-free subjects who were also all Spanish (Caucasian) men, either inactive undergraduate students living in Madrid or recreational runners (n = 6). Their age and VO$_{2\text{max}}$ were 19–32 years and 45.0±5.5 [mean±SD] ml/kg/min, respectively. In addition, MAF from Iberian (IBS) population from 1000G was used as a Spanish population reference. All subjects provided written informed consent and the study protocol was approved by the Ethical Review Board of Universidad Pablo Olavide. DNA sampling was done in Madrid. 3) Polish short track speed skaters (n = 18). For the purpose of current study, from the local cohort study of professional athletes from Poland, the elite group of short track speed skaters was selected. Studied subjects were members of the Polish national team who at least once were finalists of the Winter Olympic Games or World/European championships. Their age was 18-28 years, and VO$_{2\text{max}}$ 68.5±4.2 [mean±SD] ml/kg/min. Polish non-athlete controls were also collected (n=24), 15 males and 9 females with age 19 - 33 years (VO$_{2\text{max}}$ was not measured). For DNA isolation purposes, whole blood samples were collected. All subjects provided written informed consent and the study protocol was approved by the Ethics committee of the Medical University of Bialystok (Poland) - approval number: R-I-002/286/2013. 4) Jamaican sprinters (n = 116) comprising national- and international-level athletes including athletes of the highest level (i.e., medalists or world record holders). The Jamaican cohorts comprised 311 control subjects (male = 156, female = 155) from throughout the island and 116 athletes (male = 60, female = 56) who had participated in sprint events up to 400 m, in jump events, and in throw events (100-200 m, n = 71; 400 m, n = 35; jump and throw, n = 10). Athletes could be subdivided further into categories defined by their level of performance: national-level athletes (n = 28), who were competitive at national-level competition in Jamaica and the Caribbean, and international-level athletes (n = 86), who had competed at major international competition for Jamaica. Forty-six of these international athletes had won medals at major international competitions or held sprint world records. All subjects provided written informed consent to participate in the study, which was approved by the Ethics Committee of University of West Indies. Genotyping of the variants achieving genome-wide significance was done and the allele frequencies were compared to the corresponding non-athlete controls of Swedish, Spanish, Polish and Jamaican population allele frequencies.

Phenotypes

In the ULSAM, MM, MEI and MSAT cohorts, muscle biopsies were taken with a 6 mm Bergström needle and frozen in liquid nitrogen, and serial sections (10 or 16 μm) were cut using a cryostat at -20 °C. Capillaries were stained using the double staining method$^{34}$, calculated as the number of capillaries divided by the number of fibers observed within a given section. Myofibrillar ATPase histochemistry was performed by preincubation at pH 4.4, 4.6, and 10.3 to identify muscle fiber types$^{35}$, the proportion of fiber types (i.e. type I, IIa or IIx) were calculated as the number of each fiber type, divided by the total number of fibers in the section, multiplied by 100. For ULSAM, computer image analysis was performed using an image analysis equipment (Multisync II, BIO-RAD SA, Richmond, CA, USA) and for MM, MEI and MSAT (BioPix IQ 2.0.16 software, BioPix AB, Sweden). The phenotypes and sample sizes for each phenotype are reported in Supplementary Table S10. Spearman correlation matrices for all phenotypes, stratified by cohort, are presented in Supplementary Tables S1-S3. Within each cohort, all phenotypes were pre-adjusted for age and BMI, and inverse rank-normal transformed using the Rankit equation (Bliss 1967) to provide normal data distributions.

Genotyping
Genotyping of the ULSAM was done using the Illumina Omni 2.5M array and Illumina CardioMetabochip (Illumina, California, USA). Genotyping of the MM and MEI cohorts were done using the HumanOmniExpress 12v1 C chips (Illumina, California, USA); genotypes were called using the Illumina Genome studio software, as described in detail elsewhere. Individuals were excluded on the basis of: call rate <95%; gender mismatch; relatedness; and non-European ancestry. Variants were excluded on the basis of: call rate <95% (<99% for MAF <5%); exact p-value for Hardy-Weinberg equilibrium <10^{-8}; and MAF <1%. Genotyping of the different cohorts of imputed genome-wide significant variants from the meta-analysis were done using either TaqMan PCR or Sequenom platforms according to the manufacturer's instructions. For TaqMan an ABI Prism Sequence Detection System ABI 7900HT (Applied Biosystems) was used for post-PCR allelic discrimination by measuring allele-specific fluorescence. The results were in Hardy-Weinberg equilibrium (p=1x10^{-4}).

**Imputation & meta-analysis**

Genotype data in each cohort were imputed up to 35 million variants from the 1000 Genomes reference panel (all ancestries, March 2012). Prephasing of haplotypes and imputation were performed using ShapeIT and IMPUTE2, respectively. The association within each cohort was performed using SNPTEST frequentist score additive model test. Genetic variants that had poor imputation quality inferred by an info score ≤0.4 and/or had high standard error (SE>10) were excluded from meta-analyses. Fixed effects meta-analyses were undertaken on the summary statistics obtained from the three cohorts using GWAMA, with phenotypes inverse rank-normal transformed and pre-adjusted for age and BMI. Post-meta-analysis QC included removing variants that: a) were detectable only in one cohort and b) had total minor allele count <10. To validate the quality of imputation, all imputed genome-wide significant SNPs were directly typed in the MM and MEI cohorts. The imputation error rates vary from variant to variant (i.e., from 0-4% for common homozygotes, 0-86% for heterozygotes), and these error rates were inversely related to the imputation info score. To assure that findings were not biased by factors related to genotype imputation, genetic association analyses were re-run using directly genotyped data, which did not indicate any major influence of imputation on the results. Results from the directly genotyped, rather than imputed, data are generally presented in the manuscript. Analyses conditioning on variants associated with the index traits at a level of genome-wide significance (P=5x10^{-8}) were undertaken in order to discover possible secondary association signals in corresponding regions. In case more than one variant located on the same haplotype was found to be associated at the genome-wide significance level, and the associated variants were not in a strong LD with each other, we tested for secondary / independent association. For this purpose, an analysis was performed, where association analysis of all genetic variants from the haplotype region was re-run adjusting for the genome-wide significant variants sequentially one-by-one. If the association with P=5x10^{-8} remained after adjusting for another neighboring genome-wide significant variant, we concluded that the two genome-wide significant variants represented independent genetic signals. Otherwise, their association was capturing the same genetic signal. The 'Metafor' command in R software as used to make Forest plots, and the LocusZoom browser was used for creating regional association plots.

**Testing allele frequencies in athletes versus population controls**

The five genome-wide significant index variants were genotyped in the four cohorts of elite athletes from different disciplines and populations, i.e. Swedish cross-country skiers, Spanish athletes of different disciplines, Polish skaters and Jamaican sprinters (up to 400 m and in jump and throw events) and in corresponding population-matched controls of non-athlete status. Since the athletes originated from different world populations (Sweden, Spain, Poland, and Jamaica) we normalized the G-allele frequency at rs115660502 in the athletes by G-allele frequency in the control (non-athletes) group from the corresponding population. For example, the ratio of the G-allele frequency in Swedish skiers to G-allele frequency in Swedish controls was ~2.5, while the corresponding ratio for Jamaican sprinters was ~0.5, suggesting that the G-allele of rs115660502 is over-represented in endurance sports (skiing) and under-represented in power sports (sprint). To test for significance, resampling with replacement by randomly drawing equal number of controls as athletes and calculating ratios 100,000 times was done to build confidence intervals. The same resampling procedure was repeated for building the empirical distribution of G-allele frequencies in all populations. Next, statistical significance was tested by counting the number of times the re-sampled G-allele frequency of the controls was different or equal to the respective athletes. We defined athletes competing in cross country skiing and triathlon as extreme endurance athletes (i.e., disciplines with >1.5 h total uninterrupted physically demanding performance time) and athletes competing in sprint running as extreme power athletes (i.e., disciplines with <60 s total performance time). Other disciplines were defined as not being extreme in their physical profile, e.g., skating, basketball, and judo. No difference in frequencies were found for any of the variants associated with Type IIX fiber distribution. However, the capillary-increasing G-allele at rs115660502 has a frequency of 4.8% and 3.3% in the Swedish and Spanish non-athlete population controls respectively but had 10% and 6.3% in the Swedish cross-country skiers and Spanish triathletes respectively. By contrast, the G-allele had a 1.1% frequency in the non-athlete Jamaican population controls, which was significantly higher compared to the 0.4% frequency of the Jamaican sprinters (p = 0.03). Athletes of non-extreme or mixed sports had a similar G-allele frequency as their corresponding population controls.

**eQTL analysis**

eQTL analyses were done using Affymetrix microarray expression data from the MEI (n=39; 23,941 probe sets) and MM (n=38; 22,283 probe sets) cohorts. Genome-wide significant variants from the meta-analysis for each phenotype were tested for linear associations with gene expression levels using the ‘Matrix_eQTL’ command in R. cis-eQTL (within a 1 Mb window) was performed separately for MEI and MM with the corresponding False Discovery Rate (FDR) correction for multiple testing (FDR<5%). To be consistent with the primary GWAS analyses, we used age and BMI as covariates for eQTL analysis. The results of the eQTL analysis were meta-analyzed with Stouffer's z-score method. The R commands “hgu133a.db” and “nugohs1a520180.db” were used for probeset-gene annotation.
HMEC-1 cells (CRL-3243 (Lot: 64334159), ATCC, VA, USA) were grown at 37°C and 5% CO₂ in growth medium (MCDB131 medium (Thermo Fischer Scientific (10372019)) supplemented with 10 ng/ml EGF (Thermo Fischer Scientific (PHG0314)), 1 µg/ml hydrocortisone (Sigma (H8988-1G)), 10 mM GlutaMAX (Thermo Fischer Scientific (35050038)), 10% FBS (Sigma (F7524)) and Antibiotic-Antimycotic solution (Thermo Fisher Scientific, Gibco (15240062)) on tissue culture plates precoated with Attachment Factor (Thermo Fisher Scientific (S-006-100)). Cells were passaged at around 70-80% confluence. Testing for mycoplasma infection was done routinely.

Quantitative real-time PCR (QPCR)

Verification of a sustained CRISPR/Cas9-mediated knockdown of RAB3GAP2 in HMEC cells, and analysis of the expression of RAB3GAP2 mRNA in human skeletal muscle was done using QPCR. RNA was isolated with the RNeasy Plus Mini Kit (Qiagen) and concentration and purity were measured using a NanoDrop ND-1000 spectrophotometer (A260 / A280 > 1.8 and A260 / A230 > 1.0) (NanoDrop Technologies, USA). Reverse transcription of 500 ng RNA was performed using the Quantitect Reverse Transcription kit (Qiagen). QPCR was run on a Viia 7 real-time PCR system (Thermo Fisher Scientific) with 2 ng cDNA in 10 µl reactions and TaqMan Expression PCR Master Mix according to the manufacturer's instructions (Applied Biosystems, Thermo Fisher Scientific). Samples for each gene were analyzed in triplicates on the same 384 well plate with 3 endogenous controls (POL2A, HPRT1 and PPIA). The expression levels were calculated and normalized by geometric averaging of the endogenous controls as previously described. Assays used: RAB3GAP2 (exon 24-25 boundary) (Hs00202700_m1) and RAB3GAP2 (exon 3-4 boundary) (Hs01073425_m1). Endogenous control assays: POLR2A (Hs00172187_m1), HPRT1 (4326321E, VIC-MGB) and PPIA (4326316E, VIC-MGB).

RNA sequencing

HMEC-1 cells grown at a confluence of ~70-80%, were washed in PBS, detached using TrypLE (Thermo Fisher Scientific) and counted. RNA was extracted from ~500k cells lysed in RLT plus lysis buffer (Qiagen), using the RNeasy Plus Mini Kit (Qiagen). RNA concentration was determined using a NanoDrop ND-1000 spectrophotometer (A260 / A280 > 1.8 and A260 / A230 > 1.0) (NanoDrop Technologies, Wilmington, DE, USA) and RNA integrity was verified using the 2200 TapeStation instrument (Agilent Technologies, CA, US). RNA sequencing was performed on 500ng input RNA. Library preparation was made using the TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Human/Mouse/Rat Set A (Illumina) and sequencing was performed on a NextSeq instrument using the NextSeq® 500/550 High Output Kit v2 (150 cycles) (Illumina). Transcript expression quantification was performed with Salmon v1.3.0. Differential gene expression analysis was carried out using DESeq2 v1.26.0.

HMEC-1 secretion experiments

HMEC-1 cells seeded at a confluence of ~60-70% were allowed to grow for 24h at 37°C and 5% CO₂. Cells were then washed in growth medium (GM) and new GM was added after which the cells were allowed to incubate for an additional 2-3h at 37°C and 5% CO₂. Cell medium was collected and centrifuged at 200g for 5 min, and the supernatant was frozen at -20°C. Cells were washed in PBS, detached using TrypLE (Thermo Fisher Scientific) and centrifuged at 200g for 5 min and the pellet was frozen at -20°C. Total DNA was extracted from the cells using the DNeasy Blood & Tissue Kit (Qiagen) and DNA concentration was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Secreted factors were measured in 40µl of supernatant using the Olink Target 96 immunoassay panels for protein biomarker discovery (Olink, Uppsala, Sweden). Panels used in the analysis were the Cardiometabolic, Oncology II, Cardiovascular III and the Immuno-Oncology panel where each panel offers the simultaneous detection of 92 biomarkers (for a full list see: https://www.olink.com/products/target/). Relative quantification of biomarkers is made generating normalized protein expression (NPX) values in arbitrary units. Results are then divided by cell DNA content to adjust for cell density.

Measurement of von Willebrand factor (VWF)

HMEC-1 cells were seeded on a culture flask (800k cells / 25cm²) and were allowed to grow for 24h at 37°C and 5% CO₂. Cells were washed with PBS and lysed in an ice-cold buffer (MCDB131 growth medium (without FBS), Triton-X100 (0.5% vol/vol) (MP Biomedicals (807426)), 1 mM EDTA (Sigma (E6758-100G))) and protease inhibitors (Pierce, Thermo Fisher Scientific (A32955)) for 20 min at 4°C on a shaking table. The buffer was collected and centrifuged at 1000g for 5 min, and the supernatant was subsequently analyzed for VWF. Plasma and intracellular levels of VWF were measured using a VWF Human ELISA Kit (Thermo Fisher Scientific (EHV/WF)).

Immunohistochemistry (IHC) and Immunofluorescence (IF) on frozen human muscle biopsies

5 µm thick sections of human skeletal muscle tissue was fixed in ice-cold acetone, rinsed in PBS and permeabilized in 0.5 % TritonX-100. For Sections for IHC, sections were blocked with 0.5 % hydrogen peroxide solution. Sections were blocked with 10 % normal goat serum or 10 % bovine serum albumin followed by incubation with 0.1 µg/ml for IHC and 0.5 µg/ml for IF rabbit anti-RAB3GAP2 antibody (HPA026273, Atlas antibodies, Bromma, Sweden), washed in PBS and incubated with biotinylated goat anti-rabbit IgG (BA-1000, Vector Laboratories, Burlingame, CA) diluted 1:1000 or with Cy-5 conjugated donkey anti-rabbit IgG, (711-117-152, Jackson Immuno Research, Cambridge, UK, cat. #) diluted 1:400. Non-immune rabbit IgG (ab27478, Abcam, Cambridge, UK) was used as negative control. For the IHC sections an ABC kit, (PK-6100, Vectastain Elite, Standard, Vector Laboratories, Burlingame, CA) and immPACT-DAB (SK-4105, Vector Laboratories) were used. After counter staining with Mayers hematoxylin, sections were mounted and imaged using an Aperio ScanScope CS scanner. Sections for IF were counterstained with Rhodamine labeled Ulex Europaeus Agglutinin I (UEA I), (RL-1062, Vector Laboratories) 10 µg/ml, washed, mounted, and imaged in a Zeiss LSM 5 Pascal laser scanning confocal microscope.
HUEVC or HMEC cells grown on coverslips were fixed in ice-cold acetone, rinsed in PBS, permeabilized in 0.5 % TritonX-100 and blocked with 10 % normal bovine serum followed by incubation with 0.5 µg/ml rabbit anti-RAB3GAP2 antibody or 0.5 µg/ml non-immune rabbit IgG followed by incubation with a Cy-5 conjugated donkey anti-rabbit IgG diluted 1:400. SYTOX Green nuclear stain, (S7020, Molecular Probes, Eugene, OR), diluted 1:3000, was used as counterstain. After rinsing with PBS the coverslips were mounted and imaged using a Zeiss LSM 5 laser scanning confocal microscope.

**IF on HUVEC (fig. 3 F)**

The HMEC cells with or without RAB3GAP2 knockdown were seeded on glass bottomed dishes and cultured until 70-80% confluence. The cells were fixed by 3% PFA and permeabilized with Perm buffer (BD). Primary anti-VWF antibody (ab6994, Abcam) was diluted (1:200) in blocking solution (5% donkey serum in PBS) and incubated with the cells for overnight at 4°C. The cells were then washed three times and incubated with cy2 conjugated secondary antibody (Jackson ImmunoResearch, Cat 711-225-152, 1:300) for 2 hours in room temperature. The images were acquired by Meta510 confocal system (Zeiss, Germany) with excitation wavelength 488 nm and emission filter between 500-530 nm. The quantification of VWF intensity performed by ZEN2012 software on the cells under same conditions.

**Immunoblotting**

The following antibodies were used for immunoblotting: rabbit anti-RAB3GAP2 (PA555296, ThermoFisher), mouse anti-α-tubulin (Ab7291, Abcam), rabbit anti-GAPDH (Ab181602, Abcam). Secondary HRP-conjugated antibodies (goat anti-rabbit, goat anti-mouse) were from Bio-Rad. For tissue lysate blots, a pre-blotted membrane with different human tissue lysates was purchased (INSTA-Blot NBP2-31378 from Novusbio). For analyses of CRISPR/Cas9 knockdown efficiency on RAB3GAP2 the cells were lysed in 2% SDS in PBS with protease inhibitor cocktail (Roche), passed through QiAamp DNA miniprep columns (Qiagen) to reduce the viscosity, and the protein content was quantified using BCA Protein Assay Kit (ThermoFisher). 20 mg of protein was run on a 4-20% Mini-Protean TGX SDS-PAGE gel (Bio-Rad) and electroblotted onto nitrocellulose membrane (Bio-Rad). From this point, for both human tissue lysate blots and the CRISPR knockdown analysis, the same blotting protocol was used. The membrane was rinsed in TBS, blocked in 5% bovine serum albumin in TBS, for 1 h at room temperature, then transferred to antibody incubation buffer (1% bovine serum albumin in TBS with 0.05% Tween-20) with anti-RAB3GAP2 diluted 1:2000, and incubated overnight at 4°C. Next, the membrane was washed 4 x 5 min in wash buffer (TBS with 0.05% Tween-20) and incubated with HRP-conjugated anti-rabbit diluted 1:10000 in antibody incubation buffer, for 1 h, at room temperature. The membrane was washed 4 x 5 min in wash buffer and developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (ThermoFisher) and a CCD camera (Bio-Rad). For the second blotting step the membrane was stripped in Restore Western Blot Stripping Buffer (ThermoFisher), then the above-described blotting protocol was used, but using mouse anti-tubulin (diluted 1:5000) or rabbit anti-GAPDH (diluted 1:10000) antibodies, and species-specific secondary HRP-conjugated antibodies.

**CRISPR/Cas9 genomic editing of HMEC cells**

Synthetic sgRNA containing the sequence ATCTCCAACCATGATCTTTA complementary to the protein-encoding sequence in RAB3GAP2 exon 3 (cut site at Leu-88) was obtained from Integrated DNA Technologies (USA). Electroporation of CRISPR/Cas9-sgRNA complex into HMEC cells: 135 pmol Cas9 protein (IDT) was combined with 150 pmol sgRNA in total volume of 5 µl and incubated for 15 min. Next, the ribonucleoprotein complex was combined with 1 million HMEC cells suspended in 100 µl nucleofection buffer kit V (Lonza) and electroporated using T-020 program on Amaxa nucleofector IIb (Lonza), and the cells were seeded in 6-well plates. To evaluate the CRISPR/Cas9 editing efficiency genomic DNA was extracted from the cells five days post-nucleofection using DNeasy Blood & Tissue Kit (Qiagen) and used to PCR amplify the genomic region surrounding the sgRNA-targeted locus. The primers used were: GGAGAAAGAGGAGAATGGAGAG and CCCACAGGAAGAAGGAAATA. PCR was performed using AmpliTaQ Gold 360 Master Mix (ThermoFisher) following manufacturer's protocol, but with annealing temperature of 58 °C. PCR amplicons were then purified using GeneJET PCR Purification Kit (ThermoFisher) and submitted for Sanger sequencing (Eurofins Genomics) using the primer AAACATCCTGGCCTCACAGTTG. Additionally, the CRISPR/Cas9 editing efficiency was evaluated using Alt-R Genome Editing Detection Kit (IDT) that employs T7 endonuclease to detect indel-caused heteroduplexes in re-hybridized PCR amplicons. The edited cells were then expanded, and RAB3GAP2 protein content was detected using immunoblotting. To detect possible off-targeting effects caused by CRISPR/Cas9 editing, we PCR-amplified five loci with the highest off-targeting potential, as assessed by IDT CRISPR-Cas9 guide RNA design checker. AmpliTaQ Gold 360 Master Mix (ThermoFisher) was used for the PCR, according to the manufacturer's protocol. The amplicons were purified using GeneJET PCR Purification Kit (ThermoFisher) and submitted for Sanger sequencing (Eurofins Genomics). No traces of Cas9-induced indels were detected in the sequencing files. The selected genomic loci (hg38) and the corresponding primers were: chr16:61669411 (TGAGGCAAGCCACCAAT and TGTATGGGAGCAGAGACTAGAA); chr6:66820388 (ATCAGGCAATCCTTCAAC and AGGCCCATATAGAAGAAGATA); chr13:107837835: (CATTCATACCACCTCCCACTATC and TGTATGGGAGCAGAGACTAGAA); chr14:40316144 (GTTCTGTGAAGGCTAAGAGAGG and CCTTCGGGCTTTGTTGTTTC); chr1:34200298 (ATGGCTCTGCTCTTCAAC and AGGCCCATATAGAAGAAGATA). PCR was performed using AmpliTaq Gold 360 Master Mix (ThermoFisher) following manufacturer's protocol, but with annealing temperature of 58 °C. PCR amplicons were then purified using GeneJET PCR Purification Kit (ThermoFisher) and submitted for Sanger sequencing (Eurofins Genomics) using the primer AAACTTCCTGGGCTCCAAGATTG. Additionally, nucleofection using DNeasy Blood & Tissue Kit (Qiagen) and used to PCR amplify the genomic region surrounding the sgRNA-targeted locus. The primers used were: GGAGAAAGAGGAGAATGGAGAG and CCCACAGGAAGAAGGAAATA. PCR was performed using AmpliTaq Gold 360 Master Mix (ThermoFisher) following manufacturer's protocol, but with annealing temperature of 58 °C. PCR amplicons were then purified using GeneJET PCR Purification Kit (ThermoFisher) and submitted for Sanger sequencing (Eurofins Genomics) using the primer AAACTTCCTGGGCTCCAAGATTG. Additionally, the CRISPR/Cas9 editing efficiency was evaluated using Alt-R Genome Editing Detection Kit (IDT) that employs T7 endonuclease to detect indel-caused heteroduplexes in re-hybridized PCR amplicons. The edited cells were then expanded, and RAB3GAP2 protein content was detected using immunoblotting. To detect possible off-targeting effects caused by CRISPR/Cas9 editing, we PCR-amplified five loci with the highest off-targeting potential, as assessed by IDT CRISPR-Cas9 guide RNA design checker. AmpliTaQ Gold 360 Master Mix (ThermoFisher) was used for the PCR, according to the manufacturer's protocol. The amplicons were purified using GeneJET PCR Purification Kit (ThermoFisher) and submitted for Sanger sequencing (Eurofins Genomics). No traces of Cas9-induced indels were detected in the sequencing files. The selected genomic loci (hg38) and the corresponding primers were: chr16:61669411 (TGAGGCAAGCCACCAAT and TGTATGGGAGCAGAGACTAGAA); chr13:107837835: (CATTCATACCCTCCACCATC and CACAGGGAACCTCCATCAA); chr2:220422562: (AGGGTACCCACCACTTTC and TGAATCCTCATACCCCTTCTC).

**Matrigel Tube-Forming Assay**

The Matrigel tube-forming assay is a specific and extensively used tool for studying angiogenesis in vitro, and a detailed experimental protocol is previously described[20]. Briefly, HMEC-1 cells, cultured to 80-90% confluence in growth medium, were seeded on 12-well plates precoated with Matrigel at a density of 50k cells / well, and were allowed to incubate for 24h at 37°C in 5% CO₂. Images of the tubular network formation (between 3-4 images / well) were captured using a brightfield microscope at regular time intervals during the 24 h period. Network formation was analyzed and quantified using the online automated image analysis platform myWim (Wimasis Image Analysis, Cordoba, Spain). The identity of the different images was not disclosed during the analysis.

**In vivo angiogenesis plug assay**
All animal experiments were approved by the local ethical committee for animal care in Lund (M167-15). Matrigel plugs were prepared according to Nowak-Sliwinska et al.\textsuperscript{51}. Harvested HMECs in PBS were mixed with ice cold, overnight thawed Matrigel (GFR, Phenol Red-free, BD Biosciences), at a final concentration of 5 mg/ml Matrigel, without additional supplements and kept on ice prior to injection. Eight to eighteen-week-old female NSG mice were anaesthetized with isoflurane, the groin was shaved and cleaned before injection of 100µl cold Matrigel solution containing 1×10^6 HMECs per plug into the groin area. After 30 seconds gelling time, the needle was removed, and mice were removed from anesthesia. Following 7 days, mice were anaesthetized with 2.5% Avertin (12.5 mg/kg body weight; Sigma-Aldrich), plugs were removed and fixed overnight before paraffin embedding. Sections were stained with the connective tissue stain, Trichrome (Abcam ab150686).

**Declarations**

**Acknowledgements**

We thank all participants of the different studies for their crucial contributions and professor Tom Carter at University of London for his kind contribution of VWF-mRFP plasmids and advice. This study was supported by the Swedish Research Council, through Strategic Research Area EXODIAB (Dnr 2009-1039), project grants (Dnr 2018-02635, 2018-02837 and 2018-03086) and Linnaeus grant (Dnr 349-2006-237), the Swedish Foundation for Strategic Research (Dnr IRC15-0067), the Swedish Heart-Lung Foundation (Dnr. 20190470), Crafoord foundation, Wallenberg foundation, Novo Nordisk foundation, Magn. Bergvall foundation, Pålsson foundation, Diabetes Wellness, Swedish Diabetes Research Foundation, Hjelt foundation, European Research Council (CoG-2015_681742_NASCENT) and Region Skåne. Tugce Karaderi is supported by the Novo Nordisk Foundation Data Science Investigator grant (NNF20OC0062294). Maria Sabater-Lleal is supported by a Miguel Servet contract from the ISCIII Spanish Health Institute (CP17/00142) and co-financed by the European Social Fund. We thank Olga Kotova and Esa Laurila (Lund University Diabetes Center) for technical assistance. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**References**

1. Bouchard, C. et al. Familial aggregation of VO2(max) response to exercise training: results from the HERITAGE Family Study. *J Appl Physiol* (1985) **87**, 1003-1008 (1999).
2. Bouchard, C. & Rankinen, T. Individual differences in response to regular physical activity. *Med Sci Sports Exerc* **33**, S446-451; discussion S452-443 (2001).
3. Mittleman, M. A. et al. Triggering of acute myocardial infarction by heavy physical exertion. Protection against triggering by regular exertion. Determinants of Myocardial Infarction Onset Study Investigators. *N Engl J Med* **329**, 1677-1683, doi:10.1056/NEJM199312023292301 (1993).
4. Willich, S. N. et al. Physical exertion as a trigger of acute myocardial infarction. Triggers and Mechanisms of Myocardial Infarction Study Group. *N Engl J Med* **329**, 1684-1690, doi:10.1056/NEJM199312023292302 (1993).
5. Maron, B. J. The paradox of exercise. *N Engl J Med* **343**, 1409-1411, doi:10.1056/NEJM200011093431911 (2000).
6. Franklin, B. A. et al. Exercise-Related Acute Cardiovascular Events and Potential Deleterious Adaptations Following Long-Term Exercise Training: Placing the Risks Into Perspective-An Update: A Scientific Statement From the American Heart Association. *Circulation* **141**, e705-e736, doi:10.1161/CIR.0000000000000749 (2020).
7. Nelson, M. R. et al. The support of human genetic evidence for approved drug indications. *Nat Genet* **47**, 856-860, doi:10.1038/ng.3314 (2015).
8. Komi, P. V. et al. Skeletal muscle fibres and muscle enzyme activities in monozygous and dizygous twins of both sexes. *Acta Physiol Scand* **100**, 385-392 (1977).
9. Simoneau, J. A. & Bouchard, C. Genetic determinism of fiber type proportion in human skeletal muscle. *FASEB J* **9**, 1091-1095 (1995).
10. Costill, D. L., Fink, W. J. & Pollock, M. L. Muscle fiber composition and enzyme activities of elite distance runners. *Med Sci Sports* **8**, 96-100 (1976).
11. Harber, M. & Trappe, S. Single muscle fiber contractile properties of young competitive distance runners. *J Appl Physiol* (1985) **105**, 629-636, doi:10.1152/japplphysiol.00995.2007 (2008).
12. Widrick, J. J., Trappe, S. W., Costill, D. L. & Fitts, R. H. Force-velocity and force-power properties of single muscle fibers from elite master runners and sedentary men. *Am J Physiol* **271**, C676-683, doi:10.1152/ajpheart.1996.271.2.C676 (1996).
13. Saltin, B., Henriksen, J., Nygaard, E., Andersen, P. & Jansson, E. Fiber types and metabolic potentials of skeletal muscles in sedentary man and endurance runners. *Ann N Y Acad Sci* **301**, 3-29 (1977).
14. Pillon, N. J. et al. Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity. *Nat Commun* **11**, 470, doi:10.1038/s41467-019-13869-w (2020).
15. Rosenbloom, K. R. et al. ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res* **41**, D56-63, doi:10.1093/nar/gks1172 (2013).
16 Sakane, A. et al. Rab3 GTPase-activating protein regulates synaptic transmission and plasticity through the inactivation of Rab3. *Proc Natl Acad Sci U S A* **103**, 10029-10034, doi:10.1073/pnas.060304103 (2006).

17 Nagano, F. et al. Molecular cloning and characterization of the noncatalytic subunit of the Rab3 subfamily-specific GTPase-activating protein. *J Biol Chem* **273**, 24781-24785 (1998).

18 Knop, M., Aareskjold, E., Bode, G. & Gerke, V. Rab3D and annexin A2 play a role in regulated secretion of vWF, but not tPA, from endothelial cells. *EMBO J* **23**, 2982-2992, doi:10.1038/sj.emboj.7600319 (2004).

19 Zografou, S. et al. A complete Rab screening reveals novel insights in Weibel-Palade body exocytosis. *J Cell Sci* **125**, 4780-4790, doi:10.1242/jcs.104174 (2012).

20 McCormack, J. J., Lopes da Silva, M., Ferraro, F., Patella, F. & Cutler, D. F. Weibel-Palade bodies at a glance. *J Cell Sci* **130**, 3611-3617, doi:10.1242/jcs.208033 (2017).

21 Sapp, R. M. et al. The effects of moderate and high-intensity exercise on circulating markers of endothelial integrity and activation in young, healthy men. *J Appl Physiol (1985)* **127**, 1245-1256, doi:10.1152/japplphysiol.00477.2019 (2019).

22 van Loon, J. E., Sonneveld, M. A., Praet, S. F., de Maat, M. P. & Leebeek, F. W. Performance related factors are the main determinants of the von Willebrand factor response to exhaustive physical exercise. *PLoS One* **9**, e91687, doi:10.1371/journal.pone.0091687 (2014).

23 Gill, J. C., Endres-Brooks, J., Bauer, P. J., Marks, W. J. & Montgomery, R. R. The effect of ABO blood group on the diagnosis of von Willebrand disease. *Blood* **69**, 1691-1695 (1987).

24 Sabater-Lleal, M. et al. Genome-Wide Association Transethnic Meta-Analyses Identifies Novel Associations Regulating Coagulation Factor VIII and von Willebrand Factor Plasma Levels. *Circulation* **139**, 620-635, doi:10.1161/CIRCULATIONAHA.118.034532 (2019).

25 Abdel-Hamid, M. S. et al. Micro and Martsolf syndromes in 34 new patients: Refining the phenotypic spectrum and further molecular insights. *Clin Genet* **98**, 445-456, doi:10.1111/cge.13825 (2020).

26 Chen, J. & Chung, D. W. Inflammation, von Willebrand factor, and ADAMTS13. *Blood* **132**, 141-147, doi:10.1182/blood-2018-02-769000 (2018).

27 Rietveld, I. M. et al. High levels of coagulation factors and venous thrombosis risk: strongest association for factor VIII and von Willebrand factor. *J Thromb Haemost* **17**, 99-109, doi:10.1111/jth.14343 (2019).

28 Miyamoto-Mikami, E. et al. Gene expression profile of muscle adaptation to high-intensity intermittent exercise training in young men. *Sci Rep* **8**, 16811, doi:10.1038/s41598-018-35115-x (2018).

29 Hedstrand, H. A study of middle-aged men with particular reference to risk factors for cardiovascular disease. *Ups J Med Sci Suppl* **19**, 1-61 (1975).

30 Mootha, V. K. et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273 (2003).

31 Elgzyri, T. et al. First-Degree Relatives of Type 2 Diabetic Patients Have Reduced Expression of Genes Involved in Fatty Acid Metabolism in Skeletal Muscle. *Journal of Clinical Endocrinology & Metabolism* **97**, E1332-E1337, doi:10.1210/jc.2011-3037 (2012).

32 Ekman, C. et al. Less pronounced response to exercise in healthy relatives to type 2 diabetic subjects compared with controls. *Journal of Applied Physiology* **119**, 953-960, doi:10.1152/japplphysiol.01067.2014 (2015).

33 Bar-Oz, O. The Wingate anaerobic test. An update on methodology, reliability and validity. *Sports Med* **4**, 381-394, doi:10.2165/00007256-19870406-00001 (1987).

34 McGawley, K. et al. No Additional Benefits of Block-Over Evenly-Distributed High-Intensity Interval Training within a Polarized Microcycle. *Front Physiol* **8**, 413, doi:10.3389/fphys.2017.00413 (2017).

35 Rankinen, T. et al. No Evidence of a Common DNA Variant Profile Specific to World Class Endurance Athletes. *PLoS One* **11**, e0147330, doi:10.1371/journal.pone.0147330 (2016).

36 Qu, Z., Andersen, J. L. & Zhou, S. Visualisation of capillaries in human skeletal muscle. *Histochem Cell Biol* **107**, 169-174 (1997).

37 Brooke, M. H. & Kaiser, K. K. Three "myosin adenosine triphosphatase" systems: the nature of their pH lability and sulfhydryl dependence. *J Histochem Cytochem* **18**, 670-672 (1970).

38 Keldso, S. et al. Expression of Phosphofructokinase in Skeletal Muscle Is Influenced by Genetic Variation and Associated With Insulin Sensitivity. *Diabetes* **63**, 1154-1165 (2014).

39 Delaneau, O. & Zagury, J. F. Haplotype inference. *Methods Mol Biol* **888**, 177-196, doi:10.1007/978-1-61779-870-2_11 (2012).
Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529, doi:10.1371/journal.pgen.1000529 (2009).

Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-913, doi:10.1038/ng2088 (2007).

Mägi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288, doi:10.1186/1471-2105-11-288 (2010).

Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. 2010 **36**, 48, doi:10.18637/jss.v036.i03 (2010).

Pruim, R. J. et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-2337, doi:10.1093/bioinformatics/btq419 (2010).

Scott, R. A. et al. ACTN3 and ACE genotypes in elite Jamaican and US sprinters. *Med Sci Sports Exerc* **42**, 107-112, doi:10.1249/MSS.0b013e3181ae2bc0 (2010).

Shabalin, A. A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353-1358, doi:10.1093/bioinformatics/bts163 (2012).

Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, RESEARCH0034 (2002).

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419, doi:10.1038/nmeth.4197 (2017).

Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

Brown, R. M., Meah, C. J., Heath, V. L., Styles, I. B. & Bicknell, R. Tube-Forming Assays. *Methods Mol Biol* **1430**, 149-157, doi:10.1007/978-1-4939-3628-1_9 (2016).

Nowak-Sliwinska, P. et al. Consensus guidelines for the use and interpretation of angiogenesis assays. *Angiogenesis* **21**, 425-532, doi:10.1007/s10456-018-9613-x (2018).

### Table 1

| Phenotype                      | SNP               | Reference allele | Non-reference allele | Chr | Genome position (build 37) | Effect estimate (M) | SE (M) | Nearest Gene | Direction of effect (ULSAM, MM, MEI) | MAF | N         | p-value          |
|-------------------------------|-------------------|------------------|----------------------|-----|---------------------------|--------------------|--------|--------------|---------------------------------------|-----|-----------|-----------------|
| Type IIX (%)                  | rs145631867       | A                | G                    | 2   | 226410525                 | -1.97              | 0.35   | NYAP2        | - x -                                 | 0.02| 525       | 1.94x10^-8      |
| Type IIX (%)                  | rs7776803         | T                | C                    | 7   | 89526223                  | -0.40              | 0.07   | STEAP2-AS1   | - - -                                 | 0.27| 651       | 2.25x10^-8      |
| Type IIX (%)                  | rs149081100       | T                | C                    | 5   | 159364127                 | 1.68               | 0.30   | ADRA1B       | + + x                                 | 0.01| 608       | 3.03x10^-8      |
| capillary-to-fiber ratio (nr/fiber) | rs115660502     | G                | A                    | 1   | 220291773                 | 0.77               | 0.14   | IARS2        | + + x                                 | 0.05| 603       | 2.73x10^-8      |
| capillary-to-fiber ratio (nr/fiber) | rs191465330     | A                | G                    | 1   | 220101936                 | 1.92               | 0.35   | SLC30A10     | + + x                                 | 0.01| 603       | 2.90x10^-8      |

"-" and "x" indicates direction of effect for the non-reference allele, "x" = either phenotype or genotype not measured in the dataset

**Figures**

Page 12/16
(a) Manhattan plot showing the (−log10) p-values by genomic position for capillary-to-fiber ratio and (b) QQ-plot of associations with capillary-to-fiber ratio from the meta-analysis in the Malmö Men and Uppsala Longitudinal Study of Adult Men cohorts (ntot = 603). The blue and red lines in a indicate the thresholds for suggestive (p < 1 x 10^{-6}) and genome-wide significance (p < 5 x 10^{-8}), respectively. (c) LocusZoom plot of the region on chromosome 1 harbouring genetic variants associated with capillary-to-fiber ratio. The purple circle represents the genetic variant (rs115660502) with the lowest p-value (left y-axis) in the region. Other variants (circles) are colored according to their linkage disequilibrium with this variant, i.e. from highest (red) to lowest (dark blue). Estimated recombination rates are shown on the right y-axis. (d) Distribution of capillary-to-fiber ratio in the ULSAM and Malmö Men cohorts.
Figure 2

(a) Expression of RAB3GAP2 mRNA in skeletal muscle after 6-weeks high-intensity intermittent training. \( p_{\text{Wilcoxon}} = 0.002, n = 11 \). (b) Expression of RAB3GAP2 mRNA across rs115660502 genotypes in the MM (n = 33 vs 5) and MEI studies (n = 36 vs 3), \( p_{\text{Meta-FDR}} = 0.007 \). (c) Plot of RAB3GAP2 mRNA expression versus capillary-to-fiber ratio in the Malmö Men cohort. \( r = 0.38, p_{\text{Spearman}} = 0.03, n = 32 \). (d) Lower and (e) higher magnification of RAB3GAP2 protein localisation (brown) in human skeletal muscle. (f) Negative control using non-immune IgG with same concentration as in (e). Nuclei stained with hematoxylin (blue) in d-f. (g) Lower and (h) higher magnification of confocal immunofluorescence images of RAB3GAP2 (red) and the endothelial marker lectin (green) of human skeletal muscle, demonstrating RAB3GAP2 localization to the endothelium (green). (i) Confocal immunofluorescence of human skeletal muscle stained for RAB3GAP2 (red) and lectin (green), demonstrating RAB3GAP2 localization to capillaries, but not to large vessels. ** \( p < 0.01 \)
Figure 3

(a - d) Tube formation assay using HMECs with or without CRISPR-Cas9-mediated knockdown of RAB3GAP2. n = 11 (wildtype) and 13 (RAB3GAP2 KD) independent experiments at 6 and 24 hours. (a) Number of loops, pMann-Whitney, 6 h = 8.0 x 10^-7 and pMann-Whitney, 24 h = 7.1 x 10^-5. (b) Number of branching points, pMann-Whitney, 6 h = 8.8 x 10^-6 and pMann-Whitney, 24 h = 7.6 x 10^-4. (c) Total tube length, pMann-Whitney, 6 h = 8.0 x 10^-7 and pMann-Whitney, 24 h = 2.4 x 10^-5. (d) Representative images of tube formation in HMECs. (e) In vivo angiogenesis plug assay in NGS mice using HMECs with or without RAB3GAP2 knockdown. Percentage of Masson's Trichrome staining of total Matrigel area (pMann-Whitney = 0.036, n = 3-5). (f) Representative images of Masson's Trichrome staining. * p < 0.05, *** p < 0.001
Figure 4

(a - h) Amount of protein in the cell medium after culturing of HMECs using Olink panels with or without CRISPR-Cas9-mediated knockdown of RAB3GAP2. (a) CD70, (b) TNC, (c) TNXB, (d) COL1A1, (e) TFPI2, (f) IGFBP3, (g) PLAT and (h) CCL2, n = 6. Inserts show corresponding RNA expression by RNA sequencing in HMECs from independent experiments, n = 8. (i) Intracellular amount of VWF after RAB3GAP2 knockdown in HMECs. Characterized as fluorescence intensity per cell (left, n = 3 with 20-24 cells per experiment) and by ELISA (right, pWilcoxon = 7.8 x 10^{-3}, n = 8 independent experiments). (j) Representative images of VWF fluorescence. (k) Circulating levels of VWF before and after a 30s Wingate cycling test (pWilcoxon = 7.3 x 10^{-12}, n = 39).

Cluster of differentiation 70 (CD70), TNC (Tenascin C), TNXB (Tenascin XB), COL1A1 (Collagen, type I, alpha 1), TFPI2 (Tissue factor pathway inhibitor 2), IGFBP3 (Insulin like growth factor binding protein 3), PLAT (Plasminogen activator, tissue type, or tPA), CCL2 (C-C motif chemokine ligand 2, or MCP-1), VWF (von Willebrand factor). * p < 0.05, ** p < 0.01, *** p < 0.001

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.doc
- SupplementTableS5.xlsx
- SupplementaryFigures.doc