The Athlome Project Consortium: A Concerted Effort to Discover Genomic and other "OMIC" Markers of Athletic Performance.

Authors: Yannis P. Pitsiladis[1*], Masashi Tanaka[2], Nir Eynon[3], Claude Bouchard[4], Kathryn N. North[5], Alun G. Williams[6]; Malcolm Collins[7]; Colin N. Moran[8], Steven L. Britton[9], Noriyuki Fuku[10], Euan A. Ashley[11], Vassilis Klissouras[12], Alejandro Lucia[13], Ildus I. Ahmetov[14], Eco de Geus[15], Mohammed Alsayrafi[16]

Affiliations: [1]FIMS Reference Collaborating Centre of Sports Medicine for Anti-Doping Research, University of Brighton, Eastbourne, United Kingdom; [2]Department of Longevity and Health, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan; [3]Institute of Sport, Exercise, and Active Living (ISEAL), Victoria University, Melbourne, Australia; [4]Human Genomics Laboratory, Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, United States; [5]Murdoch Childrens Research Institute, Royal Children's Hospital, Melbourne, Australia; [6]Department of Exercise and Sport Science, Manchester Metropolitan University, Crewe, United Kingdom; [7]Division of Exercise Science and Sports Medicine, Department of Human Biology, University of Cape Town, Cape Town, South African; [8]Health and Exercise Sciences Research Group, University of Stirling, Stirling, Scotland; [9]Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, United States; [10]Graduate School of Health and Sports Science, Juntendo University, Chiba, Japan; [11]Stanford University Medical Center, Stanford, United States; [12]Ergophysiology Research Laboratory, Department of Sport Medicine and Biology of Physical Activity, University of Athens, Athens, Greece; [13]School of Doctorate Studies & Research, Universidad Europea de Madrid, Madrid, Spain; [14] Sport Technology Research Centre, Volga Region State Academy of Physical Culture, Sport and Tourism, Kazan, Russia; [15]VU University and VU Medical Centre, Amsterdam, Netherlands; [16]Anti-Doping Lab Qatar (ADLQ), Doha, Qatar.

***Corresponding Author:** Professor Yannis P. Pitsiladis
Professor of Sport and Exercise Science
FIMS Reference Collaborating Centre of Sports Medicine for Anti-Doping Research
University of Brighton
Eastbourne, BN20 7SN, United Kingdom
Email: y.pitsiladis@brighton.ac.uk
Telephone: +44 (0) 1273 643707 (Reception)
Telephone: +44 (0) 1273 643612 (Office)
Fax: +44 (0) 1273 643704

Running Head: The Athlome Project Consortium

Key Words: Genetics, Performance

The Athlome Project Consortium: A Concerted Effort to Discover Genomic and other "OMIC" Markers of Athletic Performance.

The Athlome Project Consortium*

Despite numerous attempts to discover genetic variants associated with elite athletic performance, injury predisposition and elite/world-class athletic status, there has been limited progress to date. Past reliance on candidate gene studies predominantly focusing on genotyping a limited number of single nucleotide polymorphisms (SNPs) or the insertion/deletion variants in small, often heterogeneous cohorts (i.e., made up of athletes of quite different sport specialties) have not generated the kind of results that could offer solid opportunities to bridge the gap between basic research in exercise sciences and deliverables in biomedicine. A retrospective view of genetic association studies with complex disease traits indicates that transition to hypothesis-free genome-wide approaches will be more fruitful. In studies of complex disease, it is well recognized that the magnitude of genetic association is often smaller than initially anticipated and, as such, large sample sizes are required to identify the gene effects robustly. Thus, alternative large-scale, collaborative efforts involving well-phenotyped male and female cohorts from which high-resolution genome-wide data is generated and interrogated using advanced bioinformatics approaches are necessary for meaningful progress to be made. Accordingly, a symposium was held in Athens and on the Greek island of Santorini from 14-17th May 2015 (http://celebratorysymposium.net) to review the main findings in exercise genetics and genomics and to explore promising trends and possibilities. The symposium also offered a forum for the development of a position stand (the Santorini Declaration). Among the participants, many were involved in ongoing collaborative studies (e.g., GAMES, Gene SMART, GENESIS and POWERGENE). A consensus emerged among participants that it would be advantageous to bring together all current studies and those recently launched into one new large collaborative initiative, which was subsequently named the *Athlome Project Consortium*.

At the outset, the Athlome Project aims to collectively study the genotype and phenotype data currently available on elite athletes, in adaptation to exercise training (in both human and animal models) and on exercise-related musculoskeletal injuries from individual studies and from consortia worldwide. To achieve this, several steps are set out:

1. To establish an ethically sound international research consortium (Athlome Project Consortium) and biobank resource systematically across individual centres;
2. To discover genetic variants associated with exercise performance, adaptive response to exercise-training, and skeletal-muscle injuries using the genome-wide association study (GWAS) approach, targeted sequencing or whole genome sequencing, where possible;
3. To validate and replicate the genetic markers from the discovery phase across sex and ethnicity; and
4. To conduct functional investigations following replicated findings (e.g., study the replicated SNPs and their linkage disequilibrium regions, *in vitro* expression studies and knockouts of nearby genes) to better understand the associated biology.

During the development of the initial phase of the Athlome Project in determining the genetic variations related to elite athletic performance and injury predisposition, epigenomic, transcriptomic and proteomic analyses need also be carefully planned to strengthen the understanding of gene functions. Linking these findings with metabolic profiling (the end products of the cellular processes) is also a future aspiration of the Athlome Project. Another challenge is to be able to efficiently integrate the multiple "omics" datasets generated from the different approaches. The ultimate goal of the Athlome Project Consortium is to generate the ethically sound environment, interest and capacity needed to develop the specialist knowledge to inform personalized training and injury prevention, as well as doping detection. The following individual or collaborative studies have agreed to work together in the global partnership that constitutes the Athlome Project Consortium. The participating cohorts and the focus of each are depicted in Figure 1.

Eastern Europe population studies (The Russian and Belarusian cohorts, GELAK, GELAV, and GUAP)

The Russian and Belarusian cohorts, the Genetics and Epigenetics of Lithuanian Athletes from Kaunas (GELAK) and Vilnius (GELAV), and the Genome of Ukrainian Athletes Project (GUAP) have consolidated to identify genetic and epigenetic variations associated with high-level sports performance. The cohort comprises East Europeans (from Belarus, Lithuania,

Russia, and Ukraine; in total n = 8,228 athletes and n = 4,121 controls). The athletes are grouped into international (including participants in Olympics and World Championships), national, regional, or local/non-competitive categories. These include biathletes, distance runners, cyclists, triathletes, kayakers, rowers, canoers, modern pentathletes, orienteers, skiers, speed skaters, short-trackers, walkers, weightlifters, bodybuilders, powerlifters, strongmen, sprint runners (≤ 400 m), sprint swimmers (50 - 100 m), decathletes, heptathletes, combat athletes, field athletes, bobsleigh athletes, rhythmic and artistic gymnasts, figure skaters, fencers and team ball-sport players. A portion of the participants have been evaluated with a variety of quantitative performance- and health-related assessments, including strength/power-related measurements, agility/speed-related measurements, balance, flexibility and coordination measurements, endurance-related measurement, skeletal muscle biopsy, and health-related measurements.

Principal Investigators: Ildus I Ahmetov (Volga Region State Academy of Physical Culture, Sport and Tourism, RUS), Svitlana B Drozdovska (National University of Physical Education of Ukraine, UKR), Colin N Moran (University of Stirling, GBR), Valentina Ginevičienė (Vilnius University, LTU), Andrei A Gilep (Institute of Bioorganic Chemistry NASB, BLR).

ELITE *elite.stanford.edu*

The Exercise at the Limit – Inherited Traits of Endurance (ELITE) consortium is a global initiative with the main objective to map the role that genetics plays in athletic ability versus environmental factors, such as training. Study participant (n > 500) selection is based on a physiological variable relevant for both health and sport performance, i.e., maximum oxygen uptake ($\dot{V}O_2max$). The main inclusion criterion is $\dot{V}O_2max$ > 75 ml/kg/min for men and > 63 ml/kg/min for women, respectively. The consortium is continuously expanding and is recruiting athletes from all over the globe (with main focus on Caucasians, North East Africans, East Asians and South Americans) who are successful in endurance sports (running, cycling, cross country skiing, triathlon, and rowing). Analyses currently include enhanced whole exome sequencing and GWAS (1.7 million SNPs). The combination of analytic methods will enable findings and differentiation between common variants with small effects and novel rare variants with larger effects. The aim is also to investigate gender and ethnic differences.

Principal Investigators: Euan A Ashley, C Mikael Mattsson, Matthew Wheeler, Daryl Waggott (Stanford University, USA).

Elite East African athlete cohort

The consortium also aims to study the East African running success by analyzing data from previously recruited subjects: (i) 76 endurance runners (64 men) and 38 sprint and power event athletes (18 men) from the Ethiopian national athletics teams, 315 controls from the general Ethiopian population (281 men), 93 controls from the *Arsi* region of Ethiopia (80 men) and (ii) 291 elite Kenyan endurance athletes (232 men) and 85 control participants (40 men). Seventy (59 men) Kenyan athletes had competed internationally and achieved outstanding success.

Principal Investigators: Yannis Pitsiladis (University of Brighton, GBR), Robert Scott (University of Cambridge, GBR).

GAMES

An international consortium (GAMES) was established to compare allele frequencies between elite endurance athletes and ethnicity-matched controls. GWASs were undertaken on two cohorts of elite endurance athletes (GENATHLETE and Japanese endurance runners) and their respective controls, from which a panel of 45 candidate SNPs was identified. These markers were tested for replication in seven additional cohorts of endurance athletes and controls from Australia, Ethiopia, Japan, Kenya, Poland, Russia and Spain. The study is based on a total of 1,520 endurance athletes (835 of them had competed in World Championships or Olympic Games) and 2,760 controls.

Principal Investigators: Claude Bouchard, Tuomo Rankinen (Pennington Biomedical Research Centre, Louisiana State University, USA), Noriyuki Fuku (Juntendo University, JPN), Yannis Pitsiladis (University of Brighton, GBR), Bernd Wolfarth (Humboldt University, DEU), Alejandro Lucia (Universidad Europea de Madrid, SP).

GENATHLETE

The study was launched in 1993 with the aim of identifying DNA variants that are present at

different frequencies between elite endurance athletes and sedentary controls. Male endurance athletes and controls were recruited from Canada, Finland, Germany and the USA. The cohort assembled to date includes 315 elite endurance athletes and 320 matched controls. Selection criteria for the all-male endurance athlete sample include that they had to be athletes of national or international caliber with a $\dot{V}O_2$max of at least 75 ml/kg/min. The mean value for the 315 athletes is currently 79 ml/kg/min while the mean for the 320 control subjects reached 40 ml/kg/min. Multiple candidate genes have been studied using the resources of GENATHLETE. A genome-wide screen for common variants has been performed on GENATHLETE (see GAMES cohort above) and further studies are focusing on nuclear and mitochondrial DNA sequencing.

Principal Investigators: Claude Bouchard, Tuomo Rankinen (Pennington Biomedical Research Centre, Louisiana State University System, USA), Bernd Wolfarth (Department of Sports Medicine, Charite Medical School, Berlin, Germany), Louis Perusse (Laval University, Quebec, Canada), Rainer Rauramaa (University of Eastern Finland, Kuopio, Finland).

GENESIS

The GENetics of Elite Status In Sport (GENESIS) consortium aims to identify molecular genetic characteristics associated with successful sports performance. The cohort (current n > 1,200) is mainly composed of UK athletes. Sports include marathon running and other track-and-field athletics, cycling and team sports (e.g. soccer). The RugbyGene Study is a major subcomponent of GENESIS and focuses on rugby (both union and league codes). Objectives of GENESIS are: (i) to increase current cohort size substantially; (ii) to apply hypothesis-free approaches to identify molecular genomic markers; (iii) to expand GENESIS from genomics to other "omics"; and (iv) to combine the "omics" data with athlete health and performance data to maximize practical impact of GENESIS.

Principal Investigators: Alun G Williams, Stephen H Day, Georgina K Stebbings (Manchester Metropolitan University, GBR), Robert M Erskine (Liverpool John Moores University, GBR), Hugh E Montgomery (University College London, GBR).

Gene SMART Study *www.vu.edu.au/speed-gene*

The Gene SMART (Skeletal Muscle Adaptive Response to Training) study aims to identify the gene variants that predict the skeletal muscle response to both a single bout and 4 weeks of High-Intensity Interval Training (HIIT) in three different training centres. While the lead training and testing centre is located in Victoria University, Melbourne, two other centres have been launched at Bond University, Australia and the University of Sao Paulo, Brazil. A fourth centre (University of Brighton, UK) will focus on the omics analyses. The cohort is comprised of moderately-trained, healthy male participants (aged 20-45 years, body mass index $\leq$ 30 kg/m^2). Participants are undergoing similar exercise testing and exercise training in three different laboratories. Dietary habits are assessed by questionnaire and nutritionist consultation. Activity history is assessed by questionnaire and current activity level is assessed by activity monitoring. A number of muscle and blood analyses are to be performed, including genotyping, mitochondrial respiration, transcriptomics, proteomics, and enzymes activity before, during and after training, where appropriate. Currently ~40 participants have finished the study and the aim is to train a total of 250 participants. The Gene SMART also includes baseline and post-training testing and sampling for all participants.

Principal Investigators: David Bishop, Nir Eynon (Victoria University, AUS).

GOINg

The recently established Genomics Of INjuries (GOINg) consortium aims to identify DNA variants that modify the risk of anterior cruciate ligament (ACL) injuries. It is the only consortium within the Athlome Project to specifically investigate exercise-associated musculoskeletal injuries. The plan is to screen current known loci for ACL injury susceptibility in larger data sets in an attempt to determine if they remain as susceptibility loci across all populations using the hypothesis-driven candidate gene case-control study design. Care will be taken to use the same criteria to accurately phenotype, with respect to ancestry, sporting and occupational details, injury profile and mechanism(s) of injury, other injury history and family history, as well as, other appropriate medical history and medication use. The actual functional significance of the identified variants will also be investigated. This initial phase will be followed by sequencing and the research objectives will be eventually expanded to include other

231 "omics". Thus far, ACL rupture consortium has collected DNA samples and clinical, as well as
232 physical and occupational activity information from subjects from South Africa, Poland,
233 Australia, Russia and Italy.
234
235 *Principal Investigators:* Malcolm Collins, Alison September, Michael Posthumus (University of
236 Cape Town, ZAF), Nir Eynon (Victoria University, AUS), Pawel Cieszczyk (University of
237 Szczecin, POL).
238
239 **J-HAP**
240 The Japanese Human Athlome Project (J-HAP) focuses on the study of genes associated with
241 physical performance and its related phenotypes (e.g., muscle mass, muscle fiber type, $\dot{V}O_2$max).
242 The cohort is comprised of Japanese athletes (currently > 2,400, mainly international and
243 national levels) and healthy Japanese controls (currently > 1,000). These athletes are mainly
244 track-and-field athletes and swimmers competing in endurance- and sprint/power-oriented events.
245 Multiple "omics" approaches will be used to determine genes in talent identification in the
246 Japanese population. Among the collected Japanese athletes' and controls' samples,
247 approximately 200 muscle biopsies were obtained from both athletes and controls in order to
248 investigate genetic variants associated with muscle fibre type distribution.
249
250 *Primary Investigators*: Noriyuki Fuku (Juntendo University, JPN), Naoki Kikuchi (Nippon Sport
251 Science University, JPN), Eri Miyamoto-Mikami (The National Institute of Fitness and Sports in
252 Kanoya, JPN).
253
254 **NTR**
255 The Netherlands Twin Register (NTR) is a population-based cohort recruiting both newborn and
256 adult multiples and their family members with continuous longitudinal data collection. In the
257 past 25+ years, around 40% of all twins and multiples in the Netherlands have taken part in the
258 NTR research projects. Family members and spouses of twins also took part, leading to a total of
259 over 185,000 participants across multiple research projects. The longitudinal information that has
260 been collected extends from genotype to biomarkers, gene expression to rich behavioral
261 information including biennial reports on (competitive) sports participation and performance

level and on injuries related to sports. In its sports research track, NTR aims to understand the interplay between genetic and environmental factors shaping individual differences in sports participation and performance. In the NTR, participants are recruited as newborns and followed into young adulthood, 520 have played competitively at a regional and 189 at a national level. Main sports that Dutch adolescents/young adults engage in are swimming, tennis, bicycling, soccer and field hockey. The longitudinal data collection of the NTR is ongoing and securely funded for the next 5 years.

Principal Investigators: Eco de Geus, Meike Bartels (VU University and VU medical centre, NLD).

POWERGENE

The POWERGENE consortium aims to characterise the elite sprint/power athlete genotype. The internationally competitive (Olympic/Word championship qualifiers) sprint/power athletes are from: Australia, Belgium, Greece, Italy, Jamaica, Japan, Lithuania, Poland, Spain, the U.S.A., Brazil, and Russia. They will be compared with sub-elite athletes (national qualifiers), endurance athletes, team athletes and controls. The current cohort consists of female (n = 264) and male (n = 481) specialist power athletes across three major ethnicities (i.e., European, West African and East Asian ancestries). Sprint/power athletes include those individuals competing in track (≤ 800 m) and field (jump, throw) events, cycling (track), swimming (≤ 200 m), gymnastics (artistic), weightlifting, judo, speed-skating and power lifting. Endurance athletes (n = 586) include track and road running specialists (> 800 m), rowers, cyclists, swimmers (> 200 m), triathletes and ironmen. Team sports (n = 862) include football (soccer), cricket, hockey, volleyball and basketball.

Principal Investigators: Yannis Pitsiladis (University of Brighton, GBR), Kathryn North (Murdoch Childrens Research Institute, AUS), Nir Eynon (Victoria University, AUS).

Super-athletes: Genes and Sweat

The study aims to (i) identify genetic variants associated with elite athletic performance, (ii) study potential ethnic differences, and (iii) study the functional significance of the identified

variants. A GWAS will be carried out in 3,000 consented elite athletes, tested negative for doping substances at the Anti-Doping Laboratories, Federazione Medico Sportiva Italiana (FMSI) and Anti-Doping Lab Qatar (ADLQ), using Illumina genotyping technologies. Examining genotype frequency distribution of elite athletes from European countries (where most of FMSI samples will be obtained) against those from South Asian and African countries (where most of ADLQ samples are expected to be obtained) would help to identify potential ethnic differences in the genetic predisposition to athletic performance. Subsequently, urine metabolome in a subset of these athletes (1,000 subjects) will be performed, and will be related to the athlete's sporting discipline.

Principal Investigators: Mohamed El-Rayess**,** Costas Georgakopoulos, Mohammed Alsayrafi (ADLQ, QAT), Francesco Botre (FMSI, ITA), Karsten Suhre (Weill Cornell Medical College in Qatar, QAT), Mike Hubank (University College London, GBR).

Epigenetics of Elite Athletic Performance

It is clear from animal and human studies that epigenetic marks play a role in the modulation of gene expression in relevant tissues. There also are indications that epigenetic marks can be altered by acute and chronic exercise in skeletal muscle and adipose tissue where they have been studied. Thus individual differences in any exercise-related traits can potentially be explained by not only the impact of DNA sequence variation on biology and behavior but also by the effects of epigenomic signaling on gene expression. We are formulating the hypothesis that elite athletic performance is influenced by epigenomic alterations, facilitating morphological, physiological, metabolic, cognitive, emotional and behavioral changes that empower the athlete to push performance beyond existing boundaries. We envisage testing this hypothesis by recruiting twin athletes competing at the Olympic or World Championship levels.

Principal Investigators: Vassilis Klissouras (University of Athens, GRC), Yannis Pitsiladis (University of Brighton, GBR).

Rat models of exercise and health (LCR-HCR rat model)

The purpose of the Low Capacity Rats-High Capacity Rats (LCR-HCR) model is to serve as a resource for the in-depth study of rat models to resolve the extremes of exercise and health. By connecting clinical observation with a theoretical base, the working hypothesis is that: *variation in capacity for energy transfer is the central mechanistic determinant between disease and health* (*energy transfer hypothesis*). As an unbiased test of this hypothesis, this study showed that two-way artificial selective breeding of rats for low and high intrinsic endurance exercise capacity also produced rats that differed for numerous disease risks, including the metabolic syndrome, premature aging, fatty liver disease, obesity, and Alzheimer's disease. Exercise capacity is a result of intrinsic capacity plus adaptation to all aspects of physical activity. To capture this biology, rats for low and high response to 8 weeks of treadmill running exercise were selectively bred. Thus, the study has models that represent the 4 "corners" of exercise capacity. These contrasting animal model systems may prove to be translationally superior relative to more widely used simplistic models for understanding disease conditions. The rat models may be deeply explored to discover causal mechanisms and develop effective therapeutics. These rats are being studied at over 50 institutions in 11 countries.

Principal Investigators: Steven Britton, Lauren Koch (University of Michigan, USA).

1000 Athlome Project

The 1000 Athlome Project aims to sequence 1000 genomes of sprinters and distance runners of West and East African descent. Phase 1 of the project has already commended and involves the sequencing of 12 sprinters and 12 distance runners of the highest level (i.e. world record holders, Olympians and World Champions). Phase 2 (2016-2018) will involve increasing the sample size for sequencing to 100 genomes. The pool of the runners to be sequenced will be expanded to 1000 by 2020 (Phase 3). An important aim of this sequencing project is to document the genotype distribution of elite east and west African athletes. The large amount of genotype data to be generated from the 1000 Athlome project will serve 1) as a reference panel for future performance studies, and 2) to guide other extreme phenotype studies in medical science.

Principal Investigators: Masashi Tanaka (Tokyo Metropolitan Institute of Gerontology, JPN), Yannis Pitsiladis (University of Brighton, GBR).

Ethical Principles for Athlome biobanking

The rise of biobanking has brought about a whole range of issues that are not all wholly relevant to the Athlome project. Nevertheless, certain key principles must be noted here that will inform the governance framework for Athlome: (i) the consortia are global in reach but there is no universal agreement on the precise nature of ethically justifiable governance for biobanking; (ii) given the globality of the consortia, no single regional (e.g., European, American) framework ought to be adopted; (iii) a general framework drawing on widely shared principles should be discussed and adopted. Chief among the concerns, but only one among several, is the problem of consent.

Each of the projects that comprise Athlome are existing bio-guardians with a duty to protect the rights of participants who have contributed their samples to the individual projects noted above. The collection, storage, access to and use by researchers of those samples has been approved by relevant regulatory authorities (e.g., IRBs, RECs National Health Services Research Ethics Services) appropriate to the lead institution of the individual projects/consortia. Existing procedures do not currently extend to the sharing of samples beyond the study, since consent models are prospective (i.e. they guide future actions of researchers) and typically entail a form of specificity and the specific consent obtained varies between project partners. No retrospective consent is feasible and this is a widely shared problem for biobank development. Since the form of collaboration Athlome envisages was not laid out before participants gave their consent, it might be concluded that the sharing of data beyond the original research group and its stated purposes invalidates that consent. The problem for Athlome is not an uncommon one for biobank collaborations since it seeks retrospective extension of the consent model.

An ethical solution to this problem and related consent problems for new participants is to consider the use of a technique such as “broad consent”. The nomenclature here is important since this notion is variously described as “broad consent”, “blanket consent”, “future consent”, “hypothetical consent”, “passive/tacit/silent consent”, or “waived consent” (4,5). This would entail asking participants to agree to future unspecified uses of their data that are und(er)determined in the consent process and relevant forms (6). Without sufficient grasp of the

uses of the data or with whom it might be shared, this process fails the test of "comprehension" a user must understand sufficiently what they are agreeing to (3). Another possibility going forward would be "meta-consent" where consent is sought for broad categories of unspecified future research (7,8). Others have argued with respect to biobanking that the ethical issues entailed (e.g., privacy, confidentiality, ownership of access to the data) may be sufficiently assuaged by rigorous anonymization (1) and associated practices of data storage, though this is far from universally agreed upon (2).

The Athlome project will develop principles and protocols for safeguarding participants rights to access, confidentiality, privacy of data, and assurances that there is no significant mission drift of the kind of which is permitted under some conceptions of broad consent (or its similes). This would, for example, prohibit commercialization of participants' data. In order to preserve the integrity of this process and the principles, rigorous anonymisation processes will be developed by a partner institution that does not have any direct role in data collection, storage or analysis. This will assure independence and integrity to the process. This is especially important in this case since some of the research participants are public figures, which increases the likelihood that someone might be interested in re-identifying their data and genomic sequences. The independent institution would also have an oversight of each new proposal for the Athlome project going forward in order to ensure compliance with those principles and protocols.

In conclusion, by presenting the main study cohorts and projects that are currently included in the Athlome consortium it is our intention to show a global view not only of the main studies and initiatives that will be performed in the foreseeable future in the field of sports genomics (and that are likely to provide new exciting findings); we also wish to motivate potential collaboration initiatives with other research groups worldwide. International collaborations are likely to go well beyond the study of sports performance per se. Indeed, the Athlome consortium presents a unique chance to study the biology of the best elite athletes across most ethnicities, which is profoundly interesting from a medical point of view. World-class athletes represent the actual end-point of the human continuum of fitness-related phenotypes. In this regard, there is growing evidence (coming from both human and rodent study approaches – such as those included in the consortium) that not only physical activity levels but also individual fitness levels (a trait which

has a strong genetic component independent of activity levels) are inversely associated with the risk of major cardiometabolic diseases of western civilization, several cancer types and Alzheimer's disease. Thus, studying the genes of elite athletes offers a unique chance to gain insight into important medical, including genetic predisposition (or resilience) to chronic disease. Indeed, the "rare-common" strategy, underpinned by ethically sound research governance, is a valuable approach model to examine general mechanisms of disease pathophysiology, with world-class athletes representing the "rare" ("super-fit") human phenotype. Finally, identifying genetic markers of exercise capacity, adaptation to exercise programmes and in the predisposition to injury is certain to provide useful information to prescribe personalised exercise interventions in the context of 21st century medicine, which should not be based only on identifying new drug targets but also on implementing lifestyle interventions for disease prevention at the individual level.

Acknowledgements

Eastern Europe population studies (The Russian and Belarusian cohorts, GELAK, GELAV, GUAP) were supported by the grants from the Federal Agency for Physical Culture and Sport of the Russian Federation and the Ministry of Education and Science of the Russian Federation (contract number 02.522.11.2004), the Federal Medical-Biological Agency of the Russian Federation ("Sportgen project"), Republic of Belarus (State program of development of physical culture and sports for 2011-2015) and Royal Society International Joint Project grant from the United Kingdom (code F-90014). **GELAV** (Epigenetics of Lithuanian Athletes from Vilnius) project was developed by the Lithuanian National Olympic Committee and Lithuanian Olympic Sports Centre, while actual research was carried out at the Vilnius University, where Lithuanian athletes DNA samples are stored. We would like to thank Prof. Vaidutis Kučinskas from the Department of Human and Medical Genetics, Faculty of Medicine, Vilnius University, Lithuania, for providing ideas and support for accessing the control samples.

ELITE is supported by SAP/Stanford Sequencing Initiative and Women's Heart Health at Stanford.

GAMES was partially funded by the Prince Faisal Prize awarded to Drs C. Bouchard, T. Rankinen, M. Sarzynski and B. Wolfarth. CB is partially funded by the John W Barton Jr Chair in Genetics and Nutrition. The study in Russia was supported by a grant from the Federal Medical-Biological Agency ("Sportgen project"), http://fmbaros.ru/en/fmba/infor/. The Spanish group is funded by *Cátedra Real-Madrid Universidad Europea, and Fondo de Investigaciones Sanitarias and Fondos Feder* (grant # PI12/00914). This work in Japan was supported in part by grants from the programs Grants-in-Aid for Challenging Exploratory Research (24650414 to NF) from the Ministry of Education, Culture, Sports, Science and Technology; and by a grant-in-aid for scientific research from the Ministry of Health, Labor, and Welfare of Japan (to MM). No specific funding for this work was received for the studies performed in Australia, Poland, Kenya and Ethiopia.

Gene SMART Study is partly supported by Dr Eynon's Australian Research Council Early career Fellowship (ARC DECRA) DE#140100864, and by the Victoria University Central Research Grant Scheme (CRGS).

GOINg is supported by funds from the National Research Foundation (NRF) of South Africa and the Thembakazi Trust.

J-HAP project was supported by JSPS KAKENHI (Grant Numbers 21680050, 11J04771, 24650414, 26882041, 15H03081, and 15K16467) and by MEXT- commissioned projects.

NTR's funding was obtained from the Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 912-10-020, 904-61-193, 480-04-004, 463-06-001, 451-04-034, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192, Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007). VU University's Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the European Science Foundation (ESF, EU/QLRT-2001-01254), the European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-

F4-2007-201413); the European Research Council (ERC Advanced, 230374, ERC Starting grant 284 167), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R01D0042157-01A, MH081802; R01 DK092127-04, Grand Opportunity grants 1RC2 MH089951 and1RC2 MH089995). Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO.

Super-athletes: Genes and Sweat is partly funded by Qatar National Research Foundation NPRP grant (NPRP 7-272-1-041).

POWERGENE GWAS genotyping in Jamaicans, African-Americans (the U.S.A. cohort) and Japanese was funded by JSPS KAKENHI (Grant Number 21680050 and 24650414).

Rat models of exercise and health (LCR-HCR rat model) was funded by the Office of Research Infrastructure Programs/OD grant R24OD010950 and by grant R01DK099034 (to LGK and SLB) from the National Institutes of Health. We acknowledge the expert care of the rat colony provided by Molly Kalahar and Lori Heckenkamp. Contact: LGK lgkoch@umich.edu or SLB brittons@umich.edu for information on the LCR and HCR rats: these rat models are maintained as an international resource with support from the Department of Anesthesiology at the University of Michigan, Ann Arbor, Michigan.

1000 Athlome Project is funded by the Grants-in-aid for Scientific Research (KAKENHI) (A)-15200051, (A)-22240072, (A)-25242062 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and from Japan Society for the Promotion of Science (JSPS).

References

1. **Caplan, A. L.** What no one knows cannot hurt you: The limits of informed consent in the emerging world of biobanking. In: *The Ethics of Research Biobankin*, edited by J. H. Solbakk, S. Holm, B. Hofmann. Springer US, 2009, p. 25-32.

2. **Greely, H. T.** The uneasy ethical and legal underpinnings of large-scale genomic biobanks. *Annu Rev Genomics Hum Genet* 8: 343-364, 2007.

3. **Hansson, M. G.** Ethics and biobanks. *Br J Cancer* 100: 8-12, 2009.

4. **Hansson, M. G., Dillner, J., Bartram, C. R., Carlson, J. A., Helgesson, G.** Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol* 7: 266-269, 2006.

5. **Hofmann, B.** Broadening consent—and diluting ethics? *J Med Ethics* 35: 125-129, 2009.

6. **Hofmann, B., Solbakk, J. H., Holm, S.** Consent to Biobank Research: One Size Fits All? In: *The Ethics of Research Biobanking* edited by J. H. Solbakk, S. Holm, B. Hofmann. Springer US, 2009, p. 3-23.

7. **Holm, S., Ploug, T.** Patient choice and preventive genomic sequencing—more trouble upstream. *Am J Bioeth* 15: 24-26, 2015.

8. **Ploug, T., Holm, S.** Meta consent: a flexible and autonomous way of obtaining informed consent for secondary research. *BMJ* 350: h2146, 2015.

***The Athlome Project Consortium**

Corresponding author: Yannis P Pitsiladis[1]

Steering Committee: Yannis P Pitsiladis (Chair), Nir Eynon, Claude Bouchard, Kathryn N North, Alun G Williams, Malcolm Collins, Colin N Moran, Steven L Britton, Noriyuki Fuku, Eco de Geus, Vassilis Klissouras, Euan A Ashley, Alejandro Lucia, Ildus I Ahmetov, Mohammed Alsayrafi and Masashi Tanaka.

Consortium Group: University of Brighton, GBR: Yannis P Pitsiladis[1] (Principal Investigator), Nick Webborn[1], Guan Wang[1]; **Victoria University, AUS:** Nir Eynon[2] (Principal Investigator), David J Bishop[2] (Principal Investigator), Ioannis Papadimitriou[2], Xu Yan[2], Oren Tirosh[2], Jujiao Kuang[2]; **Pennington Biomedical Centre, USA:** Claude Bouchard[3], Tuomo Rankinene[3]**,** Mark Sarzinsky[3]; **Stanford University, USA:** Euan A Ashley[4], C Mikael Mattsson[4], Matthew Wheeler[4], Daryl Waggott[4]; **Bond University, AUS:** Nuala M Byrne[5]; **University of Sao Paulo, BRA:** Guilherme G Artioli[6]; **University of Cape Town, ZAF:** Malcolm Collins[7] (Principal Investigator), Alison September[7], Michael Posthumus[7], Willem van der Merwe[7]; **Gdansk University of Physical Education and Sport, POL; University of Szczecin, POL:** Pawel Cieszczyk[8,9] (Principal Investigator), Agata Leonska-Duniec[8,9], Krzysztof Ficek[9], Agnieszka Maciejewska-Karlowska[9], Marek Sawczuk[9], Marta Stepien-Slodkowska[9]; **Epworth Healthcare Melbourne, AUS:** Julian Feller[10]; **Aspetar, QAT:** Paul Dijkstra[11]; **Ural State University of Physical Culture, RUS:** Aleksandr M Chmutov[12], Dmitry A Dyatlov[12], Evgeniy F Orekhov[12], Yuliya E Pushkareva[12], Irina A Shvedkaya[12]; **University of Cagliari, ITA:** Myosotis Massidda[13], Carla M Calò[13]; **Manchester Metropolitan University, GBR; University College London, GBR; Liverpool John Moores University, GBR:** Alun G Williams[14,15] (Principal Investigator), Stephen H Day[14], Georgina K Stebbings[14], Robert M Erskine[15,16], Hugh E Montgomery[15]; **Murdoch Childrens Research Institute, AUS:** Kathryn N North[17] (Principle Investigator), Fleur C Garton[17], Peter Houweling[17]; **Ghent University, BEL:** Wim Derave[18], Audrey Baguet[18]; **Universidad Europea de Madrid, ESP:** Alejandro Lucia[19], Carlos A Muniesa[19]; **University of Foggia, ITA:** Francesco Sessa[20], Annamarie Petito[20]; **Nottingham Trent University, GBR:** Craig Sale[21]**,** David C Hughes[21]; Ian Varley[21] (Principal Investigator)**; VU University Amsterdam, NLD:** Eco de Geus[22] (Principal Investigator), Dorret Boomsma[22], Meike Bartels[22], Gareth E Davies[22] **Vilnius University, LTU; Lithuanian**

Olympic Sports Centre, LTU: Valentina Ginevičienė[23,24] (Principal Investigator), Audronė Jakaitienė[23], Vaidutis Kučinskas[23], Linas Tubelis[24], Algirdas Utkus[23]; **Lithuanian University of Educational Sciences, LTU:** Kazys Milašius[25] (Principal Investigator), Linas Tubelis[25]; **University of Stirling, GBR:** Colin N Moran[26] (Principal Investigator); **Lithuanian Sports University, LTU:** Tomas Venckunas[27], Albertas Skurvydas[27], Arvydas Stasiulis[27]; **University of Glasgow, GBR:** Dalia Malkova[28], Richard Wilson[28]; **University of Michigan, USA:** Steven L Britton[29] (principal investigator), Lauren G Koch[29] (principal investigator); **Juntendo University, JPN:** Noriyuki Fuku[30] (Principal Investigator), Hirofumi Zempo[30], Hisashi Naito[30], Noriko Ichinoseki-Sekine; **Nippon Sport Science University, JPN:** Naoki Kikuchi[31]; **National Institute of Fitness and Sports in Kanoya, JPN:** Eri Miyamoto-Mikami[32], **National Institute of Health and Nutrition, NIBIOHN, JPN:** Haruka Murakami[33], Motohiko Miyachi[33]; **Japan Institute of Sports Sciences, JPN:** Hideyuki Takahashi[34], Nao Ohiwa[34], Takashi Kawahara[34]; **Toyo University, JPN:** Hiroyasu Tsuchie[35]; **University of Nagasaki, JPN:** Takuro Tobina[36]; **The Open University of Japan, JPN:** Noriko Ichinoseki-Sekine[37]; **Fukuoka University, JPN:** Hiroaki Tanaka[38]; **Waseda University, JPN:** Koji Kaneoka[39]; **Nippon Sport Science University, JPN:** Koichi Nakazato[40]; **Sport Technology Research Centre Kazan, RUS; Kazan State Medical University, RUS**: Ildus I Ahmetov[41,42] (Principal Investigator), Emiliya S Egorova[42], Leysan J Gabdrakhmanova[41,42], Alina A Arkhipova[42], Alyona V Borisova[42], Rashid T Gabbasov[42], Albina A Stepanova[41], Ravil I Kashapov[41]; **St Petersburg Research Institute of Physical Culture, RUS:** Victor A Rogozkin[43] (Principal Investigator); Irina V Astratenkova[43], Anastasiya M Druzhevskaya[43], Olga N Fedotovskaya[43], Natalya D Golberg[43], Albina M Hakimullina[43]; **Research Institute for Physical-Chemical Medicine Moscow, RUS:** Elena S Kostryukova[44], Dmitry G Alexeev[44], Edward V Generozov[44], Dmitry S Ischenko[44], Nickolay A Kulemin[44], Andrey K Larin[44], Elena A Ospanova[44], Alexander V Pavlenko[44], Vadim M Govorun[44]; **The National Academy of Sciences of Belarus, BLR:** Andrei A Gilep[45] (Principal Investigator), Irina L Gilep[45], Irina V Haidukevich[45], Irina L Rybina[45]; **National University of Physical Education and Sport of Ukraine, UKR:** Svitlana B Drozdovska[46] (Principal Investigator), Victor E Docenko[46], Vladimir N Ilyin[46]; **M.Akmullah Bashkir State Pedagogical University, RUS:** Eugeniy Lekontsev[47]; **Moscow Department of Physical Culture and Sport, RUS:** Egor B Akimov[48]; **Anti-Doping Lab Qatar, QAT:** Mohamed El-Rayess[49], Costas Georgakopoulos[49], Mohammed Alsayrafi[49], **FMSI, ITA:** Francesco Botre[50], **Weill Cornell**

Medical College in Qatar, QAT: Karsten Suhre[51], **University College London, GBR:** Mike Hubank[52]; **University of Athens, GRC:** Vassilis Klissouras[53]; **University of Humboldt and Charite Medical University, DEU:** Bernd Wolfarth[54,55]; **Army Recruiting and Training Division, GBR:** Julie P. Greeves[56]; **Canadian Sport Institute Pacific, CAN:** Trent Stellingwerff[57]; **Cardiff Metropolitan University, GBR:** Craig Ranson[58]; **University of East Anglia, GBR:** William D Fraser[59,60]; **Institute of Health and Biomedical Innovation, Queensland University of Technology, AUS:** Rebecca Grealy[61] , Lyn Griffiths[61]; **University of Cambridge, GBR:** Robert Scott[62]; **The Swedish School of Sport and Health Sciences, SWE:** C Mikael Mattsson[63]; **Tokyo Metropolitan Institute of Gerontology, JPN:** Masashi Tanaka[64,65]; **Physiology Laboratory of the Urals Research Center for Radiation Medicine of the Federal Medical-Biological Agency of Russia, RUS:** Vladimir P Pushkarev[66].

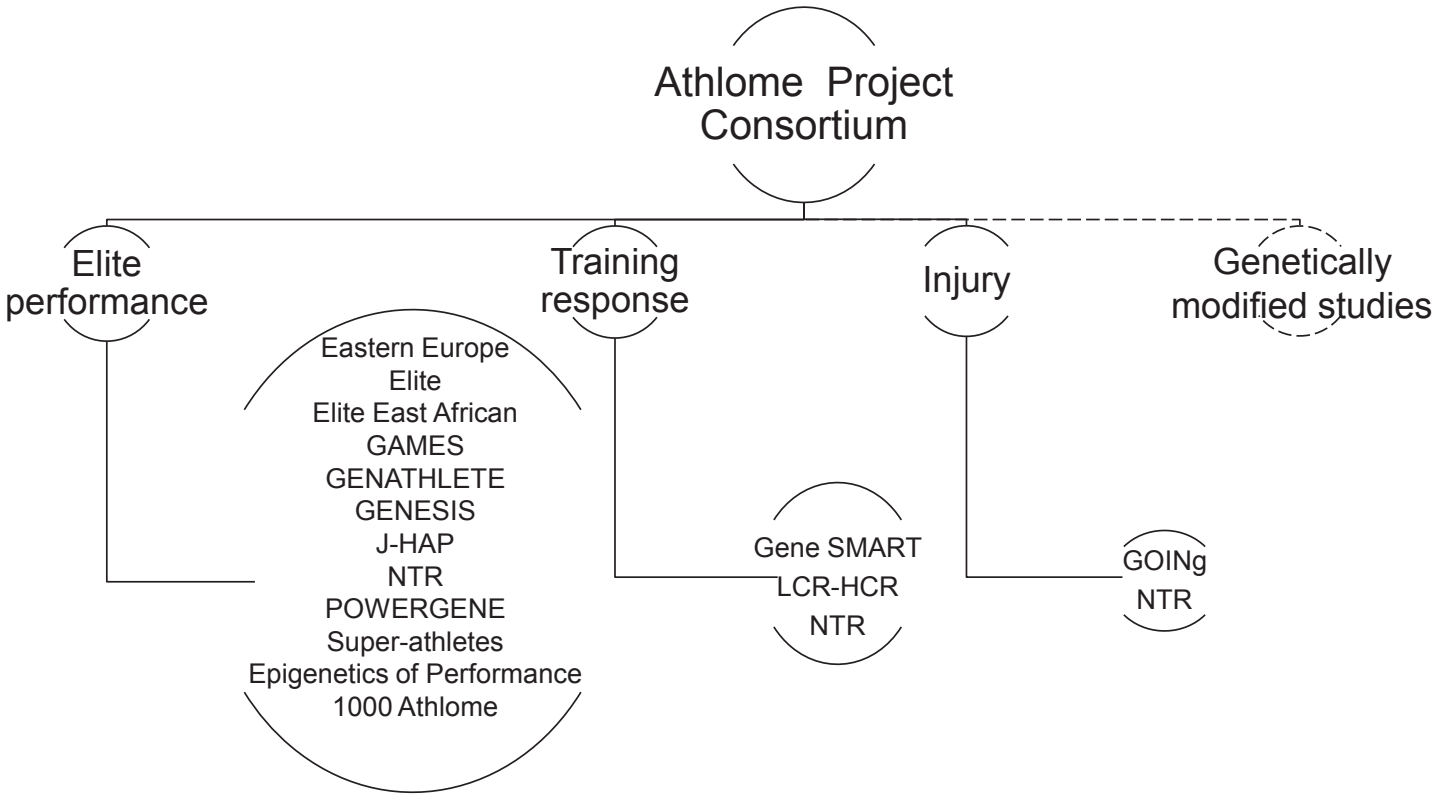
Affiliations: [1] FIMS Reference Collaborating Centre of Sports Medicine for Anti-Doping Research, University of Brighton, Eastbourne, United Kingdom; [2] Institute of Sport, Exercise, and Active Living (ISEAL), Victoria University, Melbourne, Australia; [3] Human Genomics Laboratory, Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, United States; [4] Stanford University Medical Center, Stanford, United States; [5] Bond Institute of Health and Sport (BIHS), Bond University, Gold Coast, Australia; [6] Laboratory of Applied Nutrition and Metabolism, School of Physical Education and Sport, University of Sao Paulo, Sao Paulo, Brazil; [7] Division of Exercise Science and Sports Medicine, Department of Human Biology, University of Cape Town, Cape Town, South African; [8] Department of Tourism and Recreation, Academy of Physical Education and Sport, Gdansk, Poland; [9] Department of Physical Culture and Health Promotion, University of Szczecin, Szczecin, Poland; [10] Epworth Healthcare, Melbourne, Australia; [11] Department of Sport Medicine, Aspetar, Doha, Qatar; [12] Genetic laboratory, Ural State University of Physical Culture, Chelyabinsk, Russia; [13] Department of Environmental and Life Science, University of Cagliari, Cagliari, Italy; [14] Department of Exercise and Sport Science, Manchester Metropolitan University, Crewe, United Kingdom; [15] Institute of Sport, Exercise and Health (ISEH), University College London, London, United Kingdom; [16] Research Institute for Sport and Exercise Sciences (RISES), Liverpool John Moores University, Liverpool, United Kingdom; [17] Murdoch Childrens Research Institute, Royal Children's Hospital, Melbourne, Australia; [18] Department of Movement and Sports Sciences,

Ghent University, Ghent, Belgium; [19]School of Doctorate Studies & Research, Universidad Europea de Madrid, Madrid, Spain; [20] Medical Genetics Unit, Department of Medical Sciences and Department of Clinical and Experimental Medicine, University of Foggia, Foggia, Italy; [21] Musculoskeletal Physiology Research Group, Sport, Health and Performance Enhancement Research Centre, Nottingham Trent University, Nottingham, United Kingdom; [22] VU University and VU Medical Centre, Amsterdam, Netherlands; [23] Department of Human and Medical Genetics, Vilnius University, Vilnius, Lithuania; [24] Lithuanian Olympic Sports Centre, Vilnius, Lithuania; [25] Department of Sport Teaching Methods, Lithuanian University of Educational Sciences, Vilnius, Lithuania; [26] Health and Exercise Sciences Research Group, University of Stirling, Stirling, Scotland; [27] Lithuanian Sports University, Kaunas, Lithuania; [28] School of Medicine, University of Glasgow, Glasgow, Scotland; [29] Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, United States; [30] Graduate School of Health and Sports Science, Juntendo University, Chiba, Japan; [31] Sports Training Center, Nippon Sport Science University, Tokyo, Japan; [32] Department of Sports and Life Science, National Institute of Fitness and Sports in Kanoya, Kagoshima, Japan; [33] Department of Health Promotion and Exercise, National Institute of Health and Nutrition, Tokyo, Japan; [34] Department of Sports Science /Medical Centre, Japan Institute of Sports Sciences, Tokyo, Japan; [35] Faculty of Law, Toyo University, Tokyo, Japan; [36] Faculty of Nursing and Nutrition, University of Nagasaki, Nagasaki, Japan; [37] Faculty of Liberal Arts, The Open University of Japan, Chiba, Japan; [38] Institute for Physical Activity, Fukuoka University, Fukuoka, Japan; [39] Faculty of Sports Sciences, Waseda University, Saitama, Japan; [40] Department of Exercise Physiology, Nippon Sport Science University, Tokyo, Japan; [41] Sport Technology Research Centre, Volga Region State Academy of Physical Culture, Sport and Tourism, Kazan, Russia; [42] Laboratory of Molecular Genetics, Kazan State Medical University, Kazan, Russia; [43] Sports Genetics Laboratory, St Petersburg Research Institute of Physical Culture, St Petersburg, Russia; [44] Research Institute for Physical-Chemical Medicine, Moscow, Russia; [45] Institute of Bioorganic Chemistry NASB, Minsk, Belarus; [46] National University of Physical Education and Sport of Ukraine, Kiev, Ukraine; [47] M. Akmulla Bashkir State Pedagogical University, Ufa, Russia; [48] Centre for Sports Innovation Technologies and National Teams of the Moscow Department of Physical Culture and Sport, Moscow, Russia; [49] Anti-Doping Lab Qatar (ADLQ), Doha, Qatar; [50] Anti-Doping Lab, Federazione Medico Sportiva Italiana (FMSI), Italy; [51] Department of

Physiology and Biophysics, Weill Cornell Medical College in Qatar, Doha, Qatar; [52] Centre for Translational Omics, University College London, London, United Kingdom; [53] Ergophysiology Research Laboratory, Department of Sport Medicine and Biology of Physical Activity, University of Athens, Athens, Greece; [54] Department of Sports Sciences, University of Humboldt, Berlin, Germany; [55] Laboratory of Sports Medicine, Charite Medical University, Berlin, Germany; [56] Department of Occupational Medicine, Headquarters Army Recruiting and Training Division, United Kingdom; [57] Canadian Sport Institute Pacific, Pacific Institute for Sport Excellence, Victoria, British Columbia, Canada; [58] Sports injury Research Group, Cardiff School of Sport, Cardiff Metropolitan University, Cardiff, Wales, United Kingdom; [59] Norwich Medical School, University of East Anglia, United Kingdom; [60] Norfolk and Norwich University Hospital, Norwich, United Kingdom; [61] Genomics Research Centre, Institute of Health and Biomedical Innovation (IHBI), Queensland University of Technology, Brisbane, Australia; [62] MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom; [63] Åstrand Laboratory of Work Physiology, The Swedish School of Sport and Health Sciences, Stockholm, Sweden; [64] Department of Longevity and Health, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan; [65] Department of Clinical Laboratory, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan; [66] Physiology Laboratory of the Urals Research Center for Radiation Medicine of the Federal Medical-Biological Agency of Russia, Chelyabinsk, Russia.

Figure Legends:

Figure 1. The Athlome Project Consortium. Genomic, epigenomic, transcriptomic, proteomic and metabolomic studies are being conducted by the participating centres to address questions in the three main research areas: elite performance, training response, and injury. Future investigations planned include genetically modified studies.


Athlome Project Consortium
Elite performance
Training response
Injury
Genetically modified studies
Eastern Europe Elite
Elite East African
GAMES
GENATHLETE
GENESIS
J-HAP
NTR
POWERGENE
Super-athletes
Epigenetics of Performance
1000 Athlome
Gene SMART
LCR-HCR
NTR
GOINg
NTR