

## Genomics DNA Profiling in Elite Professional Soccer Players: A Pilot Study.

Kambouris M<sup>1</sup>, Del Buono A<sup>2</sup>, Maffulli N<sup>3,4</sup>

<sup>1</sup> Shafallah Medical Genetics Center, Doha, Qatar

<sup>2</sup> Department of Orthopaedic and Trauma Surgery Hospital Antonio Cardarelli, 86100 Contrada Tappino, Campobasso, Italy

<sup>3</sup> Department of Musculoskeletal Disorders, University of Salerno, Salerno, Italy

<sup>4</sup> Queen Mary University of London, Center of Sports and Exercise Medicine, London, UK

Corresponding author: Nicola Maffulli. (e-mail: n.maffulli@qmul.ac.uk)

**Abstract** - Functional variants in exonic regions have been associated with development of cardiovascular disease, diabetes and cancer. Athletic performance can be considered a multi-factorial complex phenotype. Genomic DNA was extracted from buccal swabs of seven soccer players from the Fulham football team. Single nucleotide polymorphism (SNPs) genotyping was undertaken. To achieve optimal athletic performance, predictive genomics DNA profiling for sports performance can be used to aid in sport selection and elaboration of personalized training and nutrition programs. Predictive DNA profiling may be able to detect athletes with potential or frank injuries, or screening and selection of future athletes, and can help them to maximize utilization of their potential and improve performance in sports. The aim of this study is to provide a wide scenario of specific genomic variants that an athlete carries, to implement which measures should be taken to maximize the athlete's potential.

**Keywords:** athletes, performance, SNP variants

### I. INTRODUCTION

Human athletic performance is a highly complex phenotype considered a multi-factorial polygenic trait [1]. Environment, physiological and psychological factors are involved, but there is also evidence that some genes may be determinant in physical fitness and performance phenotype<sup>1</sup>. Some variations in DNA sequences have been associated with specific phenotypes involved in athletic performance, including endurance capacity, muscle performance, susceptibility to injuries, body mass composition, and psychological aptitude, and provide information about advantages and genetic barriers that reflect the athletic performance phenotype [2]. Therefore, it is possible that performance can be improved working on the genetic advantages an individual has been endowed with. In this way, the physical fitness and sports performance phenotype would be optimized. The fact that functional variants in exonic regions have been associated with development of cardiovascular disease, diabetes, cancer could induce to focus on genetic predispositions [3]. However, given the multi-factoriality of the complex phenotype of athletic performance, lifestyle should be taken into account. Quality of nutrition,

smoking and alcohol abuse, for example, may influence the risk of onset of some conditions. This study assessed the genetic patterns of some elite professional soccer players, and discussed their implications.

### II. METHODOLOGY

This study was approved by the Institutional Review Board and Ethical Committee of the Shafallah Medical Genetics Center, Doha, Qatar. Seven soccer players from the Fulham football team were enrolled after they had given a written informed consent. All assessments were performed on saliva samples after DNA extraction. We tested all athletes for genomic variants, independently of their age. Some genes have been analyzed for the presence of some conditions or traits, and mutations (polymorphisms). Biological effects and genotypes of each gene were investigated. The influence of each genotype has been described, reporting whether certain genotypes have been associated with increased, average, or reduced risk compared to the general population [1].

The general conclusion for each gene category was drawn according to the relative importance of each genotype. This means that even, if there are many patterns with average or positive influences, the conclusion for the specific section of the test may show increased risk compared to the relative risk of general population.

#### **DNA extraction**

Genomic DNA was extracted from buccal swabs (Whatman's Sterile *Omni Swab*) using the Genra Puregene Tissue Kit (cat#158422, QIAGEN) according to the manufacturer guidelines.

#### **SNP genotyping**

The method used for SNPs genotyping was based on the Sequenom MassARRAY platform. The assay implied a primer extension at a two level specificity. A standard PCR reaction was first performed. followed by a specific primer extension reaction (iPLEX assay). This assay allows an oligonucleotide primer to anneal immediately upstream/downstream of the polymorphic site desired to be genotyped. The primer and amplified target DNA were then incubated with mass-modified dideoxynucleotide

terminators (ddNTPs). The primer extension was made according to the sequence of the variant site: it was a single complementary mass-modified base. The mass of the extended primer was determined using a MALDI-TOF mass spectrometry. The extension primer indicated which alleles were present at the polymorphic site of interest. Sequenom provides software (SpectroTYPER) which automatically translates the mass of the observed primers into a genotype for each reaction. Sequenom massARRAY designer software automatically designed PCR and extension primers (probes) for each SNP. The designer grouped all SNPs under investigation in several plexes based on the nature of targeted sequences and the masses of extension primers.

PCR primer mixes (Integrated DNA Technologies, Inc.) were prepared and PCR reactions were carried out in the thermocycler (9700 PCR System, Applied Biosystems). Concentrations and conditions of reactions were followed according to Sequenom guidelines.

Treatment with SAP-Shrimp alkaline phosphatase (Sequenom) was performed to remove non-incorporated dNTPs from amplification products. SAP cocktail mix was prepared using the instructions provided by Sequenom.

SAP cocktail was dispensed into each individual well of the 384-well plate with MassARRAY® Liquid Handler Station (Matrix) (Sequenom).

The sample plate was placed into the thermocycler (9700 PCR System, Applied Biosystems) for incubation. iPLEX extend primer mixes (Integrated DNA Technologies, Inc.) were prepared using linear adjustment method. Extension primers were divided in 4 groups from low to high mass and mixes. iPLEX cocktail mixes were prepared according to instructions provided by Sequenom and were dispensed to each well of the 384-well sample plate and mixed. At a next stage SpectroCLEAN (Sequenom) was added directly to primer extension reaction products to remove salts such as Na<sup>+</sup>, K<sup>+</sup>, and Mg<sup>2+</sup> ions. This cationic slurry resin pretreated with acid reagents is necessary for ions removal to avoid high background noise in the mass spectra. If not removed, these ions could result in high background noise in the mass spectra. Experimental procedure was followed according to Sequenom instructions.

Reactions were dispensed to a SpectroCHIP, using a MassARRAY™ Nanodispenser RS 1000 Fusio (Sequenom). SpectroCHIP run on a MassARRAY Typer Workstation Compact (Sequenom), with settings for iPLEXGold (iPLEX.par) for both FlexControl and SpectroAcquire.

### III. RESULTS

We report the general conclusions and executive summary for the main investigated fields. Single genes, effects, genotypes and specific medical, training and nutritional guidelines are reported in separate Tables.

*Sport performance - Endurance capacity* Five athletes presented an overall genetic profile not associated with increased endurance capacity. Two had a genetic profile associated with increased endurance capacity, supposed being associated with increased abilities in sports requiring endurance.

*Sport performance – Muscle performance* The overall genetic profile was associated with increased muscle performance in 6 athletes, suspected to have increased abilities in sports requiring strength, and only 1 with a genetic profile not associated with increased muscle performance.

*Susceptibility to tendon injuries* The genetic profile was not associated with increased susceptibility to tendinopathies in 5 subjects. The 2 athletes with increased susceptibility to tendinopathies were advised to consult a Sports Physician.

*Susceptibility to bones injuries* The overall genetic profile was not associated with a negative effect in terms of risk for bone fractures in 5 subjects. In 2 athletes, the overall genetic profile was associated with impaired use of Vitamin D and reduced absorption of calcium, with negative effect on bones health (Bone Mineral Density) and increased risk for bone fractures.

*Body Mass Index (BMI)* The overall genetic profile was not associated with disruptions in regulation of body mass index in 4 athletes. Three athletes had a profile associated with increased fat accumulation. Three athletes had an overall genetic profile associated with normal lipids metabolism; in 4, the genetic profile was associated with higher risk for disruptions in lipids metabolism.

*Nutrigenomics - folic acid metabolism general conclusion* The profile was associated with efficient homocysteine removal and sufficient Vitamin B metabolism in only 1 subject, and with impaired homocysteine removal.

*Nutrigenomics – iron absorption and storage* This overall genetic profile was not associated with higher risk for iron overload in all the athletes.

*Nutrigenomics - inflammatory response* The overall genetic profile was not associated with increased susceptibility to inflammatory conditions in 1 athlete, and with increased susceptibility in 6 athletes.

*Nutrigenomics - antioxidation ability* Anti-oxidant protection was considered satisfactory in 6 athletes.

*Nutrigenomics - detoxification ability* The overall genetic profile was associated with satisfactory detoxification ability in 4 athletes, and with reduced detoxification ability in 3.

*Nutrigenomics – salt sensitive hypertension* The genetic profile was not associated with increased risk for hypertension, which is mostly dependent on salt intake through diet, in 3 athletes, and was associated with increased risk in 4.

*Nutrigenomics – alcohol metabolism* Four athletes had an overall genetic profile associated with fast metabolism of alcohol. In 3 athletes, the overall genetic profile was associated with slow metabolism of alcohol.

*Nutrigenomics – caffeine metabolism* The overall genetic profile was associated with slow metabolism of caffeine in 4 subjects. Three subjects had an overall genetic profile associated with fast caffeine metabolism.

*Nutrigenomics – gluten tolerance* The overall genetic profile was not associated with increased risk for gluten intolerance (celiac disease) in 2 subjects. The 5 subjects with increased risk for gluten intolerance (celiac disease) were advised to consult a gastroenterologist, especially if they had experienced symptoms after intake of nutrients, like abdominal pain and bloating, diarrhea, constipation, pale, loose and greasy stool (steatorrhea), weight loss or failure to gain weight, fatigue. In the meantime they were recommended to avoid consumption of gluten containing products (all types of wheat, barley and rye).

#### IV. DISCUSSION

Predictive genomics DNA profiling does not detect or determine superior athletic performance, but can predict abilities and weaknesses associated with sports performance, and detect low penetrance sequence variation not proven to be causative, but probably contributing, in a cumulative fashion, to the relevant phenotypic manifestations [1]. Predictive genomics testing is of a qualitative rather than quantitative nature, even though it is not possible to quantify the predisposition to specific physiological traits and the overall predisposition for a certain profile. The way in which overall genetic profiles indicate certain abilities or weaknesses relevant and related to the sports performance phenotype may be known, but it does not mean that the trait under investigation will be certainly manifested if a specific gene or genetic profile is present. Hence, these information may outline some individual tendencies, not describing or predicting what will happen. This study evaluates traits and conditions, possibly influenced by the external environment, to individuate genetic traits the disadvantages of which may be modified, without any interference to medical diagnosis or status. Genomic variants of genes ACE [4], BDKRB2 [5], NOS3 [6], HIF1 [7], and VEGF [8] may influence maximal oxygen consumption and energy supply of aerobic metabolism, enhance oxygen supplies to muscle tissues, and increase endurance performance (Table 1). Polymorphisms of the gene HBB [9], involved in the production and function of red blood cells, may increase oxygen supply to muscles

whereas CHRM2 genomic variants genetically influences cardio-respiratory fitness and heart rate recovery ability [10].

The gene ACTN3 is involved in athletic performance and function of skeletal muscle fibers [11]. In fact, muscle strength and performance (Table 2) may be differently influenced by genes involved in supplying oxygen to muscles, including NOS3 [6], HIF-1 [6], ACE [4], and ACTN3 [12]. In addition, a nonsense variant in the gene AMPD1 is associated with acute exercise intolerance. This is a genetic barrier to intense muscle performance and, in sedentary patients, may predispose to cramps, easy fatigability, and myalgia after exercise [13]. Specifically, muscle fatigue depends on removal of lactic acid, which is influenced by a variant in the MCT-1 gene, determinant for lactate transport capability and intensity of performance [2]. On the contrary, a gene variant in DIO1 affects positively anaerobic exercise phenotypes, enhancing muscular strength [13].

The panels of genes associated with endurance capacity and muscle performance provide information that can be used to choose sport or athletic activity. Several factors, including anthropological and biochemical measurements, technical skills, and first and foremost social and economic factors have to be considered, and genetic advantages in either endurance or power/speed or both can be used to select individuals who may have better or even elite performance in certain types of sports. In case an individual carries the genetic variants favoring endurance performance, the athlete may be more likely to excel in sports requiring high aerobic power. If an individual's genetic profile favors muscle power and anaerobic metabolism, this individual would exhibit extra advantage in sports requiring increased speed and strength performance.

Based on specific genomic variants that an athlete carries, measures should be taken to maximize the athlete's potential. Depending on the biological effect of certain polymorphisms, we provide sample recommendations to improve endurance and power performance, respectively based on indicative DNA profile (Table 1 and 2), which reflect a 'common sense' medical and training corrective action approach. In addition, nutritional adjustments based on nutrigenetic / nutrigenomic DNA variants can impact markedly sports performance.

Genetic predisposition and increased risk for tendinopathies and sports-related injuries is conferred by the genetic variants in genes encoding various collagen types (COL1A1 [14], COL5A1 [15]), since collagen is the main structural component of tendons and ligaments. The specific variant in the MMP3 gene encoding a matrix metalloproteinase interacts with COL5A1 [16], and is associated with increased risk to develop Achilles tendinopathy, which is further increased in individuals who have unfavorable variants in both COL5A1 and MMP3 genes.

Although the association between genomic variants and susceptibility, predisposition and increased risk for bone injury is limited, multiple association studies link Bone

Mineral / Mass Density to functional genomic variants in the Vitamin D receptor gene (VDR which mediates calcium absorption) [17] (Table 3).

Physical activity is beneficial not only for athletes but also for the population at large in terms of body weight management and BMI reduction. Depending on the DNA profile for genes involved in body mass composition (Table 4), physical activity alone might not be sufficient to achieve optimal lean body mass, maintain a good BMI and improve sports performance, but proper nutrition and nutritional supplementations may be useful for individuals with at-risk DNA profiles.

The main purpose of nutritional genomics is to determine for each individual their personalized and individualized nutritional needs, essential and beneficial to promote optimal health and well-being [18]. The vitamin B complex has a key role (folic acid, vitamin B6, vitamin B12) in cell metabolism and especially DNA synthesis and repair processes, as well as many variants in genes responsible for the metabolism of these vitamins (MTHFR, MTR, MTRR, etc.) (Table 5) [19]. Often increased iron intake and supplementation are needed, especially in athletes involved in endurance sports, but genetic profiles may be indicative of increased risk for hemochromatosis, when specific variants of the HFE gene are present (Table 6) [20]. In these subjects, an increased iron intake may lead to toxic effects on tissues.

Intensive exercise may increase the risk for inflammation, which is further intensified by functional genomic variants in genes such as TNF- $\alpha$ , IL-6 and CRP) (Table 7) [21], with the need to enhance anti-inflammatory bioactive compounds in their diet.

Accumulation of variants in genes affecting anti-oxidation and detoxification (among other GSTT1, GPX1, CAT) [22] may compromise the defense against free radical damage and oxidative stress insults (Tables 8,9). Nutritional programs for athletes are traditionally designed to accommodate the nutritional demands of the sport of choice on various biochemical markers that indicate specific nutritional needs. A well-structured dietary program, meeting the dietary and energy requirements for the particular athlete and sport is crucial to success, and DNA profiling may be taken into account when designing and implementing these programs.

## V. CONCLUSIONS

To achieve optimal athletic performance, predictive genomics DNA profiling for sports performance should be used to aid in sport selection and elaboration of personalized training and nutrition programs. The purposes of predictive DNA profiling are not to detect athletes with potential or frank injuries, or screening and selection of future athletes, but to help them to maximize utilization of their potential and improve performance in sports.

Additional variants that may affect this complex phenotype have to be investigated. Predictive genomics

DNA profiling for sport and athletic performance allows to understand what genetic advantages have to be exploited, and which genetic barriers have to be overcome. We are aware that the validity of the conclusions could be biased because of methodological shortcomings: the sample size is too small to draw definitive conclusions, and no inferential analysis was performed. However, we provide descriptive findings probably useful for future investigations.

## REFERENCES

- [1] Kambouris M, Ntalouka F, Ziogas G, Maffulli N. Predictive Genomics DNA Profiling for Athletic Performance. *Recent Pat DNA Gene Seq* 2012; 6:229-39.
- [2] Cupeiro R, Benito PJ, Maffulli N, Calderon FJ, Gonzalez-Lamuno D. MCT1 genetic polymorphism influence in high intensity circuit training: a pilot study. *J Sci Med Sport* 2010; 13: 526-30.
- [3] Del Buono A, Denaro V, Maffulli N. Genetic susceptibility to aseptic loosening following total hip arthroplasty: a systematic review. *Br Med Bull*; 101: 39-55
- [4] Myerson S, Hemingway H, Budget R, Martin J, Humphries S, Montgomery H. Human angiotensin I-converting enzyme gene and endurance performance. *J Appl Physiol* 1999; 87: 1313-6
- [5] Van Guilder GP, Pretorius M, Luther JM, Byrd JB, Hill K, Gainer JV, Brown NJ. Bradykinin type 2 receptor BE1 genotype influences bradykinin-dependent vasodilation during angiotensin-converting enzyme inhibition. *Hypertension* 2008; 51: 454-9.
- [6] Saunders CJ, Xenophontos SL, Cariolou MA, Anastassiades LC, Noakes TD, Collins M. The bradykinin beta 2 receptor (BDKRB2) and endothelial nitric oxide synthase 3 (NOS3) genes and endurance performance during Ironman Triathlons. *Hum Mol Genet* 2006; 15: 979-87.
- [7] Mason SD, Rundqvist H, Papandreou I, Duh R, McNulty WJ, Howlett RA, Olfert IM, Sundberg CJ, Denko NC, Poellinger L, Johnson RS. HIF-1 $\alpha$  in endurance training: suppression of oxidative metabolism. *Am J Physiol Regul Integr Comp Physiol* 2007; 293: R2059-69.
- [8] Kraus RM, Stallings HW, 3rd, Yeager RC, Gavin TP. Circulating plasma VEGF response to exercise in sedentary and endurance-trained men. *J Appl Physiol* 2004; 96: 1445-50.
- [9] He Z, Hu Y, Feng L, Lu Y, Liu G, Xi Y, Wen L, Xu X, Xu K. Polymorphisms in the HBB gene relate to individual cardiorespiratory adaptation in response to endurance training. *Br J Sports Med* 2006; 40: 998-1002.

- [10] Hautala AJ, Rankinen T, Kiviniemi AM, Makikallio TH, Huikuri HV, Bouchard C, Tulppo MP. Heart rate recovery after maximal exercise is associated with acetylcholine receptor M2 (CHRM2) gene polymorphism. *Am J Physiol Heart Circ Physiol* 2006; 291: H459-66.
- [11] Niemi AK, Majamaa K. Mitochondrial DNA and ACTN3 genotypes in Finnish elite endurance and sprint athletes. *Eur J Hum Genet* 2005; 13: 965-9.
- [12] MacArthur DG, Seto JT, Chan S, Quinlan KG, Raftery JM, Turner N, Nicholson MD, Kee AJ, Hardeman EC, Gunning PW, Cooney GJ, Head SI, Yang N, North KN. An Actn3 knockout mouse provides mechanistic insights into the association between alpha-actinin-3 deficiency and human athletic performance. *Hum Mol Genet* 2008; 17: 1076-86.
- [13] Bray MS, Hagberg JM, Perusse L, Rankinen T, Roth SM, Wolfarth B, Bouchard C. The human gene map for performance and health-related fitness phenotypes: the 2006-2007 update. *Med Sci Sports Exerc* 2009; 41: 35-73.
- [14] Collins M, Posthumus M, Schwelanus MP. The COL1A1 gene and acute soft tissue ruptures. *Br J Sports Med* 2010; 44: 1063-4.
- [15] September AV, Cook J, Handley CJ, van der Merwe L, Schwelanus MP, Collins M. Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations. *Br J Sports Med* 2009; 43: 357-65.
- [16] Raleigh SM, van der Merwe L, Ribbans WJ, Smith RK, Schwelanus MP, Collins M. Variants within the MMP3 gene are associated with Achilles tendinopathy: possible interaction with the COL5A1 gene. *Br J Sports Med* 2009; 43: 514-20.
- [17] Nakamura O, Ishii T, Ando Y, Amagai H, Oto M, Imafuji T, Tokuyama K. Potential role of vitamin D receptor gene polymorphism in determining bone phenotype in young male athletes. *J Appl Physiol* 2002; 93: 1973-9.
- [18] Boehl T. Emerging science raises questions: what to tell your clients about nutritional genomics. *J Am Diet Assoc* 2007; 107: 1094-6.
- [19] Lu C, Xie H, Wang F, Shen H, Wang J. Diet folate, DNA methylation and genetic polymorphisms of MTHFR C677T in association with the prognosis of esophageal squamous cell carcinoma. *BMC Cancer* 2011; 11: 91.
- [20] Coppin H, Bensaid M, Fruchon S, Borot N, Blanche H, Roth MP. Longevity and carrying the C282Y mutation for haemochromatosis on the HFE gene: case control study of 492 French centenarians. *BMJ* 2003; 327: 132-3.
- [21] Yamin C, Duarte JA, Oliveira JM, Amir O, Sagiv M, Eynon N, Amir RE. IL6 (-174) and TNFA (-308) promoter polymorphisms are associated with systemic creatine kinase response to eccentric exercise. *Eur J Appl Physiol* 2008; 104: 579-86.
- [22] Palmer CN, Doney AS, Lee SP, Murrie I, Ismail T, Macgregor DF, Mukhopadhyay S. Glutathione S-transferase M1 and P1 genotype, passive smoking, and peak expiratory flow in asthma. *Pediatrics* 2006; 118: 710-6.