THE GENETIC PROFILES OF ELITE ATHLETES

S.J. LOCKEY

PHD 2017

THE GENETIC PROFILES OF ELITE ATHLETES

S.J. LOCKEY

This thesis is submitted in partial fulfilment of the requirements of the Manchester Metropolitan University for the degree of Doctor of Philosophy

Department of Exercise and Sport Science

Manchester Metropolitan University

2017

I. ACKNOWLEDGEMENTS

This PhD achievement was two attempts and 12 years in the making. Although there were breaks in scholarship over the 5 years of this thesis production, my mind never left its contents. For those of you reading this and thinking 'I'm not sure I can make it' – you can, and you will…keep going. This girl thought she was crazy enough to make it – and she did…with the unwavering support of the wonderful people below.

Firstly, I would like to thank my supervisory team Dr Steve Day, Dr Rob Erskine and Dr Alun Williams. At one time or another you have all been my director of studies, offered your advice, support and encouragement in the planning, analysis and writing up of this project. Thanks are due to my examination team Dr Gladys Pearson, Dr Colin Moran and Dr Chris Morse for offering their time, constructive comments and supportive questioning. The work described in this thesis was supported by a research studentship from the Institute of Performance Research (IPR) at Manchester Metropolitan University. My thanks are also extended to all colleagues and friends at the IPR for their advice, moral support and assistance, inclusive of:

Dr Charlie Pedlar and Dr Courtney Kipps, your introductions and support for the project allowed for the extensive data collection from the English Institute of Sport athletes and those competing at an elite level in the London Marathon. I and grateful to Mr John Beeden for granting access to athletes at the London Marathon Expo for three consecutive years. Mr Richard Hall and Mr Mike Neighbour, of the London Marathon Expo volunteer team for their support in identifying elite athletes of interest to the project.

Dr Shane Heffernan for ‘nocturnally sharing’ an office and laboratory with me throughout the project, supporting me with transportation and data collection and the occasional Disney sing along

Dr Adam Herbert for the hours spent out collecting and inputting data, cataloguing, extracting and loading plates in the lab. It was great to have your company on those very long days and someone to bounce ambitious ideas off as to where to recruit next.

I would also like to thank those who supported the data collection for this project (not mentioned already), Dr Jamie McPhee, Dr Georgina Stebbings, Jamie Guney, Brandon Foster, James Gavin and Josh Neale, Dr Emma Bostock, Nicola Fisher and David Egbosimba. Of course, none of the experiments would have been possible without the research participants, who were willing to give their time and DNA, required to complete this project.

There are several personal close friends who have had to live this project with me. These include ‘My Girls’ Becca, Kelly, and Claire, thanks for helping me to pick up the pieces, reorganise, make sense and insist I have another go usually over a glass of ‘pink feet wine’ and lots of chocolate. To Steph, the tiara wobbled but it didn’t fall thank you for your friendship through all the ups and down. To the fabulous half of ‘Team Phys’, Alan, thank you for your support, guidance, candid conversation and cocktail making abilities. To Jamie, thank you for your love, introducing me to self-care, for the grounding walks through Keele Campus, and the most amazing lasagne dinners after a 14 hour day in the lab. To Alice, thank you for your motherly love, your hospitality, and your friendship. To Kirk, ‘this is all your fault’, without your support I would have settled. Thank you for loving me enough to light a fire under me, make me crazy enough to think I could do it and showing me ‘fear is not real’.

Lastly, Thank you to my family, Mum, Dad, Andy, Anne, thank you for your endless love, for living the triumphs, for listening patiently to the frustrations, for giving me the time and space to work it out and grow. To Grandma thank you for being my lay audience and to Uncle Paul thank you for the support with my GCSE maths homework it informed the stats in this project and I couldn't have done it without each of you. Thank you x

II. ABSTRACT

Elite endurance athletes are strongly suspected to have differing genetic profiles from sub-elite endurance athletes and non-athletes. This thesis will contribute to the developing knowledge in this area, providing a more detailed analysis of the genetic profile of elite endurance athletes in the sport of marathon running. Identifying 'advantageous' genetic characteristics would be a significant development. The insight provided about the underlying physiological mechanisms may have implications for both sport, exercise and for the prevention and treatment of disease.

Numerous physiological systems detailing a complex phenotype are required for elite endurance performance therefore it is likely that 'elite status' is polygenic. Eight 'endurance' alleles have previously demonstrated discrete associations with elite endurance athlete status. The human *ACE* gene contains a restriction fragment length polymorphism consisting of the presence (insertion, I) or absence (deletion, D) of a 287 base pair Alu repeat sequence in Intron 16. The renin-angiotensin aldosterone system (RAAS) plays a homeostatic role in the human circulation. Renin catalyses the conversion of Angiotensinogen (AGT) to angiotensin I. Angiotensin I-converting enzyme is responsible for the breakdown of vasodilator kinins while catalysing the formation of the vasoconstrictor angiotensin II. Angiotensin II stimulates adrenal aldosterone release, leading to salt and water retention. These two elements maintain blood pressure and volume before, during and after a marathon competition and would therefore influence aerobic power, $\dot{V}O_2$ kinetics

The alpha actinins cross-link with actin at the z- lines of skeletal muscle and are therefore major contributory structural components. ACTN3 is responsible for the stabilisation of the contractile apparatus of the sarcomere during exercise. However, knock out mice have shown enhanced enzyme expression associated with oxidative capacity and superior

endurance running performance and improved recovery time. Extrapolation of this information lead to the hypothesis that the *ACTN3* XX genotype may confer some advantage to endurance athletes based on an enhanced oxidative capacity and preferential skeletal muscle fibre type proportion to compete in endurance events such as marathon running.

PPARGC1A is thought to indirectly mediate the regulation of several genes encoding key enzymes involved in fatty acid oxidation, and mitochondrial biogenesis through its interaction with specific transcription factors such as nuclear receptor PPARɣ, nuclear respiratory factors 1 and 2 and MEF2, PARAGC1A is thought to influence the fatty acid substrate availability during the later stages of a marathon and its conversion to ATP, to directly fuel skeletal muscle contraction during a marathon and will therefore influence a runners running economy and lactate threshold.

The uncoupling proteins regulate the coupling of oxidative phosphorylation to ATP production used in propulsion during a marathon. Their role is not fully understood however they have been linked to thermogenesis and the uncoupling of respiration from ATP production both important factors in the successful completion of a marathon on race day.

Three hundred and ninety-nine Caucasian marathon athletes donated DNA samples for analysis. In addition, DNA was collected from 676 non-athlete research participants. Of those 1075 samples collected, all 1075 samples were genotyped for actinin, alpha 3 (gene/ pseudogene) *(ACTN3)* (399 athletes and 676 non-marathon controls, 932 samples (399 athletes and 533 non-marathon controls) were genotyped for Angiotensin I Converting Enzyme (*ACE*), 673 samples (364 athletes) were genotyped for angiotensinogen (*AGT)* and 705 samples (399 athletes) peroxisome proliferator receptor 1

alpha (*PPARGC1A*) as well as uncoupling protein 3 (*UCP3*). For uncoupling protein 2 (*UCP2*) rs660339 702 samples were genotyped (396 athletes). Finally, for *UCP* r659336, 578 samples were genotyped (272 athletes). Three hundred and six non-marathon controls were genotyped for *AGT, PPARGC1A, UCP2* rs659336 and rs660339, and *UCP3*.

In addition, the collected samples contributed to an investigation into whether genetic characteristics differ at different levels of 'eliteness'. We compared the genotype and allele frequency distributions in 'elite' and 'sub- elite' marathon runners with those of a non-athlete population. Marathon personal best times (PBs) were verified and used to determine elite (males <2.5 h; females <3 h) or sub-elite (males 2.5-3 h; females 3-3.5 h) status. Chi-squared analysis was used to compare genotype and allele frequency distributions between athletes and non-marathon controls, while a genotype-dependent difference in marathon PB was investigated using a one-way analysis of variance for both males and females.

Analysis of the *AGT* rs699 polymorphism revealed over-representation of the TT genotype and T allele in athletes compared to non-marathon controls. This over-representation of the TT genotype and T allele was also noted when sub-elite athletes were compared to non-marathon controls.

The *PPARGC1A* rs8192678 polymorphism analysis showed the A allele tended to be more frequent in athletes than non-marathon controls ($\chi^2 = 2.988$, $p = 0.084$). The minor A-allele was over represented 9.2% in the elite male marathon athletes when compared to non-athlete controls ($\chi^2 = 6.871$, $p = 0.03$). An association was also reflected in the male elite marathon cohort towards the minor AA genotype ($\chi^2 = 6.890$, $p = 0.04$) when compared to non-marathon controls. Further to this, a tendency towards the minor A allele

was seen when the male elite marathon group was compared to the male sub-elite marathon group (χ^2 = 2.986, p = 0.084). In the female cohort, there was a 7.8% higher AA genotype frequency in sub-elite marathon athletes when compared to non-marathon controls (χ^2 = 7.193, p = 0.04)

Tendency for a higher AA frequency in sub-elite vs. elite marathon athletes (χ^2 = 5.425, p = 0.066). When considering PB, in women the *PPARGC1A* GG genotypes ran the marathon approximately 5 min 38 s faster than other genotypes (p = 0.022), which is generally consistent with previous literature.

UCP2 rs660339 analysis revealed A genotype apparent difference was recorded when male elite and sub-elite athletes were compared to non-marathon controls independently (elite χ^2 = 11.173, p = 0.001; sub-elite χ^2 = 17.584, p = 0.01) via Pearson's-Chi squared. In the female athletes, a genotype association was observed when compared to non-marathon controls (genotype χ^2=8.376, p = 0.02)

The female elite athletes also reflected a genotype association when compared to non-marathon controls (genotype χ^2 = 8.942, p = 0.02)

Our findings suggest that the *AGT* rs699, *PPARGCIA* and *UCP2* rs660339 polymorphisms are associated independently with marathon performance. In addition, it is reported that *ACE* I/D, *ACTN3* R577X, *UCP2* rs659366 and *UCP3* rs1800849 polymorphisms are not associated with elite or sub- elite marathon performance when either analysed a whole cohort or individually as males and females. TGS analysis revealed difference in the combined polygenic profile between athletes and controls (t = 4.130 p = 0.000041).

III. TABLE OF CONTENTS

IV. TABLE OF FIGURES

V. TABLE OF TABLES

VI. SUPERVISORY TEAM

Dr Stephen Day

Dr Robert M. Erskine

Dr Alun G. Williams

VII. ABBREVIATIONS

ACE	Angiotensin Converting Enzyme
ACTN2	α-actinin 2
ACTN3	α-actinin 3
AGT	Angiotensinogen
ANGI	Angiotensin I
ANGII	Angiotensin II
ATP	Adenosine Triphosphate
BMI	Body Mass Index
bp	Base Pairs
COOH	Carboxylic Acid
DAG	Diacylglycerol
DNA	Deoxyribonucleic Acid
DZ	Dizygotic
EDTA	Ethylenediaminetetraacetic Acid
GDP	Guanosine Diphosphate
GTP	Guanosine Triphosphate
GWAS	Genome Wide Association Study
HTA	Human Tissue Authority
IAAF	International Association of Athletics Federation
IMP	Inosine Monophosphate
IP_3	Inositol trisphosphate
IPR	Institute of Performance Research
LBM	Lean Body Mass

LIF	Leukaemia Inhibitory Factor
LV	Left Ventricular
MLSS	Maximal Lactate Steady State
MSTN	Myostatin
MZ	Monozygotic
NH_3	Ammonia
PB	Personal Best
PBS	Phosphate Buffered Solution
PCR	Polymerase Chain Reaction
P	Phosphate
PIP_2	Phosphatidylinositol 4,5-Bisphosphate
PO_2	Partial Pressure of Oxygen
RAAS	Renin Angiotensin Aldosterone System
SNP	Single Nucleotide Polymorphism
TGS	Total Genotype Score
UK	United Kingdom
$\dot{V}O_2$	Maximal Oxygen Uptake
W	Watts

VIII. GENES OF INTEREST

GENE NAME	ENCODES FOR:
ACE	Angiotensin I converting enzyme
ACTN3	α-Actinin, alpha 3 (gene/ pseudogene)
AGT	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)
PPARAGC1A	Peroxisome proliferator activated receptor gamma, coactivator 1 alpha
UCP2	Uncoupling protein 2 (mitochondrial, proton carrier)
UCP3	Uncoupling protein 3 (mitochondrial, proton carrier)

**** According to convention, in this thesis, gene abbreviations are shown in* italics*, proteins are not italicised.***

1 INTRODUCTION

1.1 A BRIEF HISTORY OF MARATHON AND THE MODERN MARATHON OF 26.2 MILES.

This review will give a brief history of the battle of marathon and recite a selection of the modern history of the marathon race focussing on those races that feature in the world marathon majors. The world marathon majors, founded in 2006, are series of six races (Tokyo, Boston, London, Berlin, Chicago, and New York) (Carter, 2015). The Majors also include the World Championships and the Olympics in the appropriate years of competition (Boston Athletic Association, 2015c). The world marathon majors offers a $1 million prize purse split equally between the top male and female marathoners (Carter, 2015).

The battle of marathon 490BC is a pivotal point in European history and is credited with the inception of the marathon race. The legend of the Greek messenger Pheidippides running from Marathon to Athens with news of the Greek victory over the Persians became the inspiration for this athletic event. Tragically, the legend tells of Pheidippides death after the race due to exhaustion (Perros, 2001, Martin and Gynn, 2000).

The marathon race, introduced at the first Modern Olympics in Athens in 1896, was originally run between Marathon and Athens on a course of approximately 25miles/ 40km in celebration of the ancient glory of Greece (Martin and Gynn, 2000, Perros, 2001, Clark, 2003, Ostapuk, No Year). The distance was extended to 26 miles 385 yards to cover the distance from Windsor Castle to the Olympic Stadium in White City, London, for the 1908 London Olympics (A&E Television Networks, 2014, Ostapuk, No Year). In Geneva, 1921,

the International Association of Athletic Federations (IAAF) standardized the marathon distance to 26 mile 385 yards /42.195 km based on the 1908 event held in London (Martin and Gynn, 2000)

1.1.1 **Boston Marathon**

The Boston marathon is the oldest of the city road races, inspired by the introduction of the marathon race at the modern Olympics in 1896. On April 19th 1897, the Boston Athletics Association Games concluded its athletic competition with a marathon of 24.5 miles (39.4 km). The Boston Marathon is run on ‘Patriots Day’ annually (Boston Athletic Association, 2015a). The first race in 1897 started with 15 runners, in 2015, 30,333 runners crossed the starting line (Boston Athletic Association, 2015b, Boston Athletic Association, 2015e). In 1924 the race was lengthened to the IAAF standard of 26 miles 385 yards (Boston Athletic Association, 2015b). Women were not permitted to enter the Boston Marathon officially until 1972. Roberta "Bobbi" Gibb was recognized retrospectively, by the race organizers, as the first woman to run the entire Boston Marathon in 1966 (Boston Athletic Association, 2015b). In 1967, Kathrine Switzer, who had registered as "K. V. Switzer", was the first woman to run and finish with a race number. She finished despite race official Jock Semple trying to rip off her numbers and eject her from the race (Carter, 2015, Boston Athletic Association, 2015b). Because the Boston course drops 140 meters/459 feet from the start to the finish, and the start is west of the finish, allowing for a tailwind, the Boston Marathon does not satisfy two of the criteria necessary for the ratification of world records (IAAF, 2011). This means that the 2:03:02 course record time by Geoffrey Mutai of Kenya On April 18, 2011 was the fastest marathon time ever (Boston Athletic Association, 2015d). However, the IAAF did not recognize this impressive time as a world record. Rita Jeptoo of Kenya holds the current women’s course record set in 2014 of 2:18:57 (Boston Athletic Association, 2015d).

1.1.2 **New York Marathon**

The New York Marathon takes place on the first Sunday in November. Fred Lebow founded the race in 1970 (New York Road Runners, 2015). The first race consisted of 127 competitors running four laps around the Park Drive of Central Park. Gary Muhrcke won the race in 2:31:38. Only 55 runners crossed the finish line (Fitzgerald and Fraioli, 2014). To commemorate the race and intended as a one-time event it was proposed in 1976 that the race be run through five boroughs of New York (Staten Island, Brooklyn, Queens, Bronx and Manhattan) this has remained the course route since (New York Road Runners, 2015). The 2014 New York marathon had grown to a record number of 50869 starters and 50564 finished the race (Dalek, 2014). Geoffrey Mutai of Kenya, who ran 2:05:06 in 2011, holds the New York Marathon, men's course record. In 2003, Margaret Okayo set the current women's course record of 2:22:31(Fitzgerald and Fraioli, 2014, Cryer, 2014).

1.1.3 **Berlin Marathon**

In 1974, a group of runners from one of Germany's most prestigious athletics clubs, SC Charlottenburg founded the Berlin Marathon. There were 244 finishers of the first race. Günter Hallas of West Germany won the men's race in 2:44:53 (GmbH, 2015a, bytepark GmbH, 2015a). In 1981, the race moved to the city centre of West Berlin from the Grunewald. Although the Berlin wall collapsed in November 1989 the city was not officially reunited. On the 30TH September 1990, three days before the reunification of the city, the course of the Berlin Marathon led through Brandenburg Gate and both parts of Berlin (GmbH, 2015a, bytepark GmbH, 2015a). The 2014 race saw 28,946 competitors, complete the Berlin Marathon. Seven men's world records have been set at the Berlin Marathon(bytepark GmbH, 2015a). Dennis Kimetto of Kenya set the men course record and current world record of 2:02:57 in 2014 (GmbH, 2015b). The women's course record is 2:19:12 set by Japans Mizuki Noguchi in 2005 (GmbH, 2015b).

1.1.4 **Chicago Marathon**

A marathon in Chicago was first run in 1905. Twenty runners registered for the first race, 15 actually started the race, and seven finished. However, Lee Flaherty founded the first 'modern' Chicago marathon in 1977 with 4,200 starting and 2,128 crossing the finish line. Dan Cloeter won the first men's race and Dorothy Doolittle the women's in 2:17:52 and 2:50:47 respectively. In 2014, The Chicago marathon had swelled to 40,801 finishers. The current course records are 2:03:45 held by Dennis Kimetto of Kenya set in 2013 and 2:17:18 set in 2003 by Paula Radcliffe of the UK (Bank of America Corporation, 2015).

1.1.5 **London Marathon**

John Didsley and Christopher Brasher founded the London Marathon in 1981 after Brasher travelled to New York to compete in the marathon and write about his experience (London Marathon Ltd, 2015). The first race had 7055 starters. This grew dramatically in the following years and at the 35th annual London marathon in 2015, 37,675 runners crossed the finish line (bytepark GmbH, 2015a). Many of the runners were hand in hand to commemorate the winners of the first race in 1981, Dick Beardsley from the USA and Inge Simonsen of Norway in 2 hours 11minutes 48 seconds. Joyce Smith of the UK was the first woman to cross the line in 2 hours 29 minutes 57 seconds (London Marathon Ltd, 2015). The current course record for the London marathon is held by Wilson Kipsang Kiprotich of Kenya who finished in a time of 2:04:29 in 2014. Paula Radcliffe is the women's course record holder. In 2003 Radcliffe completed the London marathon in 2:15:25 to set a new world record (bytepark GmbH, 2015a).

1.1.6 **Tokyo Marathon**

In 2007, Shintaro Ishihara former governor of Tokyo and, Yokei Kono former Japan Association of Athletics Federation president established the Tokyo Marathon and the Tokyo Marathon Foundation (Tokyo Marathon, 2015, bytepark GmbH, 2015b). The Tokyo course became part of the world marathon majors in 2012. At the inaugural race 25,000

runners started. In 2015, 35,556 started and 35,310 finished the race (bytepark GmbH, 2015b). The winner of the first Tokyo marathon Daniel Njenga of Kenya completed the race in 2:09:45. The current course record for men and women were both set in 2014. Dickson Chumba of Kenya won the race in 2:05:42 and Tirfi Tsegaye of Ethopia finished in 2:22:23(bytepark GmbH, 2015b).

The current world records for marathon recognised by the IAAF are 2:02:57, for men's marathon set by Dennis Kimetto of Kenya on September 28, 2014 at the Berlin Marathon (IAAF, 2015a). The IAAF recognizes two world records for women both held by Paula Radcliffe of the UK. A "Mixed Gender" record of 2:15:25, set on 13th April 2003, at the London Marathon, and a "Women Only" record of 2:17:42 on 17th April 2005, also at the London Marathon (Baldwin, 2011, September 20, IAAF, 2015b).

The marathon record progression towards a two hour marathon is the result of combined technological developments that allow improved training regimens including: nutritional strategies, improved footware, and specific year round training facilities such as environmental chambers that provide favorable ambient conditions to induce explicit training outcomes and, the birth of individuals with favorable genetic architecture according to Williams et al., (2011).

1.2 INTER-INDIVIDUAL VARIABILITY IN HUMAN PERFORMANCE

There is considerable difference in performance phenotypes between elite athletes and non-athletes. According to Bouchard et al., (1997), three factors contribute to inter-individual variability in observed human performance: environmental factors such as diet and habitual physical activity levels, genetic variation and experimental error. The individual and combinatory contribution of these traits to human physical performance is considered complex. Complex traits are phenotypes affected by both multiple genetic and

environmental factors. These genetic factors comprise multiple genes and perhaps even multiple polymorphisms within those genes contributing in an additive effect to complete a polygenic profile (Bouchard et al., 1997, Williams and Folland, 2008, Ahmetov and Fedotovskaya, 2012). Gene-gene interactions (the effect of one variant being dependent, in part, on the genetic context in a given individual) add further complexity. The environmental influences, as well as gene-environment interactions (the effect of one variant being dependent, in part, on the environmental exposure an individual experiences), provide further factors to be considered (Bouchard et al., 1997). Nevertheless, careful selection of important and robust phenotypes and evidence-based selection of candidate genes can provide a solid basis on which to base hypothesis-based studies of genetics in human performance.

1.3 THE GENETIC CONTRIBUTION TO HUMAN PERFORMANCE

Detailed, human variation investigations were classically compared using twin pairs. Twin pairs usually share very similar or identical environments and thus the environmental selection pressures are minimised. Therefore, monozygotic twins (who share an almost identical genomic profile) usually show a higher correlation in certain phenotypes than dizygotic twin pairs (whose genomic profiles are non-identical), thus demonstrating the heritability, or genetic component in determining that trait (Bouchard et al., 1986a, Bouchard et al., 1986b, De Moor et al., 2007). Differences record between twin pairs are usually attributed to environments, rather than genetics. Though, recent studies have shown that many environmentally induced differences are reflected in the epigenome. Bouchard et al., (1986a) demonstrated in 106 monozygotic twin pairs, of both sexes, that the maximal rate of oxygen uptake ($\dot{V}O_{2\,max}$) showed less variation than $\dot{V}O_{2\,max}$ data collected in 66 dizygotic twins of both sexes and 42 brothers. The monozygotic twin data was used to report the genetic variance effects. $\dot{V}O_{2\,max}$ was reported to have a genetic inheritance of 40%, indicating a significant genetic component for a key factor of endurance performance (Bouchard et al., 1986a). A more recent study that assessed

sporting achievement directly in female twins in the UK, demonstrated a clear dependence on multiple phenotypes including $\dot{V}O_2$ max, in which the heritability estimate was 66% (De Moor et al., 2007). Thus, twin studies have provided the initial evidence for a genetic contribution to sporting achievement, exercise capacity and trainability. The advent of DNA sequencing subsequently provided a tool to sequence the human genome with 99.99% accuracy (Schmutz et al., 2004). This meant that detailed investigation of genotype was then possible. The candidate gene approach to genetic association studies concentrates on a phenotype of interest and associations between genetic variants within genes that are thought to contribute to that phenotype. Suitable candidate genes for human performance are selected for investigation based on their already known physiological and or functional relevance to the phenotype of interest. This approach is limited as it relies on the theoretical or known physiology of the phenotype of interest. The rationale to focus on areas specific loci of the genome that are known to be biologically relevant to a phenotype is that mutations may directly alter the function of a gene and thus be causative in the phenotype of interest. The candidate gene approach to identifying the genetic contribution to human performance started with the ACE gene in 1998 (Montgomery et al). Further advances in sequencing technology and other laboratory tools such as gene chips, plus the development of large data sets such as biobanks and the HapMap Project (Gibbs et al., 2003, The International HapMap Consortium, 2005, Frazer et al., 2007) allowed completion of genetic investigations on a mass scale of several hundred polymorphisms at a time, facilitating the investigation by sports scientists into the genetic variability in human performance. As such, the field of human performance and exercise has moved on from twin studies and now uses a variety of more precise methods in molecular biology. Genetic case-control studies that compare athlete populations with non-athlete controls and cross-sectional studies completed in non-related individuals attempt to associate genotype with phenotype. An example of such a project investigating the genetic contribution to physical fitness is the Health, Risk factors, Exercise Training and Genetics (HERITAGE) study (Bouchard et al., 1995). In HERITAGE, 700 sedentary individuals from almost 200 families participated in a multi-centre collaboration to collect

data before and after a period of controlled endurance training. Sedentary behaviour is defined as any waking behaviour characterized by an energy expenditure ≤1.5 METs while in a sitting, reclining or lying posture (Tremblay et al., 2017). The aim of the HERITAGE project was to investigate the role of likely polymorphisms and their genetic contribution to the cardiovascular, metabolic and hormonal responses to endurance training. These studies range in their genomic depth from single point mutations to full genome scans (Bouchard et al., 1998, Bouchard et al., 1999, Bouchard et al., 2000) (not equivalent to whole genome sequencing) and provide insight into several candidate genes thought to be associated with human performance.

1.4 IDENTIFYING CANDIDATE GENES

The main challenge for genetics researchers in human performance is to determine which gene or combination of genes is associated with the variation in the phenotype of interest. In essence, the aim is to pinpoint the precise genetic location where allele variation exists that influences the nature or extent of expressed protein, and thus phenotype (Bouchard et al., 1997, Botstein and Risch, 2003) or to understand the activity and expression of protein–coding genes and their modulation by the regulome. There are generally two approaches used to identify candidate genes these are outlined in the following sections.

1.4.1 The Candidate Gene Approach

The initial approach used by sports scientists and geneticists to identify sporting genotype-phenotype associations was the 'candidate gene approach' (Bouchard et al., 1997). The candidate gene approach uses well-established physiological theory to identify key metabolic or structural proteins that can be used to identify candidate genes. Screening of the identified genes encoding these key proteins ascertains common genetic sequence variations or polymorphisms that alter translation of the gene and thus protein production - the physiological significance of this polymorphism is then investigated (Lander and Schork, 1994). An example of this in humans is a rare mutation in the myostatin (*MSTN*) gene. Myostatin is a protein produced by muscle cells that acts to inhibit myogenesis and growth, thus inhibition of myostatin leads to muscle hypertrophy (Schuelke et al., 2004). The mutation is located at the splice site in the first intron and alters splicing that attempts to translate the first 108 bp of the sequence in that intron into an amino acid sequence. This completely alters the protein produced, effectively knocking out the gene because a premature stop codon in the first intron is read, instead of spliced out. In 2004, a 4-year old German boy with greater muscle mass than his peers of similar age, and who was considerably stronger than other children his age had the rare mutation in both copies of the myostatin gene that caused muscular hypertrophy (Schuelke et al., 2004). The myostatin mutation was identified in that individual using existing physiological

knowledge – i.e. the candidate gene approach. Further genetic variations, relevant to human physical performance, that have been identified using the candidate gene approach, include the frequently studied *ACE* and *ACTN3* gene polymorphisms. This approach to candidate gene selection makes understanding and extrapolating the polygenic nature of human physical performance rather difficult as gene polymorphisms are often investigated individually rather than in combination. Though, because the physiology of the phenotype of interest has often been mapped in advance those proteins can be traced back to genes and SNIPs that are likely to have a causative effect. Although it should be noted, should a list of identified target genes or SNIPs not include the causative variant then this method will fail to detect an association. In an attempt to fully characterise the polygenic nature of human performance, increasingly sports scientists are favouring another method the Genome Wide Association Study (GWAS).

1.4.2 **Genome-wide Association Studies**

The GWAS is a second approach used to identify candidate genes. This method relies on a sufficiently large (and thus statistically powered) sample in which to detect an association between the genotype and the phenotype of interest. In contrast to the candidate gene approach, GWAS do not begin with hypotheses about associations between specific genetic variants and relevant phenotypes. Rather, GWAS studies are 'hypothesis-free' and simultaneously consider possible associations between large numbers of genetic variations and a given phenotype. Thus, GWAS identifies chromosomal regions of interest or loci associated with each other and correlated with the phenotype of interest (Visscher et al., 2012). Recently, an international consortium of researchers in exercise genomics has been established (Tanaka et al., 2016) which plans to, in time, use GWAS and other approaches (e.g. whole genome sequencing) to address relevant research questions. However, effective use of the GWAS approach requires many thousands of participants to provide the necessary statistical power to identify true genotype-phenotype associations from a mass of probable false positive associations – the generally accepted level of significance is 5×10^{-8} (www.ebi.ac.uk/gwas/docs/about).

Currently, cohorts of hundreds of thousands of research participants (whether elite athletes or other individuals with relevant phenotype data) do not exist. Further to this funding for this type of research is likely to be of limited availability for sporting purposes and most likely to be provided to the medical and health sectors to investigate pathology, this means the feasibility of doing such a study may be limited as the causal variants for most phenotypes of interest to human sporting performance are unlikely to be available in array-based SNIP datasets and will need to be produced at considerable cost. Therefore, the candidate gene approach remains a valuable method to investigate genotype-phenotype associations in exercise science, without the need for extremely large cohorts and commensurate levels of funding.

1.5 SELECTING CANDIDATE GENES

In the candidate gene approach, it is important to have a physiological rationale for the proposed association between a candidate gene and the resulting phenotype of interest (Bouchard et al., 1997). The method of selection of the phenotype must be consistent across the test population and inclusion criteria must be met by all research participants. It is important when selecting a test population to ensure factors such as gender, age ethnicity and athlete status are taken into consideration should the research sample not be homogenous (Ahmetov and Fedotovskaya, 2012). This reduces experimental error and the contribution of environmental factors. Usually research participants are screened against pre-determined criteria to ensure validity, reliability and reproducibility of any genetic associations.

1.6 SUMMARY

Identifying genetic polymorphisms within individual candidate genes that account directly for the phenotype of interest is a challenging process. Sporting performance is a combination of several traits, each of which themselves could be considered complex.

The extent of the individual SNIP contribution to genetic variation and inter-individual variability especially in a complex trait such as human performance continues to be debated due to a lack of clarity about the associated physiology of individual polymorphisms or groups of polymorphisms of interest and conflicting reported data in groups of athletes with limited definition or description. The noted variation may be attributed to the likely polygenic nature of an athlete's genetic profile (Williams and Folland, 2008, Ahmetov et al., 2009), though the athletes and phenotypes of interest need to be well defined. In addition, it should be noted, because of the complex nature of the phenotypes associated with human performance, there is difficulty in quantifying the genetic contribution alone.

The work described in this thesis investigates genetic characteristics potentially associated with human performance, specifically marathon running. Endurance running including marathon running is both a mass-participation sport amongst non-elite athletes as well as a highly competitive elite sport on an international scale. Due to the scale of this project the candidate gene approach was adopted to investigate selected polymorphisms that were chosen based on the existing literature.

In Chapter 2, a review of the literature is presented. It first addresses the physiological characteristics of elite marathon runners with specific focus on aerobic power, running economy, lactate threshold, $\dot{V}O_2$ kinetics and skeletal muscle fibre proportion. A review of the current evidence regarding genetic associations with endurance phenotypes and elite endurance athlete status is further presented in Chapter 2, which includes reference to marathon performance where appropriate. Lastly, the evidence regarding genetic associations with the aforementioned phenotypes in eight selected gene polymorphisms is also reviewed. In Chapter 3, some of the core methods that are common to each of the following experimental chapters are described. Chapters 4-7 each consist of an investigation into the genotype of elite marathon runners and their individual performance

capabilities (personal best competitive marathon times). Chapter 4 addresses two genes that form part of the renin angiotensin aldosterone system (RAAS), namely *ACE* and *AGT*. Chapters 5, and 6 address a single gene each (*ACTN3*, and *PPARGC1A*, respectively). Chapter 7 addresses three polymorphisms in two genes encoding for uncoupling proteins (*UCP2*, *UCP3*). Chapter 8 incorporates the aforementioned genes into a Total Genotype Score analysis. The final chapter (Chapter 9) integrates the findings and conclusions from the preceding five chapters and draws conclusions based on the original data contained in this thesis.

The aims of this research programme were to (1) compare genetic characteristics of elite athletes, sub-elite athletes and non-athletes; (2) compare personal best competitive marathon running performances between runners with differing genetic characteristics. The objectives were therefore to address aims (1) and (2) for specific variations in the genes identified above, namely *ACE*, *AGT*, *ACTN3*, *PPARGC1A, UCP2 and UCP3*.

2 LITERATURE REVIEW

2.1 THE PHYSIOLOGY OF MARATHON RUNNING

2.1.1 Introduction

The important determinants of the elite marathon runner's endurance performance are multifactorial and thus complex. This is evidenced by the large body of knowledge published from the 1970's until the present day. The most common determinants of endurance performance were summarised by Bouchard et al. (1997) who suggested that each was actually a heading for a larger group of factors, traits and determinants rather than a single characteristic. These included, physiological traits, submaximal exercise tolerance and efficiency, thermoregulation, nutritional status, biomechanical factors, body composition, $\dot{V}O_{2\,max}$, social factors and a group termed 'others'. In particular, however, running the marathon distance of 42.195 km (26 miles 285 yards) on a seriously competitive basis undoubtedly requires an extremely high level of aerobic physiological function. Thus, elite athletes provide a unique model in which to base scientific research into human physiology at one extreme of the performance-disease continuum. Elite marathon runners have typically undergone extensive training to produce large physiological adaptations (e.g. in left ventricle size, the oxidative capacity of skeletal muscle, etc.) that are probably a result of an interaction between the exercise training stimulus and favourable genetic characteristics (Ruiz et al., 2009). World marathon records (as at June 2017) are 2 h 2 min 57 s for men and 2 h 15 min 25 s for women (http://www.iaaf.org/records/by-category/world-records). The history of the progression of marathon records (and other elite performances close to those records) to their current levels has been described in detail, with interesting conclusions drawn regarding the influence of BMI, geographic ancestry and ambient environmental conditions such as temperature and humidity (Marc et al., 2014). Sophisticated mathematical modelling predicted further improvements of up to 3% from the 2008 records, with rather larger improvements anticipated for men than for women, although the record for men was still

expected to remain in excess of 2 h (Denny, 2008). Indeed, since that paper in 2008, the marathon record for men (2 h 3 min 59 s at that time) has been improved on three separate occasions to its current level, while the record for women has not changed. It is beyond the scope of this piece to extensively review the contribution to marathon running of all of its determinants. However, the primary physiological factors will be addressed in the following text (see Figure 1).

2.1.2 **Important Measurable Physiological Parameters**

The performance of an endurance athlete in a sport such as marathon running is influenced by a number of well-documented factors. These include physique, biomechanical, physiological, behavioural, psychological and social characteristics (Bouchard et al., 2008). Classical exercise physiology knowledge, synthesised in several respected textbooks (Astrand and Rodahl, 1986, McArdle et al., 2009) from extensive original research, confirms that maximal performance during a physical effort in excess of 2 h that is required to complete a marathon requires high rates of aerobic metabolism to be sustained and used efficiently during running. More specifically, high aerobic power relative to body mass (maximal rate of oxygen uptake; $\dot{V}O_{2\,max}$), the ability to sustain a high percentage of that aerobic power for prolonged periods (strongly related to the lactate threshold concept, but also influenced by any sustained rise in oxygen uptake beyond that predicted from the relationship between VO_2 and exercise intensity at lower intensities ($\dot{V}O_2$ kinetics) and an efficient conversion of the energy derived from energy substrates for ATP resynthesis and subsequent ATP breakdown into horizontal motion during running (running economy) are the key parameters that determine marathon running performance (Joyner, 1991, Jones and Carter, 2000). These concepts are discussed in more detail in the following sections, in addition to skeletal muscle fibre type proportion that has relevance to each of the other concepts discussed.

2.1.2.1 Aerobic power

Aerobic power, also known as the maximal rate of oxygen uptake or $\dot{V}O_{2\,max}$, is defined as the maximal rate at which oxygen can be taken up and utilised by the body during intense exercise at normal barometric pressure (Bassett and Howley, 2000). In humans, $\dot{V}O_{2\,max}$ is usually limited by the ability of the cardiorespiratory system to deliver oxygen to the exercising muscles although a high capacity for those activated muscle fibres to effectively use the oxygen delivered (thus Type I fibres are more advantageous) is also required (Bassett and Howley, 2000, McPhee et al., 2009). $\dot{V}O_{2\,max}$ is a clearly definable, measurable trait that has high validity and reliability in predicting endurance performance

(Bassett and Howley, 2000, Costill, 1970, Costill et al., 1971b, Saltin and Astrand, 1967). Traditionally $\dot{V}O_{2\,max}$ has been the best laboratory measure for understanding long distance competitive running performance (Foster, 1983). Elite runners have exceptionally high $\dot{V}O_{2\,max}$ usually above 70 ml kg-1 min-1 (Costill, 1970, Costill et al., 1971b, Saltin and Astrand, 1967) a much higher level than those reported in untrained individuals and thought to be the result of training. Elite marathon runners complete a marathon in just over 2 hours and these amazing performances require exercise at a sustained intensity of 80%-90% of an athlete's $\dot{V}O_{2\,max}$ (González-Alonso, 2007). The repeated correlation between fast running performance and high $\dot{V}O_{2\,max}$ (Farrell et al., 1979, Hagan et al., 1981, Foster, 1983, Hagan et al., 1987) supports the theory that a high $\dot{V}O_{2\,max}$ is essential to compete in marathon at an elite level (Foster, 1983).

2.1.2.2 Economy

Running economy is considered a good predictor of racing performance. In trained runners with similar values of $\dot{V}O_{2\,max}$, running economy correlates more strongly with performance than does $\dot{V}O_{2\,max}$ itself (Saunders et al., 2004, Conley and Krahenbuhl, 1980). Running economy is defined as the energy demand for a given velocity of submaximal running (Lacour and Bourdin, 2015). Running economy is determined by measuring the consumption of oxygen ($\dot{V}O_2$) of an individual at a given (practically relevant) running speed (Maud and Foster, 1995). In the marathon, athletes who are able to consume less oxygen while running at a given velocity have a better running economy than those who require more oxygen at the same velocity. Running economy may therefore be influenced by running style learned through training. All else being equal, a more economical runner will be able to either run faster than a less economical competitor at a given $\dot{V}O_2$, or run at the same velocity as a competitor at a lower $\dot{V}O_2$ and thus a lower percentage of $\dot{V}O_{2\,max}$ (thus probably reducing heart rate, lactate production and the utilisation of carbohydrate as an energy substrate). Accordingly, when compared with other marathon runners, one world champion marathon runner demonstrated little aerobic power superiority to other elite runners (Costill et al., 1971b). This suggests marathon

running success is influenced by running economy and the ability to utilise a large fraction of a well-developed $\dot{V}O_{2\,max}$ (Costill et al., 1971b). Saunders et al. (2004) reported that, in highly trained or elite athletes, running economy was influenced a number of physiological and biomechanical factors including increased mitochondrial density and function and the activity of various oxidative enzymes. Proliferation of mitochondria and changes in mitochondrial enzyme activity have long been known to occur in skeletal muscle in response to endurance exercise training (Holloszy, 1967). Bassett and Howley (2000) confirmed that endurance training causes an increase in the activity of mitochondrial enzymes that, for any given $\dot{V}O_{2,}$ will improve endurance performance by enhancing fat oxidation and decreasing lactic acid production. This metabolic adaptation of skeletal muscle towards a Type I fibre phenotype is thought to be crucial for improvements in economy, efficiency and submaximal endurance performance (Bassett and Howley, 2000, Coyle et al., 1992), of which marathon running is an extreme but obviously most pertinent example. Jones and Carter summarise evidence for training-induced improvements in running economy (Jones and Carter, 2000).

2.1.2.3 Lactate Threshold

As explained expertly by Jones and Carter (Jones and Carter, 2000), lactate threshold (the exercise intensity corresponding to an increase in the concentration of blood lactate above resting levels) is another key, innate and readily-measured parameter related to endurance running performance. Exercise above the lactate threshold produces a nonlinear increase in metabolic, respiratory and perceptual stress (Katch et al., 1978, Simon et al., 1983). Sustained exercise above the lactate threshold is associated with metabolic acidosis (accumulation of H^+ in muscle tissue and/or blood) which, at the level of the skeletal muscle fibre, inhibit phosphofructokinase (the rate limiting enzyme in glycolysis) and inhibit the binding of Ca^{2+} to troponin, thus contributing to local muscular fatigue (Sahlin, 1992). On the other hand, the accumulation of H^+ is indicative of increased muscle glycogen utilisation at the expense of deriving energy for ATP resynthesis from the breakdown of fatty acids (Boyd et al., 1974), and higher rates of glycogen utilisation during

a marathon are strongly related to premature fatigue (Burke, 2007, Costill et al., 1971a, Costill et al., 1971b). It is important to note, however, that an exercise intensity corresponding to lactate threshold (as defined at the beginning of this paragraph) is actually exceeded by many competitive athletes during competition and some training sessions (Jones and Carter, 2000). The concept of the maximal lactate steady state (MLSS) is therefore probably more practically relevant than lactate threshold per se to competitive athletes. MLSS may be defined as the highest exercise intensity that can be maintained with a stable or only a small increase (< 1 mmol L^{-1}) in blood lactate concentration during the period 10-30 min after the onset of constant intensity exercise (Jones and Doust, 1998). Thus, an increase of blood lactate concentration above resting levels (lactate threshold) is not as accurate an indicator of fatigue related to acidosis as MLSS. Nevertheless, the lactate threshold and MLSS are very closely related in conceptual and physiological terms, and lactate threshold is much more easily measured in a single, short laboratory test. Elite endurance athletes typically have lactate threshold values approaching ~80% of $\dot{V}O_{2\,max}$ (and MLSS a little higher), which is considerably higher than other athletes and non-athletes (Jones and Carter, 2000). Elite marathon runners typically have $\dot{V}O_{2\,max}$ values ranging from ~70 to ~85 $ml \cdot kg^{-1} \cdot min^{-1}$.

2.1.2.4 VO_2 Kinetics

At constant exercise intensities below the lactate threshold, ventilatory and pulmonary gas exchange responses respond quite rapidly to the onset of exercise to attain a new steady state within 2-3 min (Whipp and Wasserman, 1972). However, at exercise intensities above MLSS there is an additional, relatively slow increase in ventilation rate and $\dot{V}O_2$ and that probably reflects the innate recruitment of an increasing recruitment of faster, larger, less efficient motor units as fatigue develops within muscle fibres of the motor units recruited nearer the initial onset of exercise (Barstow et al., 1996). That secondary increase in $\dot{V}O_2$ during constant load exercise above MLSS has become known as the $\dot{V}O_2$ slow component (Jones and Carter, 2000). All else being equal, a smaller $\dot{V}O_2$ slow component will allow a marathon runner to complete a marathon more quickly than a

competitor even if $\dot{V}O_{2\,max}$ and running economy at intensities below MLSS are comparable between the two runners. Therefore, $\dot{V}O_2$ kinetics, although related to $\dot{V}O_{2\,max}$, lactate threshold (or MLSS) and running economy at certain intensities, has relatively recently been added to those other three concepts as an important parameter that (with appropriate real-time measurement of pulmonary gas exchange) is now widely considered one of the four main measurable components of running performance during laboratory assessment of endurance runners.

2.1.2.5 Skeletal Muscle Fibre Type

Skeletal muscle fibres are divided into two sub groups based on specialized contractile and metabolic properties and on distinctive patterns of muscle gene expression (Naya et al., 2000). Type I fibres are smaller in diameter and produce force as a relatively slow 'twitch' in response to brief electrical stimulation, but have high levels of mitochondria and capillarity to support oxidative metabolism, are more efficient, and are therefore rightly considered fatigue-resistant fibres. Type II fibres can be denoted IIa or IIx in humans while IIb (faster contractile properties) also exists in rodents. Type II fibres produce a relatively large and fast 'twitch' response to electrical stimulation, have a larger diameter for increased contraction strength and are generally found in larger motor units (Folland and Williams, 2007). Type II fibres rapidly release calcium from an extensive sarcoplasmic reticulum, rely on their large glycolytic enzyme capacity for rapid release of energy from glycogen and thus have lower levels of mitochondria and oxidative enzymes than Type I fibres and are less mechanically efficient (Boron and Boulpaep, 2012).

It is well-established that skeletal muscle fibre composition differs between elite athletes in different sports (Ahmetov et al., 2011). Endurance training like that completed by marathon runners causes various innate changes to the activated skeletal muscle as a result of the high frequency, low force output, activity (Brooke et al., 2005). This leads to increased endurance capacity and fatigue resistance primarily via an increase in the expression of genes that are components of mitochondrial proteins and oxidative

enzymes and resultant increases in fatty acid oxidative capacity (Brooke et al., 2005). It is thought that these changes in gene expression, in response to prolonged endurance training and repeated stimuli for activation of some genes and suppression of others, slowly induce the myoplasticity of Type 2 fast twitch glycolytic fibres towards Type 1 slow twitch, high-oxidative fibres (Malisoux et al., 2007, Pette and Staron, 1997). Accordingly, high proportions of Type I muscle fibres are observed in the trained muscle of endurance athletes (Gollnick et al., 1972, Gollnick et al., 1973, Ahmetov et al., 2011) such as marathon runners. It is thought that in both rodent and human skeletal muscle, fibre type characteristics are controlled by calcium-dependent signalling via the calcium, calmodulin-dependent protein phosphatase, calcineurin (Naya et al., 2000, Seto et al., 2013). Calcineurin has been shown to stimulate slow fibre specific gene promotors in *in vitro* cultured rodent skeletal muscle cells (Naya et al., 2000). Calcineurin, once activated, can (via a promoter) upregulate slow twitch Type I fibre gene expression effectively reprogramming skeletal muscle fibres from the fast twitch Type II phenotype to the slow twitch Type I phenotype, thus increasing the proportion of Type I fibres in the muscle. This may positively affect skeletal muscle mitochondrial function and consequently marathon performance, via influences on the parameters more amenable to measurement than muscle fibre type composition itself, namely aerobic power, economy, lactate threshold and $\dot{V}O_2$ kinetics, as described in the preceding sections.

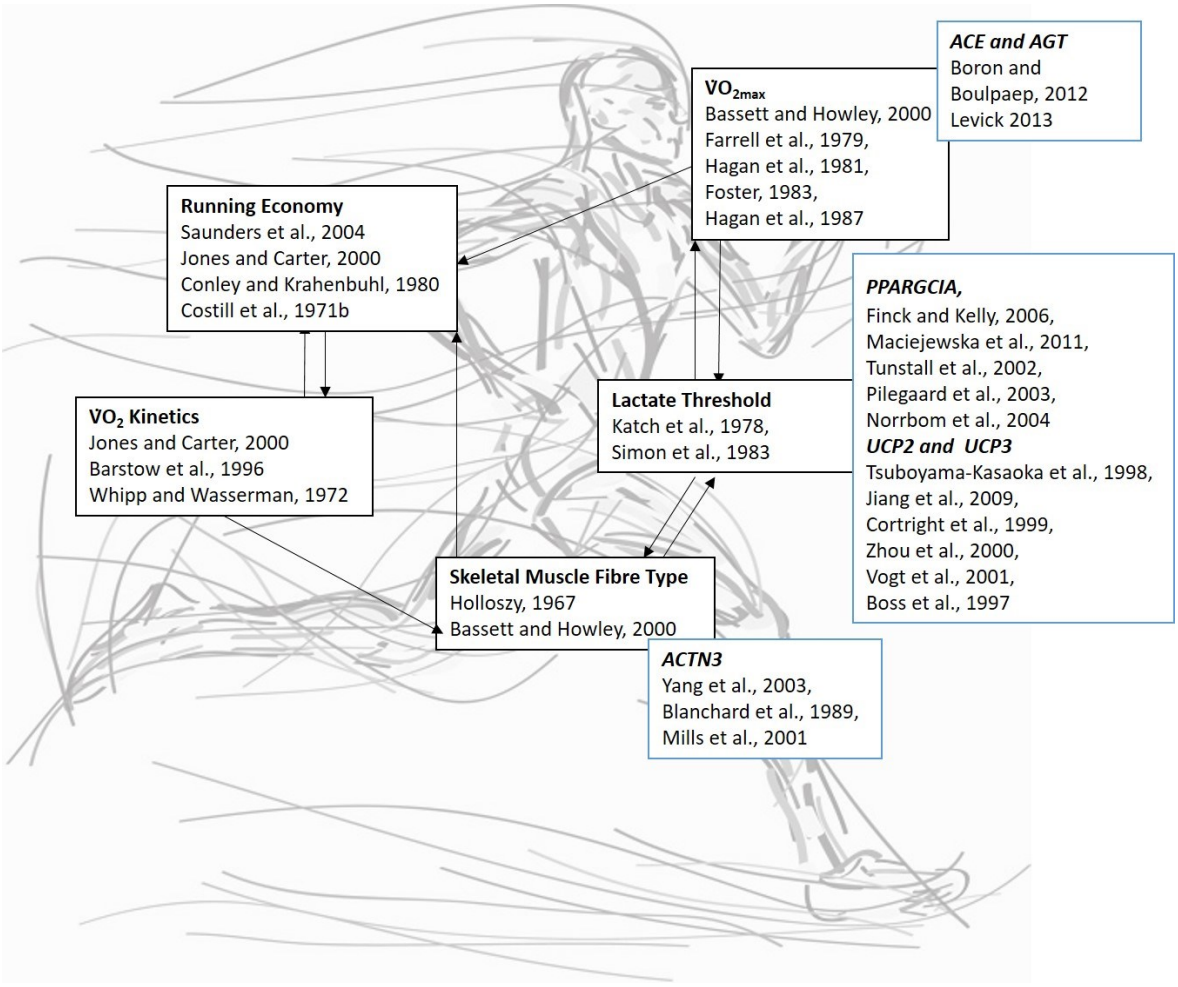

Figure 1 Important measurable parameters of marathon performance

2.1.3 **Conclusion**

To briefly conclude this section on the physiology of marathon running and bring the focus back to real-world competitive performances once again, the effort from athletes, coaches and scientists to develop a male athlete sufficiently to complete the marathon in under 2 h receives much attention (www.sub2hrs.com). For example, a special issue in the Journal of Applied Physiology (Joyner et al., 2011) with contributions from many authors in many separate opinion pieces provides several novel views and perspectives on this fascinating topic. It is illuminating to read the views of some eminent scientists about the future of sport performance in this historic event, especially when they do not appear to agree. Of particular interest in the context of the present thesis is the article in that special issue by Williams regarding the importance of genetics in the potential progression of the male

marathon record towards or even below the 2 h ‘barrier’ (Williams, 2011). In this article Williams commented on the prominence of the genetic contribution to marathon performance and proposed a TGS approach to describing the combined genetic association of several gene variants with elite marathon running performance.

2.2 GENETIC ASSOCIATIONS WITH ENDURANCE PERFORMANCE AND ELITE ATHLETE STATUS

2.2.1 Introduction

As outlined in the previous sections the primary determinants of endurance performance in elite marathon running are multifactorial, polygenic in nature and thus complex. Although complex in nature, scientists have attempted to elucidate the genetic contribution to endurance performance and elite athlete status for decades. This originated in 1974, with an association study between athlete participation in the 1968 Mexico City Olympic Games and allelic variation in red blood cell antigens or enzymes, although no association was reported (De Garay et al., 1974). A further study was completed during the 1976 Montreal Olympic Games to search for markers of aerobic performance in a group of Caucasian endurance athletes and controls. Red blood cell antigens and four erythrocyte enzymes were compared between athletes and controls, although again there were no reported differences between the groups (Chagnon et al., 1984, Couture et al., 1986) the lack of association may have resulted in a lack of sensitivity of the methods used at the time. Between the first publication of a genetic association study in an athlete population and the present day there has been an explosion of interest in the genetic contribution to human performance and elite athlete status. The increasing number of publications in the field year on year reflect this.

Between 2000 and 2007, common genetic variants associated with at least one performance phenotype were reported in a human gene map for performance and health-related fitness phenotypes (Rankinen et al., 2001, Rankinen et al., 2002, Pérusse et al., 2003, Rankinen et al., 2004, Wolfarth et al., 2005, Rankinen et al., 2006, Bray et al., 2009). A more recent literature review by Ahmetov and Fedotovskaya (2012) identified 79 gene polymorphisms associated with elite athlete status. Of these, 59 were associated with endurance performance. In the section that follows, the evidence for selected nuclear

and mitochondrial candidate genes will be reviewed for their individual contribution to endurance performance with a focus on elite endurance athlete status and marathon.

2.2.2 **Renin Angiotensin Aldosterone System (RAAS)**

ACE was the first gene to be considered a human physical performance gene (Montgomery et al., 1998, Puthucheary et al., 2011, Gayagay et al., 1998). The encoded protein is a key element in the Renin Angiotensin Aldosterone System (RAAS). RAAS is predominantly characterised as an endocrine system although paracrine, autocrine and intracrine elements of RAAS influence tissues locally and at a cellular level (Dzau, 1988, Dzau, 1989, Paul et al., 2006). As an endocrine system RAAS is a key regulator of blood pressure homeostasis (Bae et al., 2007) amongst other physiological processes that may influence endurance performance through $\dot{V}O_{2\,max,}$ $\dot{V}O_2$ kinetics, Running economy and lactate threshold. The RAAS is a complex molecular pathway initiated by the detection of low blood pressure (hypotension) by the juxtaglomerular cells baroreceptors in the afferent arteriole of the nephron causing the release of the aspartyl protease renin.

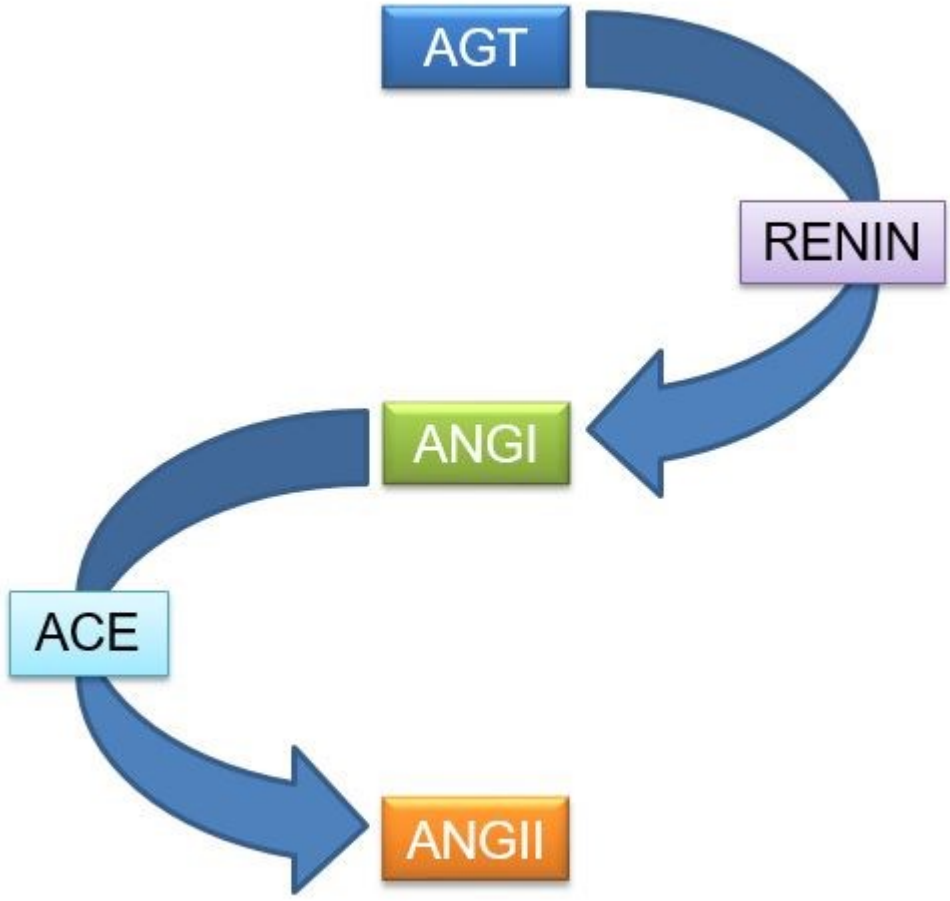

Figure 2: A simplified schematic of the RAAS pathway to indicate the roles of AGT and ACE

Angiotensinogen (AGT) synthesized in the liver (see Figure 2), in response to hypovolemia and thus blood pressure or a decrease in sodium concentrations, cause the renin dependent catalyses of the α-2 globulin AGT to the inactive decapeptide angiotensin I (ANGI). ANGI is transported in the blood to the pulmonary circulation where the dipeptidyl carboxypeptidase angiotensin I converting enzyme (ACE) cleaves the dipeptides from the C-terminal His-Leu dipeptide to give the vasoactive peptide angiotensin II (ANGII) (Woods et al., 2000). ANGII causes hypertension via the AT-1 receptor in three ways. Firstly, ANGII directly stimulates the AT-1 receptor in vascular smooth muscle cells which causes vasoconstriction of the efferent arterioles. This causes a reduction in blood flow and an increase in hydrostatic pressure increasing the total

peripheral resistance of the vasculature and thus the mean arterial pressure. In the kidney the increase in hydrostatic pressure favours filtration in an attempt to regulate the blood pressure. Secondly, ANGII binds to expressed membrane bound AT-1 receptors of vascular endothelium causing a hypertensive response via a direct reduction in the synthesis of the vasodilator nitric oxide. ACE further catalyses the inactivation of bradykinin (Erdös and Skidgel, 1987, Woods et al., 2000) required for the synthesis of nitric oxide causing a further reduction in its availability. The reduced availability of nitric oxide and the AT-1 receptor activation by ANGII stimulates the release of aldosterone from the zona glomerulosa of the adrenal gland initiating the third cause of hypertension. Aldosterone affects the sodium potassium ATPase pump on principal cells influencing the reabsorption of both sodium and water (Myerson et al., 1999). This increased plasma volume can cause hypervolaemia and lead to hypokalaemia and hypertension if not regulated (Boron and Boulpaep, 2012). Further to this aldosterone effects the ATP pump of alpha intracalated cells in the cortical collecting tubule, causing the increased excretion of hydrogen ions (that acidify the urine) and the reabsorption of water. These two pathways result in hypervolaemia and hypertension. Negative feedback resulting from the hypertension is detected by the juxtaglomerular baroreceptors and causes the kidney to reduce the production of renin regulating blood pressure.

As a paracrine system the RAAS mediated release of ANGII causes cardiac myocyte hypertrophy by increasing DNA synthesis, protein synthesis and cell number via the membrane expressed AT-1 receptor, (Sadoshima and Izumo, 1993). ANGII initiates a g_q-protein coupled cascade causing the activation of phospholipase C. Phospholipase C hydrolyses phosphatidylinositol 4,5,-bisphosphate (PIP2) to produce diacylglycerol (DAG) and inositol 1,4,5, triphosphate (IP3). IP3 acts on the sarcoplasmic reticulum to release intracellular calcium stores that increases cardiac inotropy and stoke volume (Levick, 2013, Boron and Boulpaep, 2012) to meet the oxygen demand of the respiring skeletal muscle. This can result in a normal remodelling of the cardiac muscle to the increased physiological demand for oxygen by the skeletal muscle as a result of the large volumes

aerobic cardiovascular exercise commonly completed by endurance athletes as part of their training and competition program (Levick, 2013). This adaptation is commonly known as Athletic Heart Syndrome. Athletic heart syndrome causes the heart ventricles to enlarge as a result of eccentric and concentric hypertrophy of the muscle wall (Levick, 2013). During eccentric hypertrophy new sarcomeres are added in series to those already constructing the cardiac wall causing the wall to extend further (Levick, 2013). Concentric hypertrophy causes sarcomeres to be added in parallel to those already in the wall causing the wall to thicken further increasing cardiac inotropy, stroke volume and thus cardiac output and blood pressure (Levick, 2013). The physiological mechanisms outlined make the key RAAS components *ACE* and *AGT* attractive as candidate genes for blood pressure homeostasis, oxygen delivery to respiring skeletal muscle during endurance exercise and thus endurance performance in competition.

2.2.2.1 Angiotensin I Converting Enzyme - ACE

The structural organisation of *ACE* gene was determined in 1991 (Hubert et al., 1991). The *ACE* gene is located on chromosome 17q23 and consists of 26 exons. The insertion deletion (I/D) polymorphism in this gene refers to an *Alu* repetitive sequence 287 base pairs (bp) long, in intron 16, resulting in three genotypes, DD and II homozygotes and ID heterozygotes (Rigat et al., 1992). The insertion I allele has been reported to cause lower ACE activity in serum and cardiac tissue (Rigat et al., 1990, Danser et al., 1995).

The *ACE* I and D alleles are both favourably associated with human performance in elite athlete populations. *ACE* I allele associations with endurance performance are reported in elite endurance runners (Myerson et al., 1999, Alvarez et al., 2000, Scanavini et al., 2002, Hruskovicova et al., 2006), Ironman triathletes (Collins et al., 2004), elite rowers (Gayagay et al., 1998, Jelaković et al., 2000, Cięszczyk et al., 2009), elite cyclists (Alvarez et al., 2000, Scanavini et al., 2002), elite swimmers (Tsianos et al., 2004) and elite mountaineers (Montgomery et al., 1998).

Myerson et al. (1999) reported an I allele association with competitive running distance in British Olympians. Athletes running longer distances had higher I allele frequencies than those running shorter distances, healthy control participants and athletes in 19 other sporting disciplines where endurance performance was not a primary determinant of success. Alvarez et al. (2000) genotyped 60 elite Spanish athletes including 20 long distance runners (marathon and cross-country), and similarly to Myerson et al. (1999) reported an increased frequency of the I allele in the elite athlete population compared to 400 controls. Scanavini et al. (2002) selected an elite group of 52 athletes from their initial athlete cohort. Olympic Games participation and $\dot{V}O_{2\,max}$ values of 65-80 ml / kg / min $^{-1}$ for aerobic athletes determined elite athlete status, while elite anaerobic athletes had $\dot{V}O_{2\,max}$ values of 40-55 ml / kg / min $^{-1}$. Again, the authors reported a higher frequency of the II genotype in the 33 Olympic aerobic athletes than in the 19 Olympic anaerobic athletes or controls, with the most marked difference between track and field runners and controls. Increased ACE I/I genotype frequency was determined amongst marathon runners placed 1st to 150th (classified by performance based on marathon competition results) ($p = <0.01$) of the 215 elite marathon runners participating in the study suggesting an association between the ACE I/D polymorphism and elite marathon performance (Hruskovicova et al., 2006).

Hagberg et al. (2001) reported that the increased frequency of the *ACE* I allele or the II genotype in endurance based athletes may be partially explained by an association of the *ACE* II genotype with higher $\dot{V}O_{2\,max}$ (Hagberg et al., 1998, Hagberg et al., 2001) . Furthermore, Hagberg et al. (2002) reported elite female runners with the *ACE* II genotype had a 25% greater cardiac output than *ACE* DD genotype women runners, with corresponding genotype-dependent differences in stroke volume observed in both physically active and athletic women. Hagberg et al. (2001) further hypothesised based on the work of Montgomery et al. (1997) that *ACE* II genotype carriers (due to less LV

hypertrophy) were superior at matching cardiac afterload to cardiac output and thus more efficient at modulating cardiovascular function during physical exertion. Gayagay et al. (1998) went on to hypothesize that the *ACE* I allele and II genotype may result in reduced cardiac afterload and thus enhanced ventriculo-vascular coupling efficiency during exercise. Hagberg et al. (2002) reported after measuring cardiovascular haemodynamics during maximal exercise that *ACE* genotype had not influenced stroke volume or heart rate and thus cardiac output but that female runners with the II genotype had considerably higher maximal arteriovenous O_2 difference than those with the ID and DD genotypes. The authors concluded that peripheral vascular mechanisms that determine $\dot{V}O_{2\,max}$ and thus endurance performance may be influenced by *ACE* genotype. These peripheral vascular mechanisms may in part be related to a higher proportion of slow-twitch type 1 fibres in human skeletal muscle with the *ACE* I allele (Zhang et al., 2003, Ahmetov and Rogozkin, 2009). The II genotype is also related to a genotype-dependent improvement in skeletal muscle mechanical efficiency of ~9% with aerobic training (Williams et al., 2000).

In contrast, there are investigations that have found no significant association between the *ACE* I allele and endurance performance and related phenotypes (Rankinen et al., 2000c, Scott et al., 2005, Oh, 2007, Papadimitriou et al., 2009, Ash et al., 2011). For example, the HERITAGE study found no association between baseline $\dot{V}O_{2\,max}$ and *ACE* genotype (Rankinen et al., 2000b), while in one subgroup training-induced changes in $\dot{V}O_{2\,max}$ were associated with the *ACE* DD genotype. Similarly, Day et al. (2007), having found that the *ACE* genotype was strongly associated with circulating ACE activity, found no association to support a role for systemic ACE activity or *ACE* genotype in the regulation of endurance performance in females with low physical activity.

In those studies that included athlete populations, elite participants were included. However, they varied in endurance sporting discipline. As endurance sports are so diverse (varying in duration, skill and physical demand) different athletic characteristics

are required to succeed (Nazarov et al., 2001). In addition, in some studies the ethnic ancestry of the participants was mixed. It should be noted that allele frequencies and genotype distributions for this gene vary across different populations (Eleni et al., 2008). For example, the genotype distribution in European Caucasian adults is approximately DD 25% ID 50% II 25% (Myerson et al., 1999) and in Jamaicans with recent west-African ancestry DD 36% ID 47% II 17% (Scott et al., 2010).

In contrast, there are investigations that have found significant association between the *ACE* D allele and endurance performance. A study of 121 elite Israeli marathon runners demonstrated a higher frequency of the *ACE* D allele and *ACE* DD genotype than in sprinters (Amir et al., 2007). Muniesa et al. (2010) concurred with the results of Amir et al. (2007) and found that elite endurance runners had a significantly higher proportion of DD genotype. This may suggest a positive association between the *D* allele and elite endurance athletic performance in some ethnic groups (Amir et al., 2007). Indeed Puthucheary et al. (2011) suggest that this atypical association may be explained by the heterogeneity of the Israeli Caucasian Jewish population. However, those results could also be artefacts of the small population sizes in those studies.

In summary, despite some contradictory evidence, the *ACE* genotype-dependent associations with elite endurance running performance cited in the preceding text imply that the I allele may enhance the favourable adaptations to endurance training and thus competitive performance. That conclusion is supported by a thorough meta-analysis that reported an odds ratio of 1.35 for association between *ACE* II genotype and endurance athlete status (Ma et al., 2013).

Some research groups have suggested that other genes encoding key RAAS elements may further our understanding of the influences of the RAAS pathways on elite endurance performance. One of these RAAS components of interest is Angiotensinogen.

2.2.2.2 Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)- AGT

The *AGT* gene is located on chromosome 1q42.2 and consists of 5 exons (Gaillard et al., 1989). The *AGT* Met235Thr polymorphism refers to a C to T nucleotide substitution at (rs699) position 4072, in exon 2 culminating in the conversion of Methionine to Threonine in the angiotensinogen protein. This gives three genotypes, CC and TT homozygotes and CT heterozygotes. The threonine variant is encoded by the C allele and has been associated with 10–30% higher plasma angiotensin concentration in men and women (Jeunemaitre et al., 1992).

The *AGT* rs699 polymorphism has been associated with the response to endurance training, cardio-respiratory endurance and blood pressure (Rankinen et al., 2000a). Males carrying the TT and CT genotypes have greater reductions in diastolic blood pressure at 50 W in response to 20 weeks of endurance training (Rankinen et al., 2000a). However, TT genotype males also carrying the D allele of the *ACE* gene showed no response to training. Females showed no genotype-dependent differences in either systolic or diastolic blood pressure or training response. This lack of genotype-dependent training response was further suggested by Bae et al. (2007) who found no association of the *AGT* rs699 polymorphism with the response to endurance training. In that study, 17 Korean women completed endurance training for 12 weeks, then tested for ventilatory response, glucose, body composition, total cholesterol, triglyceride concentrations, blood pressure, $\dot{V}O_{2\,max}$ and BMI. No *AGT* rs699 genotype or allele associations were reported though the findings of this study should be regarded with caution as the study was underpowered.

It has also been suggested that *AGT* rs699 polymorphism influences the variability in left ventricular hypertrophy as a result of endurance training in elite endurance athletes

(Karjalainen et al., 1999). Eighty endurance athletes (long distance runners, orienteers, cross country skiers and triathlon competitors) from the Finnish national teams were genotyped and completed echocardiography. The CC homozygotes had greater left ventricular mass than TT homozygotes in both males and females (Karjalainen et al., 1999). It was also noted that there was a sex-dependent association between heart mass and *AGT* genotype. In males, those with the TC genotype were similar to the homozygous CC genotype and the reverse was observed in females (Karjalainen et al., 1999). In contrast, a study of 83 Caucasian, male international and national endurance athletes showed no association of left ventricular mass with the *AGT* rs699 polymorphism (Diet et al., 2001). However, an increase in left ventricular mass was recorded in individuals carrying both the *ACE* DD and *AGT* TT genotypes (Diet et al., 2001).

Although some studies mentioned in the preceding paragraphs genotyped *AGT* rs699 polymorphism in elite athlete cohorts, they did not directly assess genotype distributions in relation to elite athlete status. Those studies that have assessed associations of *AGT* rs699 polymorphism with elite athlete status give mixed results. The *AGT* rs699 polymorphism genotype distributions of 63 power athletes, 100 endurance athletes and 119 Caucasian male controls revealed a higher CC genotype in the power group when compared to the endurance and control groups. No difference in genotype frequency was noted between the endurance and control groups (Gomez-Gallego et al., 2009). Similarly, Zarebska et al. (2013) reported no association with elite endurance athlete status. It should also be noted that among both the power athlete genotype distributions of the Spanish and Polish studies there was a lack of Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium is a law which states that allele and genotype frequencies will remain constant from generation to generation. This requires the maintenance of the following 6 parameters:

A large breeding population

A large breeding population size helps to ensure that the genetic equilibrium is not disrupted by chance alone. A sharp reduction in population is called a population bottleneck and can lead to inbreeding due to lack of sexual selection or the founder effect. For example, in a small population, it may be that only a few copies of an allele exist. If by chance the organisms with that allele do not reproduce successfully, the allelic frequency will change. This random, nonselective change is what happens with genetic drift. In contrast, it may be that alleles are equally represented in the parent population but that identical individual's mate through lack of population numbers and new populations lacking genetic variety are established. As a result of the loss of genetic variation, the new population may be distinctively different, both in genotype and phenotype, from the derived parent population. The population sample of this study has not been restricted or isolated from the general population in anyway before during or after inclusion in this study. Therefore, it is highly unlikely that it is subject to genetic drift, the founder effect or population bottlenecking.

Random mating

In a population that conforms to Hardy Weinberg equilibrium, mating between individuals must be random to ensure no departure in the allelic frequency from the equilibrium. In assortative mating were individuals select partners that are similar to themselves fewer heterozygotes are observed though this does not alter the allelic frequencies observed when compared to a population where mating is random. It is assumed mating between individuals is random for those offspring included in this study.

No change in allelic frequency due to mutation

For a population to maintain the law of Hardy-Weinberg equilibrium there can be no alteration in allelic frequency due to mutation. Gene mutation would alter the allelic

balance introducing new alleles into the gene pool and altering the composition of the gene pool from generation to generation.

No immigration or emigration

For a population to exhibit equilibrium the allelic frequency must remain constant. Therefore, no new alleles can be introduced to the population and no alleles may be lost. Both immigration and emigration may alter the allelic frequency by introducing or removing alleles to and from a population. All athletes included in this study were from the continent of Europe, none were known to have emigrated or immigrated to Europe.

No natural selection

In a population at equilibrium no alleles exhibit preferential selection over any other alleles present. Should a selection pressure be introduced, those alleles that are preferentially selected will become more common. For example in the case of bacteria if resistance to particular antibiotic allows bacteria to thrive when an individual is treated with that particular antibiotic, the allele for resistance may become more prevalent in the population of bacteria. No known selection pressures were influencing the alleles chosen.

Lab error

Genotype frequencies should comply with HWE proportions. Deviation from these proportions can be caused by many factors, one of which is genotyping error. Genotype mistakes can lead to increased random error and bias in gene-phenotype associations. All samples were genotyped in duplicate to confirm the genotype group status of each sample. There were no discrepancies between duplicate samples.

Chance

Genotyping frequencies may in some cases deviate from the HWE proportions just through chance.

If anyone of these assumptions is not met than the population of interest will not be in Hardy-Weinberg Equilibrium. Deviation from Hardy-Weinberg equilibrium may indicate a selection pressure or evolution of the allele frequencies from one generation to the next. This maybe the result of mutation, non-random mating, gene flow, genetic drift and natural selection.

Based on the contribution of *ACE* and *AGT* to cardio-respiratory endurance, blood pressure regulation and left ventricular hypertrophy, it is hypothesised that the *ACE* I/D and AGT rs699 polymorphisms may contribute to the elite endurance performance of a marathon runner.

2.2.3 **Actinin Alpha 3 - ACTN3**

Skeletal muscle fibre type plays and important role in elite marathon performance. Skeletal muscle fibre type has been shown to alter in rodent and human populations (detailed below) as a result of *ACTN3*.

The α-actinins are major structural components of the Z line in skeletal muscle where they crosslink the actin thin filaments to the Z-line maintaining sarcomeric integrity (Blanchard et al., 1989, Mills et al., 2001, Yang et al., 2003) as shown in Figure 3.

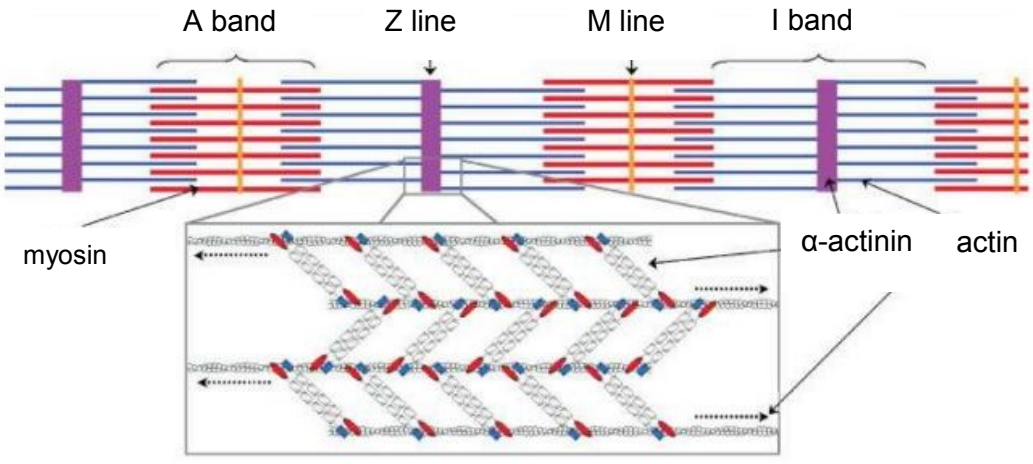

Figure 3 Localisation and domain structure of the sarcomeric α-actinins. The sarcomeric α-actinins anchor thin filaments to the Z-line. Figure adapted from (MacArthur and North, 2004).

Two genes encode the skeletal muscle α-actinins: actinin alpha 2 (*ACTN2*) and *ACTN3*. ACTN2 expression is ubiquitous to skeletal muscle fibres whereas ACTN3 has restricted expression to Type II fibres. *ACTN3* is located on chromosome 11q13.1. A common nonsense polymorphism (rs1815739) in exon 16, codon 577, results in the conversion of arginine (CGA; R allele) to a stop codon (TGA; X allele) (North et al., 1999). The X allele homozygote is completely prevented from producing functional ACTN3 protein. Both alleles are common within the general population with 18% of European population and around 1 billion world-wide presenting as X allele homozygotes (Yang et al., 2003). The absence of ACTN3 protein (XX genotype) has no known pathological consequence in

humans. The lack of pathology in X allele homozygotes may be explained by the conserved nature of the α-actinins. The ACTN2 protein is 80% identical to ACTN3 with a further 10% that is very similar (Beggs et al. 1992). The ACTN3 protein is comprised of three domains these are termed the NH_3 terminal actin binding domain, a 4x repeat of 122 amino acid motifs, that share homology with spectrin, known as the central rod domain and a COOH terminal domain (Berman and North, 2010).

The SNP transition from C>T at position 1747 (codon 577) in the *ACTN3* coding sequence results in three genotypes - homozygous RR and XX and the heterozygous RX (North et al., 1999). The 577R allele has been associated with strength in both rodent and human populations (MacArthur et al., 2008, Clarkson et al., 2005, MacArthur and North, 2007).

Multiple case controlled studies have reported that the RR genotype is represented more frequently than that of the XX genotype in strength athletes and sprinters when compared to a control population (Niemi and Majamaa, 2005, Roth et al., 2007, Druzhevskaya et al., 2008). An Australian study reported the XX genotype was significantly reduced in elite Australian power athletes and completely absent from the Australian Olympic female power athletes (Yang et al., 2003).

These findings have been replicated in several independent case controlled studies including elite Finnish sprint athletes (Niemi and Majamaa, 2005), elite Greek track and field athletes,(Papadimitriou et al., 2008) elite strength athletes from the United States of America (USA) (Roth et al., 2007) and Russian power athletes (Ahmetov and Rogozkin, 2009). The R allele conferring some athletic advantage in strength based sport was further evidenced by a study that reported the percentage surface area and number of type IIx (fast twitch glycolytic) fibres was greater in the RR than the XX genotype group of young healthy men (Vincent et al., 2007, Ahmetov and Rogozkin, 2009).

In addition, ACTN3 knockout mice had smaller muscle fibre cross sectional area, lower *in vivo* strength and lower *in vitro* force generating capacity, and longer twitch half relaxation times when compared to wild type mice (MacArthur et al., 2008). The ACTN3 knockout mice showed enhanced endurance running performance and better recovery from fatiguing muscle contractions compared to the wild type mice (MacArthur et al., 2007, MacArthur et al., 2008). Lastly the ACTN3 knockout mice showed higher expression of enzymes associated with oxidative capacity (MacArthur et al., 2008). It is not thought that ACTN3 protein infers any extra protection from contraction induced muscle damage or Z line damage as similar levels of muscle damage were observed in single muscle fibres following eccentric contractions from ACTN3 knockout mice (MacArthur et al 2007; Chan et al 2008; MacArthur et al 2008).This evidence would suggest that *ACTN3* 577R may confer some advantage in strength-based events but this has not been examined in a British population.

Yang et al. (2003) hypothesized a competitive *ACTN3* 577X advantage in athletes competing in endurance based events. They studied 194 elite Australian endurance athletes 18 of whom had competed at Olympic level. The endurance group included long-distance cyclists, rowers, swimmers with a competition distance ≥400 m, cross country skiers and 15 track athletes who ran distances ≥5000 m. Yang et al. (2003) found a tendency for a higher XX genotype frequency in female athletes though this was not noted in male athletes. A large European study was conducted later, comprising 633 Caucasian, Polish, Russian and Spanish mixed sporting discipline athletes (278 elite endurance athletes) and 808 non-athletic controls (Eynon et al., 2012). Athletes were considered elite if they had competed at a national or international level. The authors found that the XX genotype was over-represented in the endurance athlete population when compared to the power athletes (OR 1.88) and those athletes of elite endurance athletic status were

~3.7 times more likely to have the XX genotype compared with the RX and RR genotypes when compared to the national-level athletes (Eynon et al., 2012).

Triathlon is an endurance sport requiring good ability in prolonged swimming, cycling and running. Two studies have investigated whether the *ACTN3* R577X polymorphism was associated with triathlon performance or triathlon athlete status. In a cohort of 196 elite endurance athletes racing in the 2008 Kona Ironman championship triathlon, the 577X allele was not associated with performance time (Grealy et al., 2012). Similarly, in a large study of Caucasian male triathletes the *ACTN3* R577X polymorphism was not associated with performance time, nor was there and difference in genotype frequency between athletes and non-athlete controls (Saunders et al., 2007). It is unknown whether a study of athletes who excel in one of the triathlon disciplines alone (e.g. prolonged endurance running) would give similar or differing results.

Other studies have also noted no *ACTN3* XX genotype or X allele associations in their endurance athlete populations when compared to controls. In a small Finnish population of 52 athletes including 20 endurance track athletes who ran distances between 800 m and Marathon or race walkers, there were no differences in the *ACTN3* 577X genotype or allele frequency between athletes and controls this may be explained by the variation in competitive distances completed by the athletes. In another small study involving Spanish Olympic level runners (52), professional cyclists (50) and 123 male controls with low physical activity (Lucia et al., 2006), *ACTN3* genotype conferred no advantage for elite endurance performance (Lucia et al., 2006). This finding was similar to that reported in a larger study of 316 male endurance athletes from six different sports comprising the Genathlete cohort, where there was no difference in the frequency of the XX genotype between athletes and controls (Döring et al., 2010). Finally, a meta- analysis reviewing the association of human sporting performance with the *ACTN3* R577X polymorphism found no association of the 577X allele or genotype with endurance performance (OR, 1.03;

95% CI, 0.92–1.15) (Ma et al., 2013). However, the authors concluded that the varieties of sex, ethnicity, and sporting discipline in the studies may explain the lack of an observed association.

The association of the *ACTN3* 577X allele and XX genotype with elite endurance performance in the literature is therefore contentious. The clearly expressed endurance phenotype in the mouse and associated plausible biological pathways do suggest strong biological rationale for an association between the *ACTN3* 577X allele and human endurance performance. The studies in humans to date, however, are limited by relatively small sample sizes and/or a lack of clearly defined endurance athlete status and discipline athlete phenotypes. Therefore, there remains sufficient prior evidence to hypothesise that the *ACTN3* R577X polymorphism may contribute to marathon runner elite performance.

2.2.4 **Peroxisome proliferator activated receptor gamma, coactivator 1 alpha - PPARGC1A**

Mitochondrial oxidative phosphorylation provides the vitally important ATP, required for skeletal muscle propulsion during marathon running. PARAGC1A is thought to influence the fatty acid substrate availability during the later stages of a marathon and its conversion to ATP, to directly fuel skeletal muscle contraction during a marathon and will therefore influence a runners running economy and lactate threshold.

PPARγ coactivator 1α (PGC1α) is a nuclear protein that regulates gene expression by enhancing transcriptional factor activation (Puigserver et al., 1998, Esterbauer et al., 1999), coordinating gene expression in signalling pathways responsible for improved mitochondrial oxidative phosphorylation (OXPHOS), especially in brown fat and skeletal muscle (Arany et al., 2005). Total mitochondrial volume and density, influenced by mitochondrial biogenesis, in turn regulates OXPHOS (Puigserver et al., 1998, Wu et al., 1999), transport and oxidation of glucose and lipid (Tunstall et al., 2002) and skeletal muscle fibre type (Lin et al., 2005, Lin et al., 2002).

Coordination of mitochondrial and nuclear genomes regulates mitochondrial biogenesis (Baar, 2004). The mitochondrial genome encodes 13 proteins associated with the electron transport chain (approximately 10% of the genes required for mitochondrial biogenesis) (Baar, 2004). The nuclear genome encodes proteins required for the replication and expression of mitochondrial DNA (mtDNA), the remaining electron transport proteins and those associated with OXPHOS of fatty acids, ketones and pyruvate (totalling approximately 90% of the genes required for mitochondrial biogenesis). PGC1α regulates OXPHOS gene expression via interaction with specific transcription factors, such as nuclear receptor PPARγ, nuclear respiratory factor 1 (NRF1) and muscle specific transcription factors such as MEF2 (Lin et al., 2002), Interaction of PGC1α with transcription factor targets is linked to specific tissues as shown in Figure 4.

PGC1α knock out (KO) murine skeletal muscle cells show lower expression of genes influencing mitochondrial function and reduced expression by 30-50% of the genes implicated in OXPHOS, fatty acid oxidation and ATP resynthesis in murine cardiac muscle (Arany et al., 2005). PGC1α reportedly doubles MtDNA content, leading to a 57% increase in mitochondrial number in murine skeletal myoblasts expressing *PPARGC1A* (Wu et al., 1999, Puigserver et al., 1998). Transgenic mice, over-expressing PGC1α in cardiac and skeletal muscle cells, demonstrated mitochondrial proliferation and OXPHOS gene expression (Lehman et al., 2000, St-Pierre et al., 2003). KO mouse Type 1 skeletal muscle fibres showed decreased mitochondrial numbers and respiratory capacity, while KO Type II fibres showed normal mitochondrial density and function (Leone et al., 2005), suggesting PGC1α mediated activation of mitochondrial biogenesis and OXPHOS in slow skeletal and cardiac muscle. These data demonstrate that PGC1α is a requirement for normal mitochondrial gene expression in cardiac and skeletal muscle (Arany et al., 2005).

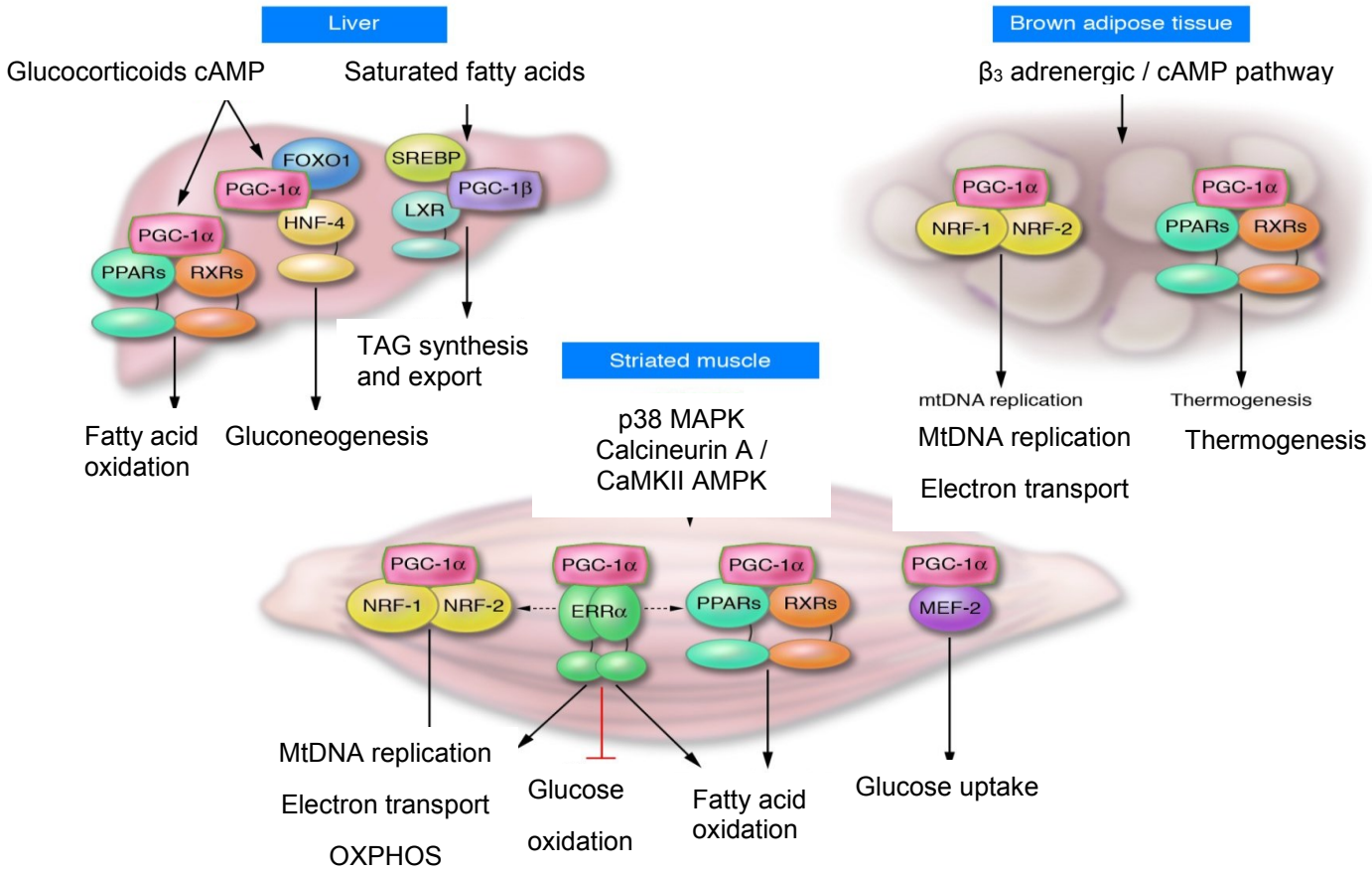

Figure 4 The PGC-1 gene regulatory cascade (Finck and Kelly, 2006).

Illustrates upstream signalling events and downstream gene regulatory actions of the inducible PGC-1 coactivators. Interaction of PGC-1α with transcription factor targets is linked to specific tissues. For example, PGC-1α coactivates PPAR nuclear receptor transcription factors to express genes involved in mitochondrial fatty acid oxidation. LXR, liver X receptor; TAG, triacylglycerol; RXR, retinoid X receptor; mtDNA, mitochondrial DNA; OXPHOS, oxidative phosphorylation. (Finck and Kelly, 2006).

Wu et al. (1999) reported that PGC1α translocates to the mtDNA D-loop to interact with NRF1 on the mitochondria transcription factor A (Tfam) promotor. Upregulation in Tfam expression further implicates PGC1α in the biogenesis of mitochondria (Shadel and Clayton, 1993, Baar, 2004). In addition, PGC1 α coactivates PPARα, shown to control transcriptional activity of mitochondrial enzymes involved in the β-oxidation of fatty acids (Gulick et al., 1994, Baar et al., 2002). Increases in mitochondrial proteins and glycolytic enzymes increasing fatty acid oxidative capacity of skeletal muscle may induce Type II fast glycolytic fibre conversion to Type I slow, high-oxidative fibres, with Type 1 fibres showing increased expression of PGC1α (Brooke et al., 2005). Lin et al. (2002) demonstrated in transgenic mice that forced expression of PGC1α at physiological levels caused an approximate 10% fibre type conversion in muscles normally rich in Type II fibres. Those muscles were observed to be redder, they showed increased expression of contractile proteins characteristic of Type I fibres, genes of mitochondrial oxidative metabolism were activated and there was increased resistance to electrically stimulated fatigue (Lin et al., 2002). Using a PGC1α knockout mouse model, Arany et al. (2005) found that PGC1α was not required for mitochondrial biogenesis, or for the differentiation of skeletal muscle fibres. However, the absence of PGC1α reduced mitochondrial function and decreased levels of ATP (Arany et al., 2005). These combined data indicate that PGC1α is likely a key regulatory factor for the determination of muscle fibre type (Lin et al., 2002) in rodents and humans (Finck and Kelly, 2006).

Due to the complex interactions of PGC1α, the exercise-induced expression pathways of PGC1α are unclear (Finck and Kelly, 2006). PGC1α is thought to interact with myocyte enhancer factor 2 through the calcineurin A and CaMK pathway (Lin et al., 2002). More recently, after endurance exercise, expression of PGC1α has been postulated to occur through the p38 MAPK and AMPK pathways (Akimoto et al., 2005) detailed in Figure 5.

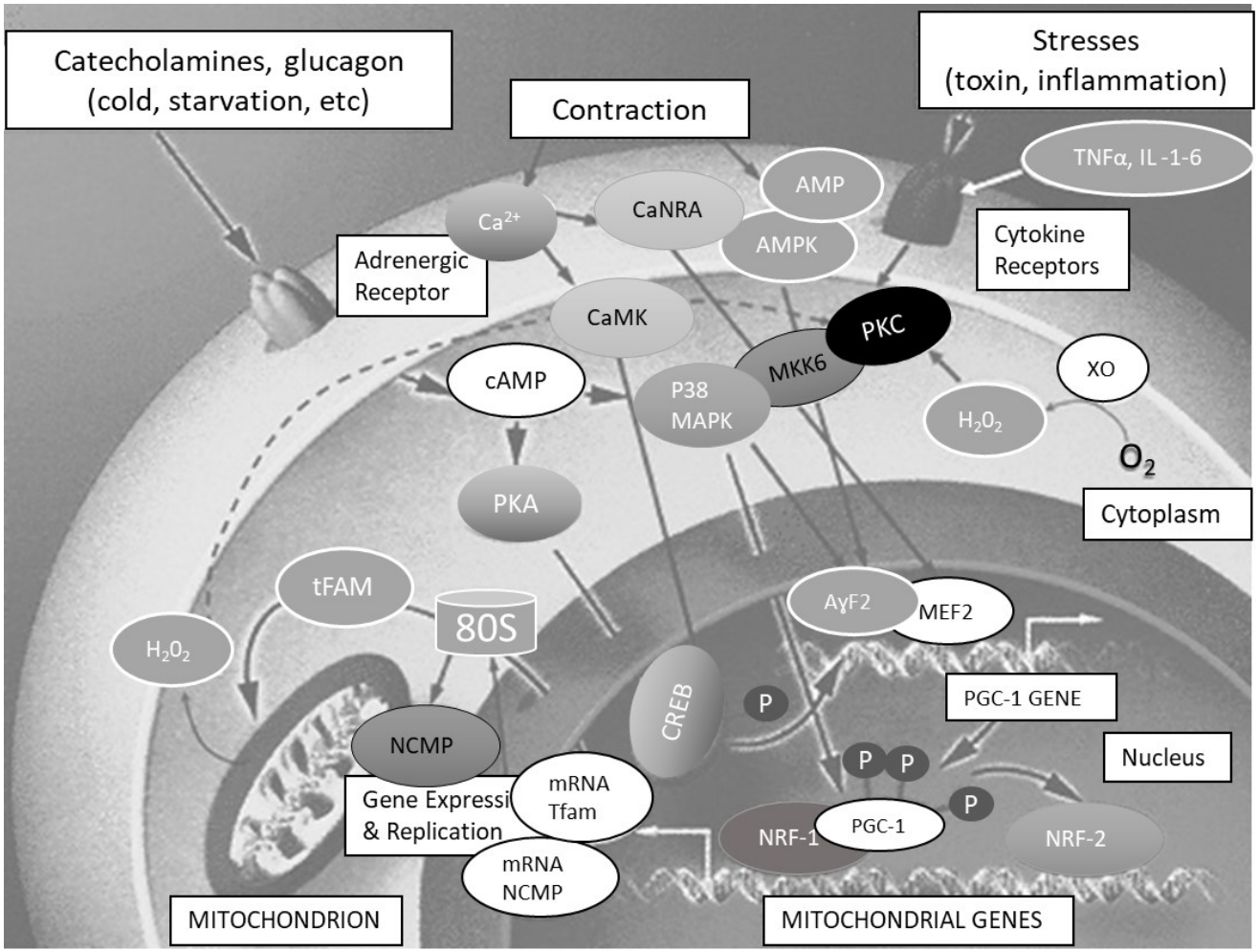

Figure 5 Regulation of PGC1α gene expression (Kang and Li Ji, 2012).

PGC-1α, peroxisome proliferator-activated receptor gamma coactivator 1-alpha; Tfam, mitochondria transcription factor A; MAPK, mitogen-activated protein kinase; MKK6, MAP kinase kinase: CaMK, Ca^{2+}/calmodulin-dependent protein kinase, CREB, cyclic AMP response element binding protein; MEF, myocyte enhancer factor; ATF, activating transcription factor; NRF, nuclear respiratory factor; ERR, oestrogen-related receptor; XO, xanthine oxidase; PKC, protein kinase C.

In rats (Baar et al., 2002, Goto et al., 2000, Terada et al., 2002) and humans (Russell et al., 2003, Pilegaard et al., 2003), acute endurance exercise and chronic endurance training increases PGC1α expression primarily in Type IIa fibres (Russell et al., 2003, Finck and Kelly, 2006). In response to endurance exercise and expression of PGC1α, there is increasing mitochondrial content, the presence of more oxidative fibres and resistance to fatigue (Wu et al., 1999, Baar et al., 2002, Terada et al., 2002, Lin et al., 2002, Russell et al., 2003). In humans, endurance training significantly increases *PPARGC1A* mRNA levels between 2-fold and 10-fold (Tunstall et al., 2002, Pilegaard et al., 2003, Norrbom et al., 2004) and thus may enhance translation and modulation of the PGC1α protein.

The *PPPARGC1A* gene of 13 exons is located on chromosome 4p15.1. The gene produces a protein with a calculated molecular mass of 91kDa (Esterbauer et al., 1999). A functional SNP (rs8192678) causes a transition substitution of the ancestral glycine (G allele) to serine (A allele) at amino acid position 482. The minor A allele is associated with reduced *PPARGC1A* expression (Ling et al., 2004, Ahmetov and Rogozkin, 2009).

PPARGC1A has been shown to influence $\dot{V}O_{2\,max}$ in both rodent (Baar et al., 2002, Lin et al., 2002, Terada et al., 2002) and human populations (Lucia et al., 2005, Norrbom et al., 2004, Russell et al., 2003). Indeed, oxygen consumption doubled in murine myotubes expressing PGC1α, treated with FCCP (a chemical uncoupler that completely uncouples mitochondria and maximizes their respiratory capacity) indicating that these cells have a higher content and/or electron transport activity of mitochondria (Wu et al., 1999). Therefore, it could be postulated those individuals carrying the *PPARGC1A* G allele may have an increased mitochondrial content in their muscle fibres and therefore an increased ATP production capacity via the electron transport chain when compared to A allele carriers. This would be a clear advantage for an individual competing in endurance sport such as marathon where OXPHOS is the primary mechanism of energy generation for

muscle contraction. However, there was no association between $\dot{V}O_{2\,max}$ and the *PPARGC1A* rs8192678 polymorphism in German, Dutch or Japanese populations (Stumvoll et al., 2004, He et al., 2008, Nishida et al., 2015). Nishida et al. (2015) observed in 112 Japanese middle-aged men that A allele carriers had a higher lactate threshold than G allele carriers. This suggests that the *PPARGC1A* A allele may be associated with a higher aerobic capacity in Japanese middle-aged men, although no association with $\dot{V}O_{2\,max}$ was observed according to an indirect estimation (using age-predicted maximum heart rate) known to have rather poor validity. Although lactate threshold is a good measure of aerobic function, it is not synonymous with $\dot{V}O_{2\,max}$ (Nishida et al., 2015). Furthermore, $\dot{V}O_{2\,max}$ is usually limited by the oxygen delivery capacity, not the ability of mitochondria to utilise oxygen in the skeletal muscle. Marathon runners are also reliant on the oxygen delivery capacity of the cardio respiratory system during a race, and therefore, both $\dot{V}O_{2\,max}$ and lactate threshold are considered important parameters of aerobic function in endurance runners (Foster, 1983).

In elite Spanish athletes, Lucia et al. (2005) suggested that *PPARGC1A* rs8192678 may be one of a collection of genetic factors to affect athletic aerobic capacity after reporting an under-representation of the A allele in their elite male endurance athletes. The A allele was further associated with lower aerobic capacity in Russian rowers (Ahmetov et al., 2007). Tural et al. (2014) reported that the A allele was associated with higher aerobic capacity in Turkish elite endurance athletes in combination with the *PPARA* G allele. A lower frequency of the A allele of *PPARGC1A* rs8192678 has been associated with elite endurance athlete status in several cohorts of athletes of mixed sporting discipline (Maciejewska et al., 2011, Lucia et al., 2005, Eynon et al., 2009c). In contrast, no association of the *PPARGCIA* rs8192678 with elite athlete status was observed in elite Polish athletes of mixed discipline (Maruszak et al., 2014). No difference in *PPARGC1A* rs8192678 genotype frequency distributions between Ironman triathletes and controls were noted by Grealy et al. (2015), nor was there any association with performance time.

In 60 elite Turkish endurance athletes, the A allele was associated with endurance athletic performance.

Although the athletes in all of these studies were reported to be at an elite level, the criteria used to define elite athlete status was variable across studies. In addition, in some of the studies, the endurance athlete groups were of a mixed sporting discipline (Tural et al., 2014). Where sporting disciplines were defined, the association of *PPARGC1A* rs8192678 with athletes/performance from the individual sports were not statistically significant, potentially due to insufficient statistical power. This makes it difficult to determine if the *PPARGC1A* rs8192678 polymorphism does indeed have an influence on performance in a single endurance discipline such as the marathon.

2.2.5 **Uncoupling protein (mitochondrial, proton carrier) – UCP**

The ability to perform at an elite level in endurance events such as marathon running requires training of the skeletal muscle to ensure the tight coupling and efficiency of oxidative phosphorylation of carbohydrate at the start of the race and fatty acids as the race progresses, to provide the energy required for skeletal muscle contraction used in propulsion during running (Jiang et al., 2009, Bruton, 2002, Coyle, 2007). In contrast, the generation of superoxide and free radical species through large oxygen flux through the electron transport chain would inhibit skeletal muscle contraction efficiency and in high concentrations may inhibit the functionality of the skeletal muscle cells (Jiang et al., 2009). The regulation of oxidative phosphorylation coupling to ATP production is thought to be the role of the uncoupling proteins. UCP function in skeletal muscle is not fully understood though they have been implicated in uncoupling respiration from ATP production and thus in the control of energy expenditure (Boss et al., 1997) and thermogenesis incredibly important factors in a prolonged endurance event such as marathon.

The UCPs are a family of intramembranous mitochondrial proteins that are responsible for the facilitation of anion transfer from the inner to the outer member of the mitochondria (Boss et al., 1997). The UCPs also transfer protons from the outer mitochondrial membrane to the inner mitochondrial membrane causing a reduction in the membrane electrical potential. Several publications have suggested further roles for UCP3 and these are addressed in Muzzin et al. (1999) and Brand and Esteves (2005). Muzzin et al (1999) propose that due to its homologous nature with UCP1, UCP3 may have a role in thermogenesis and the metabolic adaption of the mitochondria towards inproved degradation of fatty acids. Brand and Esteves (2005) report that UCPs do not mediate adaptive thermogenesis and that fatty acid transportation although reported in the literature has limited evidence, but that under specific pharmacological conditions they may have specific thermogenic properties. The uncoupling that they regulate significantly attenuates ROS production and protects against cellular damage likely to be of

importance during marathon performance due to the increased acid production and inflammation caused by prolonged exercise bouts.

The UCPs are expressed in multiple selective tissues. UCP3 is preferentially expressed in brown fat (Boss et al., 1997, Vidal-Puig et al., 1997) while UCP2 is found in most tissues (Gimeno et al., 1997, Fleury et al., 1997). Mitochondrial uncoupling proteins 2 and 3 are both expressed in skeletal muscle (Boss et al., 1997, Fleury et al., 1997, Gimeno et al., 1997) and exercise is a known stimulus for *UCP2* and *UCP3* mRNA expression. Acute exercise in vivo rapidly increases *UCP3* mRNA expression (Jiang et al., 2009, Cortright et al., 1999, Zhou et al., 2000, Tonkonogi et al., 2000) and *UCP2* expression is up regulated by exercise in rodent models (Cortright et al., 1999, Pedersen et al., 2001) and in humans (Tonkonogi et al., 2000). Increases in U*CP3* mRNA expression have been shown to accompany GLUT-4 overexpression in mice (Pedersen et al., 2001, Tsuboyama-Kasaoka et al., 1998). In addition, *UCP2* and *UCP3* mRNA expression in rat muscle shows a dose-dependent increase with insulin and muscle contraction *in vitro* (Pedersen et al., 2001). These studies suggest that changes in glucose concentration and/or fatty acid metabolism may regulate skeletal muscle *UCP* mRNA expression. Because elite marathon runners usually consume a high carbohydrate diet and consume carbohydrate during competition races insulin will be present (Burke, 2007). In addition, the repeated contraction during marathon training and competitive performance may give rise to upregulation of *UCP2* and *UCP3* mRNA expression in the skeletal muscle of marathon runners and thus increased efficiency in the coupling of respiration and ATP production (Zhou et al., 2000, Tsuboyama-Kasaoka et al., 1998, Jiang et al., 2009).

Uncoupling proteins have also been implicated in the regulation of free radicals and reactive oxygen species (ROS) by mitochondria (Brand et al., 2002) and super-oxides (Echtay et al., 2002). It is generally though that UCP2 and UCP3 require an activator to transport hydrogen ions across the mitochondrial membrane, this is evidenced by the lack

of change in the basal mitochondrial hydrogen ion conductance in isolated mitochondria of the UCP2 or UCP3 knock out mouse (Brand and Esteves, 2005, Krauss et al., 2003, Cadenas et al., 2002, Couplan et al., 2002). One suggested activator is fatty acids (Echtay et al., 2002, Brand and Esteves, 2005). Fatty acids play an important role in heavily respiring skeletal muscle in the latter stages of a marathon, therefore based on the theory put forward by Brand and Esteves (2005) that UCPs 2 and 3 attenuate the production of free radicals by mitochondria and therefore protect against ROS related oxidative damage in muscle cells (Jiang et al., 2009) this may give a competitive advantage under racing conditions when elite runners often push themselves to the limits of their skeletal muscle capabilities (Joyner, 1991, Joyner and Coyle, 2008, Joyner et al., 2011).

Several gene polymorphisms in the genes of *UCP2* and *UCP3* have been identified as markers of aerobic capacity in 1423 Russian athletes when compared to 1132 controls (Ahmetov et al., 2009, Ahmetov et al., 2008)and oxidative phosphorylation (Astrup et al., 1999, Buemann et al., 2001) which make them attractive candidate genes for elite endurance athlete status (Buemann et al., 2001). The common polymorphism rs660339 in the *UCP2* gene results in a C/T substitution and subsequent Ala55Val amino acid change. The T allele has been associated with increased metabolic efficiency in muscles (Buemann et al., 2001, Astrup et al., 1999). The *UCP2* rs659366 gene polymorphism has been associated with training related enhancement in delta efficiency in, a marker of skeletal muscle performance in 58 participants (Dhamrait et al., 2012). Carriers of the T allele had higher training induce gains in delta efficiency, and it was reported in carriers of the T allele 8.4% of the interindividuality in delta efficiency associated with endurance training was the result of *UCP2* rs659366 polymorphism (Dhamrait et al., 2012).

Relatively few studies have been conducted assessing the association of the *UCP* polymorphisms and athlete status. Those that have been completed tend to assess *UCP2* rs660339 and *UCP3* rs1800849. In fact, there were no records in the literature of prior

studies assessing *UCP2* rs659366 and elite athlete status in endurance athlete populations. *UCP2* rs659366 is thought to be in linkage disequilibrium with *UCP2* rs660339 The genotype and allele frequency distributions of the *UCP2* rs660339 polymorphism were investigated in 230 Russian rowers and a control group of 855 Russian college and high school students. The frequency of the *UCP2* rs660339 T allele, thought to be associated with endurance performance, increased in correlation with the sporting eliteness of the rowers (Ahmetov et al., 2008) furthermore, the *UCP2* rs660339 T allele was associated with higher $\dot{V}O_{2\,max}$ in male Russian rowers (Ahmetov et al., 2008) suggesting the T allele may be favourable for competitive success in endurance sports. Further evidence to support this hypothesis found the *UCP2* rs660339 T allele was recorded at a higher frequency in an elite Russian mixed sporting discipline endurance cohort of 684 participants that included 134 runners, when compared to 1132 controls (Ahmetov et al., 2009). These studies also investigated another uncoupling protein polymorphism in *UCP3*.

The *UCP3* -55C/T polymorphism at rs1800849, results in a C/T substitution 55 nucleotides before the start codon in a functional promoter region. The T allele is associated with higher protein expression (Schrauwen et al., 1999) higher aerobic potential, increased basal energy expenditure thought to be due to the increased gene expression, therefore a decreased risk of obesity (Liu et al., 2005, Schrauwen et al., 1999). Ahmetov et al. (2008) investigated the *UCP3* rs1800849 genotype and allele frequencies in a large cohort of Russian rowers and controls (outlined above) they found the *UCP3* T allele increased in frequency in association with elite athlete status. The T allele was further associated with high aerobic performance and was therefore regarded as a genetic marker of endurance performance (Ahmetov et al., 2009). This research group completed further analysis in 287 long endurance (a race duration of 5-30 min) and middle endurance (race duration of 45 s to 5 min) groups of Russian national competitive standard and controls. They found the *UCP3* rs1800849 T allele carrier were more frequent in the athletes when compared to 1132 controls. Interestingly no association was

reported in the very long endurance group (a race duration of >30 min). Hudson et al. (2004) compared 89 fastest Caucasian male ultra-endurance Ironman triathletes and of the 89 slowest to determine whether the *UCP3* gene was associated with the performance in either the 2000 and 2001 South African Ironman Triathlon. There were no differences between the groups in either *UCP3* rs1800849 C/T genotype or allele frequency. These two athlete groups were further compared to 92 Caucasian male controls who had not trained for or participated in an ultra-endurance athletic event. No difference in either genotype frequency or allele frequency of the *UCP3* -55C/T polymorphism was noted between the groups analysed (Hudson et al., 2004). The athletes were grouped together and a genotype association with race completion time was conducted, this included further analysis of the completion times of the swim, cycle and run stages independently. No genotype association was found in any of the individual sporting stages (swim, cycle or run) or when the entire race event time was analysed (Hudson et al., 2004). Therefore it was concluded that the *UCP3* -55C/T polymorphism was not associated with endurance performance in tri-triathletes. Of note and a possible explanation for the lack of association in the triathlete population is the varying sporting elements that make up a triathlon and the phenotypes of the athletes that traditionally perform well in these individually sporting elements. Elite swimmers invariably have larger volumes of fat-free mass in their upper bodies whereas elite cyclists have over developed skeletal muscle in the thigh and lower legs, neither of these attributes are favourable to be successful in endurance running where the excess weight of increased fat-free mass Is hypothesised to be hindrance to running economy. Therefore the lack of association reported in these studies must be viewed with caution. The purpose of this study therefore was to assess an association of marathon sporting performance with the *UCP2* rs659366, *UCP2* rs660339 and *UCP3* rs1800849 via a gene candidate study.

3 METHOD

The experimental method outlined in this section was applied to all samples. Gene specific methods will be further outlined in the relative chapters.

3.1 ETHICS AND GOVERNANCE

The study was approved by the ethics committee of the Department of Exercise and Sports Science, Manchester Metropolitan University, in accordance with the Declaration of Helsinki for human research of 1974 (last modified in 2000). Sample collection, storage and disposal were within the framework of the Human Tissue Authority (HTA). All research participants gave their written informed consent after an explanation of the procedure and the purpose of this study. Participants had gap period of between 10 minutes and several days depending on location of data collection. Those with the smallest gap period were collected at the London Marathon Expo. Participants approached at the London Marathon Expo were fully informed of the project to elicit interest on the expo floor and then relocated to a quiet consultation space and fully informed once in the consultation space. Participants were then given 10 minutes to discuss with peers, coaches and management staff before being asked for consent to participate. Participants were informed they could withdraw from the study at any point even if consent had previously been granted.

3.2 ATHLETE SELECTION CRITERIA

The McCain Power of 10 website (*www.thepowerof10.info/)* records athlete particulars and sporting achievements on an annual basis. Published data records for athlete rankings date back to 2006. To determine the elite selection criteria these records were accessed, to assess the recorded times of the top 100 UK male and female marathon athletes (on 31ST December 2012). The data were averaged and the resulting mean time

was rounded to the nearest 10 mins. Table 4 shows the male and female annual recorded times used to set the data collection boundaries for the elite and sub-elite groups.

Table 1 Annual recorded times for UK male and female marathon runners

	Male top 100		Female top 100	
2012		02:33:06		03:00:20
2011		02:33:59		03:01:27
2010		02:33:36		03:01:22
2009		02:33:42		03:03:20
2008		02:34:18		03:02:36
2007		02:37:14		03:04:28
2006		02:34:53		03:01:32
Mean		02:34:24		03:02:09
Threshold times for marathon runner classification	Elite Male	02:30:00	Elite Female	03:00:00
	Sub-elite Male	02:45:00	Sub-elite Female	03:15:00

The athletes meeting the selection criteria were invited to take part in the project via their primary athletics club affiliation. The author also attended national competitions such as the National Cross Country Championships (2013) and the London Marathon Expo (2013-2015) to collect samples and data from athletes meeting the selection criteria.

3.3 NON-ATHLETE CONTROL SAMPLE SELECTION

Caucasian individuals who had never competed in high level sport (i.e. at an international, national or regional competition) were recruited to participate in the project as non- athlete control participants.

3.4 DATA COLLECTION

All research participants completed a questionnaire detailing their ethnicity, height, body mass and age (see Appendix 1 for details). Athletes had their personal bests and highest placed sporting achievements recorded. The control group recorded details of their personal health to confirm the lack of pathology and that participants were apparently 'healthy'.

3.5 DNA COLLECTION

All research participants gave a sample of DNA via one of the following methods depending on their preferences for an invasive / non - invasive method of collection. Because of the nature of the collections, primarily in the 4 days in the run up to the London marathon athletes were given the choice as to the method of DNA collection. It is acknowledged that the venepuncture site does take a few days to heal completely and this may cause irritation during an extended race like marathon and may affect race day performance. Therefore, DNA was collected in accordance with the method of choice of the athlete.

3.5.1 **Whole Blood**

A phlebotomist (the author and those marked in bold listed in table 4) collected a 5 mL venous blood sample from a superficial forearm vein. The phlebotomist transferred the sample into Ethylenediaminetetraacetic acid (EDTA) treated collection tubes and stored

on ice for transportation to the laboratory. The samples were frozen at -20˚C until DNA extraction was completed. To date 679 samples have been collected via this method (452 non-marathon controls and 227 marathon athletes). The author collected 20 non-marathon controls and 227 marathon athletes.

3.5.2 **Buccal Swab**

Research participants were asked not to eat, drink, smoke or chew gum for 30 min prior to sample collection. Participants were asked to collect a buccal DNA sample from either side of the mouth using two OmniSwabs (Whatman, Sanford) as shown in Figure 6. One swab was agitated on the inside of the right cheek to slough off loose cheek cells onto the swab. Swabs with collected cells were ejected from the 'handle' into 2 mL snap cap tubes and placed on ice for transportation. The process was repeated with the second swab on the left inside cheek. Samples were stored at -20˚C until DNA extraction was completed. 236 samples have been collected via this method (10 marathon athletes and 225 non-marathon controls). The author collected 6 marathon athlete samples via this method.

Figure 6 Buccal swab cell collection using OmniSwabs. Adapted from https://www.fishersci.com/shop/products/whatman-omniswab-sterileomniswab-100-pk/09923376 accessed 24/09/15

3.5.3 **Saliva Tube**

Research participants were asked not to eat, drink, smoke or chew gum for 30 min prior to sample collection. Each participant expelled 2 mL of saliva into the Oragene® DNA OG-500 (DNA Genotek Inc. Ottawa, ON, Canada) collection tube until the saliva fluid without bubbles reached the fill line as shown in Figure 7. The top was snapped closed to release the 'stabilisation' fluid stored in the cap into the sample. The funnel was removed, and the tube sealed with the smaller screw top plastic lid (provided). Samples were inverted 10 times to ensure even distribution of the 'stabilisation' fluid throughout the saliva sample. Samples were transported at room temperature. Samples were heat treated in an incubator for 1 hour at 56°C to complete cell lysis. Heat-treated samples were stored at room temperature until DNA extraction was completed. The author collected 112 athlete samples using this method.

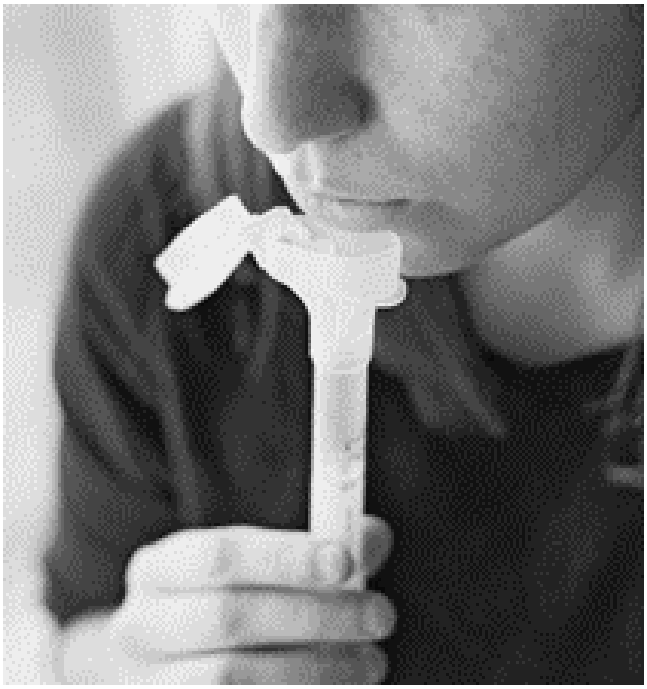

Figure 7 Saliva sample collection using Oragene® DNA OG-500. Adapted from http://www.macleans.ca/society/health/the-end-of-blood-samples/ accessed 24/09/15 original photo by Andrew Tolson.

Data collection of the non- athlete control cohort for this thesis was a collaborative effort with data included from several previous studies. These included data are attributed to their associated researchers and details of their contributions are presented in Table 5.

Table 2 Attribution details of the control samples collected by other contributors and used in the collective data set of this project.

PROJECT AUTHOR	PROJECT	PARTICIPANTS USED IN THIS PROJECT	SAMPLE COLLECTION METHOD	PRIOR ANALYSED GENE POLYMORPHISMS USED IN THIS PROJECT	GENES ANALYSED DURING THIS PROJECT
ROBERT ERSKINE	PhD	60	All Blood	*ACE I/D and ACTN3 rs*1815739	*AGT rs699, PPARGC1A rs8192678, UCP2 rs659336, UCP2 rs660339, and UCP3 rs1800349*
BRANDON FOSTER	PhD	160	All Buccal swab	*ACE I/D and ACTN3 rs*1815739	*None*
JAMES GAVIN	BSc	62	All Blood	*ACTN3 rs*1815739	*None*
JAMIE GUNEY	MSc	17	11 buccal 6 blood	*ACE I/D and ACTN3 rs*1815739	*AGT rs699, PPARGC1A rs8192678, UCP2 rs659336, UCP2 rs660339, and UCP3 rs1800349*
ADAM HERBERT	BSc and MSc	40	All Buccal swab	*None*	*ACE I/D, ACTN3 rs*1815739*, AGT rs699, PPARGC1A rs8192678, UCP2 rs659336, UCP2 rs660339, and UCP3 rs1800349*
JAMIE MCPHEE	PhD	61	All Blood	*ACE I/D, ACTN3 rs*1815739	*AGT rs699, PPARGC1A rs8192678, UCP2 rs659336, UCP2 rs660339, and UCP3 rs1800349*
GEORGINA STEBBINGS	PhD	120	22 Buccal 98 Blood	*ACE I/D and ACTN3 rs*1815739	*ACTN3 rs*1815739*, AGT rs699, PPARGC1A rs8192678, UCP2 rs659336, UCP2 rs660339, and UCP3 rs1800349*
ALUN WILLIAMS	GENESIS	67	All Blood	*ACTN3 rs*1815739	*None*

Those marked in bold are trained phlebotomists that took blood samples for this project

3.6 DNA EXTRACTION METHOD

3.6.1 Whole Blood And Heat Treated Saliva

DNA was extracted from a 200 µL sample of either peripheral, EDTA treated, anti-coagulated whole blood or heat-treated saliva using the automated QIAGEN® QIAcube spin column technique according to the manufacturer's instructions (QIAGEN, Crawley, West Sussex, UK). All buffers for DNA extraction were supplied in the QIAGEN® DNA Blood Mini kit (QIAGEN). The procedure is detailed in Figure 8.

Briefly, the 200 µL sample was heated to 56˚C for 10 min and lysed using QIAGEN Protease Enzyme. To the samples, 200 µL of 96% ethanol was added. The samples were agitated vigorously to mix. The mixture was transferred to a QIAamp mini spin column tube containing a DNA collection filter and was centrifuged at 6,000 g for 1 min. Three further wash buffer centrifugation cycles followed to wash out any remaining protein and impurities. During these wash cycles genomic DNA remains bound to the silica gel DNA filter membrane in the spin column. After each wash the filtrate was discarded. In the final buffer centrifugation, the sample genomic DNA was eluted into 100 µL of low salt buffer to provide purified genomic DNA. The purified genomic DNA was stored at 4˚C until the genotyping was performed. The author extracted 359 blood and saliva samples using this method

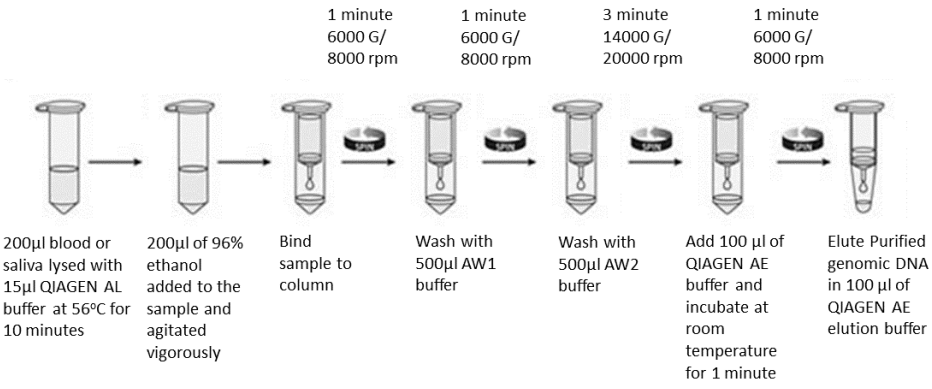

Figure 8 The QIAGEN® DNA Blood Mini procedure performed by the QIACUBE®. Image adapted from https://tools.thermofisher.com/content/sfs/gallery/high/4958.jpg accessed 24/09/15

3.6.2 **Buccal Swab**

DNA was extracted from the buccal swab using the automated QIAGEN® QIAcube spin column technique according to the manufacturer's instructions (QIAGEN). All buffers for DNA extraction were supplied in the QIAGEN® DNA Blood Mini kit (QIAGEN).

Briefly, buccal swabs were immersed in 600 µL of phosphate buffered solution (PBS). Samples were lysed using 20 µL of QIAGEN Protease Enzyme and 600 µL of buffer AL was added. The sample was briefly agitated to mix and incubated at 56˚C for 10 mins. To the samples 600 µL of 96% ethanol was added and again the samples were agitated vigorously on the mixer. Then 700 µL of the mixture was transferred to a QIAamp mini spin column tube containing a DNA collection filter and was centrifuged at 6,000 g for 1 min. Three further wash buffer centrifugation cycles followed to wash out any remaining protein and impurities. Samples were centrifuged at 20,000 g for 3 min. During these wash cycles genomic DNA remains bound to the DNA filter membrane in the spin column. After each wash cycle the filtrate was discarded. In the final buffer centrifugation at 6,000

g for 1 min, the sample genomic DNA was eluted into 100 µL of low salt buffer to provide purified genomic DNA. The purified genomic DNA was stored at 4˚C until the process of genotyping was performed. The author extracted 6 samples using this method.

3.7 DETERMINATION OF DNA PURITY AND YIELD

A Biophotometer Plus (Eppendorf UK Limited, Stevenage, UK) was used to assess the purity (whether the protein had been removed) of the extracted genomic DNA samples. Approximately 10 µL of extracted genomic DNA sample was placed into a cuvette and exposed to 260 nm of ultra violet light (optimal wavelength absorption of DNA), and then exposed to 280 nm of ultra violet light (optimal wavelength absorption of protein). The ratio of absorbance at 260 nm:280 nm was calculated and ratios within the range 1.7-2.1 were acceptable to use as PCR template (Ehli et al., 2008). Any samples that failed to achieve this level of purity were re-run through the DNA extraction process using a new QIAamp spin column. After purification and purity testing the remainder of genomic DNA sample was stored at 4˚C until genotyping analysis was completed. A total of four samples were re-run all were initially collected using the buccal swab method. It may be inferred that the dna yield was less from the buccal samples for these for samples. However, on re running the samples as outlined above they were within the acceptable range for use as template.

3.8 GENOTYPING

Genotyping was determined using real time polymerase chain reaction (rtPCR). Genotyping assay mixes including pre-designed primers and probes and Taqman® Genotyping Master Mix (Life Technologies, Carlsbad, California, USA). Genotyping was completed on the Chromo4 (BioRad Laboratories Ltd, Hertfordshire, UK), Lightcycler (Roche, West Sussex, UK) and StepOne Plus (Life Technologies, Carlsbad, California, USA) thermocycling and detection platforms. rtPCR is an *in vitro* method of producing

large amounts of specific DNA fragments from small amounts of DNA template. During a single cycle of rtPCR oligonucleotide primers hybridise with the nucleotide sequences on the complementary template strands at each end of the DNA fragment to be amplified (Figure9). A single stranded oligo probe synthesised to hybridise to the DNA sequence between the two primers binds to the template strand (Figure 11). The probe has a fluorescent tag conjugated to a terminal nucleotide. At the opposite end of the probe sequence, a quencher is tagged to the terminal nucleotide. In close proximity, the quencher rapidly absorbs any fluorescence emitted by the fluorescent tag (Figure 10). The polymerase enzyme that secures individual nucleotides to the complementary template strand also has a secondary function as an exonuclease. This means that when the enzyme encounters the double stranded DNA of the bound probe in its synthesis path, it will disassemble the strand in its way and replace all of the nucleotides in effect destroying the anchoring mechanism of the reporter and quencher tags (Figure 12). This means they are no longer bound in close proximity. In the absence of a nearby quencher, the fluorescent tag is free to emit detectable light (Figure 13). Each time another amplicon is produced another fluorescent marker is released from its neighbouring quencher. Therefore, as the number of PCR amplicons doubles during each PCR cycle so does the amount of emitted fluorescence. Cyclical polymerisation by thermostable DNA polymerase produces millions of identical copies of the DNA fragment of interest thus allowing detection of the emitted fluorescence by the fluorimeter.

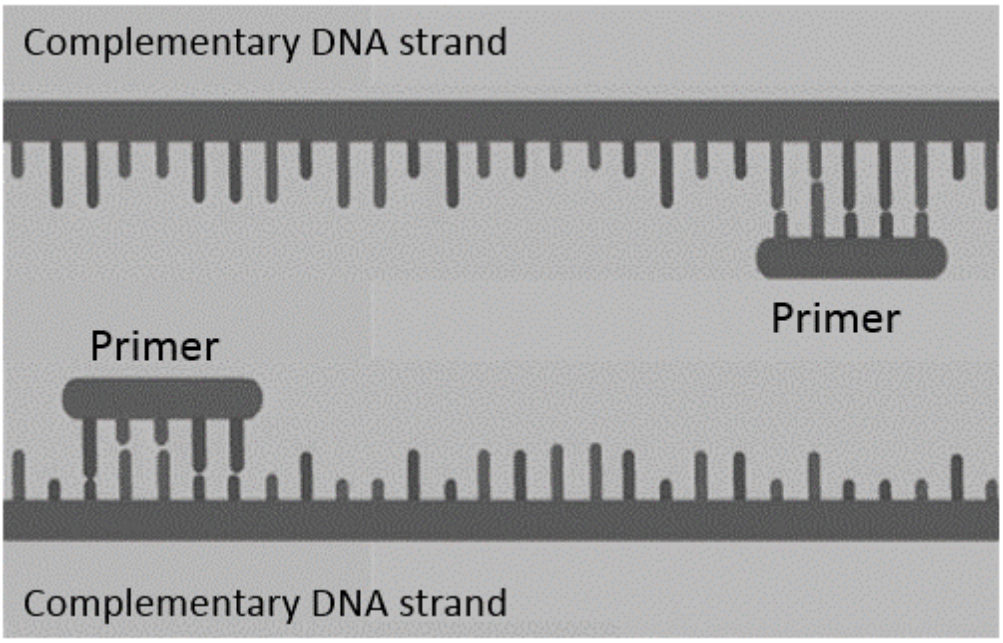

Figure 9 Oligonucleotide primers hybridise with the nucleotide sequences on the complementary template strands

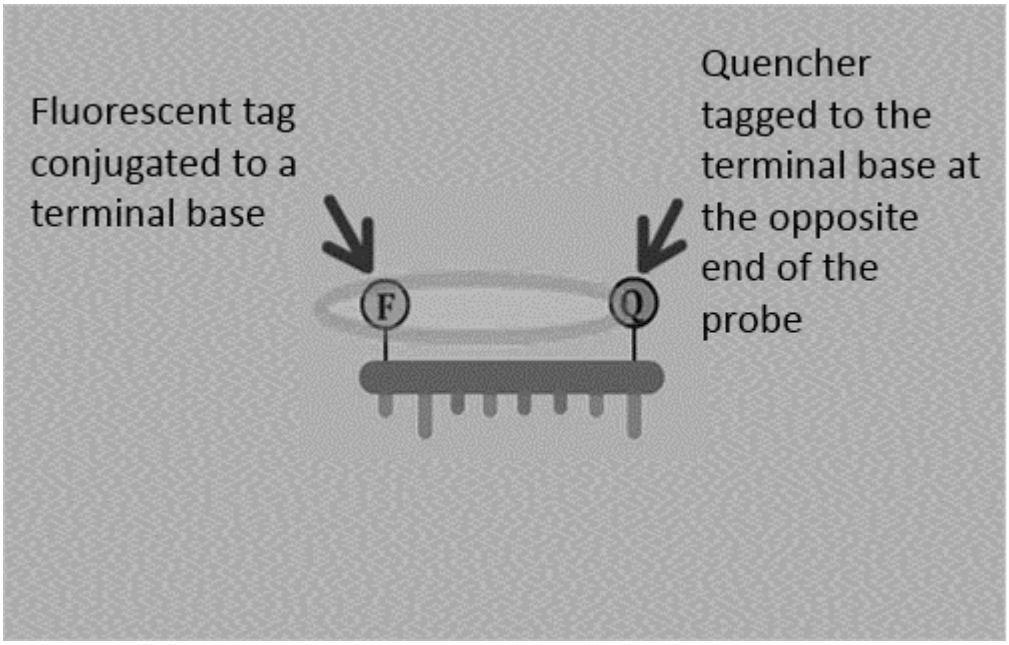

Figure 10 Oligo probe synthesised to hybridise to the DNA sequence

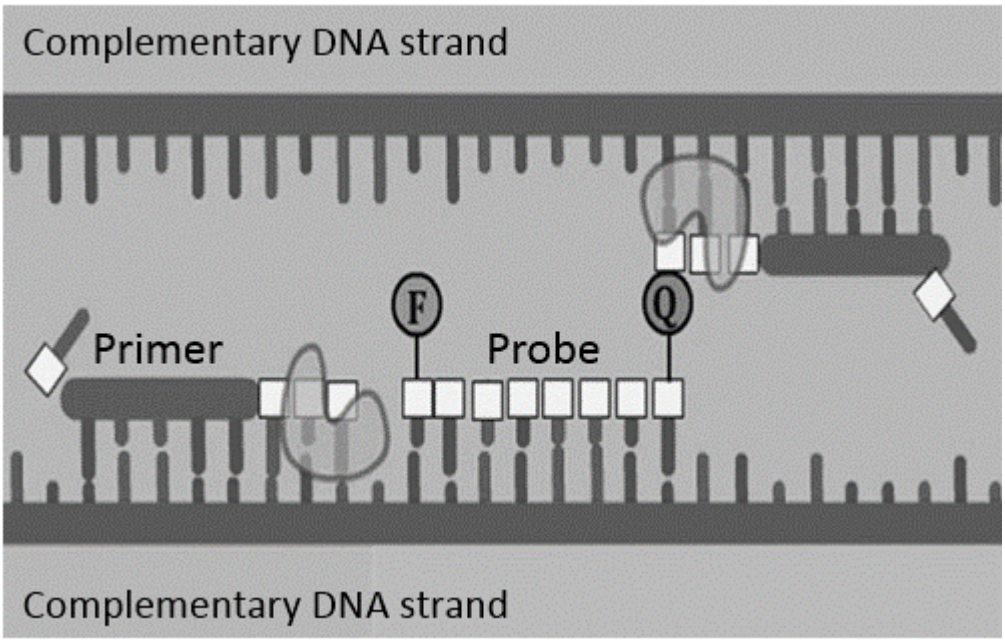

Figure 11 In close proximity, the quencher rapidly absorbs any fluorescence emitted by the fluorescent tag

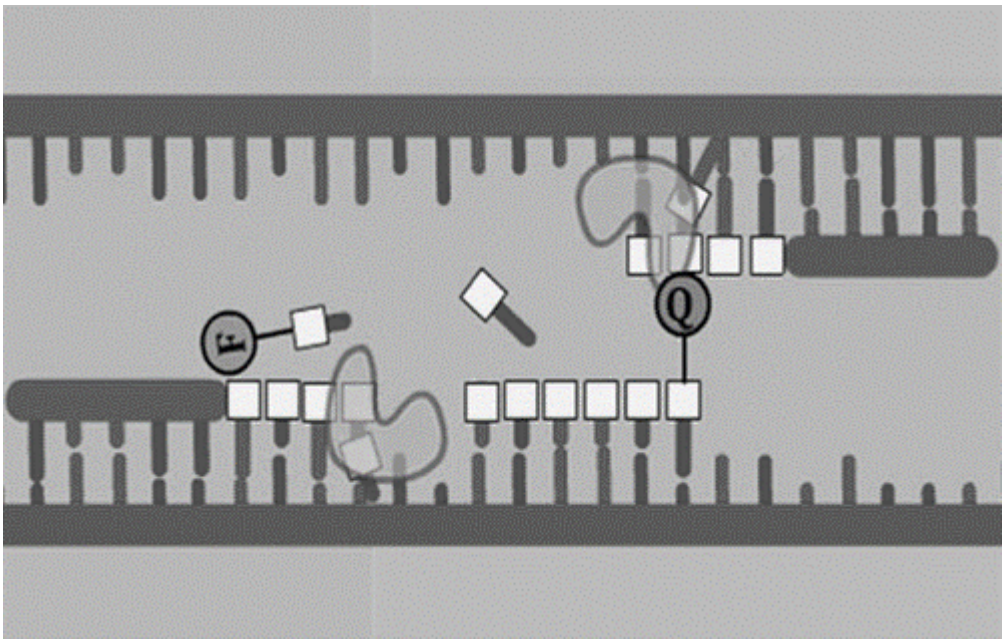

Figure 12 When the exonuclease encounters the double stranded DNA of the bound probe it digests it releasing the fluorescent tag

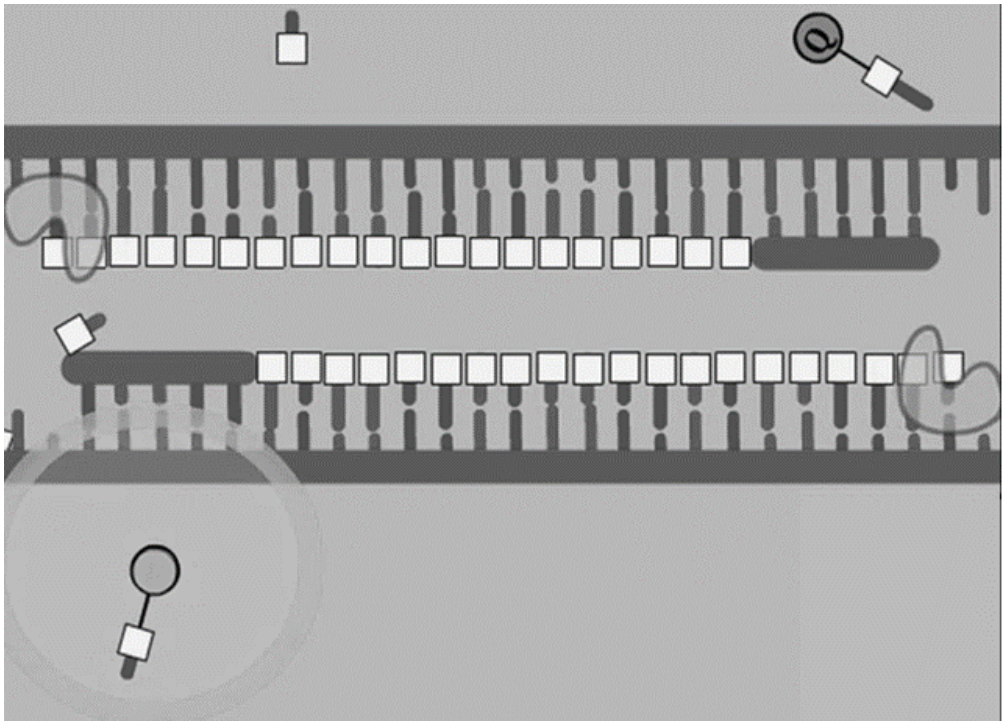

Figure 13 The fluorescent tag is free to emit detectable light. Figures 9-13 are adapted from the animation video at https://www.youtube.com/watch?v=kvQWKcMdyS4 developed by Yaw Adu-Sarkodie of Kwame Nkrumah University of Science and Technology and Cary Engleberg of University of Michigan. Copyright 2009-2010, Kwame Nkrumah University of Science and Technology and Cary Engleberg.

Reactions were all 10 µL volume. Thermocycling conditions were 95˚C for an initial 10 min, then 40 cycles of 15 s at 95˚C and 60 s at 60˚C (annealing and extension). Genotypes were determined using Opticon Monitor 3.1 software (BioRad Laboratories Ltd, Hertfordshire, UK), Lightcycler 96 sw1.1 software (Roche, West Sussex, UK) or StepOnePlus software version 2.3 (Life Technologies, Carlsbad, California, USA). All genotyping was completed in duplicate and there was 100% agreement between duplicates.

Within a single gene polymorphism, where the athlete cohort was analysed across multiple thermocycling and detection platforms, a selection of control samples representing the three genotype groups were also genotyped across those same

platforms, in duplicate to ensure continuity of the results. There was 100% agreement across the genotyping platforms.

3.8.1 **ACE I/D Polymorphism Genotyping Using The Method Outlined By Koch et al. (2005)**

Three hundred and thirteen control samples were genotyped for the *ACE* I/D polymorphism using the method detailed by Koch et al. (2005). Reactions were all 10 µL volume and contained the elements shown in Table 6. Thermocycling conditions were 95˚C for an initial 10 min, then 50 cycles of 15 s at 92˚C (denaturation) and 60 s at 57˚C (annealing and extension). All reactions were completed in duplicate. There was 100% agreement between duplicate samples. Figure 14 shows a schematic of the Koch et al. (2005) primer and probe locations. To confirm identical genotyping results between the Koch et al. (2005) and the tagSNP rs4341 method a small number of samples, including those of II, ID and DD genotype, were assayed using both methods giving identical genotype results in all cases.

Table 3 rtPCR reaction materials for the Koch et al. (2005) method

PCR Material [concentration]	Amount required for DNA derived from whole blood and saliva (µL)	Amount required for DNA derived from buccal lysate (µL)
Genotyping Master Mix	5	5
Nuclease free H_2O	1.55	0.05
I allele specific probe [150 nm]	0.9	0.9
D allele specific probe [75 nm]	0.9	0.9
ACE primer 111 [150 nm]	0.38	0.38
ACE primer 112 [150 nm]	0.38	0.38
ACE primer 113 [150 nm]	0.38	0.38
DNA solution	0.5	2

All DNA used as template was analysed for purity and yield and fell within the range of 1.7-2.1 µg/ ml

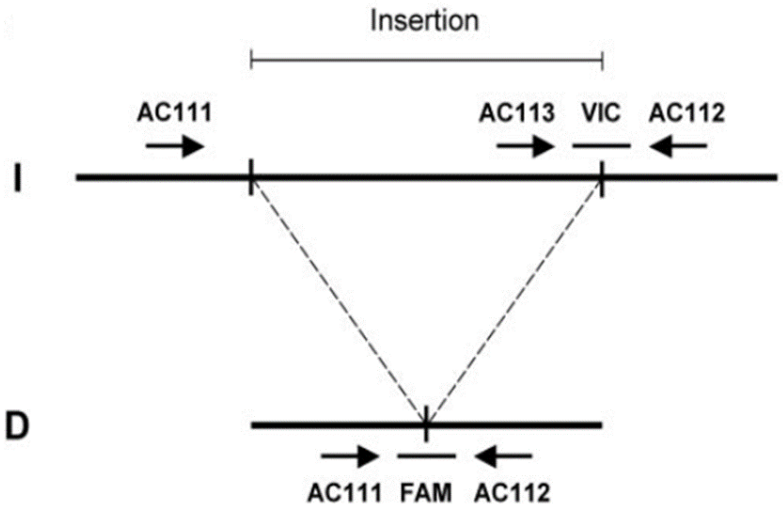

Figure 14 Schematic of the binding sites of primers ACE111, ACE112, and ACE113 and probes VIC-AC100 (VIC) and FAM AC100 (FAM) adapted from Koch et al. (2005).

3.9 STATISTICS

This section will give an overview of the tests used to analyse the data presented in each chapter of this thesis. All descriptive statistics were determined in SPSS version 21 (IBM corporation, Florida, USA). To determine differences in height and body mass distributions amongst athletes and controls independent t-tests were completed in SPSS version 21 (IBM corporation). Hardy-Weinberg Equilibrium assessment was completed using Excel (Microsoft). Calculation and comparison of genotype and allele frequencies was performed in SPSS version 21 (IBM corporation) using a Pearson's chi-squared test where Hardy-Weinberg Equilibrium was apparent. In the few instances where there was a departure from hardy Weinberg-Equilibrium a Cochran-Armitage trend test was used to assess genotype and allele comparisons. This method was chosen as it modifies the Pearson Chi squared test to incorporate a suspected ordering to the effects of the categories. This test is often used a genotype based test for case controlled genetic association studies. Benjamini Hochberg corrections for multiple testing were applied where appropriate using Excel (Microsoft). In this method of correction the p-values are ordered in descending order therefore the most likely hypothesis, those supported most strongly by the evidence are assessed first. This method of correction gives a more stringent alpha value from 0.05 by dividing the rank ordered p-values by the total number of hypothesis being tested. In this analysis each gene polymorphism with be treated a separate category.

To determine any association of genotype with personal best one way ANOVA were completed in SPSS. Positive associations were corrected using the Brown-Forsythe method in SPSS. This is because when a 1 way ANOVA is performed the distributions are assumed to have equal variance if this assumption is not valid as in the F test is invalid and thus it is necessary to use the Brown-Forsythe correction to adjust the F statistic to .the absolute deviations from the median.

4 GENES ASSOCIATED WITH BLOOD PRESSURE REGULATION DURING A MARATHON – ACE AND AGT

4.1 INTRODUCTION

Athlete fuel and oxygen delivery through adequate respiratory and haemodynamic perfusion relies on cardiac output and adequate blood pressure regulation (Boron and Boulpaep, 2012). Theoretically during a race $\dot{V}O_{2\,max}$ could be limited by the cardiac output, the respiratory systems ability to deliver oxygen to the blood or the exercising muscles ability to use oxygen. In untrained individuals it is usually cardiac output that determines $\dot{V}O_{2\,max}$. Cardiac output is directly linked to blood pressure through the following equation

Blood pressure = cardiac output x Total Peripheral Resistance

With increasing workload one determinant of cardiac output (heart rate) increases progressively until it reaches a maximum. With training stroke volume also increases though to a lesser degree. As both these elements contribute to cardiac output and thus dynamic blood pressure, blood pressure can therefore be viewed as an essential element to endurance performance at an elite level in events such as marathon. The RAAS controls the haemodynamic status of blood pressure through endocrine mediated pathways that include the two key elements AGT and ACE (Boron and Boulpaep, 2012).

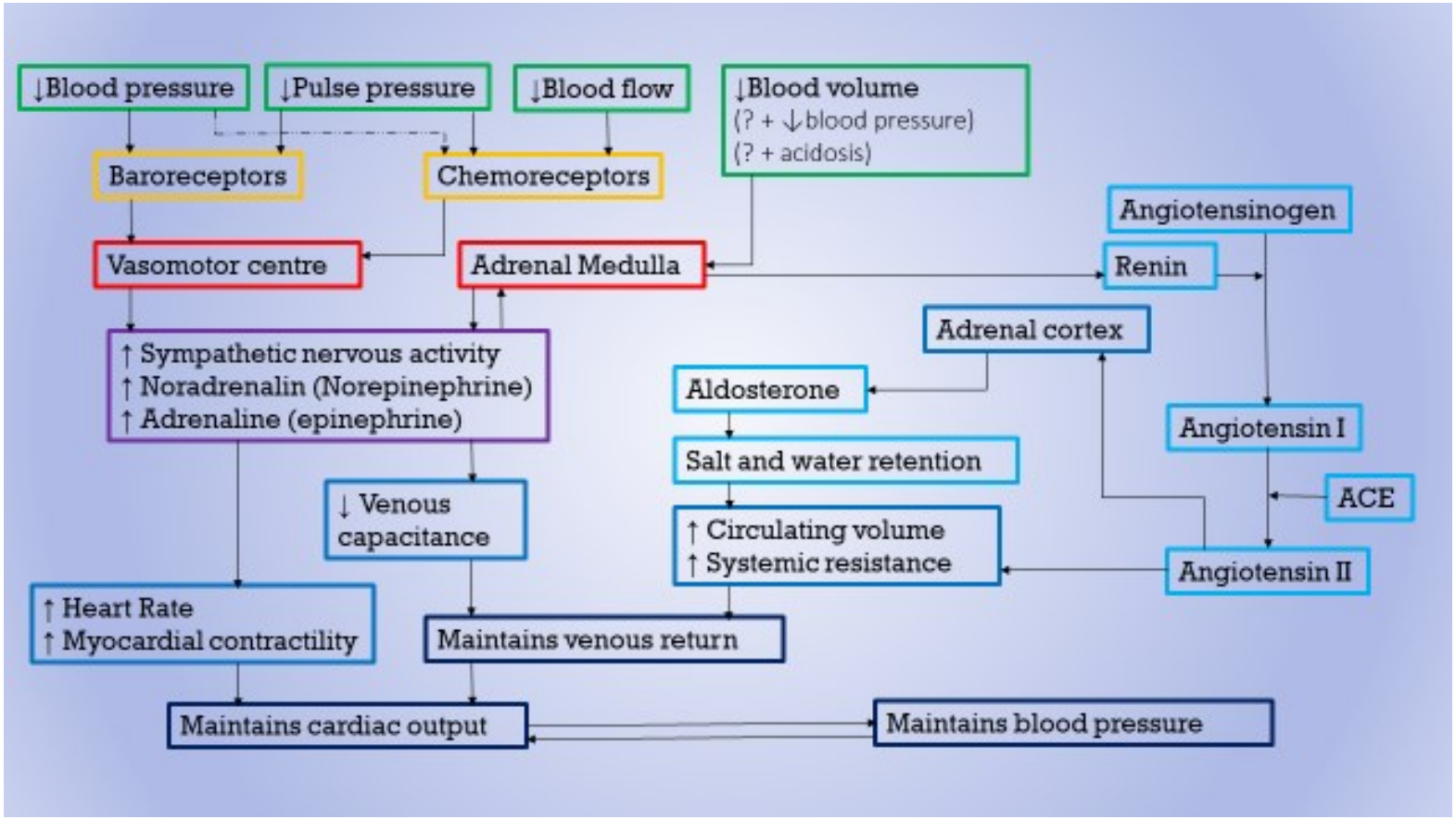

Figure 15 Adapted from Figure 25.18 The Sympatho-adrenal response to shock. Page 1151 chapter 25 Kumar and Clark. The effect of increased catecholamines is shown on the left of the diagram, and the release of angiotensin and aldosterone on the right. Both mechanisms help maintain the cardiac output and blood pressure in during a marathon.

AGT is released by the liver in response to hypovolaemia prolonged hypotension and reduced blood sodium concentrations, it is cleaved by renin to form ANGI. ANGI is further cleaved by ACE to the potent vasoconstrictive peptide ANGII (see Figure 15). ANGII mediates actions via the AT-1 receptor cause vasoconstriction via the following three mechanisms. Initially ANGII causes efferent arteriole smooth muscle constriction directly via the AT-1 receptor. Resulting in increased total peripheral resistance and hydrostatic pressure and reduced blood flow through the kidney increasing the glomerular filtration rate in an attempt to regulate blood volume (Boron and Boulpaep, 2012). Binding of ANGII to the AT-1 receptor induces the hypertensive response of a reduction in the synthesis of nitric oxide. Reduction in nitric oxide synthesis and ANGII mediated AT1-receptor activation induces aldosterone release and an increase in the reabsorption of sodium and water from the nephron lumen increasing the blood volume and thus hydrostatic pressure (Boron and Boulpaep, 2012). ANGII further mediates increases in blood pressure through the AT1-receptor in cardiac cells, causing myocyte hypertrophy and increases in cell number. The increased muscle mass increases cardiac inotropy, stroke volume, cardiac output and thus blood pressure and may be beneficial to meet the increased oxygen demand in heavily respiring skeletal muscle during training and competition in endurance events such as marathon.

The association of the *ACE* I/D polymorphism with human performance according to the literature appears unclear. The literature suggests that larger cardiac muscle mass is associated with the *ACE* D allele and *AGT* C allele, and both of these alleles have been associated with sporting performance in power based events suggesting some commonality between cardiac and skeletal muscle in mechanisms of hypertrophy (Gomez-Gallego et al., 2009, Ruiz et al., 2010, Ben-Zaken et al., 2013, Zarebska et al., 2013). An established clear link between the *ACE* I allele or *ACE* II genotype and endurance performance remains to be elucidated. Though several studies in British (Myerson et al., 1999), Italian (Scanavini et al., 2002) and Spanish (Alvarez et al., 2000) athletes have established a link there are numerous studies that have found no

association in elite endurance sporting populations (Scott et al., 2005, Oh, 2007, Papadimitriou et al., 2009, Ash et al., 2011). Puthucheary et al. (2011) were the first to review the published literature on the *ACE* I/D polymorphism (1998-2010) with associations in sporting performance. They suggested a tendency for carriers of the *ACE* I allele to have superior endurance sporting performance. In an attempt to provide a more solid evidence base for an association between the *ACE* II genotype and endurance based events a meta-analysis was recently competed. The study by Ma et al. (2013) systematically reviewed 25 journal articles addressing the association of the ACE I/D polymorphism and sporting performance. Subsequent sporting sub group analysis of endurance (17 studies) and power groups (13 studies) was completed. The main finding of the meta analysis was an association of the *ACE* II genotype with physical performance when compared to *ACE* D allele carriers (OR,1.23; CI, 1.05-1.45). Further to this, the *ACE* II genotype was positively associated with performance in endurance athletes (OR,1.35; CI, 1.17-1.55).

A lack of clear association with elite athlete status is further reflected in the published *AGT* rs699 studies. Several studies have reported an association of the *AGT* CC genotype with increased left ventricular mass in endurance athletes either when analysed alone (Karjalainen et al., 1999) or in combination with the *ACE* DD genotype (Diet et al., 2001). Other have reported no association in 75 elite endurance cyclists and 70 elite distance runners (Alvarez et al., 2000, Gomez-Gallego et al., 2009) and in 123 elite and sub-elite endurance athletes of mixed discipline including marathon runners (n=12), triathletes (n=4) and road cyclists (n=14) (Zarebska et al., 2013).

The results reported thus far in the literature are inconsistent and are difficult to interpret, with changing nomenclature in the reporting of results, particularly in the *AGT* rs699 polymorphism literature. Therefore, this study aimed to assess the difference in *ACE* I/D and *AGT* rs699 genotype and allele frequency between athletes and non-athletes. It was

hypothesised based on the literature that the *ACE* II genotype and the *AGT* rs699 CC genotype would be overrepresented in the athletes when compared to non-athletes. In addition, comparisons of both genotype and allele frequency and elite athletes status were completed. The hypothesis stated the *ACE* I allele and the *AGT* C allele would be more frequently observed in elite athletes when compared to sub-elite athletes and non-athletes. Further to this, an investigation of whether genotype was associated with personal best time was completed in male and female marathon runners.

4.2 METHOD

4.2.1 Research Participant Characteristics

Nine hundred and thirty two Caucasian adults provided written informed consent to take part in the *ACE I/D* analysis. This total cohort comprised 399 marathon runners (male, n = 243; female, n = 156) and a non-athlete (non-marathon running) cohort, comprising 337 men and 196 women (Table 7). Six hundred and seventy Caucasian adults provided written informed consent to take part in the *AGT* rs699 analysis. This total cohort comprised 364 marathon runners (male, n = 216; female, n = 148) and a non-athlete (non-marathon running) cohort, comprising 224 men and 82 women (Table 8). The marathon runners were stratified into elite and sub-elite subgroups according to their official marathon personal best performance time (http://www.powerof10.co.uk). *ACE I/D* research participant age, height and body mass are shown in Table 9. *AGT* rs699 research participant age, height and body mass are shown in Table 10

Table 4 Numbers of research participants in the ACE *I/D analyses*

	Male	Female	Total
Marathon Athletes	243	156	399
Elite Marathon Athletes	86	87	173
Sub-Elite Marathon Athletes	157	69	226
Non-Athletes	337	196	533

Table 5 Numbers of research participants in the AGT *rs699 analyses*

	Male	Female	Total
Marathon Athletes	216	148	364
Elite Marathon Athletes	75	82	157
Sub-Elite Marathon Athletes	141	66	207
Non-Athletes	224	82	306

Table 6 ACE *I/D participant characteristics (mean (standard deviation))*

		Athletes	Elite	Sub-Elite	Non-Athletes
Height (m)	Male	1.78 (0.06)	1.79 (0.06)	1.78 (0.06)	1.79 (0.07)
	Female	1.65 (0.07)	1.65 (0.07)	1.66 (0.08)	1.65 (0.07)
Mass (kg)	Male	67.0 (6.5)	67.1 (7.5)	67.0 (5.9)	77.7 (11.2)
	Female	53.6 (5.3)	52.7 (5.4)	54.8 (5.0)	66.7 (12.7)
Age (years)	Male	36 (8)	37 (10)	35 (6)	24 (9)
	Female	37 (7)	37 (8)	37 (7)	29 (15)

Table 7 AGT *rs699 participant characteristics (mean (standard deviation))*

		Athletes	Elite	Sub-Elite	Non-Athletes
Height (m)	Male	1.79 (0.06)	1.79 (0.06)	1.78 (0.07)	1.79 (0.06)
	Female	1.65 (0.07)	1.65 (0.07)	1.66 (0.08)	1.65 (0.07)
Mass (kg)	Male	67.0 (6.4)	66.6 (6.8)	67.2 (6.1)	77.0 (11.3)
	Female	53.8 (5.2)	53.0 (5.2)	54.7 (5.1)	66.3 (11.0)
Age (years)	Male	36 (8)	37 (10)	35 (7)	23 (7)
	Female	37 (7)	37 (8)	37 (7)	25 (10)

4.2.2 **DNA Collection**

Participant DNA collection is outlined in section 3.5 of the methods. In brief, 588 blood samples, 231 buccal samples and 113 saliva samples were collected from participants for the *ACE* I/D analysis. Participant DNA collection from 507 blood samples, 68 buccal samples and 95 saliva samples was completed for *AGT* rs699 analysis.

4.2.3 **DNA Isolation**

Automated DNA extraction was performed using a QIAcube (Qiagen, Crawley, UK), following the QIAamp blood mini protocol as detailed in the methods section. Concisely, 200 µL of blood or saliva was lysed with QIAGEN Protease Enzyme. To each lysed sample, 200 µL of 96% ethanol was added. Samples were agitated vigorously to mix. The sample mixture was transferred to a QIAamp mini spin column tube containing a DNA collection filter and bound by centrifugation at 6000 g for 1 min. Remaining protein and impurities were removed by three further wash buffer centrifugation cycles. Genomic DNA remained bound to the silica gel DNA filter membrane in the spin column during these wash cycles. After each wash the filtrate was discarded. In the final buffer centrifugation, the participant genomic DNA was eluted into 100 µL of molecular grade H2O to provide purified genomic DNA. The purified genomic DNA was stored at 4°C until the genotyping was performed.

4.2.4 **Genotyping**

ACE I/D genotyping of each participant was completed by real-time polymerase chain reaction (rtPCR). Three hundred and thirteen non-athletes were genotyped for the *ACE* I/D polymorphism using the method outlined by Koch et al. (2005). In brief, each 10 mL reaction contained: 5 mL Genotyping Master Mix (Applied Biosystems, Foster City, California, USA), 1.55 µL nuclease-free H_2O (Qiagen), 0.9 µL of I and D allele-specific probes and 0.38 µL of *ACE* primer 111, 112, 113 (sequences below) were combined with 0.5 µL DNA solution per well for blood and saliva. For DNA derived from buccal cells,

identical primer and probe volumes were used but 0.05 µL H_2O and 2 µL DNA solution were used.

Primers and probes produced by applied biosystems (Outlined in Figure 14)

For the direct *ACE* I/D assay (Koch et al., 2005), three primers (150 nM each) and probes VIC (150 nM) and FAM (75 nM) were used;

Primer ACE111: 5ˈ-CCCATCCTTTCTCCCATTTCTC-3ˈ

Primer ACE112: 5ˈ-AGCTGGAATAAAATTGGCGAAAC-3ˈ

Primer ACE113: 5ˈ-CCTCCCAAAGTGCTGGGATTA-3ˈ

I Allele specific probe (VIC-ACE100): VIC-5ˈAGGCGTGATACAGTCA-3ˈ-MGB

D Allele specific probe (FAM-ACE100): FAM-5ˈTGCTGCCTATACAGTCA-3ˈ-MGB

ACE I/D genotype was established for all participants by rtPCR using either a Chromo4 or StepOnePlus real-time PCR system. Briefly, there were 50 cycles of denaturation at 92°C for 15 s then annealing and extension at 57°C for 1 min. Initial analysis was performed using Opticon Monitor 3.1 software for the Chromo4 or StepOnePlus software version 2.3. There was 100% agreement within duplicates of all samples.

Five hundred and fifty nine athletes and non-athletes were *ACE* I/D genotyped using the TaqMan assay for rs4341 that contained the appropriate TaqMan primers and probes (Applied Biosystems). Each 10 µL rtPCR experiment contained 5 µL Genotyping Master Mix (Applied Biosystems), 4.3 µL nuclease-free H_2O (Qiagen), 0.5 µL *ACE* rs4341 TaqMan genotyping assay mix (Applied Biosystems), and 0.2 µL of participant DNA. For control wells, 0.2 µL nuclease-free H_2O replaced the DNA template. *ACE* rs4341 TaqMan genotyping was completed on the StepOnePlus real-time PCR system. There were 40 cycles of denaturation at 95°C for 15 s then annealing and extension at 60°C for 1 min

according to the manufacturer's instructions. StepOnePlus software version 2.3 was used for analysis. There was 100% agreement within duplicates of all samples.

AGT rs699 genotype was also established for all participants by rtPCR. The genotyping protocol was identical to that described in the preceding paragraph, except the *AGT rs699* TaqMan assay that included the appropriate TaqMan primers and probes was used (Applied Biosystems).

4.2.5 **Data Analysis and Statistics**

SPSS was used to perform independent T-tests to compare height and body mass between athletes and non-athletes. Genotype and allele frequencies in all groups were calculated and compared using chi-square analyses in SPSS (version 21, SPSS Inc). Excel (Microsoft, 2013) was used to assess Hardy-Weinberg equilibrium. Genotype associations with personal best were assessed in SPSS by one way ANOVA.

4.3 *ACE* I/D RESULTS

4.3.1 **Height and Body Mass**

Self-reported participant height and body mass are shown in Table 9. Male non-athletes were 10.7 kg heavier than male marathon athletes ($p = 1.0 \times 10^{-13}$) and female non-athletes were 13.1 kg heavier than female marathon athletes ($p = 1.0 \times 10^{-13}$). There was no difference in height between marathon athletes and non-athletes for either males ($p = 0.229$) or females ($p = 0.507$).

4.3.2 **ACE I/D Hardy-Weinberg Equilibrium**

The ACE I/D genotype distributions amongst research participants were in Hardy-Weinberg equilibrium (Marathon athlete $\chi 2 = 0.011$, Elite marathon athletes $\chi 2 = 0.212$,

Sub-Elite marathon athletes $\chi 2 = 0.299$, and Non-Athlete $\chi 2 = 1.488$, $p > 0.05$ for all groups). When stratified by sex, there was also no deviation from Hardy-Weinberg equilibrium (Males: Marathon athlete $\chi 2 = 0.053$, Elite marathon athletes $\chi 2 = 0.731$, Sub-Elite marathon athletes $\chi 2 = 0.127$, and Non-Athlete $\chi 2 = 1.124$, $p > 0.05$ for all groups), (Females: Marathon athlete $\chi 2 = 0.208$, Elite marathon athletes $\chi 2 = 0.044$, Sub-Elite marathon athletes $\chi 2 = 0.204$, and Non-Athlete $\chi 2 = 0.392$, $p > 0.05$ for all groups).

4.3.3 **ACE I/D Genotype and Allele Frequencies**

Table 11 details the ACE I/D genotype and allele frequency distributions in the marathon athlete and non-athlete cohorts. The primary analysis between the entire marathon athlete cohort and the non-athletes showed no genotype or allele association of the *ACE* I/D polymorphism and endurance marathon running (genotype $\chi^2 = 2.145$, $p = 0.342$; allele $\chi^2 = 1.143$, $p = 0.285$). Further analysis considering genotype and allele frequencies in elite marathon and sub-elite marathon athlete cohorts compared independently to non-athletes revealed no *ACE* I/D polymorphism associations (elite marathon athletes vs non-athletes genotype $\chi^2 = 2.632$, $p = 0.268$; allele $\chi^2 = 1.167$, $p = 0.280$). For completeness, the elite marathon athlete cohort was compared to the sub-elite marathon athlete cohort and again no differences were observed (genotype $\chi^2 = 0.752$, $p = 0.687$; allele $\chi^2 = 0.252$ $p = 0.616$). In essence, neither genotype nor allele associations were observed when males and females were combined in any group comparison (marathon athletes, elite marathon athletes, sub-elite marathon athletes, non-athletes).

Table 8: *ACE* I/D genotype and allele frequencies in marathon athletes and non-athletes.

	***ACE* Genotype**						***ACE* Allele**			
Total Cohort	**DD**		**ID**		**II**		**D**		**I**	
Marathon athletes	123	(30.8%)	198	(49.6%)	78	(19.5%)	444	(55.6%)	354	(44.4%)
Elite Marathon athletes	57	(32.9%)	82	(47.4%)	34	(19.7%)	196	(56.6%)	150	(43.4%)
Sub-Elite Marathon athletes	66	(29.2%)	116	(51.3%)	44	(19.5%)	248	(54.9%)	204	(45.1%)
Non-Athletes	147	(27.6%)	279	(52.3%)	107	(20.1%)	573	(53.8%)	493	(46.2%)
Males										
Marathon athletes	77	(31.7%)	118	(48.6%)	48	(19.8%)	272	(56.0%)	214	(44.0%)
Elite Marathon athletes	31	(36.0%)	38	(44.2%)	17	(19.8%)	100	(58.1%)	72	(41.9%)
Sub-Elite Marathon athletes	46	(29.3%)	80	(51.0%)	31	(19.7%)	172	(54.8%)	142	(45.2%)
Non-Athletes	94	(27.9%)	177	(52.5%)	66	(19.6%)	365	(54.2%)	309	(45.8%)
Females										
Marathon athletes	46	(29.5%)	80	(51.3%)	30	(19.2%)	172	(55.1%)	140	(44.9%)
Elite Marathon athletes	26	(29.9%)	44	(50.6%)	17	(19.5%)	96	(55.2%)	78	(44.8%)
Sub-Elite Marathon athletes	20	(29.0%)	36	(52.2%)	13	(18.8%)	76	(55.1%)	62	(44.9%)
Non-Athletes	53	(27.0%)	102	(52.0%)	41	(20.9%)	208	(53.1%)	184	(46.9%)

Independent analysis of the male marathon athlete cohort detailed in Table 11 demonstrated no difference in genotype or allele frequencies when compared to non-athletes (genotype $\chi^2 = 2.164$, $p = 0.339$; allele $\chi^2 = 0.959$, $p = 0.327$). Nor were any differences observed when comparing the elite marathon athletes and non-athlete (genotype $\chi^2 = 3.333$ $p = 0.189$; allele $\chi^2 = 1.332$, $p = 0.248$), sub-elite marathon athletes and non-athlete (genotype $\chi^2 = 0.235$, $p = 0.889$; allele $\chi^2 = 0.133$, $p = 0.716$) or elite marathon athlete and sub-elite marathon athlete groups (genotype $\chi^2 = 1.323$, $p = 0.516$; allele $\chi^2 = 0.541$, $p = 0.462$).

Also shown in Table 11 are the genotype and allele frequencies of the female participants. Similarly to the men, there was no difference in genotype or allele distribution frequency between female marathon athletes and non-athletes (genotype $\chi^2 = 0.295$, $p = 0.863$; allele $\chi^2 = 0.238$, $p = 0.626$), nor between the elite marathon athletes and non-athlete (genotype $\chi^2 = 0.232$ $p = 0.890$; allele $\chi^2 = 0.141$, $p = 0.707$), sub-elite marathon athletes and non-athlete (genotype $\chi^2 = 0.102$, $p = 0.950$; allele $\chi^2 = 0.097$, $p = 0.756$) or elite marathon and sub-elite marathon athlete groups (genotype $\chi^2 = 0.040$, $p = 0.980$; allele $\chi^2 = 0.005$ $p = 0.945$).

4.4 *AGT* RS699 RESULTS

4.4.1 Height and Body Mass

No difference in height was observed between marathon athletes and non-athletes in either the male ($p = 0.599$) or the female ($p = 0.747$) cohorts. However, non-athletes males were 10 kg heavier than male marathon athletes ($p = 1.0 \times 10^{-13}$), while female non-athletes were 12.5 kg heavier than female marathon athletes ($p = 1.003 \times 10^{-10}$), as shown in Table 10 all data are self-reported by participants.

4.4.2 **AGT rs699 Hardy-Weinberg Equilibrium**

There was no deviation from Hardy-Weinberg equilibrium in *AGT* rs699 genotype distributions amongst research participants and non-athletes (Marathon athlete $\chi^2 = 0.009$, Elite marathon athletes $\chi^2 = 0.011$, Sub-Elite marathon athletes $\chi^2 = 0.041$, and Non-Athletes $\chi^2 = 0.042$, $p > 0.05$ for all groups). When stratified by sex, Hardy-Weinberg equilibrium also existed in all groups (Males: Marathon athlete $\chi^2 = 2.307$, Elite marathon athletes $\chi^2 = 3.193$, Sub-Elite marathon athletes $\chi^2 = 0.300$, and Non-Athletes $\chi^2 = 0.059$, $p > 0.05$ for all groups), (Females: Marathon athlete $\chi^2 = 2.831$, Elite marathon athletes $\chi^2 = 3.476$, Sub-Elite marathon athletes $\chi^2 = 0.196$, and Non-Athletes $\chi^2 = 0.002$, $p > 0.05$ for all groups).

4.4.3 **AGT rs699 Genotype and Allele Frequencies**

The genotype frequency distributions of the *AGT rs699* polymorphism differed between marathon athletes and non-athletes, with a 5.6% higher TT genotype frequency in the marathon athletes ($\chi^2 = 6.248$, $p = 0.044$; Table 12, OR = 0.777, 95% CI 0.562 – 1.074. $p = 0.126$). Similarly, the T allele was 4.5% more frequent in marathon athletes compared with non-athletes ($\chi^2 = 5.961$, $p = 0.015$; Figure 16, OR = 0.832, 95% CI 0.669 – 1.034 $p = 0.097$). There was no difference in either genotype ($\chi^2 = 1.079$, $p = 0.583$) or allele ($\chi^2 = 1.060$, $p = 0.303$) frequencies between elite marathon athletes and non-athletes, although the sub-elite marathon cohort tended to contain more TT genotypes ($\chi^2 = 5.844$, $p = 0.054$) and 5.7% more T alleles ($\chi^2 = 5.480$, $p = 0.019$ allele: OR = 0.790, 95% CI 0.613 – 1.019. ***p*** = 0.069) when compared to non-athletes (shown in Figure 16). No further differences were observed when the male and female data were combined.

Table 9 AGT *rs699 genotype and allele frequencies in marathon athletes and non-athletes*

	***AGT* Genotype**						***AGT* Allele**			
Total Cohort	**TT**		**TC**		**CC**		**T**		**C**	
Marathon Athletes	131	(36.0%)	174	(47.8%)	59	(16.2%)	436	(59.9%)	292	(40.1%)
Elite Marathon Athletes	53	(33.8%)	77	(49.0%)	27	(17.2%)	183	(58.3%)	131	(41.7%)
Sub-Elite Marathon Athletes	78	(37.7%)	97	(46.9%)	32	(15.5%)	253	(61.1%)	161	(38.9%)
Non Athletes	93	(30.4%)	153	(50.0%)	60	(19.6%)	339	(55.4%)	273	(44.6%)
Males										
Marathon Athletes	83	(38.4%)	93	(43.1%)	40	(18.5%)	259	(60.0%)	173	(40.0%)
Elite Marathon Athletes	29	(38.7%)	29	(38.7%)	17	(22.7%)	87	(58.0%)	63	(42.0%)
Sub-Elite Marathon Athletes	54	(38.3%)	64	(45.4%)	23	(16.3%)	172	(61.0%)	110	(39.0%)
Non Athletes	65	(29.0%)	113	(50.4%)	46	(20.5%)	243	(54.2%)	205	(45.8%)
Females										
Marathon Athletes	48	(32.4%)	81	(54.7%)	19	(12.8%)	177	(59.8%)	119	(40.2%)
Elite Marathon Athletes	24	(29.3%)	48	(58.5%)	10	(12.2%)	96	(58.5%)	68	(41.5%)
Sub-Elite Marathon Athletes	24	(36.4%)	33	(50.0%)	9	(13.6%)	81	(61.4%)	51	(38.6%)
Non Athletes	28	(34.1%)	40	(48.8%)	14	(17.1%)	96	(58.5%)	68	(41.5%)

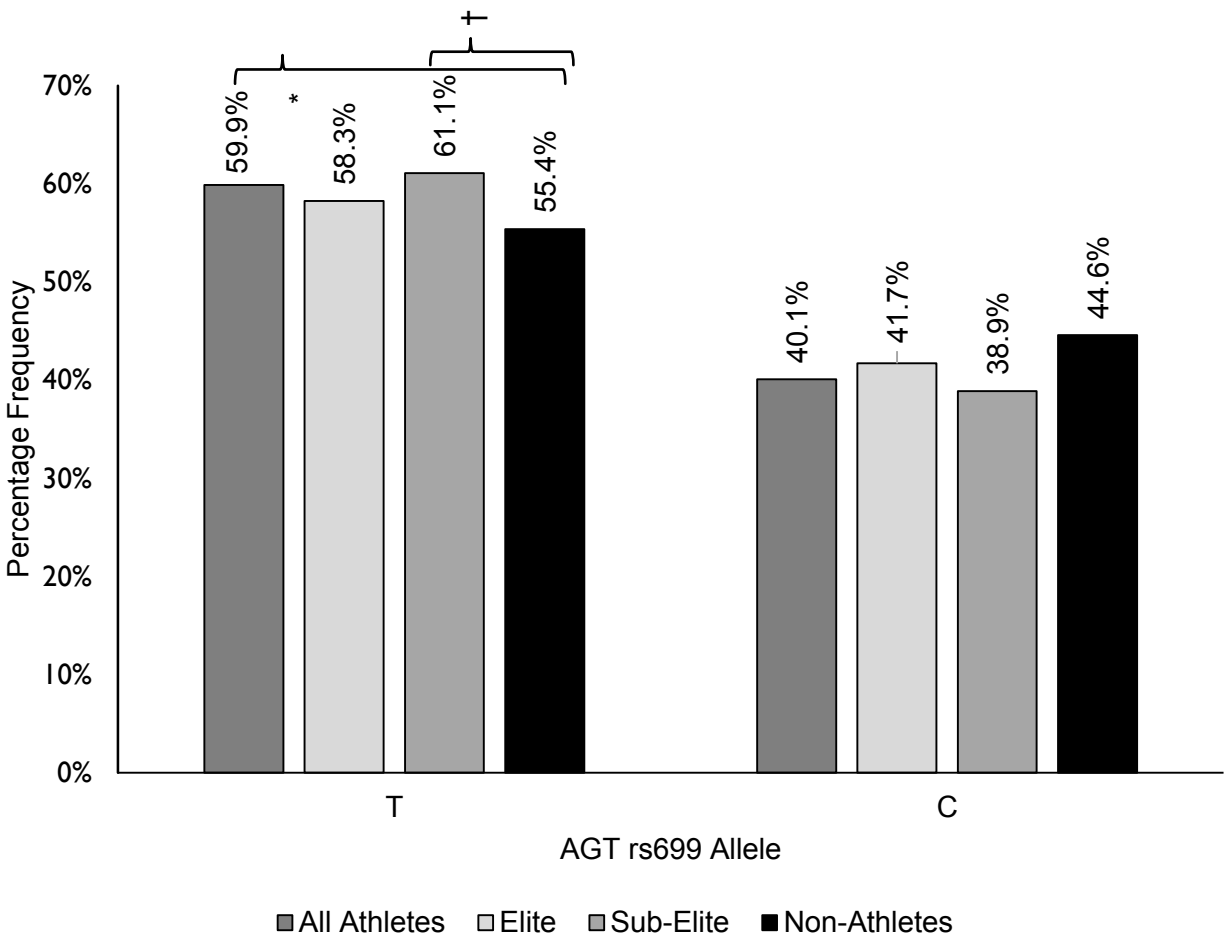

Figure 16 AGT rs699 allele frequencies in marathon athletes and non-athletes. The T allele is more frequent in marathon athletes (* p = 0.015) and sub-elite marathon athletes († p = 0.019) compared to non-athletes.

In men, marathon athletes showed a 9.7% higher TT genotype frequency than non-athletes ($\chi2 = 6.801$, p = 0.033; Table 12, OR = 0.655, 95% CI 0.440 - 0.975, p = 0.037). In agreement, there was a tendency for the T allele to be more frequent in marathon athletes than non-athletes ($\chi2 = 3.638$, p = 0.056). No differences in genotype or allele frequencies were observed between elite marathon athletes and non-athletes (genotype $\chi2 = 3.974$, p = 0.137; allele $\chi2 = 0.413$, p = 0.521). Neither were genotype frequencies different between sub-elite marathon athletes and non-athletes ($\chi2 = 4.280$, p = 0.118; Table 12), although there was a tendency for the T allele to be more frequent in the sub-elite marathon group ($\chi2 = 3.580$, p = 0.058; Figure 17). There were no differences in

genotype (χ^2 = 1.583, p = 0.453) or allele frequency distributions (χ^2 = 0.445, p = 0.505) between elite and sub-elite marathon athletes.

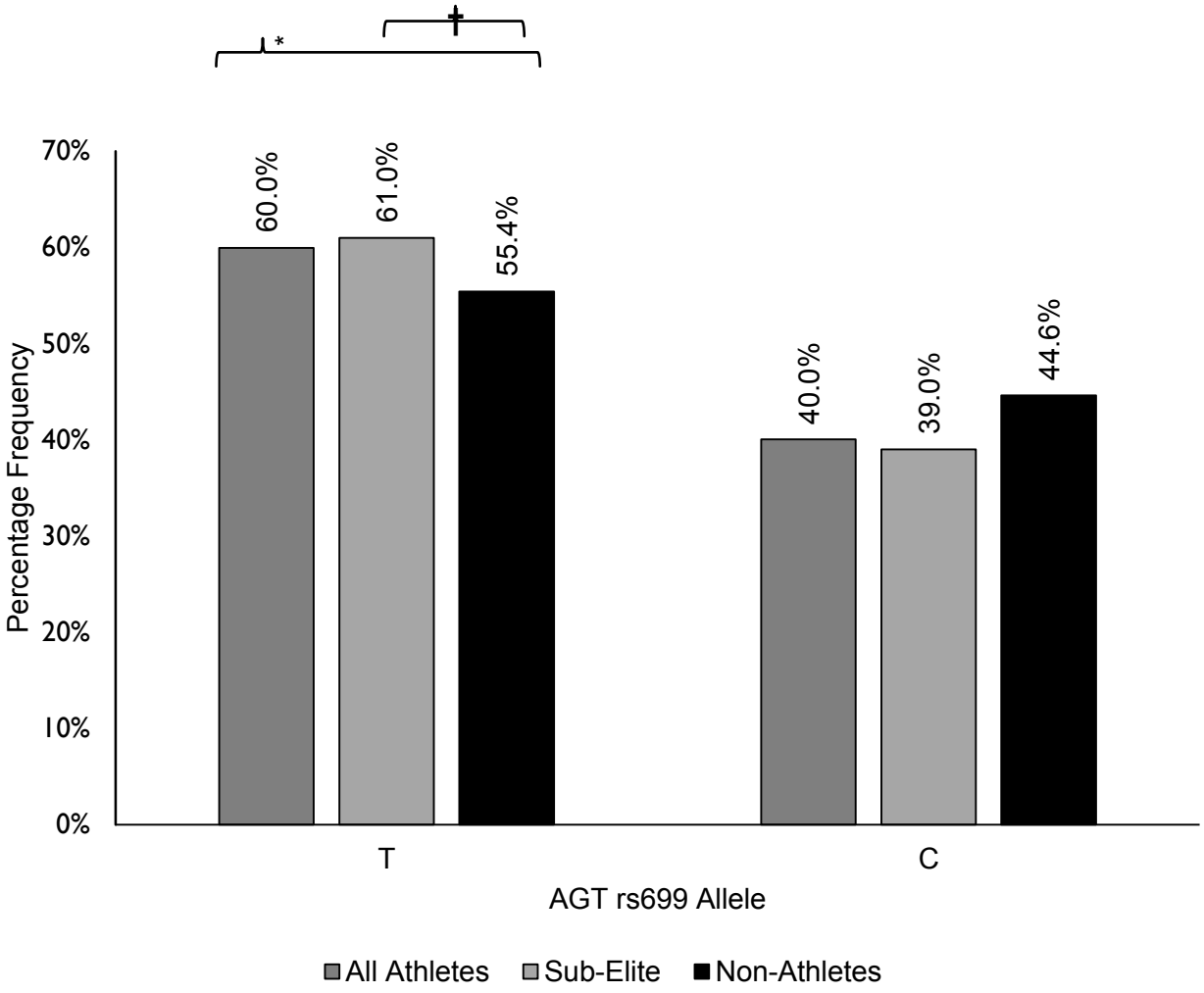

Figure 17 AGT rs699 allele frequencies in male marathon athletes and non-athletes. The T allele tended to be more common in marathon athletes (* p = 0.056) and sub-elite marathon athletes († p = 0.058) than non-athletes.

Female marathon athletes showed no difference in genotype or allele frequency distributions when compared to non-athletes (genotype χ^2 = 4.324, p = 0.115; allele χ^2 = 2.325, p = 0.127). Neither were any differences observed when comparing elite marathon athletes and non-athlete (genotype χ^2 = 3.527, p = 0.171; allele χ^2 = 0.656, p = 0.418), sub-elite marathon athletes and non-athlete (genotype χ^2 = 1.975, p = 0.373; allele χ^2 =

1.905, p = 0.168) or elite and sub-elite marathon athlete groups (genotype χ^2 = 1.114, p = 0.573; allele χ^2 = 0.325, p = 0.568).

4.4.4 **Comparison of Marathon PB and Genotype**

Performance times were not significantly different between *ACE* I/D genotype groups (males F = 0.939, **p** = 0.393; females F = 0.010, **p** = 0.990; Figure 18 and 19). Neither were performance times different between the *AGT* rs699 genotype groups (males F = 0.142 **p** = 0.868; females F = 0.315, **p** = 0.730; Figure 20 and 21).

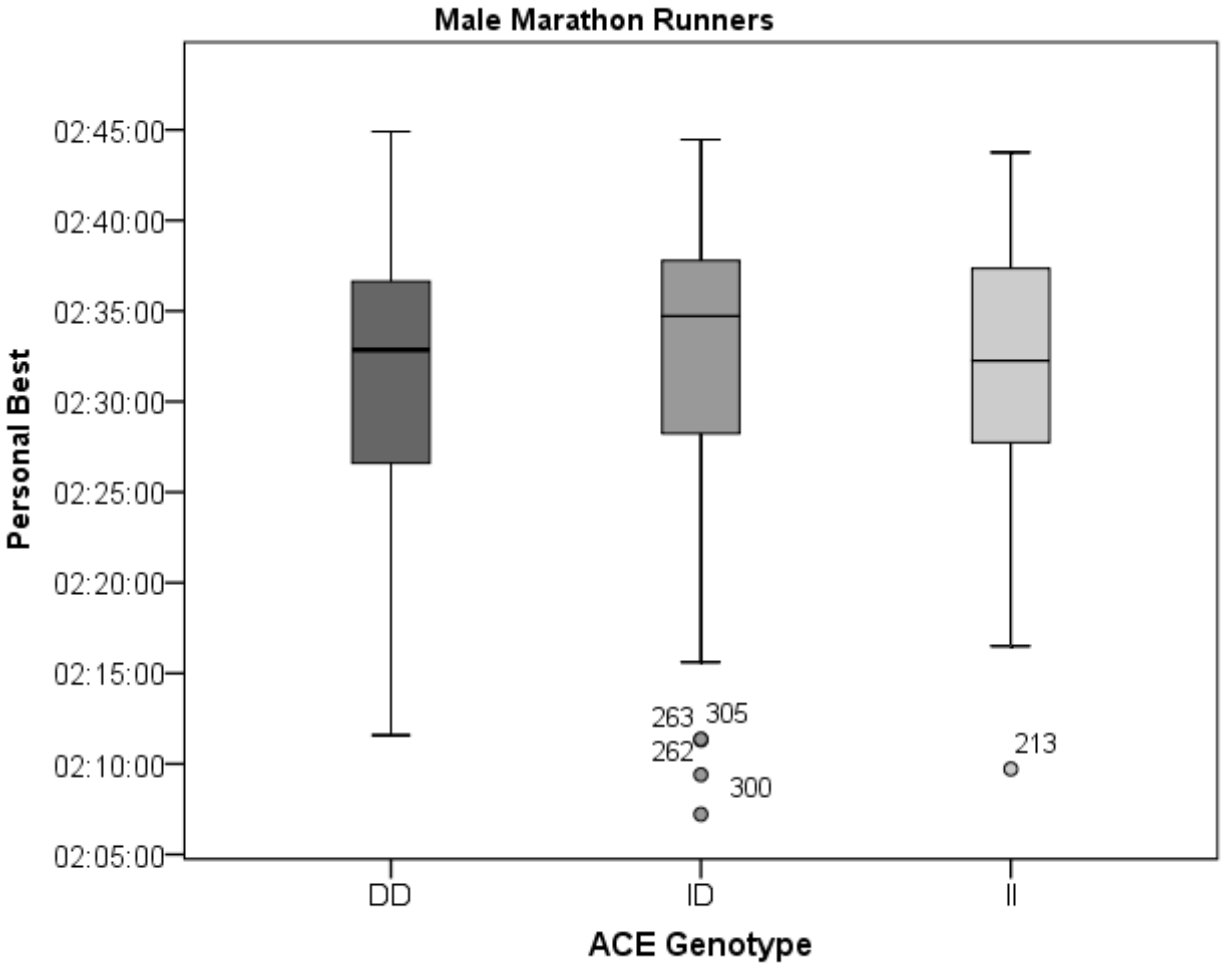

Figure 18 Marathon personal best times for males grouped by ACE I/D genotype. Data are medians and minimum and maximum.

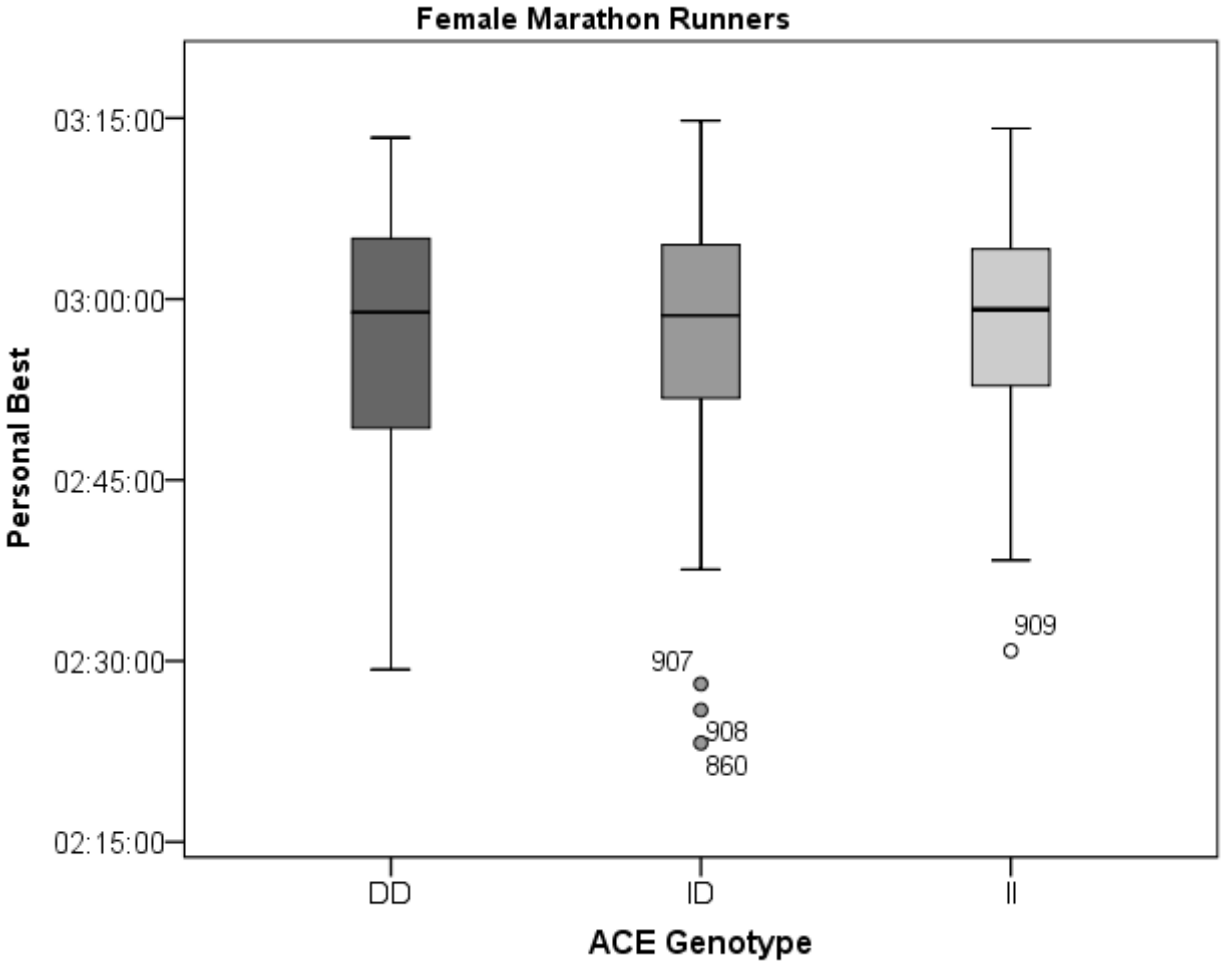

Figure 19 Marathon personal best times for females grouped by ACE I/D genotype. Data are medians and minimum and maximum.

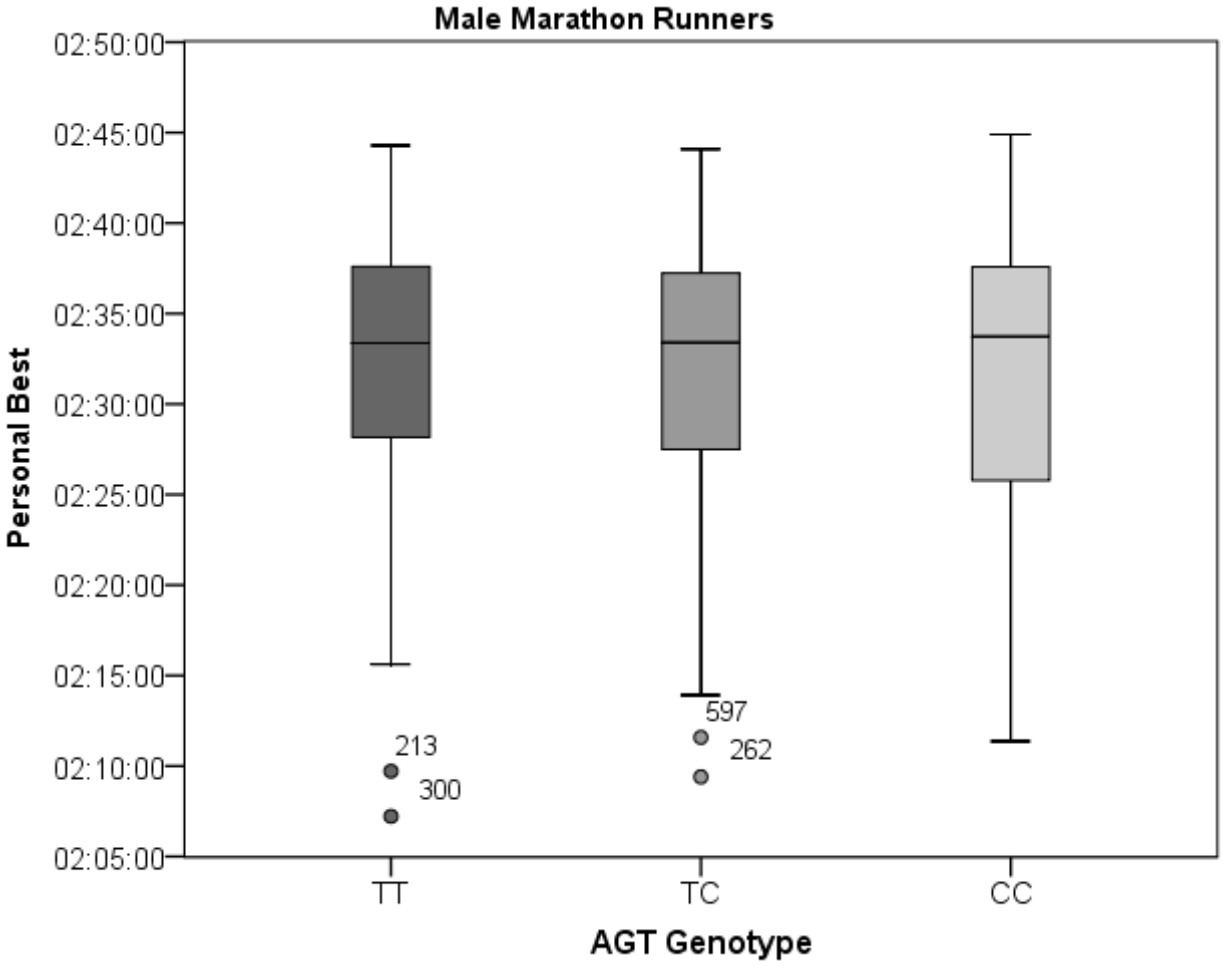

Figure 20 Marathon personal best times for males grouped by AGT rs699 genotype. Data are medians and minimum and maximum.

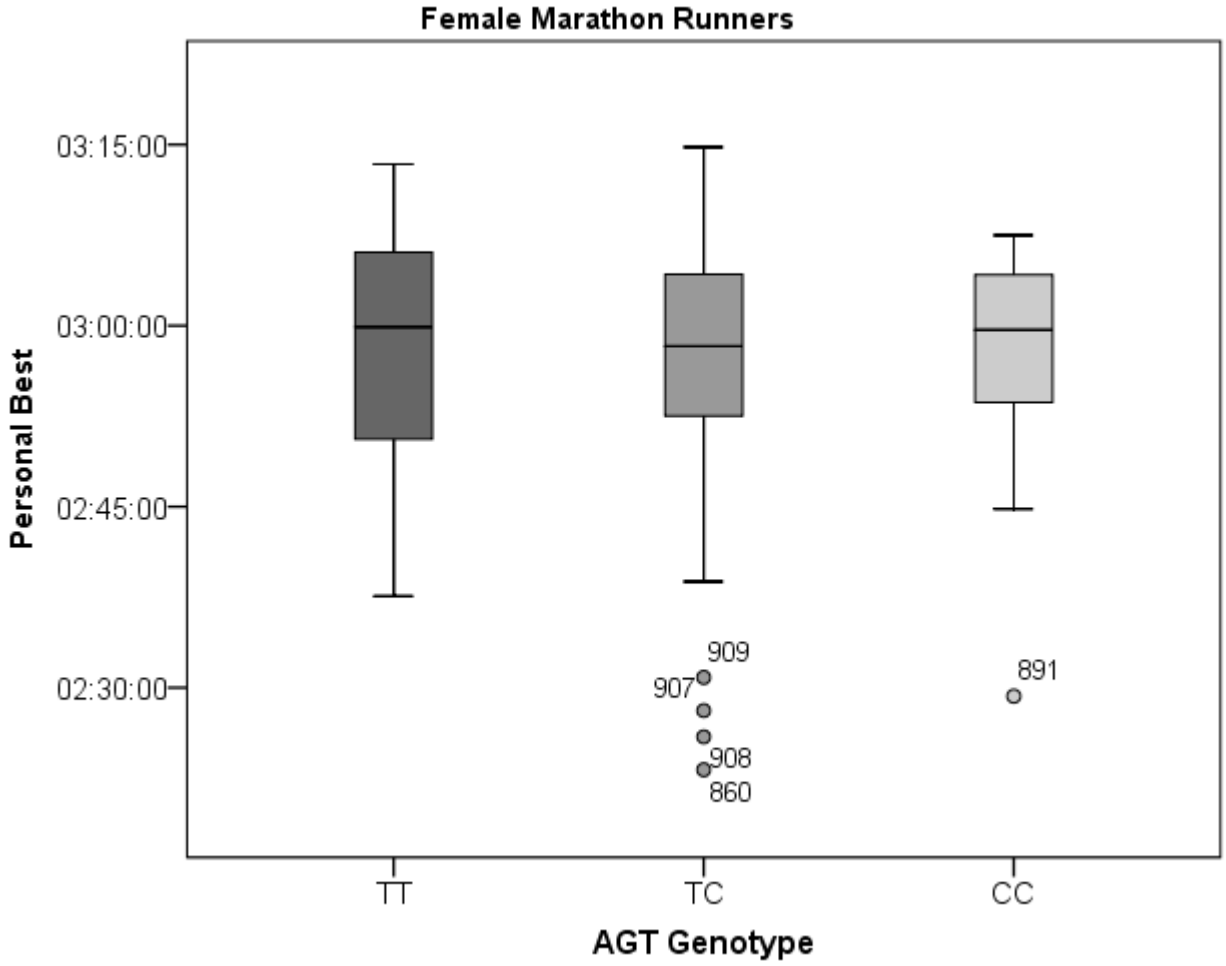

Figure 21 Marathon personal best times for females grouped by AGT rs699 genotype. Data are medians and minimum and maximum.

4.5 DISCUSSION

This study assessed the association of marathon sporting performance with the *ACE* I/D and *AGT* rs699 polymorphisms via a candidate gene approach. The main finding of this study was that the *AGT* rs699 TT genotype and T allele were associated with marathon athlete status compared to non-athletes. When males and females were combined, the sub-elite marathon athletes showed an *AGT* rs699 polymorphism TT genotype tendency ($p = 0.054$) and T allele association with marathon performance rather than the elite marathon athletes. A T allele tendency was also recorded in the whole male marathon athlete cohort when compared to non-athletes. The sub-elite male marathon athletes (those who complete a marathon in times between 2 hours 30 min and 2 hours 45 min) also showed a T allele tendency. There were no other *AGT* rs699 genotype or allele

associations with marathon performance. This result may be explained by adaptedness in the context of population genetics. It may be that the physiological phenotype fits the sub-elite marathon athletes to a greater extent than the elite athlete and the relative fitness is higher in the sub-elite marathon population than the elite marathon population. It may also be that gene flow has influenced this sample though as the laws of Hardy-Weinberg were met this is unlikely.

Elite endurance athlete status was not associated with either genotype or allele frequency distribution of the *ACE* I/D polymorphism when analysed by the same methods. There was no association between personal best marathon time in males or females marathon runners and either the *ACE* I/D polymorphism or the *AGT* rs699 polymorphism.

These results support the notion that *AGT* rs699 may signify a RAAS-dependent association with endurance performance in marathon running (Karjalainen et al., 1999, Diet et al., 2001). The *AGT* rs699 polymorphism has been reported to influence cardio-respiratory endurance, left ventricular hypertrophy and circulating ANGII levels in several cohorts (Diet et al., 2001, Karjalainen et al., 1999, Gomez-Gallego et al., 2009). The results of this study contrast those of Karjalainen et al. (1999) and Diet et al. (2001) who found CC genotype associations with increased left ventricular mass in endurance athletes and extrapolated that the CC genotype may be influential in elite endurance sporting performance due to increased oxygen delivery. However, the current study of 364 Caucasian elite and sub-elite marathon athletes suggested association with the TT genotype and T allele when compared to non-athletes. Although an increased left ventricular mass may be postulated to improve oxygen delivery to the skeletal muscle during endurance exercise it will undoubtedly also increase the oxygen demand of the cardiac muscle to execute the same workload. An increased cross-sectional area of cardiac muscle tissue may also negatively influence the volume of the left ventricular chamber, reducing end-diastolic volume and thus stroke volume. This increased cardiac

workload and possible reduction in stroke volume, in addition to an increased metabolic demand by a higher cardiac tissue mass facing an already high metabolic demand by skeletal muscle, would have a detrimental effect on an endurance athlete. Carriers of the *AGT* rs699 CC genotype exhibit higher diastolic blood pressure (Rankinen et al., 2000a) and then T allele carriers (TT and TC genotypes) after 20 weeks of endurance training (Rankinen et al., 2000a). TT and TC genotype carriers also showed a greater decrease in diastolic pressure when compared to CC homozygotes (Rankinen et al., 2000a) and this may indicate that the pressor response induced via AGT and the AT1-R is detrimental to endurance performance in marathon running - thus TT genotype and T allele carriers may be preferential for endurance competition at an elite level.

In addition, the RAAS-mediated association with endurance performance may be explained by preferential degradation of the vasodilator bradykinin by ACE over the formation of the vasoconstrictor ANGII. Higher kinin activity has been associated with the ACE I allele (Murphey et al., 2000). It could therefore be postulated that the reduced levels of AGT, a precursor substrate in the RAAS for ANGII, exhibited by the T allele carriers could cause a pathway flux towards bradykinin production. Bradykinin (a vasodilator) increases skeletal muscle perfusion and glucose uptake (Wicklmayr et al., 1983). Greater skeletal muscle substrate availability for respiration is likely to mediate positive alterations in metabolic activity and increased efficiency of oxygen utilisation during the exertion of a prolonged endurance event like a marathon.

4.6 CONCLUSION

Although this study found there was no direct *ACE* I/D polymorphism association with endurance performance, it may be that through the mechanisms outlined above RAAS mediates enhancements in marathon performance. Previous studies that reported an association between ACE activity and endurance performance in humans have typically

used cohorts containing relatively small numbers and mixed athlete groups that often lack a clearly defined performance phenotype. Often, cohort data appear to have been reused in subsequent studies and therefore there are might be some bias in the published research. While the Puthucheary et al. (2011) review and Ma et al. (2013) meta-analysis attempted to address these limitations, the current study is the first to do so with a relatively large and homogenous group of marathon athletes.

5 MUSCLE STRUCTURAL GENE - A-ACTININ 3

5.1 INTRODUCTION

The α-actinin 3 (*ACTN3)* gene has previously been associated with rodent and human physical performance and elite athlete status in both endurance (Ahmetov et al., 2010, Niemi and Majamaa, 2005, Grealy et al., 2012) and sprinting events (Yang et al., 2003, Niemi and Majamaa, 2005). The human sarcomeric α-actinins are a family of actin-binding proteins related to dystrophin. They are major structural components of the Z line in skeletal muscle. The α-actinins crosslink the actin thin filaments to the Z-line and therefore play a static role in the structural organisation of the sarcomeric myofibrils (see Figure 3). In addition, it is thought the α-actinins ensure sarcomeric integrity during rapid myofibre contraction (Yang et al., 2003, Blanchard et al., 1989, Mills et al., 2001).

In humans, two genes encode the skeletal muscle: α-actinin 2 (*ACTN2)* and *ACTN3*. α actinin-2 is expressed in all skeletal muscle fibres whereas α actinin-3 has restricted expression to fast twitch (type II) fibres (Yang et al., 2003). A common nonsense polymorphism (rs1815739) results in a transition from C>T at position 1747 in the *ACTN3* coding sequence causing a premature stop codon (TGA; X allele) at position 577 (Mills et al., 2001) leading to a non-functioning protein. Both the R (coding for a functioning protein) and X alleles are common within the general population although 18% of the European population and around 1 billion world-wide are XX homozygous and thus deficient of any functioning α-actinin-3 protein (Yang et al., 2003).

Yang et al. (2003) were the first to hypothesize that a deficiency in α-actinin-3 protein may elicit an advantage to athletes competing in endurance-based events such as the marathon, based on a study of a cohort of Australian endurance athletes. Further studies detailing α actinin-3 expression in a knockout mouse model demonstrated a decrease in

muscle fibre cross sectional area as a result of an increase in the type 1 fibre type and also increased oxidative enzyme expression in the muscle associated with oxidative capacity -indicating a move to the efficient aerobic pathway by the mouse (MacArthur et al., 2007, MacArthur et al., 2008). This lead to enhanced endurance running performance and better recovery from fatiguing muscle contractions compared to the wild type mice. The *ACTN3* knock out mouse metabolic phenotype is similar to that of skeletal muscle following aerobic endurance exercise training in the mouse. It is suspected the *ACTN3 R577X* polymorphism contributes to genetic variation in muscle fibre composition and further suggested in the literature that the α-actinin 3 deficient muscle is pre-conditioned for endurance performance (MacArthur et al., 2008, Seto et al., 2013). The hypothesis that the *ACTN3* 577X allele is associated with enhanced endurance performance is controversial as further elite endurance athlete studies have not supported of the Yang et al. (2003) hypothesis finding no association of the ACTN3 577X allele and athletic performance (Niemi and Majamaa, 2005, Lucia et al., 2006, Döring et al., 2010, Saunders et al., 2007, Yang et al., 2007, Muniesa et al., 2010) or contradictory evidence of an R allele association with performance (Ahmetov et al., 2010). The previous associations of the *ACTN3* 577X polymorphism and elite endurance performance have been primarily completed in mixed endurance athlete cohorts bar Grealy et al. (2012) who reported an association in the multiple part event ironman triathlon. Therefore, the aim of the present study was to examine the frequency of the ACTN3 R577X genotype and allele in Caucasian elite and sub elite marathon runners and to investigate whether there was an association between *ACTN3* and elite marathon performance.

5.2 METHODOLOGY

5.2.1 Research Participant Characteristics

One thousand and seventy five Caucasian adults provided written informed consent to take part in this study. This total cohort comprised 399 marathon runners (male, n = 243; female, n = 156) and a non-athlete (non-marathon running) cohort, comprising 383 men

and 293 women. The marathon athletes were stratified into elite and sub-elite subgroups according to their official marathon personal best performance time (http://www.poweroften.co.uk; Table 13). Research participant height and body mass are shown in Table 14.

Table 10 Numbers of research participants in the ACTN3 rs1815739 analyses

	Male	Female	Total
Marathon Athletes	243	156	399
Elite Marathon Athletes	86	87	173
Sub-Elite Marathon Athletes	157	69	226
Non-Athletes	383	293	676

Table 11 ACTN3 *rs1815739 participant characteristics (mean (standard deviation))*

		Athletes	Elite	Sub-Elite	Non-Athletes
Height (m)	Male	1.78 (0.06)	1.79 (0.06)	1.78 (0.06)	1.79 (0.07)
	Female	1.65 (0.07)	1.64 (0.07)	1.66 (0.08)	1.65 (0.07)
Mass (kg)	Male	67.0 (6.5)	67.1 (7.5)	67.0 (5.9)	78.4 (11.5)
	Female	53.6 (5.3)	52.7 (5.4)	54.8 (5.0)	66.3 (12.4)
Age (years)	Male	35.9 (8.0)	37.3 (10.2)	35.0 (6.4)	27.9 (15.3)
	Female	38.0 (7.4)	36.6 (7.9)	37.4 (6.9)	31.4 (18.4)

5.2.2 **DNA Collection**

Participant DNA collection is outlined in section 3.5 of the methods. In brief, 731 EDTA treated blood samples, 231 buccal samples and 113 saliva samples were collected from participants for the *ACTN3* rs1815739 analyses.

5.2.3 **DNA Isolation**

Automated DNA extraction was performed using a QIAcube (Qiagen,), following the QIAamp blood mini protocol as detailed in the methods section. Briefly, a 200 µL sample was lysed with QIAGEN Protease Enzyme. To each sample, 200 µL of 96% ethanol was added. Samples were agitated vigorously to mix. The mixture was transferred to a

QIAamp mini spin column tube containing a DNA collection filter and bound by centrifugation at 8000 rpm for 1 min. Three further wash buffer centrifugation cycles followed to wash out any remaining protein and impurities. Genomic DNA remained bound to the silica gel DNA filter membrane in the spin column during these wash cycles. After each wash the filtrate was discarded. In the final buffer centrifugation, the sample genomic DNA was eluted into 100 µL of molecular grade H_2O to provide purified genomic DNA. The purified genomic DNA was stored at 4°C until the genotyping was performed.

5.2.4 **Genotyping**

Each participant's *ACTN3* rs1815739 genotype was established using real-time Polymerase Chain Reaction (rtPCR). Genotyping of all participants was completed using the *ACTN3* rs1815739 TaqMan assay that included the appropriate TaqMan primers and probes (Applied Biosystems,). Each 10 µL PCR reaction contained: 5 µL Genotyping Master Mix (Applied Biosystems), 4.3 µL nuclease-free H_2O (Qiagen), 0.5 µL *ACTN3* rs1815739 TaqMan genotyping assay mix (Applied Biosystems), and 0.2 µL of participant DNA. For control wells, 0.2 µL nuclease-free H_2O replaced the DNA template. Genotyping was completed on the StepOnePlus (Applied Biosystems) real-time PCR system. Briefly, there were 40 cycles of denaturation at 95°C for 15 s then annealing and extension at 60°C for 1 min according to the instructions for optimal performance. StepOnePlus software version 2.3 (Applied Biosystems) was used to determine each participant's *ACTN3* rs1815739 genotype. All samples were analysed in duplicate. There was 100% agreement within all sample duplicates.

5.2.5 **Data Analysis and Statistics**

Independent T-tests to determine differences in height and weight distributions amongst athletes and non-athletes were performed in SPSS Hardy-Weinberg equilibrium assessment was completed using Excel (Microsoft, 2013). SPSS was used to calculate and compare genotype and allele frequencies between athlete groups and non-athletes

by Pearson's Chi square test. Analysis between elite marathon and sub-elite marathon athlete groups were completed using chi square test of difference in SPSS. Genotype associations with personal best were assessed by one way ANOVA in SPSS).

5.3 ACTN3 RS1815739 RESULTS

5.3.1 Height and Body Mass

Table 14 shows the self-reported height of the non-athlete and the marathon athlete groups, they were similar for both males ($p = 0.849$) and females ($p = 0.333$). The male non-athlete group body mass was 11.2 kg heavier than the marathon athlete group ($p = 1.000 \times 10^{-13}$). The female non-athlete group body mass was heavier by 10.9 kg ($p = 1.000 \times 10^{-13}$) all data are self-reported.

5.3.2 ACTN3 rs1815739 Hardy-Weinberg Equilibrium

The *ACTN3* genotype distributions amongst all participants were in Hardy-Weinberg equilibrium (Marathon athletes $\chi^2 = 0.008$, Elite marathon athletes $\chi^2 = 0.049$, Sub-Elite marathon athletes $\chi^2 = 0.093$, and Non-Athletes $\chi^2 = 1.121$, $p > 0.05$ for all groups, 1 df). Hardy-Weinberg equilibrium remained when the data were stratified into males and females.

5.3.3 **ACTN3 rs1815739 Genotype and Allele Frequencies**

Table 12 Genotype and allele frequencies of ACTN3 *rs1815739 polymorphism in marathon athletes and non-athletes.*

	***ACTN3* rs1815739 Genotype**						***ACTN3* rs1815739 Allele**			
Total Cohort	**RR**		**RX**		**XX**		**R**		**X**	
Marathon **Athletes**	119	(29.8%)	197	(49.4%)	83	(20.8%)	435	(54.5%)	363	(45.5%)
Elite Marathon Athletes	52	(30.1%)	87	(50.3%)	34	(19.7%)	191	(55.2%)	155	(44.8%)
Sub-Elite Marathon Athletes	67	(29.6%)	110	(48.7%)	49	(21.7%)	244	(54.0%)	208	(46.0%)
Non-Athletes	230	(34.0%)	317	(46.9%)	129	(19.1%)	777	(57.5%)	575	(42.5%)
Males										
Marathon **Athletes**	79	(32.5%)	112	(46.1%)	52	(21.4%)	270	(55.6%)	216	(44.4%)
Elite Marathon Athletes	30	(34.9%)	37	(43.0%)	19	(22.12%)	97	(56.4%)	75	(43.6%)
Sub-Elite Marathon Athletes	49	(31.2%)	75	(47.8%)	33	(21.0%)	173	(55.1%)	141	(44.9%)
Non-Athletes	140	(34.0%)	177	(46.9%)	66	(19.1%)	457	(57.5%)	309	(42.5%)
Females										
Marathon **Athletes**	40	(25.6%)	85	(54.5%)	31	(19.9%)	165	((52.9%)	147	(47.1%)
Elite Marathon Athletes	22	(25.3%)	50	(57.5%)	15	(17.2%)	94	(54.0%)	80	(46.0%)
Sub-Elite Marathon Athletes	18	(26.1%)	35	(50.7%)	16	(23.2%)	71	(51.4%)	67	(48.6%)
Non-Athletes	90	(34.0%)	140	(46.9%)	63	(19.1%)	320	(57.5%)	266	(42.5%)

Table 15 shows no difference in *ACTN3* R577X genotype (χ^2 = 3.209, p = 0.201 or allele frequencies (χ^2 = 0.715, p = 0.398) when marathon athletes were compared to non-athletes. When compared to non-athletes the elite marathon athlete group showed no difference in genotype (χ^2 = 1.255, p = 0.534) or allele distribution (χ^2 = 0.728, p = 0.393). There were no differences in genotype or allele frequency of *ACTN3* R577X when sub-elite marathon athletes were compared to non-athletes (genotype χ^2 = 2.225, p = 0.329; allele χ^2 = 2.250, p = 0.134). There were no differences between the elite marathon athletes and the sub-elite marathon athletes in either genotype frequency (χ^2 = 0.251, p = 0.882) or allele frequency (χ^2 = 0.118, p = 0.732).

The XX genotype (χ^2 = 0.880, p = 0.644) and allele (χ^2 = 0.729 p = 0.393) frequency distributions of the *ACTN3* rs1815739 polymorphism were not different amongst the male marathon athletes and non-athletes (Table 15). This lack of association was further seen when male elite marathon athletes when compared to non-athletes (genotype χ^2 = 0.702, p = 0.704, allele χ^2 = 0.081, p = 0.776; Table 15). The sub-elite male marathon athletes showed no difference in genotype frequency distribution (χ^2 = 0.699, p = 0.705) or allele frequency distribution (χ^2 = 0.725, p = 0.395) when compared to non-athletes (Table 3). Between group analysis of male elite marathon athletes and all sub-elite marathon athletes revealed no difference in genotype (χ^2 = 0.532, p = 0.766) or allele (χ^2 = 0.293, p = 0.588) frequency distributions (Table 15)

The genotype frequency distributions and allele frequency distribution of the *ACTN3* rs1815739 polymorphism were not different amongst the female marathon athletes and non-athletes (genotype, χ^2 =5.191, p = 0.075; allele, χ^2= 0.729, p = 0.393; Table 15). No difference was observed in the genotype frequency distributions (χ^2 = 4.182, p = 0.124) or allele frequency distribution (χ^2 = 0.846, p = 0.358) when between group analysis of the female elite marathon athlete and non-athletes was completed (Table 15). There was no difference in genotype distribution (χ^2 = 2.103, p = 0.349) or allele frequency distributions

(χ^2 = 2.103, p = 0.349) when sub-elite female marathon athletes were compared to non-athletes (χ^2 = 2.047, p = 0.153; Table 3). The female elite marathon athlete cohort showed no differences when compared all sub-elite marathon athletes (genotype χ^2 =1.016, p = 0.602; allele χ^2 =8.37 x10^{-4}, p = 0.993) as shown in Table 15.

5.3.4 **Comparison of Marathon PB and Genotype**

Median marathon performance time for each genotype group are shown in Figure 22 for males and Figure 23 for females. Median marathon performance time were not significantly different between genotype groups (males F = 0.206, **p** = 0.814; females F = 0.257, **p** = 0.774).

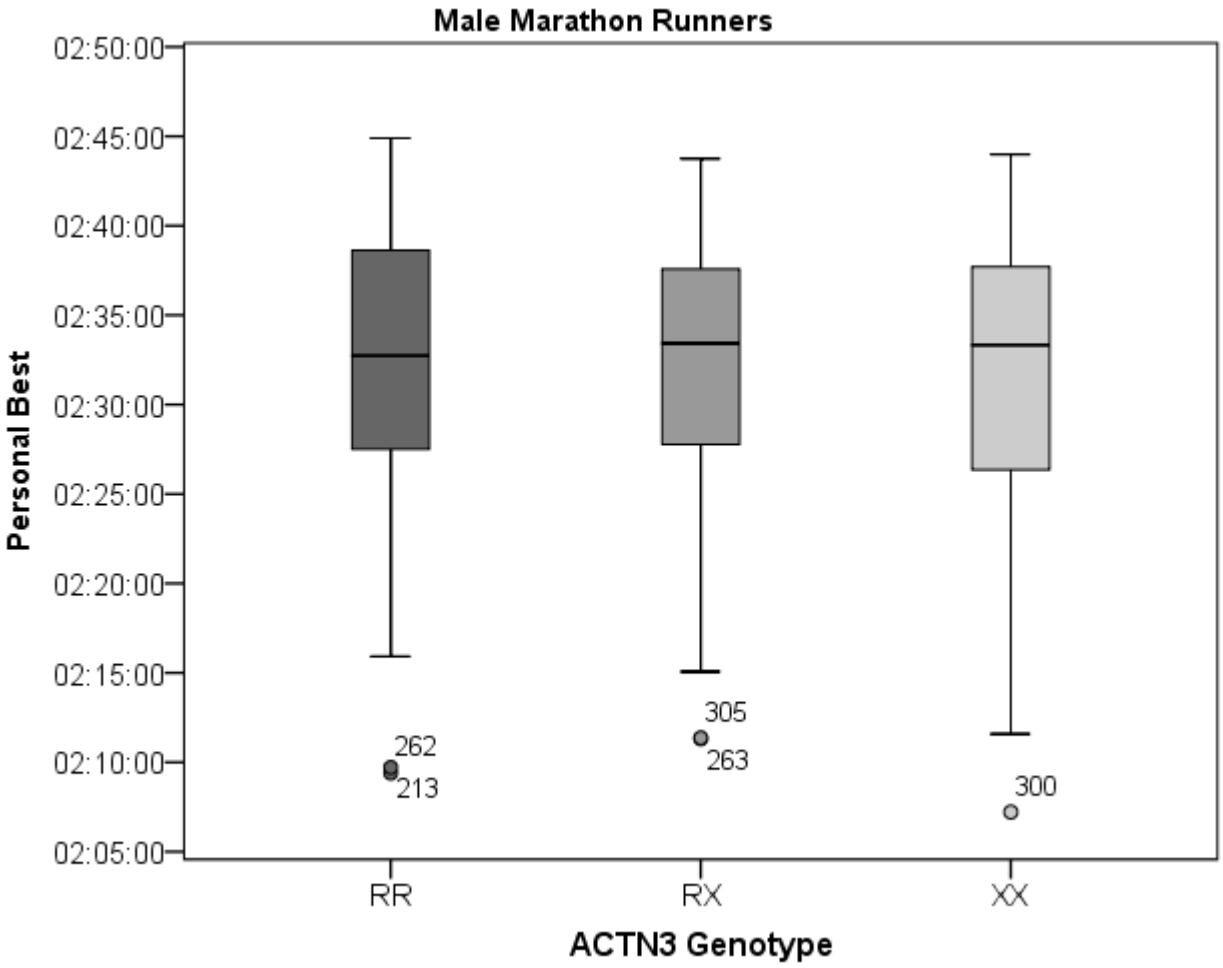

Figure 22 Marathon personal best times for males grouped by ACTN3 rs1815739 genotype. Data are medians and maximum and minimum

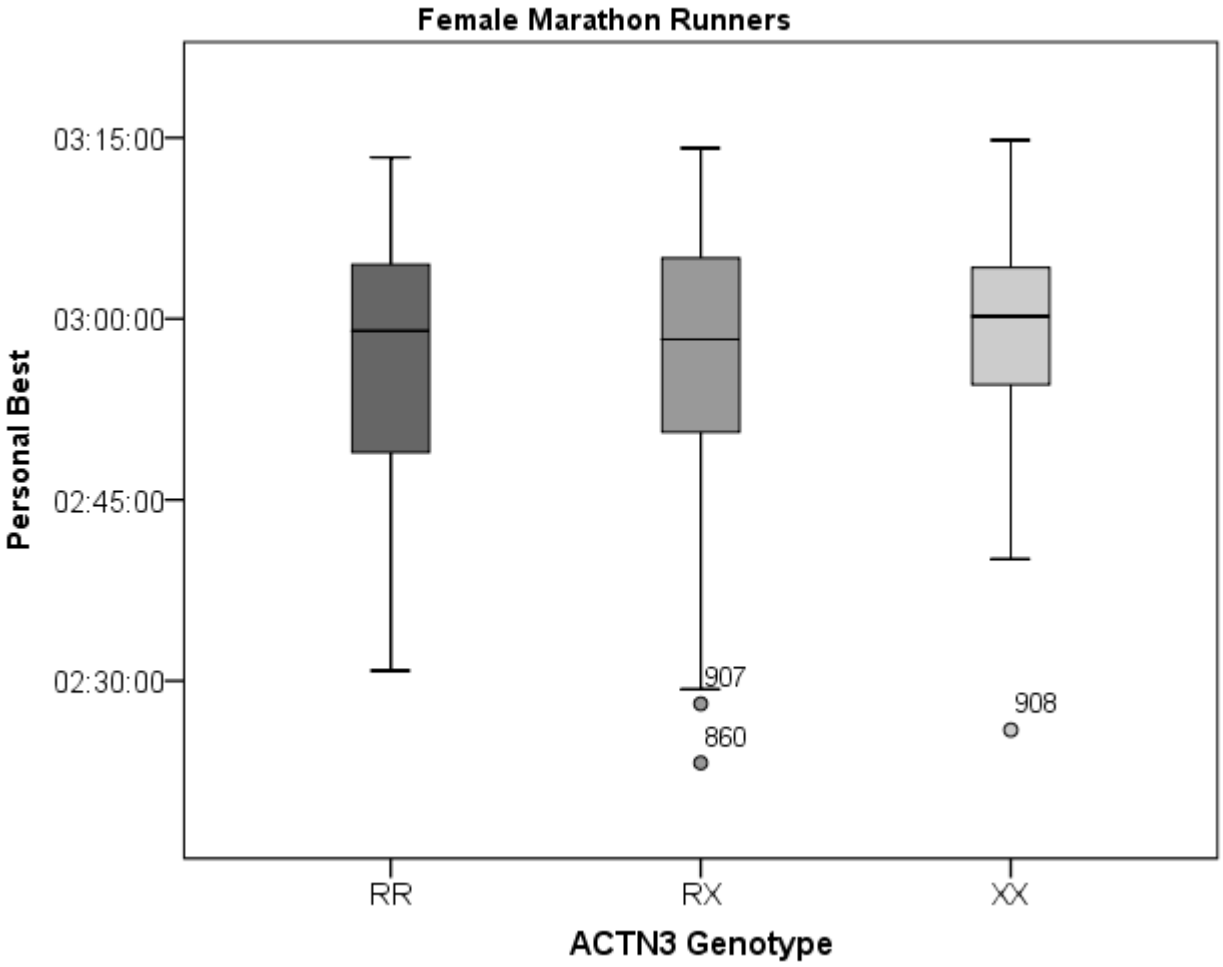

Figure 23 Marathon personal best times for females grouped by ACTN3 rs1815739 genotype. Data are medians and maximum and minimum

5.4 DISCUSSION

The purpose of this study was to comparatively analyse the frequency of the *ACTN3 R577X* polymorphism in elite and sub-elite endurance marathon athletes, and non-athletes. In addition, the study tested the association of the *ACTN3 R577X* polymorphism with personal best performance in the marathon. There was no difference in *ACTN3 R577X* genotype or allele frequency distribution between the marathon athlete population and the non-athlete population compared as whole groups or stratified by sex and elite athlete status. No genotype association with personal best time was recorded for either males or females.

The *ACTN3* genotype frequencies in the non-athlete population were similar to those reported for the Australian Caucasian population (Yang et al., 2003). However, they were different to the Finnish non-athlete population (Niemi and Majamaa, 2005) who recorded fewer XX genotypes (9.2%) and also different to the Spanish non-athlete population of Lucia et al .,(2006) who recorded 26% XX genotype in their non-athlete population. Both of these studies had approximately 120 non-athlete participants. Those in the study by Niemi and Majamaa (2005) were described only as anonymous Finnish blood donors with no description of ancestry. In the study by Lucia et al. (2006) the non-athlete population was described as 123, healthy, sedentary, unrelated, Caucasian, male controls of European ancestry though sedentary behaviour was not reported so it is likely these participants exhibited low physical activity rather than true sedentary behaviour. Both of these non-athlete cohorts are much smaller than the 676 participants used in the Caucasian non-athlete cohort in this study therefore the genotype frequencies of the non–marathon athlete cohort are more reliable.

In this population the *ACTN3* R577X, polymorphism does not appear to confer enhanced performance in marathon runners. This investigation supports the work of other groups using sporting populations (Lucia et al., 2006, Saunders et al., 2007, Yang et al., 2007, Ahmetov et al., 2010, Muniesa et al., 2010, Döring et al., 2010) who reported no association between the *ACTN3* genotype and elite endurance performance.

It was hypothesised that there may be a genotype dependant difference in marathon PB, with those of the XX genotype having a significantly faster marathon PB than those of the other two-genotype groups. This hypothesis was based on the rodent and human work of Vincent et al. (2007) and (Seto et al., 2013) who showed that the percentage surface area and number of type II fibres was higher in RR rather than XX genotypes. As marathon is predominantly an endurance event, athletes demonstrate more type 1 slow oxidative fibres (Gollnick et al., 1972, Gollnick et al., 1973) than fast-twitch therefore the XX

genotype was assumed to be favourable. In addition, Ahmetov et al., (2011) showed that *ACTN3* XX genotype was associated with an increased slow-twitch fibre proportion and with increased preferred racing distance in speed skaters. They concluded that the *ACTN3* R577X polymorphism was only probably partially associated with muscle fibre type composition and that this had an important but small influence on the ability to perform at a high level in speed skating.

We found that personal best were not genotype-dependent for either men or women ($P \geq 0.783$). This study supports the work of Grealy et al.,(2012) who concluded, based on analysis of triathlon performance times, that *ACTN3* alone does not appear to confer any advantage to endurance athletes in allowing them to sustain an extreme bout of endurance exercise.

These investigations have all used endurance athletes from multiple sporting disciplines (800 m - marathon, cycling, rowing, swimming, cross-country skiing and race walking) making it difficult to extrapolate whether the ACTN3 R577X polymorphism is associated with athlete performance in a particular sport such as marathon. In contrast, this investigation has recruited athletes from the same sporting discipline. This is advantageous over the heterogeneous sample of endurance athletes of multiple sporting disciplines as the physiological characteristics of marathon runners is likely to be more homogenous, so more confidence can be placed in the reported results. Another possible explanation for the discrepancy in the reported results between the aforementioned studies and this study is each study varies in the author's interpretation of elite athlete status. Eynon et al. (2009a) and Niemi and Majamaa (2005) determined elite athlete status by national or international representation at a track and field championships, whereas the elite athletes reported by Yang et al. (2003) had all represented Australia at an international level only. In the current study athletes were classified as elite if they had achieved a personal best marathon performance time of 2 hours 30 minutes for male

athletes and 3 hours for female athletes. This may in part explain the difference in reported results. Eynon et al. (2009a) notes that the groups used in their study were not large, due to the small number of elite athletes available in Israel. The long distance endurance group contained 77 athletes who had either competed in 10000m or marathon at national or international track and field championships. Niemi and Majamaa (2005) also had a small endurance athlete population of 52 that ranged from 800 to marathon and included race walkers. Although Yang et al. (2003) had a larger overall endurance cohort of 194 only 15 of those athletes were track athletes (≥5000m) this may have contributed to the false reporting of a positive association due to a small sample size.

5.5 CONCLUSION

In summary, the findings of this study suggest that the ACTN3 R577X polymorphism does not play a role in predisposing elite / sub-elite marathon runner status, or in determining marathon PB performance A possible explanation for the observed lack of ACTN3 R577X association with elite athlete status in marathon runners is that due to the conserved nature of the α-actinin proteins ACTN2 partially compensates for the speculated advantageous effect that ACTN3 provides in sarcomeric structural integrity during fast skeletal muscle contraction in the human. Although Seto et al. (2013) provide plausible evidence in a very small number of human research participants for the ACTN3 model proposed by the Australian research group further functional data in humans to verify the hypothesis is required.

6 GENES THAT MAY INFLUENCE ENERGY PRODUCTION - PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR GAMMA, COACTIVATOR 1 ALPHA - PPARGC1A

6.1 INTRODUCTION

Endurance events, such as marathon, induce co-ordinated changes in gene expression, which are regulated by transcription factors and co-activators. PPARγ coactivator 1α (PGC-1α) is a transcription co-activator that regulates gene expression in multiple biological responses, including mitochondrial biogenesis and oxidative phosphorylation (Lin et al., 2005). As such, in humans, tissues that catabolize fatty acids such as liver, skeletal muscle, and myocardium have high levels of PGC-1α mRNA expression, along with kidney and to a lesser extent brain, pancreas and white adipose tissue (Esterbauer et al., 1999). The interaction of PGC-1α with specific transcription factors, such as nuclear receptor PPARγ, nuclear respiratory factors 1 and 2, and muscle specific transcription factors, such as MEF2 (Lin et al., 2002), mediates PGC-1α's regulation of several genes encoding key enzymes involved in fatty acid oxidation (Finck and Kelly, 2006, Maciejewska et al., 2011) and oxidative phosphorylation (Lucia et al., 2005, Arany et al., 2005).

PGC-1α knock out mice show decreased expression of genes involved in mitochondrial function (Arany et al., 2005) and a reduction in expression of 30-50% genes implicated in oxidative phosphorylation, fatty acid oxidation and ATP synthesis cardiac muscle. Decreased mitochondrial density and respiratory capacity have been reported in the skeletal muscle of type I knock out mice (Leone et al., 2005). Therefore, PGC-1α is a key requirement for normal expression of mitochondrial genes in cardiac and skeletal muscle (Arany et al., 2005). In humans, PGC-1 α is encoded by the *PPARGC1A* gene, the expression of which is increased in response to endurance training (Tunstall et al., 2002,

Pilegaard et al., 2003, Norrbom et al., 2004). This may therefore induce enhanced skeletal muscle oxidative capacity via PPARa and PPARγ regulation of gene expression (Lin et al., 2002, Lin et al., 2005, Russell et al., 2003). In addition, enhanced mitochondrial density and oxidative capacity may promote a transition from fast to slow muscle fibre type, utilizing a primarily oxidative metabolism (Lin et al., 2002, Lin et al., 2005). Training induced adaptation of the muscle caused in part by increased PGC-1α expression may also influence the use of energy substrates in pathways such as fatty acid oxidation, the Krebs cycle and glucose transportation and oxidation (Arany et al., 2005, Baar, 2004, Baar et al., 2002), which is likely to influence endurance sporting performance.

A G>A SNP at position 23814039 in chromosome 4 in the *PPARGC1A* gene (rs8192678) results in the substitution of Gly for Ser at amino acid position 482. Elderly carriers of the A allele have been shown to have lower basal *PPARGC1A* mRNA levels compared to GG homozygotes (Ling et al., 2004). Several studies have investigated an association between the *PPARGC1A* rs8192678 SNP and elite athlete status/performance. Lucia et al. (2005) reported an association of the A allele in elite endurance athletes compared with non-athlete controls. This finding was replicated by (Eynon et al., 2009b, Eynon et al., 2009c), who also found lower numbers (0.25) of A-allele carriers in a group of Israeli endurance athletes ($p = 0.0001$. The A-allele was further reported to be associated with lower aerobic capacity in Russian rowers (Ahmetov et al., 2007). When 1423 Russian athletes of mixed sporting discipline were compared to 1,132 non-athlete controls, the *PPARGC1A* G>A SNP was associated with endurance athlete status, the proportion of slow-twitch muscle fibres and maximal oxygen consumption (Ahmetov et al., 2009). However, in 60 Turkish elite endurance athletes the AA genotype and A allele was over represented in endurance athletes. In essence, when genotype distributions were examined according to aerobic performance there was an association between *PPARGC1A* genotype and maximal oxygen consumption ($p < 0.001$) (Tural et al., 2014). This data suggested the *PPARGC1A* A-allele has a strong effect on aerobic performance in 60 Turkish elite endurance athletes ($p < 0.001$) (Tural et al., 2014).

This inconsistency in the published results raises the question of whether endurance performance influenced by the *PPARGC1A* G>A SNP. The purpose of this study was to analyse the genotype and allele frequency distribution of the *PPARGC1A* rs8192678 SNP in prima facie Caucasian (elite and sub-elite) marathon runners and a healthy non-athletic prima facie Caucasian non-athlete control populations. In addition, marathon performance was compared between *PPARGC1A* genotype and allele groups. Based on the literature it was hypothesised that the minor A-allele would be under-represented in the elite marathon runner cohort compared to a non-athlete control group, and that carriers of the A-allele would have slower marathon performance times compared to GG homozygotes.

6.2 METHOD

6.2.1 Research Participant Characteristics

The *PPARGC1A rs8192678* SNP investigation included a marathon cohort of 243 male and 156 female Caucasian marathon runners, which was divided into elite and sub-elite groups based on personal best performance times. Controls consisted of 224 males and 82 females to give a total cohort of 306 individuals, who had not competed in sport at an elite/sub-elite standard. For completeness of analysis, the marathon and control cohorts were characterised according to sex for some of the data analysis. Complete numbers in each research group for the *PPARGC1A rs8192678* study are outlined in Table 16

Table 13 Numbers of research participants in the PPARGC1A rs8192678 *analyses*

	Male	Female	Total
Marathon Athlete	243	156	399
Elite Marathon Athletes	86	87	173
Sub-Elite Marathon Athletes	157	69	226
Non-Athletes	224	82	306

Table 14 PPARGC1A rs8192678 *participant characteristics (mean (standard deviation))*

	Athlete	Elite	Sub-Elite	Control

Height (m)	Male	1.78 (0.06)	1.79 (0.06)	1.78 (0.06)	1.79 (0.07)
	Female	1.65 (0.07)	1.65 (0.07)	1.66 (0.08)	1.65 (0.07)
Mass (kg)	Male	67.0 (6.5)	67.1 (7.5)	67.0 (5.9)	77.0 (11.3)
	Female	53.6 (5.3)	52.7 (5.4)	54.8 (5.0)	66.3 (11.0)
Age (years)	Male	36 (8)	37 (8)	37 (7)	31 (18)
	Female	37 (7)	37 (8)	37 (7)	31 (18)

6.2.2 **DNA Collection**

Participant DNA collection is outlined in Chapter 3, Section 3.5 of the methods. In brief, 521 blood samples, 71 buccal samples and 113 saliva samples were collected from participants. Section 3.6 of Chapter 3 details the protocols for DNA extraction of the various samples.

6.2.3 **Genotyping**

Section 3.8 of Chapter 3 describes the DNA genotyping in full. Briefly each participants *PPARGC1A* rs8192678 genotype was detected by real-time PCR, using the TaqMan assay that included the appropriate primers and probes (Applied Biosystems). Each 10 µL pcr reaction contained: 5 µL Genotyping Master Mix (Applied Biosystems), 4.3 µL nuclease-free H_2O (Qiagen), 0.5 µL *PPARGC1A* rs8192678 TaqMan genotyping assay mix (Applied Biosystems), and 0.2 µL of participant DNA. For control wells, 0.2 µL nuclease-free H_2O replaced the DNA template. Genotyping was completed on the StepOnePlus real-time PCR machine (Applied Biosystems). Briefly, there were 40 cycles of denaturation at 95°C for 15 s then annealing and extension at 60°C for 1 min. StepOnePlus software version 2.3 (Applied Biosystems) was used to determine each participant's genotype. All samples were analysed in duplicate and there was 100% agreement within all duplicate samples.

6.2.4 **Data Analysis and Statistics**

SPSS (version 21, SPSS Inc) was used to perform the statistical analyses. Independent T-tests were used to determine differences in height and body mass between marathon

athletes and non-athlete controls. Genotype and allele frequencies between marathon athlete groups and non-athlete controls were calculated and compared by Pearson's Chi square test. Analysis of genotype and allele frequency distribution between elite and sub-elite marathon groups were completed using chi square test of difference. Excel (Microsoft, 2013) was used to complete Hardy-Weinberg equilibrium assessment. A genotype association with personal best performance time was assessed by one way ANOVA. A post hoc Brown-Forsythe correction was added to the significant one way ANOVA result to account for the group variances being unequal as calculated by the Levene's test. This meant that the one way ANVOA analysed the absolute deviations from the median rather than the mean.

6.3 *PPARGC1A* RS8192678 RESULTS

6.3.1 Height and Body Mass

The male marathon athlete and non-athlete control cohorts did not differ in height ($p = 0.427$). However, on average male marathon athletes weighed 10 kg lighter than the male non-athlete controls ($p = 1.0 \times 10^{-13}$). In females, there was no difference in the height of the marathon athletes when compared to the non-athlete controls ($p = 0.495$). However, the female non-athlete controls were on average 12.7 kg heavier than the female marathon athletes were ($p = 1.002 \times 10^{-10}$). These data are shown in Table 17 and are self-reported.

6.3.2 PPARGC1A rs8192678 Hardy-Weinberg Equilibrium

In both marathon athletes and non-athlete controls the *PPARGC1A* rs8192678 genotype met Hardy-Weinberg equilibrium (Marathon athlete $\chi^2 = 0.673$, Non-Athlete $\chi^2 = 0.002$, $p = 0.05$, 1 df) in all groups tested independently (Elite marathon athletes $\chi^2 = 0.520$, Sub-Elite marathon athletes $\chi^2 = 3.037$). Hardy-Weinberg equilibrium remained when data were stratified by sex and re-analysed (Males: Marathon athlete $\chi^2 = 0.222$, Elite marathon

athletes $\chi^2 = 0.142$, Sub-Elite marathon athletes $\chi^2 = 0.569$, and Non-Athlete $\chi^2 = 0.460$, p = 0.05 for all groups), (Females: Marathon athlete $\chi^2 = 0.523$, Elite marathon athlete $\chi^2 = 0.820$, Sub-Elite marathon athlete $\chi^2 = 3.414$, and Non-athletes $\chi^2 = 1.111$, p = 0.05 for all groups).

6.3.3 **PPARGC1A rs8192678 Genotype and Allele Frequencies**

The results of the distribution of *PPARGC1A* genotypes and alleles in the athlete and controls are presented in Table 18. The initial analysis conducted in the whole marathon athlete cohort revealed a tendency to a higher frequency of the A-allele in the marathon athlete cohort than in non-athletic controls ($\chi^2 = 2.988$, p = 0.084). When considering the frequency of the A-allele in the two separate marathon athlete groups to non-athletic controls no differences were observed. For completeness, the elite and sub-elite marathon athlete cohorts allele frequencies were also compared, again no differences were observed. There were no differences in genotype frequency distribution between any of the group comparisons (elite marathon athletes, sub-elite marathon athletes, non-athlete controls) when both male and female marathon athletes were analysed together.

Table 15 The genotype and allele frequencies of PPARGC1A *rs8192678 polymorphism in marathon athletes and non-athletes*

	***PPARGC1A* Genotype**						***PPARGC1A* Allele**			
Total Cohort	**GG**		**GA**		**AA**		**G**		**A**	
Marathon **Athletes**	179	(44.9%)	171	(42.9%)	49	(12.3%)	529	(66.3%)	269	(33.7%)
Elite Marathon Athletes	71	(41.0%)	83	(48.0%)	19	(11.0%)	225	(65.0%)	121	(35.0%)
Sub-Elite Marathon Athletes	108	(47.8%)	88	(38.9%)	30	(13.3%)	304	(67.3%)	148	(32.7%)
Non-Athletes	146	(47.7%)	131	(42.8%)	29	(9.5%)	423	(69.1%)	189	(30.9%)
Males										
Marathon **Athletes**	107	(44.0%)	106	(43.6%)	30	(12.3%)	320	(65.8%)	166	(34.2%)
Elite Marathon Athletes	30	(34.9%)	43	(50.0%)	13	(15.1%)	103	(59.9%)	69	(40.1%)
Sub-Elite Marathon Athletes	77	(49.0%)	63	(40.1%)	17	(10.8%)	217	(69.1%)	97	(30.9%)
Non-Athletes	103	(46.0%)	101	(45.1%)	20	(8.9%)	307	(68.5%)	141	(31.5%)
Females										
Marathon **Athletes**	72	(46.2%)	65	(41.7%)	19	(12.2%)	209	(67.0%)	103	(33.0%)
Elite Marathon Athletes	41	(47.1%)	40	(46.0%)	6	(6.9%)	122	(70.1%)	52	(29.9%)
Sub-Elite Marathon Athletes	31	(44.9%)	25	(36.2%)	13	(18.8%)	87	(63.0%)	51	(37.0%)
Non-Athletes	43	(52.4%)	30	(36.6%)	9	(11.0%)	116	(70.7%)	48	(29.3%)

In the male cohort, the primary analysis of *PPARGC1A* rs8192678 genotype and allele frequencies showed no association between marathon athletes and non-athlete controls. On closer inspection, the minor A-allele was over represented 9.2% in the elite male marathon athletes when compared to non-athlete controls ($\chi^2 = 6.871$, $p = 0.03$) (shown in Figure 24) An association was also reflected in the male elite marathon cohort towards the minor AA genotype ($\chi^2 = 6.890$, $p = 0.04$) when compared to non-athlete controls (OR = 0.696, 95% CI 0.383 – 1.265, $p = 0.235$). Further to this, a tendency towards the minor A-allele was seen when the male elite marathon group was compared to the male sub-elite marathon group ($\chi^2 = 2.986$, $p = 0.084$). Data for all comparisons are displayed in Table 18.

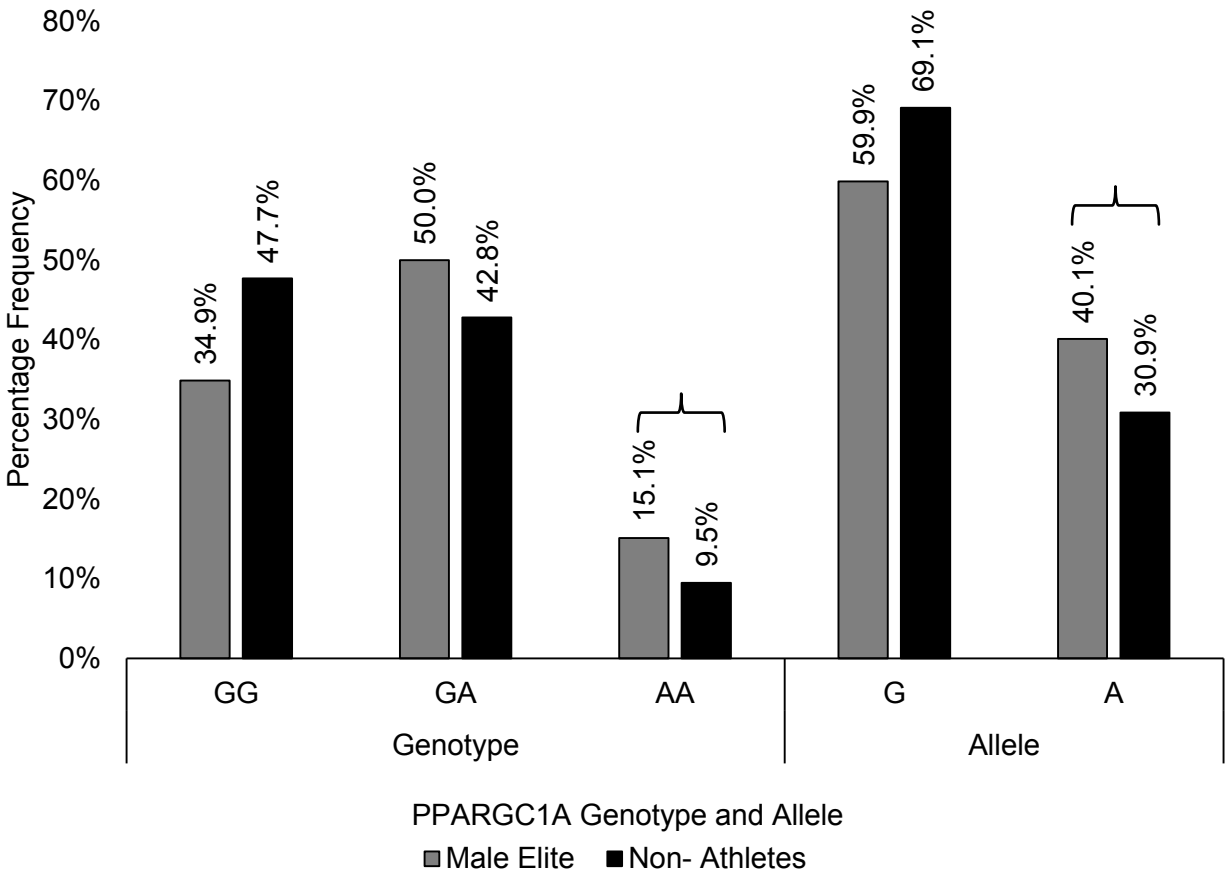

Figure 24 The PPARGC1A rs8192678 allele frequencies in male elite marathon athletes all non-athletes the AA genotype (* p = 0.032) and A allele († p = 0.009) is more frequent in elite athletes than non-athlete controls.

In the female cohort, there was a 7.8% higher AA genotype frequency in sub-elite marathon athletes compared to non-athlete controls ($\chi^2 = 7.193$, $p = 0.04$) (shown in Figure 25) (OR = 0.531, 95% CI 0.212 – 1.331, p = 0.177) and a tendency for a higher AA frequency in sub-elite vs. elite marathon athletes ($\chi^2 = 5.425$, $p = 0.066$). There were no further associations of either genotype or allele frequencies in the female cohorts (shown in Figure 25).

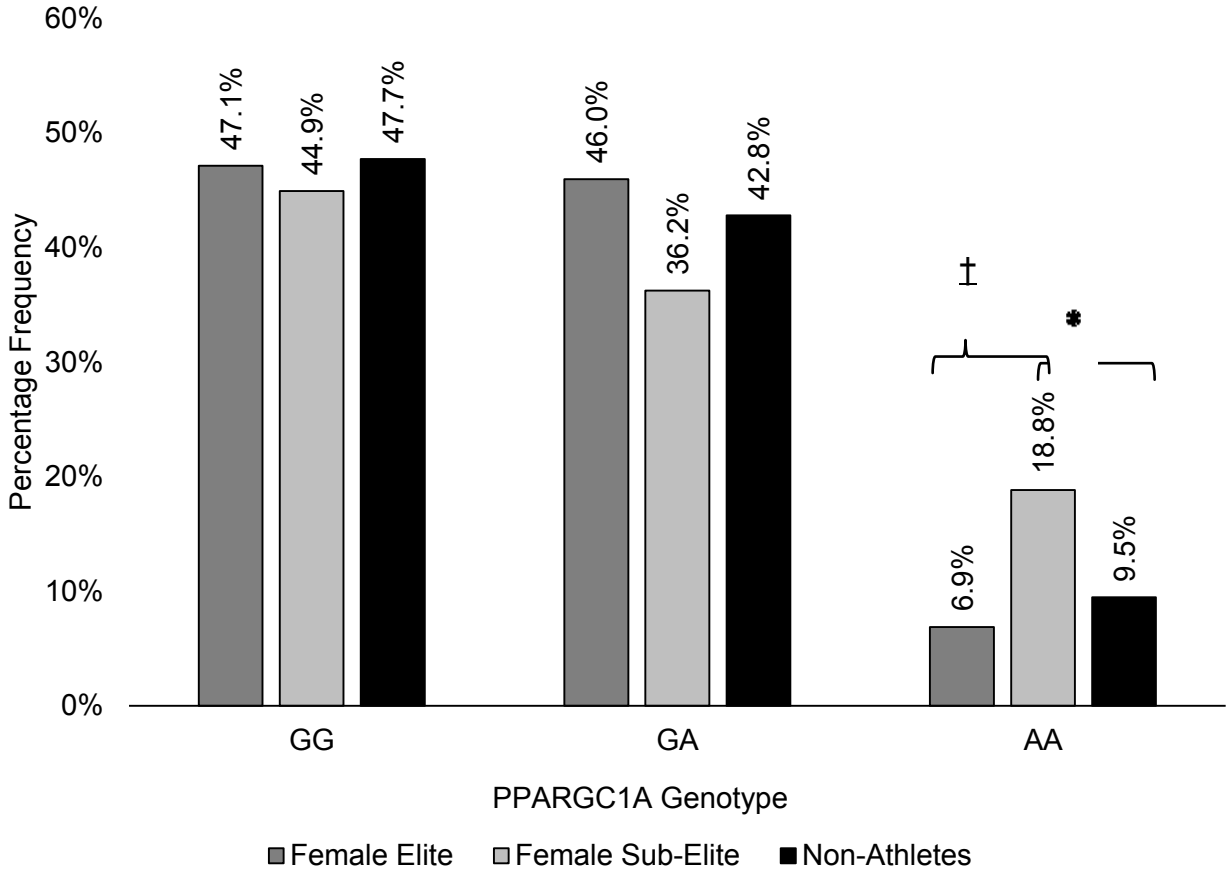

Figure 25 The PPARGC1A rs8192678 allele frequencies in female elite marathon athletes, female sub-elite marathon athletes and all non-athletes: The AA genotype is more frequent in female sub-elite marathon athletes than non-athletes (* p = 0.027). Female sub-elite marathon athletes showed a tendency towards a higher AA genotype frequency than elite marathon athletes († p = 0.066)

6.3.4 **Comparison of Marathon PB and Genotype**

Male and female median performance time for each genotype group are shown in Figure 26 and 27. Median performance times did not differ between *PPARGC1A* genotype

groups in male marathon runners (F = 2.051, p = 0.131). However, in female marathon runners, performance time was slower in AA homozygotes compared to individuals of GG and GA genotype (F = 3.136, p = 0.022 (Brown-Forsythe correction).

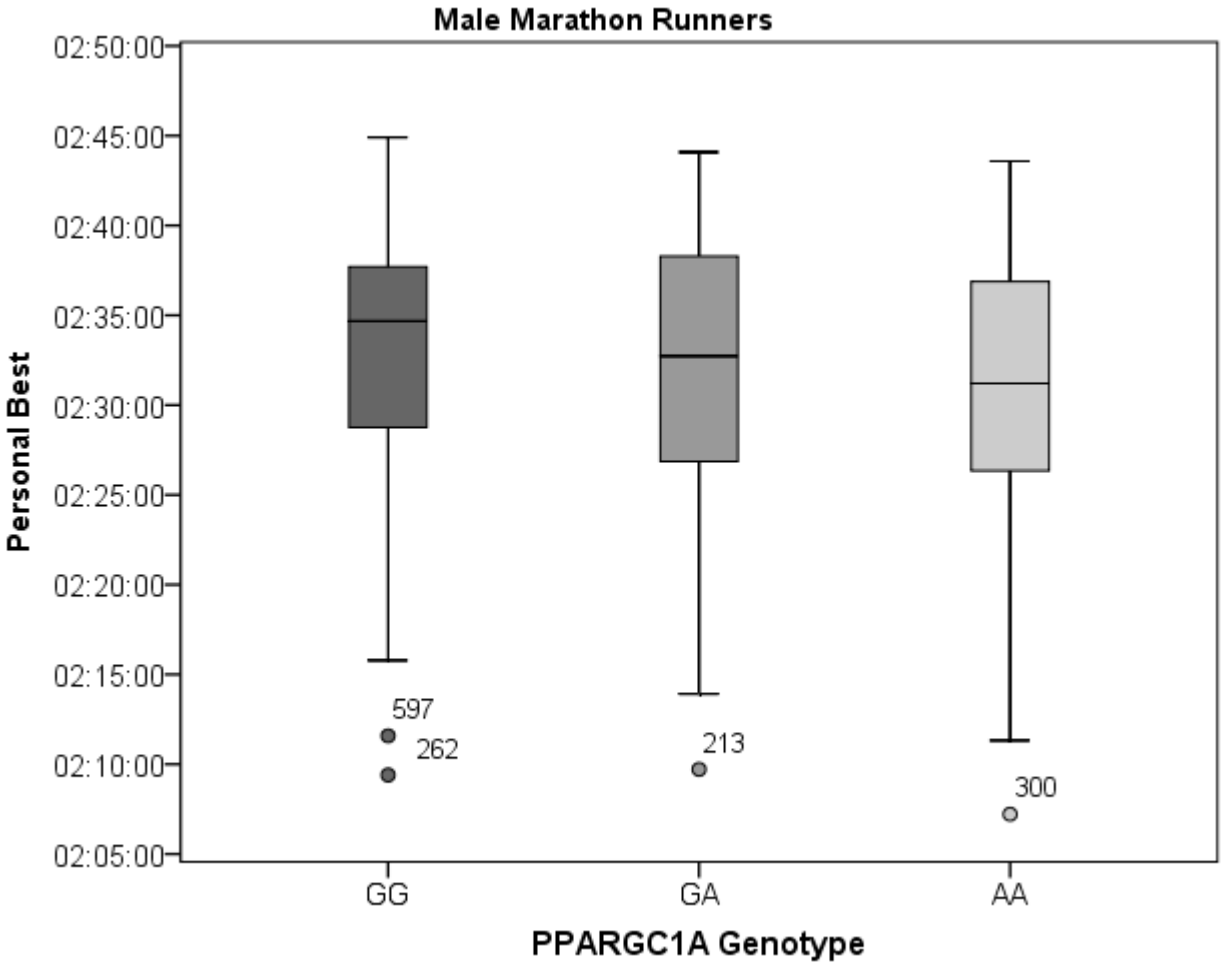

Figure 26 Marathon personal best times for males grouped by PPARGC1A rs8191678 genotype. Data are medians and maximum and minimum

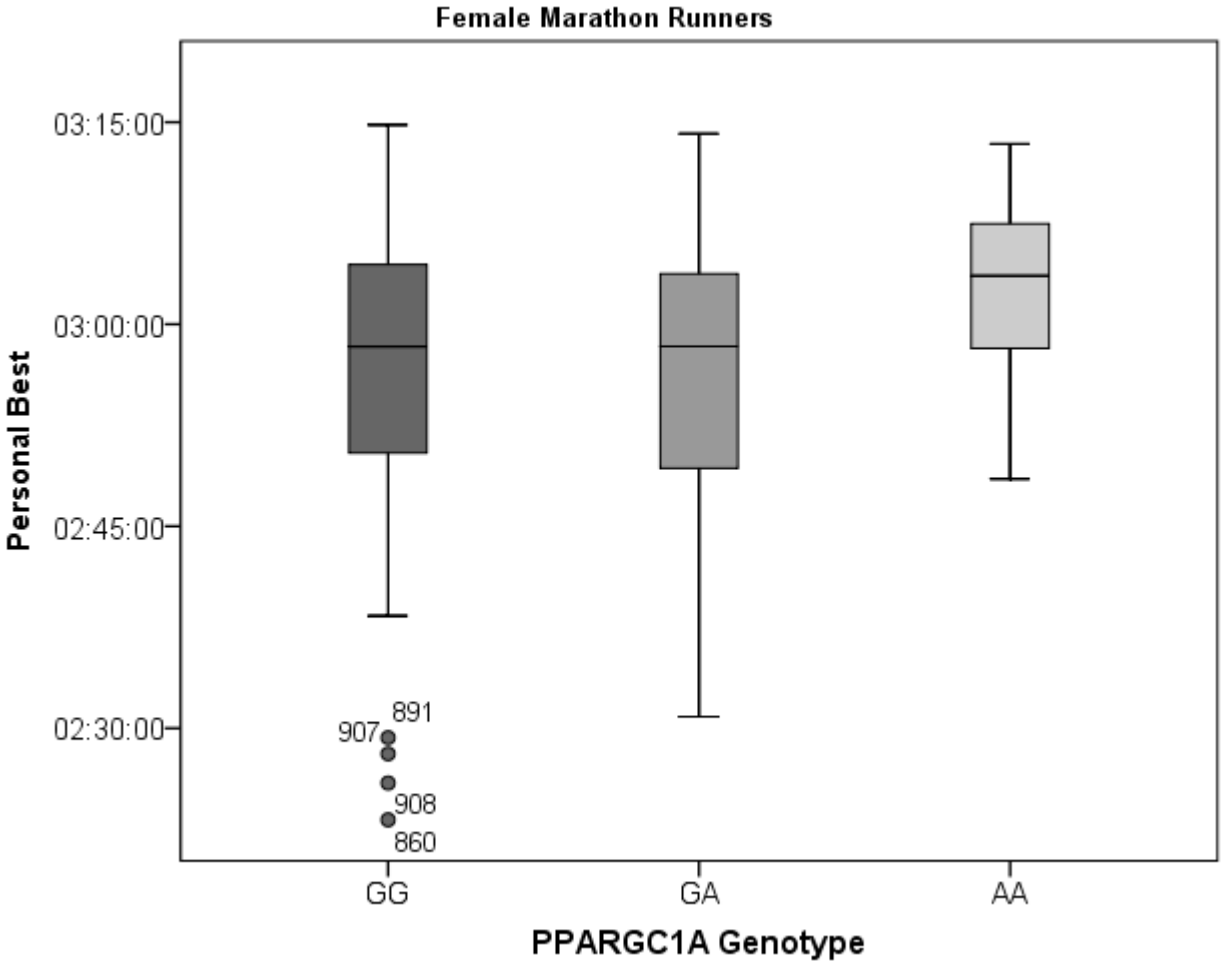

Figure 27 Marathon personal best times for females grouped by PPARGC1A rs8191678 genotype. Data are medians and maximum and minimum

6.4 DISCUSSION

The current study demonstrates that the *PPARGC1A* G>A SNP is associated with elite athlete status in male marathon runners. The minor A-allele was overrepresented in a cohort of elite male marathon runners, who complete a marathon in 2 hours 30 minutes or less, when compared to non-athlete controls. Non-significant tendencies for a greater frequency distribution of the A-allele were observed between the whole marathon athlete cohort and non-athletic controls, and the male elite marathon athletes when compared to the sub-elite male marathon athletes. Non-significant tendencies for a greater frequency distribution of the minor AA genotype were seen when male elite marathon athletes when compared to non-athlete controls and female sub-elite marathon athletes when compared to non-athlete controls and elite female marathon athletes. A genotype association with

personal best time was not observed in males but female AA homozygotes were found to be slower than their GA and GG counterparts.

The greater frequency of the minor A-allele among elite male marathon athletes compared controls are somewhat in line with a previous finding that elite endurance performance was greater in carriers of the A-allele compared to G-allele carriers in 60 Turkish elite endurance athletes, who had participated in national/international track and field championships. Tural et al. (2014) also reported that AA genotype and the A-allele were over-represented in the endurance athletes compared to non-athlete controls. Thus, these results suggest the *PPARGC1A* A-allele contributes to elite endurance athlete status and performance. The physiological mechanism(s) underpinning the genotype association with elite endurance athlete status in the present study and that by Tural et al. (2014) cannot be inferred from the present results. Although (elderly) carriers of the minor A-allele have been shown to have lower basal *PPARGC1A* mRNA levels compared to GG homozygotes (Ling et al., 2004), it is possible that the A-allele positively influences *PPARGC1A* mRNA stability, thus potentially increasing translation and skeletal muscle PGC-1α content. As well as PGC-1α being a key requirement for normal expression of mitochondrial genes in cardiac and skeletal muscle (Arany et al., 2005), it has also been found to interact with oestrogen-related receptor-α (ERRα) to increase transcription of the lactate dehydrogenase B (*LDHB*) gene in skeletal muscle (Summermatter et al., 2013). As LDH catalyses the conversion of lactate to pyruvate, it is important in lactate clearance. Hence, a combination of an increase in mitochondrial density and *LDHB* expression (via increased PGC-1α concentration/activity) should increase time to exhaustion and, therefore, endurance performance in *PPARGC1A* A-allele carriers.

In contrast to the apparent benefit of being an A-allele carrier in terms of elite male marathon athletes status, the present study demonstrated that female AA homozygotes had a slower personal best marathon time compared to their GG and GA counterparts.

This sex-dependent genotype association with endurance athlete status/performance may be linked to by PGC-1α's interaction with ERRα, as explained above (Summermatter et al., 2013).Due to the similarity between ERR-α and oestrogen receptor-α (ER-α), and the finding that the two proteins regulate similar genes (Vanacker et al., 1999), it is possible that higher circulating levels of oestrogen in females compared to males inhibits the otherwise beneficial effect of the *PPARGC1A* A-allele, thus explaining why female AA homozygotes in the present study had a slower personal best marathon performance compared to female marathon athletes of GA and GG genotype.

The slower female marathon performance in AA homozygotes compared to female athletes of GG and GA genotypes is in line with the findings of (Lucia et al., 2005), who found a higher frequency of the G-allele among a mixture of different types of elite endurance athletes (50 male world-class Spanish cyclists and 54 male middle- and long-distance) compared to controls. This association study was replicated in a large mixed cohort of Russian and Polish endurance athletes (Maciejewska et al., 2011). They determined that the minor A-allele was under represented in their mixed sporting discipline cohort of athletes when compared to a low activity control population. It should be noted that the Polish and Russian athlete cohorts did not contain any marathon runners in their endurance cohorts and therefore may not be representative of the genetic frequency distributions that may influence elite marathon runner status and marathon running performance.

This is the first study to investigate an association between the *PPARGC1A* rs8191678 G>A SNP and elite endurance athlete status and running performance in a large cohort of solely marathon runners. Previous studies investigating the association of this SNP with elite athlete status have done so in mixed athlete cohorts, which may confound any association due to increased "noise" caused by the different physiological importance of PGC-1α in different athlete groups. By examining an association in a pure athlete group,

the present study suggests that the minor A-allele is important in determining elite athlete status in male marathon runners, while this minor allele appears to be disadvantageous in determining performance in female marathon runners.

6.5 CONCLUSION

In summary, this study found that the minor *PPARGC1A* A-allele is over-represented in a large cohort of elite male marathon runners. However, female homozygotes of the A-allele were slower than other female marathon runners of GG and GA genotype.

This may be due to oestrogen inhibiting the interaction between PGC1- α and ERRα, thus limiting lactate clearance and decreasing endurance performance in females of AA genotype.

7 UNCOUPLING PROTEINS – UCPS

7.1 INTRODUCTION

The ability to complete a marathon at an elite speed requires short-term responses and long-term training adaptations to the extreme physiological demands placed upon the skeletal muscle (Bruton, 2002, Jiang et al., 2009, Stewart and Rittweger, 2006, Coyle, 2007, Gamboa and Andrade, 2012, Zhou et al., 2000). Mitochondria are a key determinant of that adaption providing the energy required for skeletal muscle contraction and thus propulsion and through uncoupling proteins, the regulation of energy expenditure and heat production during exercise (Zhou et al., 2000, Jiang et al., 2009). Uncoupling proteins are mitochondrial intramembranous proteins responsible for the anti-porting of anions and protons across the inner mitochondrial membrane (Boron and Boulpaep, 2012, Boss et al., 1997, Muzzin et al., 1999). In doing so, they regulate oxidative phosphorylation coupling to ATP production - vital in an endurance exercise event such as marathon. Several publications have suggested further roles for UCPs, though these mechanisms are not fully understood (Muzzin et al., 1999). One such suggestion is that uncoupling proteins mediate ROS production by skeletal muscle in response to exercise, caused by the large fluxes in oxygen through the electron transport chain in heavily respiring skeletal muscle and an inevitable oxygen perfusion mismatch causing localised hypoxia (Vogt et al., 2001, Boss et al., 1997). This would modulate ATP production through intracellular fatty acids and account for the hypothesised possible influences of insulin on UCP3 expression and UCP3 expression on GLUT 4 expression in mice (Pedersen et al., 2001, Astrup et al., 1999, Tsuboyama-Kasaoka et al., 1998, Nisr and Affourtit, 2014). Exercise in rodent models has been reported to increase both *UCP2* and *UCP3* mRNA expression (Tsuboyama-Kasaoka et al., 1998, Jiang et al., 2009, Cortright et al., 1999, Zhou et al., 2000) suggesting that glucose and fatty acid metabolism in response to exercise may be regulated via UCPs and that in heavily respiring muscle UCPs reduce free radical production. This would confer a competitive advantage in the marathon as elite runners often in the later stages of the race are reliant on fatty acid

based production of ATP. Increased efficiency in the coupling of ATP production and oxidative phosphorylation through the uncoupling proteins may provide a competitive advantage to endurance athletes and predispose athletes to performance success at the elite level.

Gene polymorphisms that have been associated with aerobic capacity (Ahmetov et al., 2008, Ahmetov et al., 2009), delta efficiency (Dhamrait et al., 2012) and oxidative phosphorylation (Jiang et al., 2009, Muzzin et al., 1999, Nisr and Affourtit, 2014), making them candidates for associations with elite endurance performance, include *UCP2* rs659366, *UCP2* rs660339 and *UCP3* rs1800849. In 61 healthy, adults who exhibit low physical activity, the T allele of *UCP2* rs659366 was associated with delta efficiency assessed via cycle ergometry before and after endurance training (Dhamrait et al., 2012). *UCP2* rs660339 was also associated with endurance performance in 230 Russian rowers when compared to non-athletes (Ahmetov et al., 2008). The rowers showed T allele associations with increased $\dot{V}O_{2\,max}$ and higher frequencies of the T allele in their most elite athletes. This group also investigated the frequency of the *UCP3* rs1800849 polymorphism the T allele was associated with elite performance in rowing and increased $\dot{V}O_{2\,max}$ (Ahmetov et al., 2009). Later, the same research group expanded their study to include endurance athletes from other endurance sporting disciplines including 134 runners (Ahmetov et al., 2008). Again, they assessed the *UCP2* rs660339 and *UCP3* rs1800849 polymorphisms reporting a higher frequency of both the *UCP2* rs660339 T allele and the *UCP3* rs1800849 T allele in the athlete cohort when compared to controls (Ahmetov et al., 2008). Interestingly, the *UCP3* rs1800849 T allele was not associated with endurance performance in the athletes competing in the longest race disciplines though this stratification by race distance meant there were very few athletes from each sporting discipline in each category and this may have reduced statistical power. Further association studies of the *UCP3* rs1800849 polymorphism in the sport of triathlon reported no association of either allele with race completion time (Hudson et al., 2004). These

athletes were also compared to a control cohort, where no differences were reported in either genotype or allele frequency (Hudson et al., 2004).

Thus the findings regarding *UCP* polymorphisms in endurance athletes are unclear and the athlete group's used previously either contain athletes of multiple sporting disciplines or athletes who complete in tri-part events, each with varying phenotypes. Consequently, the purpose of this study was to investigate whether there was a difference in genotype and/or allele frequency between the marathon cohort and a control group. In essence, it was hypothesised that the TT genotype would be over represented in the marathon cohort for all three *UCP* polymorphisms. Similarly it was hypothesized that the T allele would be more prevalent in the elite marathon runners when compared to sub-elite marathon and controls. Accordingly, it was hypothesised that T allele carriers would be able to complete a marathon in a faster time than those carrying the C allele in all three UCP polymorphisms.

7.2 METHOD

7.2.1 Research Participant Characteristics

7.2.1.1 UCP2 rs659366

Five hundred and seventy eight prima facie Caucasian adults provided written informed consent to take part in this study. This total cohort comprised 364 marathon runners (male, n = 216; female, n = 148) and a control (non-marathon running) cohort, comprising 224 men and 82 women. The marathon runners were stratified into elite and sub-elite subgroups according to their official marathon personal best performance time (http://www.poweroften.co.uk; Table 19). Research participant height and body mass are shown in Table 22.

Table 16 Research participant numbers for the UCP2 rs659366 *analyses*

	Male	Female	Total
Marathon Athletes	216	148	364
Elite Marathon Athletes	75	82	157
Sub-Elite Marathon Athletes	141	66	207
Non-Athletes	224	82	306

7.2.1.2 UCP2 rs660339

Six hundred and fifty eight prima facie Caucasian adults provided written informed consent to take part in this study. This total cohort comprised 396 marathon runners (male, n = 241; female, n = 155) and a control (non-marathon running) cohort, comprising 224 men and 82 women. The marathon runners were stratified into elite and sub-elite subgroups according to their official marathon personal best performance time (http://www.poweroften.co.uk; Table 20). Research participant height and body mass are shown in Table 22.

Table 17 Research participant numbers for the UCP2 rs660339 *analyses*

	Male	Female	Total
Marathon Athletes	241	155	396
Elite Marathon Athletes	86	87	173
Sub-Elite Marathon Athletes	155	68	223
Non-Athletes	224	82	306

7.2.1.3 UCP3 rs1800849

Seven hundred and five prima facie Caucasian adults provided written informed consent to take part in this study. This total cohort comprised 399 marathon runners (male, n = 243; female, n = 156) and a non-athlete control cohort, comprising 224 men and 82 women. The marathon runners were stratified into elite and sub-elite subgroups according to their official marathon personal best performance time (http://www.poweroften.co.uk; Table 21). Research participant height and body mass are shown in Table 22.

Table 18 Research participant numbers for the UCP3 rs1800849 *analyses*

	Male	Female	Total
Marathon Athletes	243	156	399
Elite Marathon Athletes	86	87	173
Sub-Elite Marathon Athletes	157	69	226
Non-Athletes	224	82	306

Table 19 UCP*2 rs659366,* UCP2 *rs660339, and* UCP3 *rs1800849 research participant characteristics (mean (standard deviation.))*

		Athlete	Elite	Sub-Elite	Control
Height (m)	Male	1.78 (0.06)	1.79 (0.06)	1.78 (0.06)	1.79 (0.07)
	Female	1.65 (0.07)	1.65 (0.07)	1.66 (0.08)	1.65 (0.07)
Mass (kg)	Male	67.0 (6.5)	67.1 (7.5)	67.0 (5.9)	77.0 (11.3)
	Female	53.6 (5.3)	52.7 (5.4)	54.8 (5.0)	66.3 (11.0)
Age	Male	36 (8)	37 (10)	35 (7)	23 (7)
	Female	37 (7)	37 (8)	37 (7)	25 (10)

7.2.2 **DNA Collection**

Participant DNA collection is outlined in in section 1.5 of the methods. In brief, 521 blood samples, 71 buccal samples and 113 saliva samples were collected from participants for the *UCP2* rs659366 analysis. Participant DNA collection from 518 blood samples, 71 buccal samples and 113 saliva samples was completed for *UCP2* rs660339 analysis. Five hundred and twenty one blood samples, 71 buccal samples and 113 saliva samples were collected from participants to complete the *UCP3* rs1800849 analyses. 5mL blood samples were taken from a superficial forearm vein and stored in an EDTA treated tube at -20°C until processing. Saliva samples were collected into Oragene DNA OG-500 collection tubes (DNA Genotek Inc., Ontario, Canada) according to the manufacturer's instructions and stored at room temperature until genotyping. Sterile buccal swabs (Omni swab, Whatman, Springfield Mill, UK) were rubbed against the cheek approximately 20

times to collect buccal cells. Tips were ejected into sterile tubes and stored at -20°C for transportation to the lab for DNA isolation.

7.2.3 **DNA Isolation**

Automated DNA extraction was performed using a QIAcube (Qiagen, Crawley, UK), following the QIAamp blood mini protocol as detailed in the methods section. Briefly, 200 µL sample was lysed with QIAGEN Protease Enzyme. To each sample, 200 µL of 96% ethanol was added. Samples were agitated vigorously to mix. The mixture was transferred to a QIAamp mini spin column tube containing a DNA collection filter and bound by centrifugation at 6000 g for 1 min. Three further wash buffer centrifugation cycles followed to wash out any remaining protein and impurities. Genomic DNA remained bound to the silica gel DNA filter membrane in the spin column during these wash cycles. After each wash the filtrate was discarded. In the final buffer centrifugation, the sample genomic DNA was eluted into 100 µL of molecular grade H_2O to provide purified genomic DNA. The purified genomic DNA was stored at 4°C until the genotyping was performed.

7.2.4 **Genotyping**

UCP genotyping of all participants was completed by rtPCR. Genotyping of all participants was completed using the TaqMan assays for *UCP2 rs659366, UCP2 rs660339* and *UCP3 rs1800849* (Applied Biosystems). Each TaqMan assay that included the appropriate TaqMan primers and probes (Applied Biosystems). Each 10 µL rtPCR reaction contained 5 µL Genotyping Master Mix (Applied Biosystems), 4.3 µL nuclease-free H_2O (Qiagen West Sussex, UK), 0.5 µL TaqMan genotyping assay mix (Applied Biosystems), and 0.2 µL of participant DNA. For control wells, 0.2 µL nuclease-free H_2O replaced the DNA template. Genotyping was completed on the StepOnePlus real-time PCR system (Applied Biosystems). Briefly, there were 40 cycles of denaturation at 95°C for 15 s then annealing and extension at 60°C for 1 min, as per Applied Biosystems guidelines. StepOnePlus

software version 2.3 (Applied Biosystems) was used for analysis. All samples were analysed in duplicate. There was 100% agreement within duplicates of all samples.

7.2.5 **Data Analysis and Statistics**

SPSS (version 21, SPSS Inc) was used to perform independent T-tests to compare height and body mass between athletes and non-athletes. Genotype and allele frequencies in all groups were calculated and compared using chi-square analyses in SPSS. Excel (Microsoft, 2013) was used to assess Hardy-Weinberg equilibrium. Genotype associations with personal best were assessed in SPSS by one way ANOVA.

7.3 **UCP RESULTS**

7.3.1 **Participant Height and Body Mass**

Table 22 shows the height and body mass for the male and female marathon athletes and non-athlete control. Although the participant numbers are slightly different between the three UCP polymorphisms analysed, the data for the mean and standard deviation height and mass data did not differ. The male marathon athletes and non-athlete controls did not differ in height ($p \geq 0.427$). However, the male non-athlete controls were on average 10 kg heavier than the male marathon athletes ($p = 1.0 \times 10^{-13}$). In females, there was no difference in height of the marathon athletes when compared to the non-athlete controls ($p \geq 0.491$). However, female marathon athletes were 12.7 kg lighter than female non athlete controls ($p \leq 1.0 \times 10^{-13}$) all data are self-reported.

7.3.2 **Hardy – Weinberg Equilibrium**

7.3.2.1 UCP2

Hardy-Weinberg equilibrium deviated in *UCP2 rs659366* and *UCP2 rs660339* amongst the whole marathon athlete cohorts (*UCP2 rs659366* $\chi^2 = 4.337$, *UCP2 rs660339* $\chi^2 = 4.427$). In all *UCP2 rs659366* further analysis Hardy-Weinberg remained. Hardy-Weinberg

equilibrium was not observed among the mixed elite marathon athletes in *UCP2 rs660339* (χ^2 = 5.127) and the total male marathon athlete cohort (χ^2 = 4.845) though in all further analysis of *UCP2 rs660339* Hardy-Weinberg equilibrium was observed.

7.3.2.2 UCP3

UCP3 rs1800849 Hardy-Weinberg equilibrium was present in all groups analysed regardless of further stratification by sex or elite marathon athlete status.

7.3.3 **UCP2 rs659366 Genotype and Allele Frequencies**

Table 20 Genotype and allele frequencies of the UCP2 rs659366 *polymorphism in marathon athletes and non-athletes*

	***UCP2 rs659366* Genotype**						***UCP2 rs659366* Allele**			
Total Cohort	**CC**		**CT**		**TT**		**C**		**T**	
Marathon **Athletes**	156	(39.1%)	202	(50.6%)	41	(10.3%)	514	(64.4%)	284	(35.6%)
Elite Marathon Athletes	65	(37.6%)	91	(52.6%)	17	(9.8%)	221	(63.9%)	125	(36.1%)
Sub-Elite Marathon Athletes	91	(40.3%)	111	(49.1%)	24	(10.6%)	293	(64.8%)	159	(35.2%)
Non-Athletes	132	(43.1%)	140	(45.8%)	34	(11.1%)	404	(66.0%)	208	(34.0%)
Males										
Marathon **Athletes**	97	(39.9%)	121	(49.8%)	25	(10.3%)	315	(64.8%)	171	(35.2%)
Elite Marathon Athletes	35	(40.7%)	43	(50.0%)	8	(9.3%)	113	(65.7%)	59	(34.3%)
Sub-Elite Marathon Athletes	62	(39.5%)	78	(49.7%)	17	(10.8%)	202	(64.3%)	112	(35.7%)
Non-Athletes	90	(40.2%)	112	(50.0%)	22	(9.8%)	292	(65.2%)	156	(34.8%)
Females										
Marathon **Athletes**	59	(37.8%)	81	(51.9%)	16	(10.3%)	199	(63.8%)	113	(36.2%)
Elite Marathon Athletes	30	(34.5%)	48	(55.2%)	9	(10.3%)	108	(62.1%)	66	(37.9%)
Sub-Elite Marathon Athletes	29	(42.0%)	33	(47.8%)	7	(10.1%)	91	(65.9%)	47	(34.1%)
Non-Athletes	42	(51.2%)	28	(34.1%)	12	(14.6%)	112	(68.3%)	52	(31.7%)

UCP2 rs659366 genotype and allele distributions are detailed in Table 23 and did not differ between the entire marathon athlete cohort and non-athlete controls (genotype χ^2=3.832, Z = 0.648, p = 0.517; allele χ^2=0.913, Z = 0.725 p = 0.451) when analysed using the Cochran-Armitage test. When analysed using the Pearson chi squared test, there was no difference in genotype distribution between elite marathon athletes and non-athlete controls (genotype χ^2 = 3.273, p = 0.195; allele χ^2 = 0.706, p = 0.401) or sub-elite athletes and non-athlete controls (genotype χ^2 = 1.040, p = 0.595; allele χ^2 = 0.285, p = 0.593). When elite marathon athletes were compared to sub-elite marathon athletes no difference in genotype distribution was observed (χ^2 = 0.477, p = 0.788; allele χ^2 = 0.770, p = 0.781).

Male marathon runners showed no difference in *UCP2* rs659366 genotype or allele distributions when compared to all controls (genotype χ^2 = 1.600, p = 0.449; allele χ^2 = 0.311, p = 0.577; Table 23). No differences in genotype or allele distribution were observed when male elite marathon runners were compared to all controls (genotype χ^2 = 0.711, p = 0.701; allele χ^2 = 0.008, p = 0.930; Table 23). There was no difference in genotype (χ^2 = 1.025, p = 0.599) or allele (χ^2 = 0.396, p = 0.529) frequency distribution observed when sub-elite marathon runners were compare to controls (Table 23). There were no differences between the elite male marathon runners and the total sub-elite marathon runners in either genotype (χ^2 = 0.147, p = 0.929) or allele frequency (χ^2 = 0.042, p = 0.838; Table 23).

Table 23 also details the genotype and allele frequencies of the female marathon runners who demonstrated no difference in genotype or allele distribution frequencies when compared to the total control cohort (genotype χ^2 = 2.423, p = 0.298; allele χ^2 = 0.692, p = 0.405). This lack of difference in genotype and allele frequency was further noted when female elite marathon runners were compared to the entire non-athletes (elite, genotype χ^2 = 3.244 p = 0.197; allele frequency χ^2 = 1.206, p = 0.272) (sub-elite, genotype χ^2 = 0.143, p = 0.931; allele χ^2 = 3.10 x10^{-4}, p = 0.986). Between group analysis of the female

elite athlete and all sub-elite athletes revealed no difference in either the genotype (χ^2 = 0.981, p = 0.612) or allele (χ^2 = 0.414, p = 0.520) frequency.

7.3.4 **UCP2 rs660339 Genotype and Allele Frequencies**

Table 21: The genotype and allele frequencies of UCP2 rs660339 *polymorphism in marathon runners and non-marathon controls*

	***UCP2 rs660339* Genotype**						***UCP2 rs660339* Allele**			
Total Cohort	**CC**		**CT**		**TT**		**C**		**T**	
Marathon **Athletes**	133	(33.6%)	210	(53.0%)	53	(13.4%)	476	(60.1%)	316	(39.9%)
Elite Marathon Athletes	56	(32.4%)	97	(56.1%)	20	(11.6%)	209	(60.4%)	137	(39.6%)
Sub-Elite Marathon Athletes	77	(34.5%)	113	(50.7%)	33	(14.8%)	267	(59.9%)	179	(40.1%)
Non-Athletes	117	(38.2%)	147	(48.0%)	42	(13.7%)	381	(62.3%)	231	(37.7%)
Males										
Marathon **Athletes**	76	(31.5%)	133	(55.2%)	32	(13.3%)	285	(59.1%)	197	(40.9%)
Elite Marathon Athletes	27	(31.4%)	49	(57.0%)	10	(11.6%)	103	(59.9%)	69	(40.1%)
Sub-Elite Marathon Athletes	49	(31.6%)	84	(54.2%)	22	(14.2%)	182	(58.7%)	128	(41.3%)
Non-Athletes	82	(36.6%)	113	(50.4%)	29	(12.9%)	277	(61.8%)	171	(38.2%)
Females										
Marathon **Athletes**	57	(36.8%)	77	(49.7%)	21	(13.5%)	191	(61.6%)	119	(38.4%)
Elite Marathon Athletes	29	(33.3%)	48	(55.2%)	10	(11.5%)	106	(60.9%)	68	(39.1%)
Sub-Elite Marathon Athletes	28	(41.2%)	29	(42.6%)	11	(16.2%)	85	(62.5%)	51	(37.5%)
Non-Athletes	35	(42.7%)	34	(41.5%)	13	(15.9%)	104	(63.4%)	60	(36.6%)

Marathon runners showed no difference in the genotype (Z = 0.648 p = 0.517) frequency distribution or the allele (Z = 0.584 p = 0.619) frequency distribution when compared to controls (Table 24). When elite marathon runners and sub-elite marathon runners were independently compared to controls (elite Z = 0.589, p = 0.556; sub-elite χ^2 = 17.099, p = 1.9 x 10^{-4}) No differences in elite (Z = 0.402 p = 0.748) or sub-elite (χ^2 = 1.084, p = 0.298) allele frequency distributions were observed when compared to the control participants. There were no differences between the elite marathon runners and the sub-elite marathon runners in either genotype (χ^2 = 1.433, p = 0.488) or allele frequency (χ^2 = 0.024, p = 0.878).

Table 24 shows in the male athletes there was no difference in the genotype distribution (Z = 0.890, p = 0.373) when compared to non-athletes via the Cochran-Armitage test. A genotype apparent difference was recorded when male elite and sub-elite athletes were compared to non-athletes independently (elite χ^2 = 11.173, p = 0.001 OR 1.130, CI 95% 0.525-2.432 p = 0.754; sub-elite χ^2 = 17.584, p = 0.01: OR 0.899, CI 95% 0.495-1.632 p = 0.726) via Pearsons-Chi squared. Between group analysis of male elite and all sub-elite athletes revealed neither a genotype (χ^2 = 0.353, p = 0.838) or an allele (χ^2 = 0.058, p = 0.810) association. No allele associations were recorded when the male athlete group was compared to the entire non-athletes group (χ^2 = 2.005, p = 0.157) nor when the elite male marathon runners and sub-elite male marathon runners were compared to non-athletes independently (elite, χ^2 = 0.412, p = 0.521; sub-elite χ^2 = 1.658, p = 0.198).

In the female athletes, a genotype association was observed when compared to non-athletes (genotype χ^2=8.376, p = 0.02: OR 1.202, CI 95% 0.568-2.546 p = 0.630) though this was not reflected in the allele analysis (allele χ^2=0.054, p = 0.816). The female elite athletes also reflected a genotype association when compared to non-athletes (genotype χ^2 = 8.942, p = 0.02: OR 1.450, CI 95% 0.598-3.519 p = 0.411), though this was not reflected in the allele distribution (allele χ^2 = 0.132, p = 0.716). A lack of association was

observed when sub-elite runners were compared to non-athletes (genotype $\chi^2 = 2.140$, $p = 0.343$; allele $\chi^2 = 0.003$, $p = 0.953$) and when female elite runners were compared to all sub-elite runners ($\chi^2 = 2.461$, $p = 0.292$; allele $\chi^2 = 0.058$, $p = 0.810$). Data for all comparisons are displayed in Table 24.

7.3.5 UCP3 rs1800849 Genotype and Allele Frequencies

Table 22: The genotype and allele frequencies of UCP3 rs1800849 *polymorphism in marathon runners and non-marathon controls*

	***UCP3 rs1800849* Genotype**						***UCP3 rs1800849* Allele**			
Total Cohort	**CC**		**CT**		**TT**		**C**		**T**	
Marathon Athletes	219	(54.9%)	159	(39.8%)	21	(5.3%)	597	(74.8%)	201	(25.2%)
Elite Marathon Athletes	94	(54.3%)	68	(39.3%)	11	(6.4%)	256	(74.0%)	90	(26.0%)
Sub-Elite Marathon Athletes	125	(55.3%)	91	(40.3%)	10	(4.4%)	341	(75.4%)	111	(24.6%)
Non-Athletes	167	(54.6%)	118	(38.6%)	21	(6.9%)	452	(73.9%)	160	(26.1%)
Males										
Marathon Athletes	135	(55.6%)	94	(38.7%)	14	(5.8%)	364	(74.9%)	122	(25.1%)
Elite Marathon Athletes	50	(58.1%)	30	(34.9%)	6	(7.0%)	130	(75.6%)	42	(24.4%)
Sub-Elite Marathon Athletes	85	(54.1%)	64	(40.8%)	8	(5.1%)	234	(74.5%)	80	(25.5%)
Non-Athletes	122	(54.5%)	89	(39.7%)	13	(5.8%)	333	(74.3%)	115	(25.7%)
Females										
Marathon Athletes	84	(53.8%)	65	(41.7%)	7	(4.5%)	233	(74.7%)	79	(25.3%)
Elite Marathon Athletes	44	(50.6%)	38	(43.7%)	5	(5.7%)	126	(72.4%)	48	(27.6%)
Sub-Elite Marathon Athletes	40	(58.0%)	27	(39.1%)	2	(2.9%)	107	(77.5%)	31	(22.5%)
Non- Athletes	45	(54.9%)	29	(35.4%)	8	(9.8%)	119	(72.6%)	45	(27.4%)

UCP3 rs1800849 genotype and allele distributions are detailed in Table 25. No difference was recorded in *UCP3* rs1800849 genotype frequency distributions (χ^2 = 1.666, p = 0.435) when the athletes were compared to non-marathon controls (Table 25). This lack of association was further reflected in the allele frequency distributions (χ^2 = 0.378, p = 0.539; Table 25). There were no differences in *UCP3* rs1800849 genotype or allele frequency distributions between elite and control participants (genotype χ^2 = 0.910, p = 0.956; allele χ^2 = 0.003, p = 0.955; Table 25). Between group analysis of sub-elite marathon runners and non-marathon controls showed no difference in genotype or allele frequency distributions (genotype χ^2 =2.150, p = 0.341; allele χ^2 = 0.589, p = 0.443). When elite and sub-elite groups were compared, *UCP3* rs1800849 genotype and allele frequency distributions showed no difference (genotype χ^2 = 0.736, p = 0.692; allele χ^2 = 0.220, p = 0.639).

UCP3 rs1800849 genotype and allele frequency distributions did not differ between the male marathon cohort and all non-marathon controls (genotype χ^2= 0.473, p = 0.789; allele χ^2=0.378, p = 0.539). There was no difference in genotype or allele frequency distribution between male elite marathon runners and the total non-marathon control group (genotype χ^2 = 0.504, p = 0.777; allele χ^2 = 0.265, p = 0.607; Table 25). Male sub-elite runners revealed no difference in genotype or allele frequency distributions when compared to all non-marathon controls (genotype χ^2 = 0.917, p = 0.632; allele χ^2 = 0.720, p = 0.788; Table 25). When male elite runners were compared to all sub-elite runners, no difference in genotype distribution was observed (χ^2 = 0.998, p = 0.607; Table 25). This lack of association was also seen in the allele frequency distribution (allele χ^2 = 0.001, p = 0.971; Table 25).

UCP3 rs1800849 genotype distributions were not different amongst the female athletes and the total control group (χ^2 = 1.688, p = 0.430). No difference in allele distribution was observed between female athletes and all non-marathon controls (χ^2 = 0.110, p = 0.741).

There was no difference in genotype distribution (χ^2 = 1.003, p = 0.605) or allele distribution (χ^2 = 0.187, p = 0.665) between the female elite marathon runners and the total control cohort. Female sub-elite marathon runners also showed no difference in genotype frequency distribution (χ^2 = 1.732, p = 0.421) or allele frequency distribution (χ^2 = 0.968, p = 0.325) when compared to all non-marathon controls. Sub group analysis between elite female marathon runners and the total sub-elite marathon runner cohort showed no difference in genotype distribution (χ^2 = 1.278, p = 0.528) or allele distribution (χ^2 = 0.608, p = 0.435).

In summary, genotype-dependent differences were recorded for *UCP2* rs660339 when the female marathon athlete cohort was compared to non-athletes (χ^2=8.376, p = 0.02). Similarly, the male and female elite marathon athlete cohorts showed genotype-dependent associations when compared to non-athletes (males χ^2 = 11.173, p = 0.001: females χ^2=8.376, p = 0.02), A genotype difference was recorded between male sub-elite athletes and controls (χ^2 = 17.584, p = 0.01). No other analysis revealed an association of either genotype or allele for *UCP2* rs659336, *UCP2* rs660339 or *UCP3* rs1800849.

7.3.6 **Comparison of Marathon PB and Genotype**

Mean performance times were not significantly different between *UCP2* genotype groups (*UCP2* rs659366 males F = 1.546, p = 0.215; females F = 1.389, p = 0.252 Figure 28 and 29; *UCP2* rs660339 males F = 0.678, p = 0.509; females F = 1.620, p = 0.201 Figure 30 and 31) or the *UCP3* genotype (males F = 0.647, p = 0.525, female F = 0.677 p = 0.510; Figure 32 and 33).

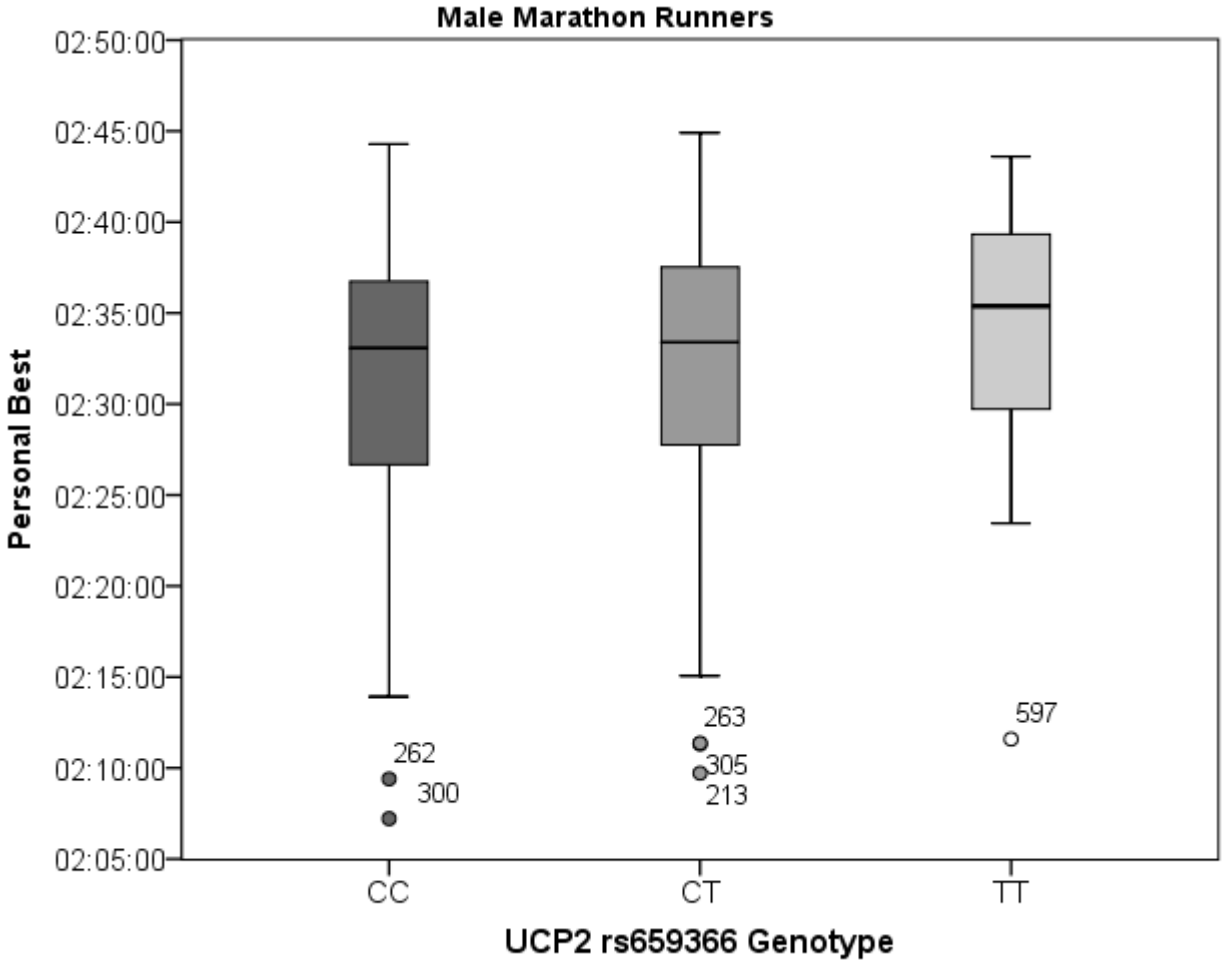

Figure 28 Marathon personal best times for males grouped by UCP rs659366 genotype. Data are medians and minimum and maximum.

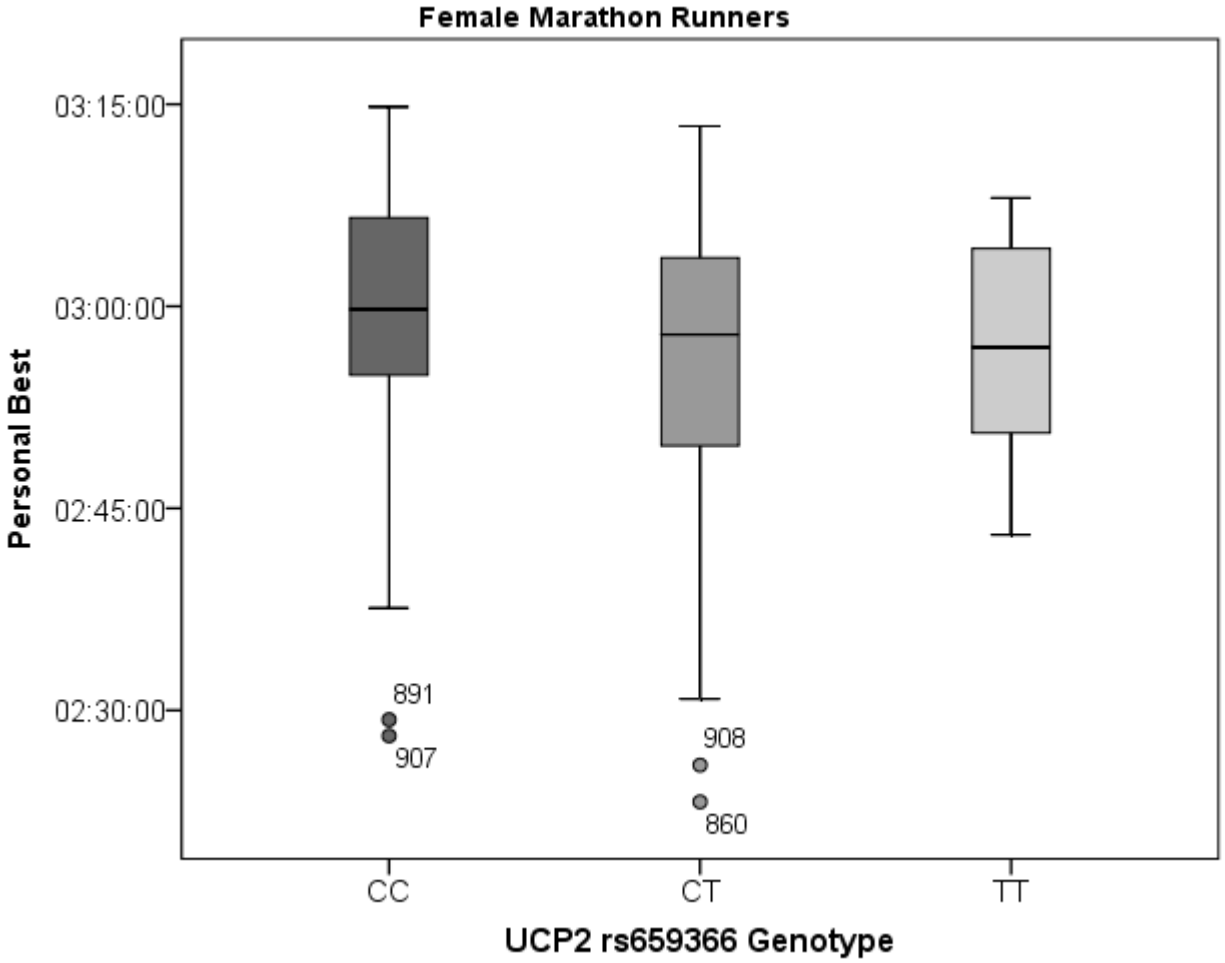

Figure 29 Marathon personal best times for females grouped by UCP2 rs659366 genotype. Data are medians and minimum and maximum.

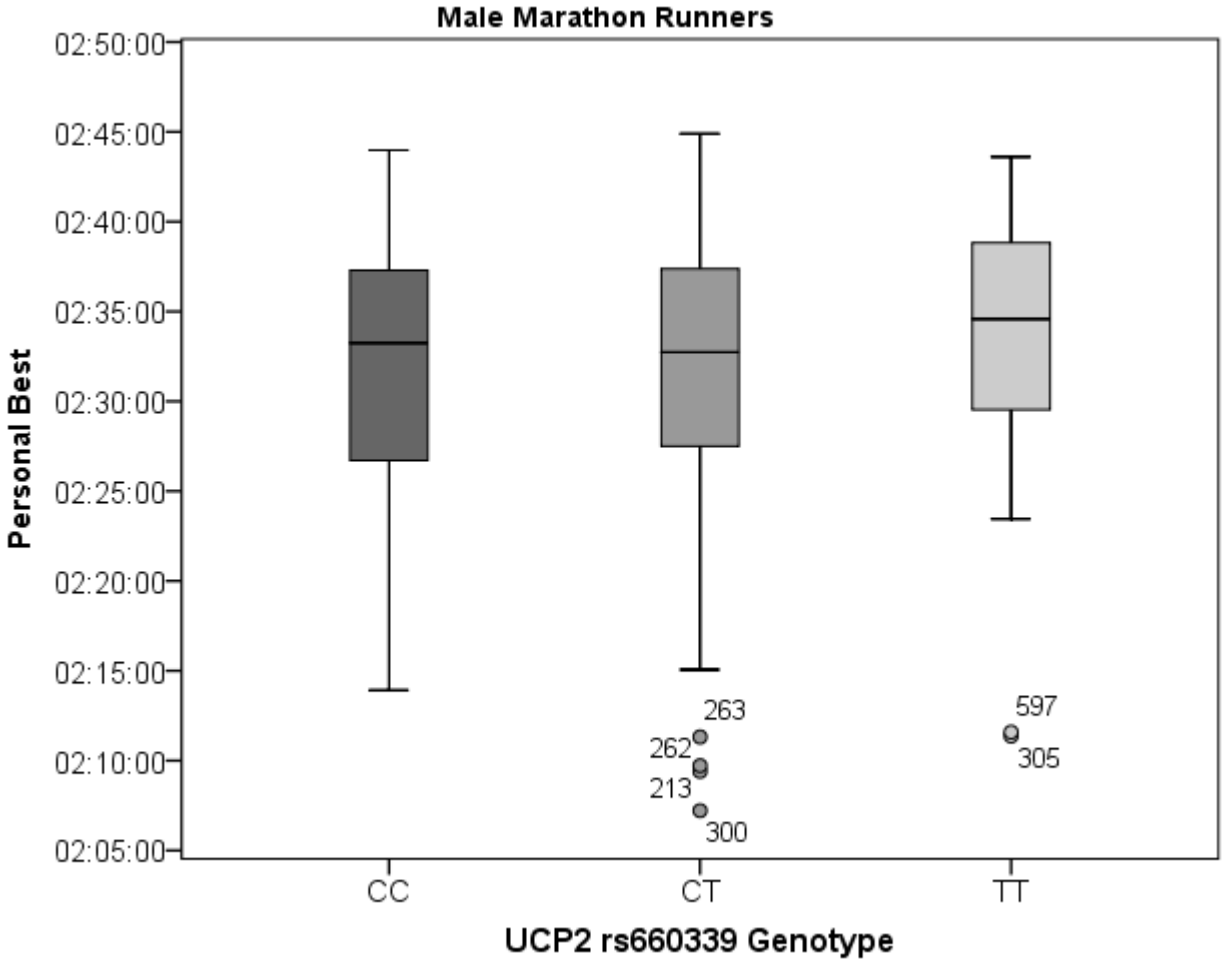

Figure 30 Marathon personal best times for males grouped by UCP2 rs660339 genotype. Data are medians and minimum and maximum.

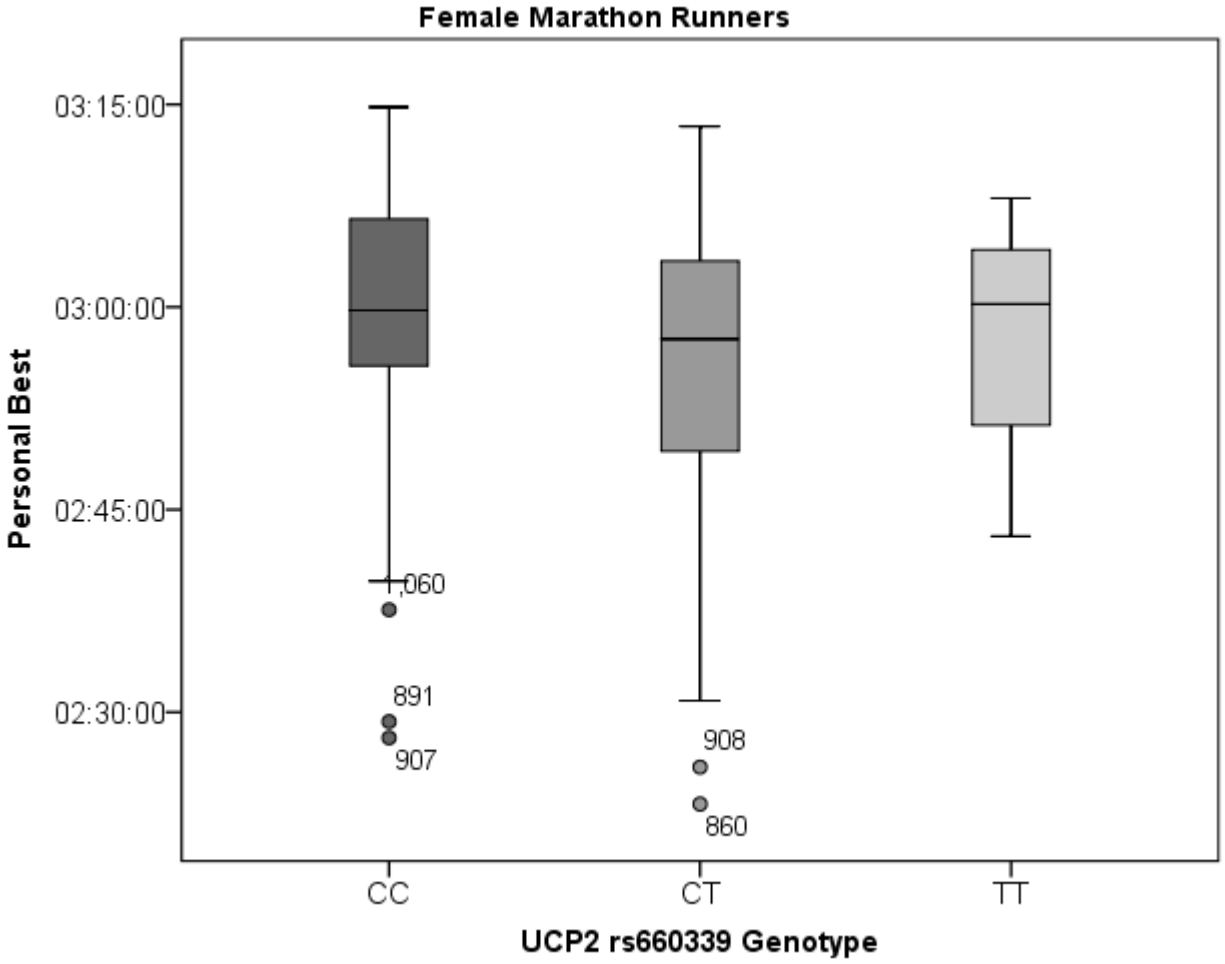

Figure 31 Marathon personal best times for females grouped by UCP2 rs660339 genotype. Data are medians and minimum and maximum.

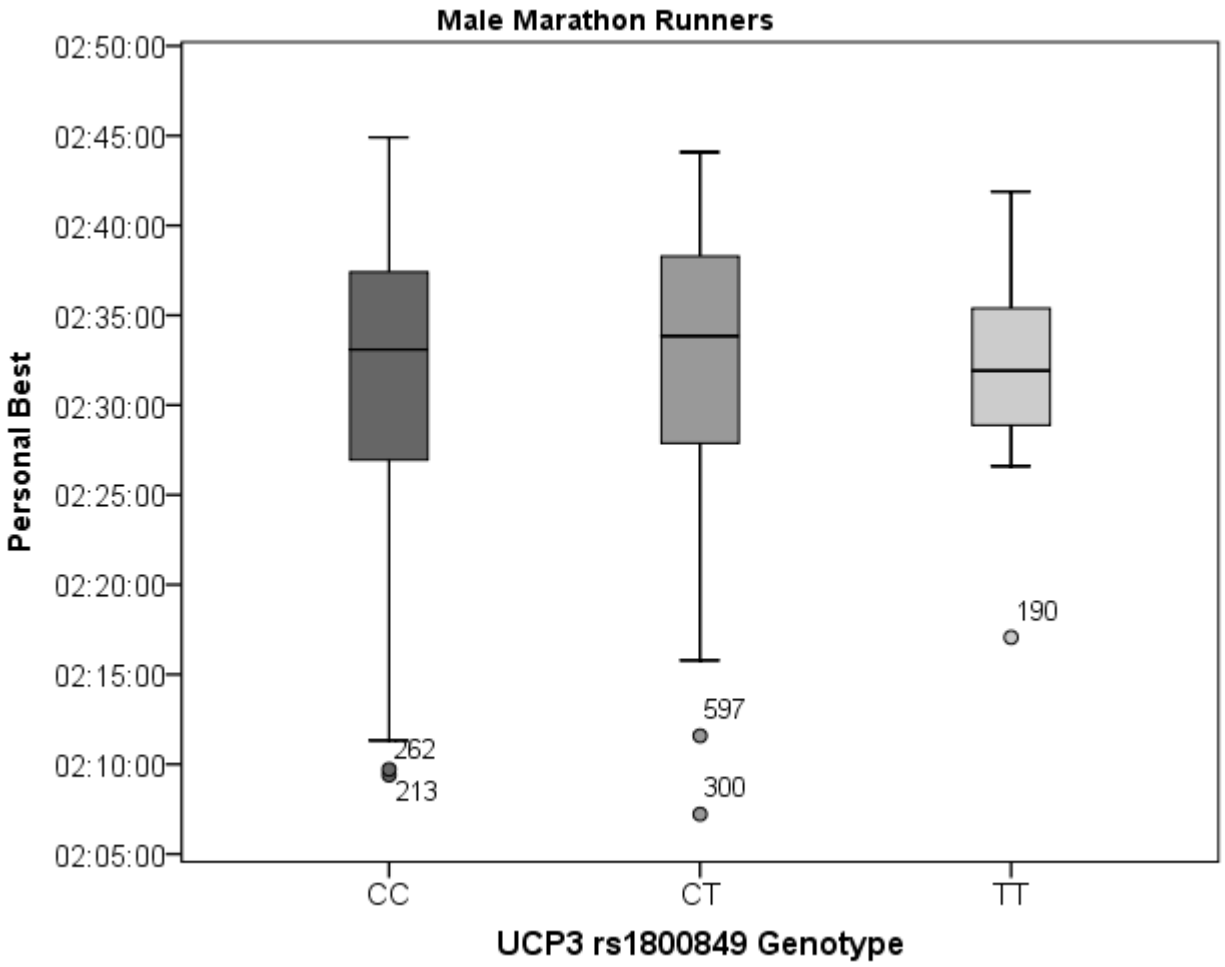

Figure 32 Marathon personal best times for males grouped by UCP3 rs1800849 genotype. Data are medians and minimum and maximum.

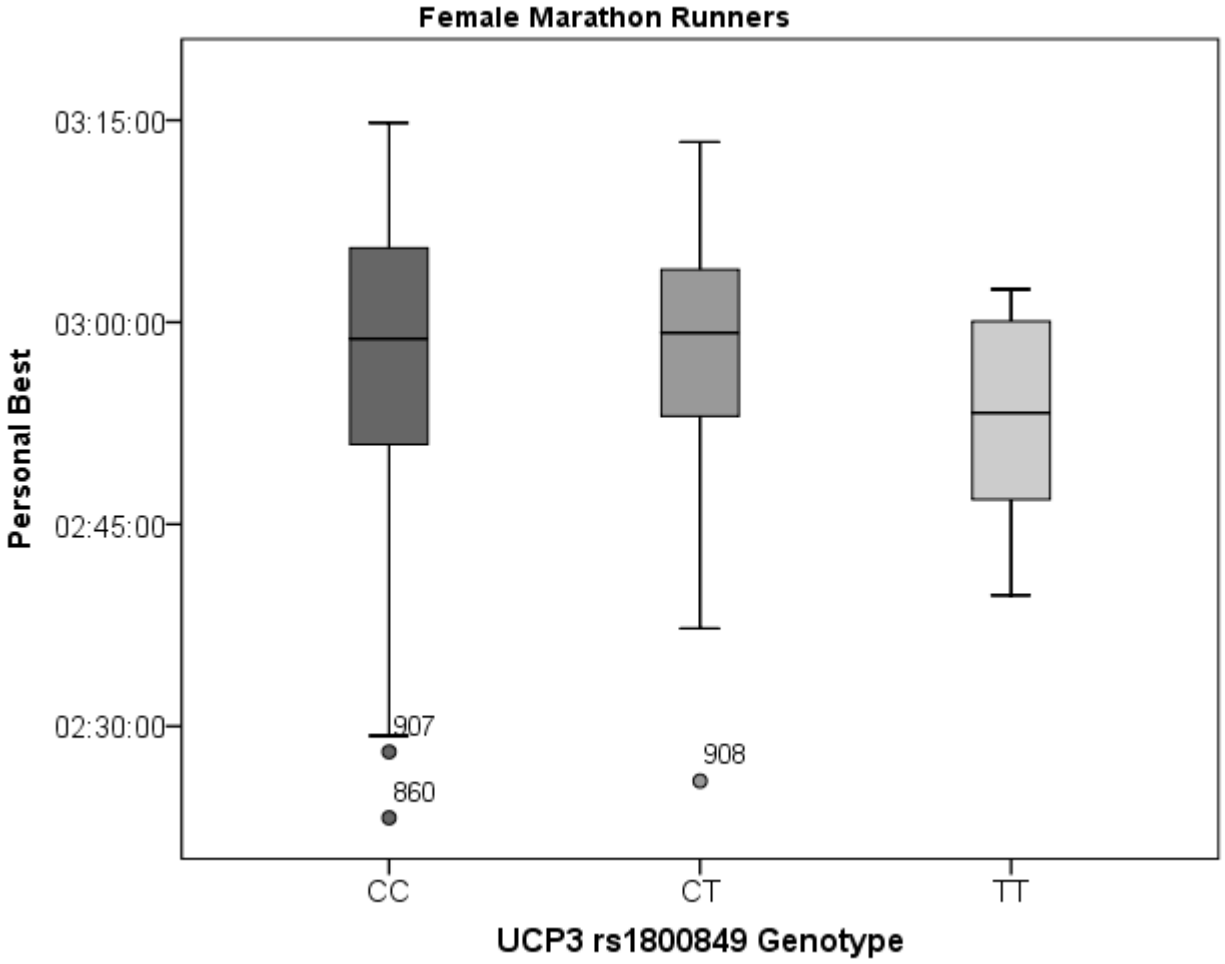

Figure 33 Marathon personal best times for females grouped by UCP3 rs1800849 genotype. Data are medians and minimum and maximum.

7.4 DISCUSSION

The study, via the candidate gene approach, assessed the association of *UCP2* rs659366, *UCP2* 660339 and *UCP3 rs1800849* polymorphisms with marathon sporting performance. A possible genotype-dependent association with personal best marathon running time was also investigated. The main finding of this study was an overrepresentation of the CT genotype of *UCP2* rs660339 in the athlete cohort when compared to non-athletes. This was further reflected in both the male and female athlete cohorts when compared to non-athlete controls.

On examination of the elite athletes there was a *UCP2* rs660339 CT genotype over-representation in the entire athlete cohort and non-athletes. Again, this association was observed in both the male and female athlete subgroups. In sub-elite athletes, the CT genotype was over-represented when compared to non-athletes and this association remained when male sub-elite athletes were compared to the complete non-athlete cohort. Interestingly, no allele associations were recorded between any of the groups when *UCP2* rs660339 was assessed. These interesting findings may partially be accounted for as a result deviation from Hardy-Weinberg equilibrium between the entire athlete cohort and non-athletes, the entire elite cohort and non-athletes and the male athlete cohort and non-athletes though this is unlikely. Deviation in Hardy–Weinberg equilibrium may result from a number of factors the most apparent is a genotyping error (Wittke-Thompson et al., 2005). All samples were genotyped in duplicate and data recorded centrally in a quality controlled database making a genotyping error unlikely. The other seven assumptions of Hardy-Weinberg equilibrium in the main have been accounted for also, leaving only the allele frequencies are equal in the sexes and there is no migration, mutation or selection (Wittke-Thompson et al., 2005). According to the literature, genotype and allele frequencies are not sex-specific for this polymorphism and genotype frequencies examined in the study population, resembled those of the European Ensembl genome browser (http://www.ensembl.org/index.html). Whilst a selection pressure may account for the deviation it would expected to result in deviation in all assessments not just selected populations of interest so therefore again is unlikely. More likely, this phenomenon is accounted for by the complexity of the endurance athletic trait, indeed departure from Hardy-Weinberg equilibrium has been noted in other studies assessing genotype associations with athletic ability (Zarebska et al., 2013, Gomez-Gallego et al., 2009, Druzhevskaya et al., 2008). No genotype or phenotype associations were observed in any of the between group analysis in either *UCP2* rs659336 or *UCP3* rs1800849.

An association between *UCP2 rs659366* genotype and personal best time was not observed in either males or females. Neither was an association between *UCP2 rs660339* genotype and personal best in male or female marathon runners. No, *UCP3* rs1800849 genotype-dependent associations with personal best time was observed in males or females.

7.5 CONCLUSION

In conclusion, this study was the first to analyse an association of endurance athlete status with the *UCP2* rs659366, *UCP2* rs660339 and *UCP3* rs1800849 polymorphisms in a large group of elite marathon runners. This study found *UCP2* rs660339 CT genotype over-representation in the athlete in the athlete cohort when compared to non-athletes, when males and female were analysed in combination and independently. These findings were echoed in the elite athletes and non-athletes when analysed as a whole cohort and when the athlete group was stratified by sex. In sub-elite athletes, an over-representation of the CT genotype was observed when athletes were compared to controls and when male athletes were compared to controls. No genotype dependant association of neither *UCP2* rs659366, *UCP2* rs660339 nor *UCP3* rs1800849 with personal best time was observed in either male or female marathon runners. This supports the hypothesis that *UCP2* rs660339 may confer a favourable advantage to endurance athletes mediated possibly in part through increased efficiency in oxidative phosphorylation coupling to ATP production in skeletal muscle during sustained endurance exercise such as marathon running.

8 TOTAL GENOTYPE SCORE

8.1 INTRODUCTION

Elite endurance athletic performance is a polygenic trait influenced by multiple genetic variants that each contribute a small amount to the observed inter-individual variability in athletic performance. Previous research has noted that elite endurance athletes carry a higher proportion of the genetic variants thought to contribute to elite endurance performance than those who compete is elite strength or power based sports and the general population. In addition, several studies have examined the combined genetic contribution to groups of athletes though it should be noted that these studies have often included athletes from multiple sporting discipline, so may not be representative of a single sporting discipline athlete population such as the one included in this study.

A number of factors including $\dot{V}O_{2\,max}$, running economy, lactate threshold skeletal muscle fibre type and $\dot{V}O_2$ Kinetics, determines elite endurance athlete status in marathon running. Genetic factors will undoubtedly influence each of these factors and therefore play a small part in genetic contribution to the inter-individual variability in performance seen in international competition and the elite endurance athlete status of a marathon runner. It may also be assumed that these small contributions to genetic performance $\dot{V}O_{2\,max}$, running economy, lactate threshold skeletal muscle fibre type and $\dot{V}O_2$ Kinetics may also confer

Despite some of the earlier chapters in this thesis reporting no association of a genetic polymorphism and endurance athlete status, or a tendency, it is possible that the contribution of these genotypes is only evident when the individual genotypes are analysed cumulatively. By examining, the cumulative effect of several non-significant associations, tendencies of association and those significant associations reported earlier it should be

possible to identify the contributory effect of the included individual polymorphisms on the elite endurance athlete status and further determine if TGS influences personal best marathon performance times.

8.2 METHODS

Participant recruitment inclusion criteria, for both marathon athletes and non-athlete controls and genotyping methods of *ACE* (rs4341), *ACTN3* (rs1815739), *AGT* (rs699), *PPARGC1A* (rs8192678), *UCP2* (rs659366), *UCP2* (rs660339) and *UCP3* (rs1800849) polymorphisms are described in detail in chapter 3 sections 3.1 to Section 3.9. Therefore, the methods are detailed in brief below.

8.2.1 Participants

Prima facie Caucasian male and female marathon runners volunteered to participate in this study. All 364 participants provided written informed consent prior to participation in the study.

8.2.2 Genotyping

Fluorophore based TaqMan® real-time PCR methods were used to determine each participants specific genotype for the following polymorphisms *ACE* (rs4341), *ACTN3* (rs1815739), *AGT* (rs699), *PPARGC1A* (rs8192678), *UCP2* (659366), *UCP2* (rs660339) and *UCP3* (rs1800849) all detection analysis was completed using StepOnePlus software version 2.3 (Applied Biosystems). All samples were analysed in duplicate. There was 100% agreement within all sample duplicates.

8.2.3 **Data Analysis**

The cumulative association of the seven polymorphisms with elite endurance athlete status was assessed using a total genotype score (TGS). The favourable genotype and allele for elite endurance athlete status was identified using data from the preceding results chapters of this thesis and where available the published literature. These are detailed in Table 26.

Table 23 Rationale and genotype score allocation for each polymorphism in relation to elite endurance athlete status.

GENE AND POLYMORPHISM	GENOTYPE SCORE (2 = OPTIMAL)	RATIONALE	ASSOCIATED PUBLICATIONS (IF AVAILABLE)
***ACE* (RS4341)**	II = 2, ID = 1, DD = 0	The II genotype and I allele are consistently reported to be overrepresented in elite endurance athlete cohorts.	Myerson et al., (1999), Scanavini et al., (2002), Alverez et al., (2000), Ma et al., (2013) ACE II genotype associated with performance in endurance athletes
***ACTN3* (RS1815739)**	XX = 2, RX = 1, RR = 0	The XX genotype and X allele are over represented, and the RR genotype and R allele are underrepresented in in elite endurance athlete cohort. The mouse and human models suggest enhanced endurance performance with the XX genotype and X allele	Grealy et al., (2012), Seto et al., (2013) MacArthur et al., (2007), MacArthur et al., (2008),
***AGT* (RS699)**	CC = 2, CT = 1, TT = 0	The CC genotype reported to be associated with increased left ventricular mass which would increase the stroke volume and help to deliver fuel and oxygen to the respiring muscle	Karjalainen et.al (1999) AGT CC genotype and left ventricular mass Diet et al (2001) AGT CC genotype and left ventricular mass when combined with ACE DD genotype
***PPARGC1A* (RS8192678)**	GG = 2, GA = 1, AA = 0	G allele has been associated with elite endurance athletes in comparison to non-athletes controls is several research studies.	Lucia et al (2005), Eynon et al (2009b) Eynon et al (2009c) Ahmetov et al., (2009)
***UCP2* (RS659366)**	TT = 2, TC = 1, CC = 0	T allele associated with delta efficiency with endurance training	Dhamarait et al.,(2012)
***UCP2* (RS660339)**	TT = 2, TC = 1, CC = 0	T allele associated with increased $\dot{V}O_{2\,max}$ in rowers and runners	Ahmetov et al .,(2008) and Ahmetov., (2009)
***UCP3* (RS1800849)**	TT = 2, TC = 1, CC = 0	T allele associated with endurance running performance	Ahmetov., (2008)

8.2.4 **Total genotype score**

The method detailed by Williams and Folland (2008) was used to determine the TGS. The TGS was identified by allocation of a 'genotype score' (GS) to each polymorphism of 0,1, or 2. Genotype score allocation was based on two assumptions firstly that homozygotes of the favourable genotype for the elite endurance phenotype were allocated a GS of 2, heterozygotes a score of 1 and the non- favourable genotype a score of 0. Secondly, that there was co-dominance of allele effect for each polymorphism of interest. The combined effects of each GS and the conversion to a total score that is transformed to a percentage (see equation 1 below for method) allowed the cumulative influence of all seven gene polymorphisms and elite athlete status to be analysed.

Equation 1: calculation of total genotype score (Williams and Folland, 2008)

$$TGS = (100/14) * (GS_{ACE\text{rs}4341} + GS_{ACTN3\text{rs}1815739} + GS_{AGT\text{rs}699} + GS_{PPARGC1A\text{rs}8192678} + GS_{UCP2\text{rs}659366} + GS_{UCP2\text{rs}660339} + GS_{UCP3\text{rs}1800849})$$

A TGS of 100 was representative of an ideal elite endurance athlete polygenic profile, whereas a TGS of 0 demonstrates as the worst hypothesised polygenic profile for elite endurance athlete status.

8.2.5 **Statistical analysis**

Hardy-Weinberg equilibrium was assessed for each of the included polymorphisms using χ^2 test as detail previously in Chapter 3. Where Hardy-Weinberg Equilibrium was observed Pearson's chi squared analysis was conducted to ascertain any associations with endurance athlete status. Where Hardy-Weinberg Equilibrium was not observed Cochrane–Armitage tests were used to confirm any association with elite endurance athlete status. Pearson's correlation was used to determine linear trend association of TGS with personal best time in the male and female marathon cohorts. All statistical

analyses were performed using SPSS version 19 and statistical significance was set at p = 0.05.

8.3 RESULTS

Genotype frequencies for polymorphisms *ACE, AGT, ACTN3, PARGC1A* and *UCP3* were in Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium deviated in *UCP2 rs659366* and *UCP2 rs660339* amongst the whole marathon athlete cohorts. In all *UCP2 rs659366* further analysis Hardy-Weinberg remained. Hardy-Weinberg equilibrium was not observed among the mixed elite marathon athletes in *UCP2 rs660339* though in all further analysis of *UCP2 rs660339* Hardy-Weinberg equilibrium was observed. One of the athletes had a minimal score of 0. This athlete was in the sub elite cohort. None of the elite participants had the minimal (0) or maximal (100) TGS, scores ranged from 7-85 in the elite cohort. In the sub-elite cohort where scores reached 93. When assessed by t- test there was a difference in TGS scores between athletes and controls (t = 4.130 p = 0.000041) there was no recorded difference between elite, sub-elite groups when assessed by Spearman's correlation (r = 0.27, p = 0.614). There was no association of TGS with personal best when assessed by Pearson's correlation (r = 0.030, p = 0.485).

8.4 DISCUSSION

The current chapter aimed to identify an association with the combine polymorphisms within: ACE, AGT, ACTN3, PPARGC1A, UCP2rs659336, UCP2 rs660339, and UCP3 and Endurance athlete status. Athletes TGS scores were different from controls when assessed by independent t-test (t = 4.130 p = 0.000041) with a mean difference of 4.686. The TGS scores ranged from 0 to 93 in the athletes with the highest frequency TGS score of 50 in athletes with a mean of 42 (SD 15.838), whereas in controls the TGS score ranged from 0-86 with the highest frequency was 36 and a mean of 37 (SD 13.047).The association of TGS and endurance athlete status was also reported by Santiago et al.,

2010 who recorded that in rowers the athletes were endowed with a more favourable combined polygenic profile than the controls group. The Santiago study included three of the polymorphisms assessed in this study ACE, ACTN3 and PPARGC1A. A further study by Ahmetov et al (2008) published results in accordance with Santiago, they found that their group of mixed discipline endurance sporting athletes had higher TGS scores for the 10 'endurance' genotypes tested in combination than controls. Gene polymorphisms that overlapped with this study include PPARGC1A, UCP2 and UCP3. An assessment by Spearman's correlation assessed elite athlete status and TGS. This revealed no difference between the elite and sub-elite groups ($r = 0.27$, $p = 0.614$). This finding is in accordance with Santiago et al., (2010) who found no association between athlete groups when stratified into national and world champions. However, Ahmetov et al., (2008) did find a positive correlation with the TGS and elite athlete status though this was in a group of mixed sporting disciplined athletes and so this result should be viewed with caution. A further association between personal best of the combined influence of the polymorphisms within: ACE, AGT, ACTN3, PPARGC1A, UCP2rs659336, UCP2 rs660339, and UCP3 was assessed via Pearson's correlation. No association was apparent between personal best time and TGS score. One athlete in the sub-elite cohort had a TGS score of the minimal 0 the highest TGS score of 93 was also recorded in the sub-elite cohort. The average TGS in the elite cohort was 41 and in the sub-elite cohort 42 was the average TGS. Seven common polymorphisms have been shown to contribute to the complex genetic profile of a marathon runner when compared to non-athlete controls. Further analysis showed elite endurance athlete status was not correlated with TGS. This data partially replicate the data presented by Ahmetov et al., 2008 and Santiage et al., 2010. Collectively the polymorphisms within: ACE, AGT, ACTN3, PPARGC1A, UCP2rs659336, UCP2 rs660339, and UCP3 are associated with endurance athlete status.

9 GENERAL DISCUSSION

9.1 RATIONALE FOR GENETIC ASSOCIATION STUDIES IN SPORT AND EXERCISE SCIENCE AND ENDURANCE PERFORMANCE

Complex traits such as endurance sporting performance are phenotypes likely affected by both multiple genetic and environmental factors. The individual and combinatory contribution of these factors to human physical performance is considered complex. Nevertheless, careful selection of important and robust phenotypes and evidence-based selection of candidate genes can provide a solid basis on which to base studies of genetics in endurance human performance. Based on the assessments completed in height it is likely that thousands of SNIPS are involved in marathon performance. These genetic factors comprise multiple genes and perhaps even multiple polymorphisms within those genes contributing in an additive effect to complete a polygenic profile (Bouchard et al., 1997, Williams and Folland, 2008, Ahmetov and Fedotovskaya, 2012).

Historically, the study of monozygotic twins (who share an almost identical genomic profile) shows a higher correlation in certain phenotypes than dizygotic twin pairs (whose genomic profiles are non-identical), thus demonstrating the heritability, or genetic component in determining that phenotypic trait (Bouchard et al., 1986a, Bouchard et al., 1986b, De Moor et al., 2007).

9.2 INVESTIGATING CANDIDATE GENES

The initial approach used by sports scientists and geneticists to identify sporting genotype-phenotype associations was the 'candidate gene approach' (Bouchard et al., 1997). Using existing physiological knowledge genetic variations, relevant to human physical performance are selected for investigation. An example of this approach in humans is a rare mutation in the myostatin (*MSTN*) gene discovered in a 4-year old German boy with greater muscle mass who was considerably stronger than others his age (Schuelke et al., 2004). Further genetic variations that have been identified as relevant to endurance performance include *ACE* and *ACTN3*. This approach to candidate gene selection makes understanding and extrapolating the polygenic nature of human physical performance rather difficult as gene polymorphisms are often investigated individually rather than in combination. In an attempt to fully characterise the polygenic nature of human endurance performance, increasingly sports scientists are favouring another method the Genome Wide Association Study (GWAS), though this is fraught with its own limitations. By the very nature of being 'elite' few athletes at this level of sporting performance are available for study. Due to the large number of genetic variants being analysed in GWAS, very large numbers of participants are required for statistical power meaning study designs of this type are difficult to achieve without international consortium between research groups. Due to the scope of this project the candidate gene approach was selected as an appropriate method of analysis. Of note when selecting a test population was the importance of homogeneity to maximize the genetic contribution to the elite endurance phenotype of interest. Each gene included had a physiological rationale determined from the literature for the proposed association between the candidate gene and elite endurance phenotype and research participants met predetermined inclusion criteria to ensure validity, reliability and reproducibility of any genetic associations. This genetic candidate gene study compared an endurance athlete populations with non-athlete controls in an attempt to associate common genotypic variants with the elite endurance phenotype of a marathon runner.

9.3 AIMS OF THESIS

This thesis aimed to (1) compare the selected genetic polymorphism of elite and sub-elite and non –athletes; (2) comparatively analyse personal best marathon completion times with selected genetic polymorphisms. The objective was to address aims (1) and (2) for specific polymorphisms in the following genes *ACE, AGT, ACTN3, PPARGC1A, UCP2* and *UCP3*.

9.4 MAIN FINDINGS AND IMPLICATIONS

9.4.1 Summary of Main Findings

AGT rs699 analysis revealed over representation of the TT genotype (5.6%, OR = 0.777, 95% CI 0.562 – 1.074. p = 0.126) and T allele of 4.5% in marathon runner's vs non-athletes (OR = 0.832, 95% CI 0.669 – 1.034 p = 0.097. The TT genotype (OR = 0.722, 95% CI 0.498 – 1.048. p = 0.086) and T allele showed further over representation of 5.6% and 5.7% respectively when sub-elite athletes were compared to non-athletes (genotype: OR = 0.722, 95% CI 0.498 – 1.048. p = 0.086: allele: OR = 0.790, 95% CI 0.613 – 1.019. p = 0.069) (See Table 27).

The A allele of *PPARGC1A* rs8192678 tended to be more frequent in athletes than non-athletes ($\chi^2 = 2.988$, p = 0.084) (Table 27). The male elite cohort when compared to non-athletes showed a 3.4% AA genotype ($\chi^2 = 6.890$, p = 0.04) over representation (OR = 0.696, 95% CI 0.383 – 1.265, p = 0.235). On comparison of male elite athletes and non-athletes the A allele ($\chi^2 = 2.986$, p = 0.084) was 9.2% more frequent (OR = 0.686, 95% CI 0.476-0.987, p = 0.042) (Table 28). Female sub-elite athletes when compared to non-athletes showed 7.8% over representation of the AA genotype ($\chi^2 = 7.193$, p = 0.04) (OR = 0.531, 95% CI 0.212 – 1.331, p = 0.177. A female elite vs all sub-elite athletes comparison showed the A allele tended to be more frequent in female sub-elite athletes ($\chi^2 = 5.425$, p = 0.066) (Table 29). On consideration of PB, in women the *PPARGC1A* GG

genotypes ran the marathon approximately 5 min 38 s faster than other genotypes (p = 0.022). *UCP2* rs660339 analysis revealed a CT genotype difference was recorded when male elite and sub-elite athletes were compared to non-athletes independently (elite χ^2 = 11.173, p = 0.001 OR 1.130, CI 95% 0.525-2.432 p = 0.754; sub-elite χ^2 = 17.584, p = 0.01: OR 0.899, CI 95% 0.495-1.632 p = 0.726) via Pearson's-Chi squared.

In the female athletes, a CT genotype association was observed when compared to non-athletes (genotype χ^2=8.376, p = 0.02: OR 1.202, CI 95% 0.568-2.546 p = 0.630). The female elite athletes also reflected a CT genotype association when compared to non-athletes (genotype χ^2 = 8.942, p = 0.02: OR 1.450, CI 95% 0.598-3.519 p = 0.411).

Table 24 Genotyping summary of findings for all marathon athletes

Genotype	***ACE***	***AGT***	***ACTN3***	***PPARGC1A***	***UCP2 rs659336***	***UCP2 rs660339***	***UCP3***
Marathon Athlete vs Non-Athletes	No association	**TT in Athletes**	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Sub-Elite Marathon Athlete vs Non-Athletes	No association	**TT in Sub-Elite**	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Sub-Elite Marathon Athlete	No association	No association	No association	No association	No association	No association	No association
Allele	**ACE**	**AGT**	**ACTN3**	**PPARGC1A**	**UCP2 rs659336**	**UCP2 rs660339**	**UCP3**
Marathon Athlete vs Non-Athletes	No association	**T in Athletes**	No association	**A in Athletes**	No association	No association	No association
Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Sub-Elite Marathon Athlete vs Non-Athletes	No association	**T in Sub-Elite**	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Sub-Elite Marathon Athlete	No association	No association	No association	No association	No association	No association	No association

Table 25 Genotyping summary of findings for male marathon athletes

Genotype	***ACE***	***AGT***	***ACTN3***	***PPARGC1A***	***UCP2 rs659336***	***UCP2 rs660339***	***UCP3***
Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	**AA in Elite**	No association	**CT in Elite**	No association
Sub-Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	**CT in Sub-Elite**	No association
Elite Marathon Athlete vs Sub-Elite Marathon Athlete	No association	No association	No association	No association	No association	No association	No association
Allele	***ACE***	***AGT***	***ACTN3***	***PPARGC1A***	***UCP2 rs659336***	***UCP2 rs660339***	***UCP3***
Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	**A in Elite**	No association	No association	No association
Sub-Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Sub-Elite Marathon Athlete	No association	No association	No association	No association	No association	No association	No association

Table 26 Genotyping summary of findings for female marathon athletes

Genotype	***ACE***	***AGT***	***ACTN3***	***PPARGC1A***	***UCP2 rs659336***	***UCP2 rs660339***	***UCP3***
Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	**CT in Athletes**	No association
Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	**CT in Elite**	No association
Sub-Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	**AA in Sub-Elite**	No association	No association	No association
Elite Marathon Athlete vs Sub-Elite Marathon Athlete	No association	No association	No association	No association	No association	No association	No association
Allele	***ACE***	***AGT***	***ACTN3***	***PPARGC1A***	***UCP2 rs659336***	***UCP2 rs660339***	***UCP3***
Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Sub-Elite Marathon Athlete vs Non-Athletes	No association	No association	No association	No association	No association	No association	No association
Elite Marathon Athlete vs Sub-Elite Marathon Athlete	No association	No association	No association	No association	No association	No association	No association

9.4.2 **Novel contributions to the literature**

A complex phenotype is required for endurance performance. Therefore, it is likely that 'elite status' is polygenic. We investigated individually, seven gene polymorphisms (*ACE rs1799752, ACTN3 rs1815739, AGT rs699, PPARGC1A rs8192678, UCP2 rs659336, UCP2 rs660339* and *UCP3 rs1800849*) and elite endurance athlete status in male and female marathon runners. Marathon runners (399 prima facie Caucasians) were stratified by personal best (PB) into elite (men <2 h 30 min, women <3 h) and sub-elite (men <2 h 45 min, women <3 h 15 min) and compared to 676 non-athlete prima facie Caucasian controls. This thesis found three polymorphisms *AGT rs699, PPARGC1A rs8192678 UCP2 rs660339* were positively associated with marathon performance in prima facie Caucasian endurance athletes.

9.5 **EXPLANATION OF VARIATION IN FINDINGS WITH THE PUBLISHED LITERATURE**

One of the challenging tasks in sporting genomics is the recruitment of appropriate athletes to investigate. Often access particularly of elite athletes is hampered by training regimes, competition, sporting management and injury. This invariably introduces sampling bias in the population of study. One way that researchers have tried to mitigate against this is to increase the sample sizes by combining athletes from multiple sporting disciplines, this introduces variation in phenotype and due to the inherent nature of phenotype likely genotype. Therefore this study recruited athletes from one sporting discipline, marathon in an attempt to limit the variability. A possible limitation of this thesis is that the findings cannot automatically be extended to other sporting disciplines, even those that are accepted as "endurance" such as long-distance swimming or cycling, because of differences in muscle recruitment patterns, event distance and duration, biomechanical factors, etc. However, analysis within a single sporting discipline eliminates those factors that are known to differ between sports, thus reducing unwanted 'noise' in the data set.

In combining athletes to increase sample sizes researchers may have included athletes from varying geographic ancestry this may inadvertently introduce the confounding through racial gene skew. Others, like this investigation, have tried to mitigate against this confounding by limiting participant inclusion to one racial group. One limitation of this is confirmation of ancestry is often self-reported rather than determined by a panel of SNPs related to geographic ancestry and therefore relies on recall, this in itself is not perfect and subject to bias though arguably this is still preferential to ensure to limit heterogeneity within the sample selected.

The control samples in this project were of a similar age this is due to the recruitment process. Most of the controls recruited were students of MMU and thus by demographic of being at university were of a similar age. The athlete cohort varied much more in age. It is recorded that the distance an athlete competes at often increases with age (www.powerof10.org). With marathon being the longest distance to run it falls that the participants should be older. Age may be a contributory factor to the advantage of an SNIP to sporting performance. It could be postulated that increased gene expression at one age may preferentially benefit an athlete on their sporting performance over another athlete of a lesser or greater age. It is known that to regulome changes epigenetically with age (Boron and Boulpaep, 2012) However, the fixed nature of the genome (i.e. genotype / allele frequency) from birth to death should not be influenced and therefore is a stable measure against sporting performance over say gene expression.

Another point of note, is the researcher determination of ‘elite’ in their athlete cohort Researchers across the field have varying definitions for ‘elite’ making comparisons between cohorts challenging. Authors have used achievement markers such as Olympic and world records, recorded participation in sporting events at national or international level, other have used physiological measurements such as $\dot{V}O_{2\,max}$ that are generally

associated with sporting performance. Each definition has its limitations and these were discussed recently in Swann et al. (2015) and in a BASES article (http://www.bases.org.uk/write/ARTICLE_P6.pdf). This thesis defined elite according to top 100 UK male and female marathon athlete rankings as defined by the McCain Power of 10 website (*www.thepowerof10.info/)* as accessed on 31ST December 2012. The website records athlete particulars and sporting achievements on an annual basis. The published data records for athlete rankings date back to 2006. The data were averaged and the resulting mean time was rounded to the nearest 10 mins. One limitation of this thesis could be considered to be the cut-off threshold times for inclusion: 2 h 30 for men and 3 h for women for "elite". Arguably, to be considered truly elite, only the top 25 should be included, or top 1%, or <2 h 20 for men, etc. However, these criteria are all equally arbitrary and may significantly influence sample size and therefore statistical power would need to be considered.

The explicit detailing of a runners $\dot{V}O_{2\,max}$, $\dot{V}O_2$ Kinetics, Running Economy skeletal muscle fibre type proportions lactate thresholds at given workloads etc. could help in identifying a more homogenous cohort of athletes. However, this will be hindered by natural variation within athlete groups and though it may be possible to group athletes more closely based on their phenotypic characteristics there will still be small amounts of variation within the groups due to the complex nature of human performance and the nature of competition. One way to try to reduce the variation noted between athlete and controls is to pair match them for characteristics such as age and ethnicity that may influence the results of genetic association.

The sample size in this thesis is larger than any other elite marathon cohort and second to another elite cohort from a single endurance event (Grealy et al., 2012). A relationship between sample size and statistical power exists where, as the effect size reduces the number of participants required to identify a statistically significant difference increases

exponentially. So there remain additional benefits to be gained, in terms of statistical power, by increasing the sample size in this thesis from its current ~400 yet further (Figure 34). It should be noted, however, that the law of diminishing returns applies, such that as the sample size increases the relative benefit obtained decreases.

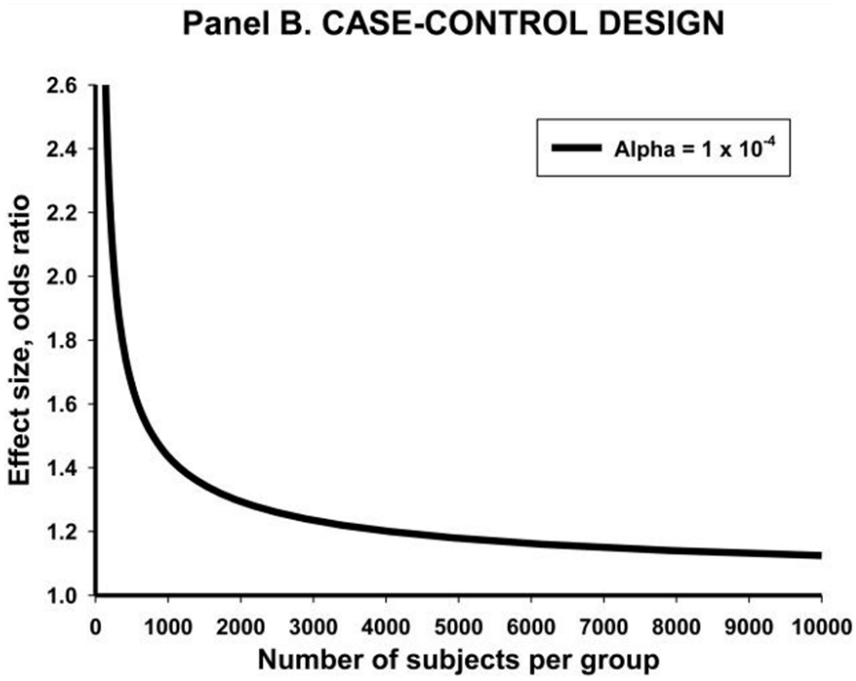

Figure 34 Number of subjects needed in each group for a given effect size (measured as an odds ratio [OR]) for a case-control design assuming 80% statistical power, a minor allele frequency of 20%, an additive model, and an alpha level of 0.0001 (500 SNPs) (Bouchard, 2011)

One possible explanation for the lack replication with data in the published literature in some of the polymorphisms tested in this thesis is due to small sample sizes in the published data there may be the possibility of a type I statistical error (false positive) - i.e. low confidence in the observed data. A type 1 error is a major uncertainty with genetic association studies particularly in those with small sample sizes where the statistical evidence is lacking due to reduced statistical power. Therefore, the expansion of the genetic investigations in a larger data set may have reduced the possibility of inadvertently reporting false positive associations. Another possible explanation for the lack of reproducibility in some of the published studies included in this thesis is publication bias. Publication bias occurs when the outcome of an experiment or research study

influences the decision to publish or not. Publication bias is an important factor particularly in elite athlete populations where sample sizes are usually small and it is difficult to determine true polymorphic associations due to the complexities of the phenotypes being studied as publishing only positive significant findings disrupts the true nature of the research field. Publication bias can make finding support for a hypothesis during a literature review challenging as positive results are far more likely to be published than those showing null results. For example, if a lab/research group assesses e.g. 10 SNPs in an athlete cohort vs non-athlete controls, and sees 1 statistically significant difference, they might be tempted to try to publish that 'more interesting' result and discount the other nine. Hence the literature could become over-populated with positive associations that is not reflective of the data the various labs have collected. Even if the lab tries to publish all 10 SNPs as separate papers, reviewers and editors are likely to be less inclined to recommend acceptance for publication the 'less interesting' 'no association' papers – i.e. another reason why the literature could become over-populated with positive associations relative to the data collected. While reviewing the literature, it was noted that results showing no association between genotype and athlete status are frequently reported as part of a broader paper that does include at least some positive association – which further adds to the notion that results showing no association are less likely to be submitted or accepted for publication on their own merits. Having acknowledged the possibility of publication bias and false positives (type I errors) and thus the resultant over-representation of positive associations in the literature there is a strong likelihood that variation will occur between the published results and those chapters in this thesis that report no association.

9.6 DIRECTIONS FOR FUTURE RESEARCH

The findings of this thesis confirms the complexity of the elite endurance athlete phenotype and likely polygenic nature of the associated contributory genetic factors. This study was limited to 7 genetic polymorphisms that were selected based on their previously reported associations, in the literature, with parameters thought to be favourable to elite

performance in an endurance sport such as marathon running. To assess the combined effect of these genetic polymorphisms with other likely candidate polymorphism to emerge in future a TGS-type analyses could be applied to groups of selected candidate gene data to simultaneously consider associations between several polymorphisms and athlete phenotypes. Those kinds of analyses could also be conducted on the output from GWAS. However, the requirements of very large participant cohorts, standardised phenotypes and sophisticated (and considerably more expensive) laboratory genomic analyses means that GWAS is not feasible at this time. Indeed Rankinen et al (2016) is the largest effort to date to identify common polymorphisms associated with endurance performance in an unbiased manner. Rankinen et al (2016) aimed to identify a panel of genetic variants responsible for the elite endurance athlete profile using GWAS. They reported that the GAMES international consortium was at this stage underpowered to identify genetic variants with small effect sizes as none of the p-values approached the 5×10^{-8} required for statistical power. They further offered, in interim solution, that some suggestive allelic traits resulting from the GWAS should be further explored in larger comparative analysis of internationally elite endurance athletes and sedentary controls it is suggested here that TGS may be a suitable method to do this.

10 REFERENCES

AHMETOV, II, POPOV, D. V., MOZHAISKAIA, I. A., MISSINA, S. S., ASTRATENKOVA, I. V., VINOGRADOVA, O. L. & ROGOZKIN, V. A. 2007. Association of regulatory genes polymorphisms with aerobic and anaerobic performance of athletes. *Ross Fiziol Zh Im I M Sechenova,* 93**,** 837-43.

AHMETOV, I. I., DRUZHEVSKAYA, A. M., ASTRATENKOVA, I. V., POPOV, D. V., VINOGRADOVA, O. L. & ROGOZKIN, V. A. 2010. The ACTN3 R577X polymorphism in Russian endurance athletes. *British Journal of Sports Medicine,* 44**,** 649-652.

AHMETOV, I. I., DRUZHEVSKAYA, A. M., LYUBAEVA, E. V., POPOV, D. V., VINOGRADOVA, O. L. & WILLIAMS, A. G. 2011. The dependence of preferred competitive racing distance on muscle fibre type composition and ACTN3 genotype in speed skaters. *Experimental Physiology,* 96**,** 1302-1310.

AHMETOV, I. I. & FEDOTOVSKAYA, O. N. 2012. Sports genomics: Current state of knowledge and future directions. *Cellular and Molecular Exercise Physiology,* 1.

AHMETOV, I. I., POPOV, D. V., ASTRATENKOVA, I. V., DRUZHEVSKAYA, A. M., MISSINA, S. S., VINOGRADOVA, O. L. & ROGOZKIN, V. A. 2008. The use of molecular genetic methods for prognosis of aerobic and anaerobic performance in athletes. *Human Physiology,* 34**,** 338-342.

AHMETOV, I. I. & ROGOZKIN, V. A. 2009. Genes, athlete status and training–An overview. *Genetics and Sport,* 54**,** 43-71.

AHMETOV, I. I., WILLIAMS, A. G., POPOV, D. V., LYUBAEVA, E. V., HAKIMULLINA, A. M., FEDOTOVSKAYA, O. N., MOZHAYSKAYA, I. A., VINOGRADOVA, O. L., ASTRATENKOVA, I. V., MONTGOMERY, H. E. & ROGOZKIN, V. A. 2009. The combined impact of metabolic gene polymorphisms on elite endurance athlete status and related phenotypes. *Human Genetics,* 126**,** 751-61.

AKIMOTO, T., POHNERT, S. C., LI, P., ZHANG, M., GUMBS, C., ROSENBERG, P. B., WILLIAMS, R. S. & YAN, Z. 2005. Exercise stimulates Pgc-1α transcription in skeletal muscle through activation of the p38 MAPK pathway. *Journal of Biological Chemistry,* 280**,** 19587-19593.

ALVAREZ, R., TERRADOS, N., ORTOLANO, R., IGLESIAS-CUBERO, G., REGUERO, J. R., BATALLA, A., CORTINA, A., FERNANDEZ-GARCIA, B., RODRIGUEZ, C., BRAGA, S., ALVAREZ, V. & COTO, E. 2000. Genetic variation in the renin angiotensin system and athletics performance. *European Journal of Applied Physiology,* 82**,** 117-120.

AMIR, O., AMIR, R., YAMIN, C., ATTIAS, E., EYNON, N., SAGIV, M., SAGIV, M. & MECKEL, Y. 2007. The ACE deletion allele is associated with Israeli elite endurance athletes. *Experimental Physiology,* 92**,** 881-886.

ARANY, Z., HE, H., LIN, J., HOYER, K., HANDSCHIN, C., TOKA, O., AHMAD, F., MATSUI, T., CHIN, S. & WU, P.-H. 2005. Transcriptional coactivator PGC-1α controls the energy state and contractile function of cardiac muscle. *Cell metabolism,* 1**,** 259-271.

ASH, G. I., SCOTT, R. A., DEASON, M., DAWSON, T. A., WOLDE, B., BEKELE, Z., TEKA, S. & PITSILADIS, Y. P. 2011. No association between ACE gene variation and endurance athlete status in Ethiopians. *Medicine & Science in Sports & Exercise,* 43**,** 590-7.

ASTRAND, P.-O. & RODAHL, K. 1986. *Textbook of Work Physiology: Physiological Bases of Exercise,* Singapore, McGraw-Hill.

ASTRUP, A., TOUBRO, S., DALGAARD, L. T., URHAMMER, S. A., SØRENSEN, T. I. A. & PEDERSEN, O. 1999. Impact of the v/v 55 polymorphism of the uncoupling protein 2 gene on 24-h energy expenditure and substrate oxidation. *International Journal Of Obesity,* 23**,** 1030.

BAAR, K. 2004. Involvement of PPARγ co-activator-1, nuclear respiratory factors 1 and 2, and PPARα in the adaptive response to endurance exercise. *Proceedings of the Nutrition Society,* 63**,** 269-273.

BAAR, K., WENDE, A. R., JONES, T. E., MARISON, M., NOLTE, L. A., CHEN, M. A. Y., KELLY, D. P. & HOLLOSZY, J. O. 2002. Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1. *The FASEB Journal,* 16**,** 1879-1886.

BAE, J., KANG, B., LEE, K. & LEE, S. 2007. Genetic variation in the renin-angiotensin system and response to endurance training. *Medical Principles and Practice,* 16**,** 142-146.

BARSTOW, T. J., JONES, A. M., NGUYEN, P. H. & CASABURI, R. 1996. Influence of muscle fiber type and pedal frequency on oxygen uptake kinetics of heavy exercise. *J Appl Physiol (1985),* 81**,** 1642-50.

BASSETT, D. & HOWLEY, E. T. 2000. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Medicine & Science in Sports & Exercise***,** 70-84.

BEN-ZAKEN, S., MECKEL, Y., NEMET, D. & ELIAKIM, A. 2013. Genetic score of power-speed and endurance track and field athletes. *Scandinavian Journal of Medicine & Science in Sports*.

BERMAN, Y. & NORTH, K. N. 2010. A Gene for Speed: The Emerging Role of α-Actinin-3 in Muscle Metabolism. *Physiology,* 25**,** 250-259.

BLANCHARD, A., OHANIAN, V. & CRITCHLEY, D. 1989. The structure and function of α-actinin. *Journal of Muscle Research and Cell Motility,* 10**,** 280-289.

BORON, W. F. & BOULPAEP, E. L. 2012. *Medical Physiology,* Philadelphia, Saunders.

BOSS, O., SAMEC, S., PAOLONI-GIACOBINO, A., ROSSIER, C., DULLOO, A., SEYDOUX, J., MUZZIN, P. & GIACOBINO, J.-P. 1997. Uncoupling protein-3: a new member of the mitochondrial carrier family with tissue-specific expression. *FEBS letters,* 408**,** 39-42.

BOTSTEIN, D. & RISCH, N. 2003. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics,* 33**,** 228-237.

BOUCHARD, C., AN, P., RICE, T., SKINNER, J. S., WILMORE, J. H., GAGNON, J., PÉRUSSE, L., LEON, A. S. & RAO, D. C. 1999. Familial aggregation of $\dot{V}O_{2max}$ response to exercise training: results from the HERITAGE Family Study. *Journal of Applied Physiology,* 87**,** 1003-1008.

BOUCHARD, C., DAW, E. W., RICE, T., PÉRUSSE, L., GAGNON, J., PROVINCE, M. A., LEON, A. S., RAO, D. C., SKINNER, J. S. & WILMORE, J. H. 1998. Familial resemblance for $\dot{V}O_{2max}$ in the sedentary state: the HERITAGE family study. *Medicine and Science in Sports and Exercise,* 30**,** 252-258.

BOUCHARD, C., LEON, A. S., RAO, D. C., SKINNER, J. S., WILMORE, J. H. & GAGNON, J. 1995. Aims, design, and measurement protocol. *Medicine and Science in Sports and Exercise,* 27**,** 721-729.

BOUCHARD, C., LESAGE, R., LORTIE, G., SIMONEAU, J.-A., HAMEL, P., BOULAY, M. R., PERUSSE, L., THERIAULT, G. & LEBLANC, C. 1986a. Aerobic performance in brothers, dizygotic and monozygotic twins. *Med Sci Sports Exerc,* 18**,** 639-46.

BOUCHARD, C., MALINA, R. M. & PERUSSE, L. 1997. *Genetics of fitness and physical performance*, Human Kinetics.

BOUCHARD, C., RANKINEN, T., CHAGNON, Y. C., RICE, T., PÉRUSSE, L., GAGNON, J., BORECKI, I. B., AN, P., LEON, A. S., SKINNER, J. S., WILMORE, J. H., PROVINCE, M. & RAO, D. C. 2000. Genomic scan for maximal oxygen uptake and its response to training in the HERITAGE Family Study*. *Journal of Applied Physiology,* 88**,** 551-559.

BOUCHARD, C., SIMONEAU, J.-A., LORTIE, G., BOULAY, M., MARCOTTE, M. & THIBAULT, M. 1986b. Genetic effects in human skeletal muscle fiber type distribution and enzyme activities. *Canadian Journal of Physiology and Pharmacology,* 64**,** 1245-1251.

BOUCHARD, C., WOLFARTH, B., RIVERA, M. A., GAGNON, J. & SIMONEAU, J.-A. 2008. Genetic Determinants of Endurance Performance. *Endurance in Sport.* Blackwell Science Ltd.

BOYD, A. E., 3RD, GIAMBER, S. R., MAGER, M. & LEBOVITZ, H. E. 1974. Lactate inhibition of lipolysis in exercising man. *Metabolism,* 23**,** 531-42.

BRAND, M. D. & ESTEVES, T. C. 2005. Physiological functions of the mitochondrial uncoupling proteins UCP2 and UCP3. *Cell Metabolism,* 2**,** 85-93.

BRAND, M. D., PAMPLONA, R., PORTERO-OTIN, M., REQUENA, J. R., ROEBUCK, S. J., BUCKINGHAM, J. A., CLAPHAM, J. C. & CADENAS, S. 2002. Oxidative damage and phospholipid fatty acyl composition in skeletal muscle mitochondria from mice underexpressing or overexpressing uncoupling protein 3. *Biochemical Journal,* 368**,** 597-603.

BRAY, M. S., HAGBERG, J. M., PÉRUSSE, L., RANKINEN, T., ROTH, S. M., WOLFARTH, B. & BOUCHARD, C. 2009. The human gene map for performance and health-related fitness phenotypes: the 2006-2007 update. *Medicine and Science in Sports and Exercise,* 41**,** 35-73.

BROOKE, G. A., FAHEY, T. D. & BALDWIN, K. M. 2005. *Exercise Physiology Human Bioenergetics and Its Applications*, McGraw Hill.

BRUTON, A. 2002. Muscle Plasticity Response to training and detraining. *Physiotherapy,* 88**,** 398-408.

BUEMANN, B., SCHIERNING, B., TOUBRO, S., BIBBY, B., SØRENSEN, T., DALGAARD, L., PEDERSEN, O. & ASTRUP, A. 2001. The association between the val/ala-55 polymorphism of the uncoupling protein 2 gene and exercise efficiency. *International Journal of Obesity & Related Metabolic Disorders,* 25.

BURKE, L. M. 2007. Nutrition strategies for the marathon : fuel for training and racing. *Sports Med,* 37**,** 344-7.

CADENAS, S., ECHTAY, K. S., HARPER, J. A., JEKABSONS, M. B., BUCKINGHAM, J. A., GRAU, E., ABUIN, A., CHAPMAN, H., CLAPHAM, J. C. & BRAND, M. D. 2002. The Basal Proton Conductance of Skeletal Muscle Mitochondria from Transgenic Mice Overexpressing or Lacking Uncoupling Protein-3. *Journal of Biological Chemistry,* 277**,** 2773-2778.

CHAGNON, Y. C., ALLARD, C. & BOUCHARD, C. 1984. Red blood cell genetic variation in Olympic endurance athletes. *Journal of Sports Sciences,* 2**,** 121-129.

CIĘSZCZYK, P., KRUPECKI, K., MACIEJEWSKA, A. & SAWCZUK, M. 2009. The angiotensin converting enzyme gene I/D polymorphism in Polish rowers. *International Journal of Sports Medicine,* 30**,** 624-627.

CLARKSON, P. M., DEVANEY, J. M., GORDISH-DRESSMAN, H., THOMPSON, P. D., HUBAL, M. J., URSO, M., PRICE, T. B., ANGELOPOULOS, T. J., GORDON, P. M., MOYNA, N. M., PESCATELLO, L. S., VISICH, P. S., ZOELLER, R. F., SEIP, R. L. & HOFFMAN, E. P. 2005. ACTN3 genotype is associated with increases in muscle strength in response to resistance training in women. *Journal of Applied Physiology,* 99**,** 154-163.

COLLINS, M., XENOPHONTOS, S., CARIOLOU, M., MOKONE, G., HUDSON, D., ANASTASIADES, L. & NOAKES, T. 2004. The ACE gene and endurance performance during the South African Ironman Triathlons. *Medicine and Science in Sports and Exercise,* 36**,** 1314-1320.

CONLEY, D. L. & KRAHENBUHL, G. S. 1980. Running economy and distance running performance of highly trained athletes. *Medicine and Science in Sports and Exercise,* 12**,** 357-60.

CORTRIGHT, R. N., ZHENG, D., JONES, J. P., FLUCKEY, J. D., DI CARLO, S. E., GRUJIC, D., LOWELL, B. B. & DOHM, G. L. 1999. Regulation of skeletal muscle UCP-2 AND UCP-3 gene expression by exercise and denervation. *American Journal of Physiology - Endocrinology and Metabolism,* 276**,** E217-E221.

COSTILL, D. L. 1970. *Metabolic responses during distance running*.

COSTILL, D. L., BOWERS, R., BRANAM, G. & SPARKS, K. 1971a. Muscle glycogen utilization during prolonged exercise on successive days. *J Appl Physiol,* 31**,** 834-8.

COSTILL, D. L., BRANAM, G., EDDY, D. & SPARKS, K. 1971b. Determinants of marathon running success. *Internationale Zeitschrift für angewandte Physiologie einschließlich Arbeitsphysiologie,* 29**,** 249-254.

COUPLAN, E., DEL MAR GONZALEZ-BARROSO, M., ALVES-GUERRA, M. C., RICQUIER, D., GOUBERN, M. & BOUILLAUD, F. 2002. No Evidence for a Basal, Retinoic, or Superoxide-induced Uncoupling Activity of the Uncoupling Protein 2 Present in Spleen or Lung Mitochondria. *Journal of Biological Chemistry,* 277**,** 26268-26275.

COUTURE, L., CHAGNON, M., ALLARD, C. & BOUCHARD, C. 1986. More on red blood cell genetic variation in Olympic athletes. *Canadian journal of applied sport sciences. Journal canadien des sciences appliquees au sport,* 11**,** 16-18.

COYLE, E. F. 2007. Physiological Regulation of Marathon Performance. *Sports Medicine,* 37**,** 306-311.

COYLE, E. F., SIDOSSIS, L. S., HOROWITZ, J. F. & BELTZ, J. D. 1992. Cycling efficiency is related to the percentage of type I muscle fibers. *Medicine and Science in Sports and Exercise,* 24**,** 782-8.

DANSER, A., SCHALEKAMP, M., BAX, W., VAN DEN BRINK, A., SAXENA, P., RIEGGER, G. & SCHUNKERT, H. 1995. Angiotensin-Converting Enzyme in the Human Heart Effect of the Deletion/Insertion Polymorphism. *Circulation,* 92**,** 1387-1388.

DAY, S. H., GOHLKE, P., DHAMRAIT, S. S. & WILLIAMS, A. G. 2007. No correlation between circulating ACE activity and $\dot{V}O_{2max}$ or mechanical efficiency in women. *European Journal of Applied Physiology,* 99**,** 11-18.

DE GARAY, A. L., LEVINE, L. & CARTER, J. E. L. 1974. *Genetic and anthropological studies of Olympic athletes*, Academic Press.

DE MOOR, M., SPECTOR, T., CHERKAS, L., FALCHI, M., HOTTENGA, J., BOOMSMA, D. & DE GEUS, E. 2007. Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. *Twin Research and Human Genetics,* 10**,** 812-820.

DENNY, M. W. 2008. Limits to running speed in dogs, horses and humans. *J Exp Biol,* 211**,** 3836-49.

DHAMRAIT, S. S., WILLIAMS, A. G., DAY, S. H., SKIPWORTH, J., PAYNE, J. R., WORLD, M., HUMPHRIES, S. E. & MONTGOMERY, H. E. 2012. Variation in the uncoupling protein 2 and 3 genes and human performance. *Journal of Applied Physiology,* 112**,** 1122-1127.

DIET, F., GRAF, C., MAHNKE, N., WASSMER, G., PREDEL, H. G., PALMA-HOHMANN, I., ROST, R. & BÖHM, M. 2001. ACE and angiotensinogen gene genotypes and left ventricular mass in athletes. *European Journal of Clinical Investigation,* 31**,** 836-842.

DÖRING, F. E., ONUR, S., GEISEN, U., BOULAY, M. R., PÉRUSSE, L., RANKINEN, T., RAURAMAA, R., WOLFAHRT, B. & BOUCHARD, C. 2010. ACTN3 R577X and other polymorphisms are not associated with elite endurance athlete status in the Genathlete study. *Journal of Sports Sciences,* 28**,** 1355-9.

DRUZHEVSKAYA, A. M., AHMETOV, I. I., ASTRATENKOVA, I. V. & ROGOZKIN, V. A. 2008. Association of the ACTN3 R577X polymorphism with power athlete status in Russians. *European Journal of Applied Physiology,* 103**,** 631-634.

DZAU, V. 1988. Circulating versus local renin-angiotensin system in cardiovascular homeostasis. *Circulation,* 77**,** I4-13.

DZAU, V. 1989. Multiple pathways of angiotensin production in the blood vessel wall: evidence, possibilities and hypotheses. *Journal of Hypertension,* 7**,** 933-936.

ECHTAY, K. S., ROUSSEL, D., ST-PIERRE, J., JEKABSONS, M. B., CADENAS, S., STUART, J. A., HARPER, J. A., ROEBUCK, S. J., MORRISON, A., PICKERING, S., CLAPHAM, J. C. & BRAND, M. D. 2002. Superoxide activates mitochondrial uncoupling proteins. *Nature,* 415**,** 96.

EHLI, E. A., LENGYEL-NELSON, T., HUDZIAK, J. J. & DAVIES, G. E. 2008. Using a commercially available DNA extraction kit to obtain high quality human genomic DNA suitable for PCR and genotyping from 11-year-old saliva saturated cotton spit wads. *BMC Res Notes,* 1**,** 133.

ELENI, S., DIMITRIOS, K., VAYA, P., ARETI, M., NORMA, V. & MAGDALINI, G. 2008. Angiotensin-I converting enzyme gene and I/D polymorphism distribution in the Greek population and a comparison with other European populations. *Journal of Genetics,* 87**,** 91.

ERDÖS, E. & SKIDGEL, R. 1987. The angiotensin I-converting enzyme. *Laboratory Investigation; A Journal of Technical Methods and Pathology,* 56**,** 345.

ESTERBAUER, H., OBERKOFLER, H., KREMPLER, F. & PATSCH, W. 1999. Human peroxisome proliferator activated receptor gamma coactivator 1 (PPARGC1) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression. *Genomics,* 62**,** 98-102.

EYNON, N., DURATE, J., OLIVEIRA, J., SAGIV, M., YAMIN, C., MECKEL, Y. & GOLDHAMMER, E. 2009a. ACTN3 R577X Polymorphism and Israeli Top-level Athletes *International Journal of Sports Medicine,* 30**,** 839.

EYNON, N., MECKEL, Y., ALVES, A. J., YAMIN, C., SAGIV, M., GOLDHAMMER, E. & SAGIV, M. 2009b. Is there an interaction between PPARD T294C and PPARGC1A Gly482Ser polymorphisms and human endurance performance? *Experimental Physiology,* 94**,** 1147-1152.

EYNON, N., MECKEL, Y., SAGIV, M., YAMIN, C., AMIR, R., GOLDHAMMER, E., DUARTE, J. A. & OLIVEIRA, J. 2009c. Do PPARGC1A and PPARα polymorphisms influence sprint or endurance phenotypes? *Scandinavian Journal of Medicine & Science in Sports,* 20**,** e145-e150.

EYNON, N., RUIZ, J. R., FEMIA, P., PUSHKAREV, V. P., CIĘSZCZYK, P., MACIEJEWSKA-KARLOWSKA, A., SAWCZUK, M., DYATLOV, D. A., LEKONTSEV, E. V., KULIKOV, L. M., BIRK, R., BISHOP, D. J. & LUCIA, A. 2012. The ACTN3 R577X Polymorphism across Three Groups of Elite Male European Athletes. *PLoS ONE,* 7**,** 1-7.

FARRELL, P. A., WILMORE, J. H., COYLE, E. F., BILLING, J. E. & COSTILL, D. L. 1979. Plasma lactate accumulation and distance running performance. *Medicine & Science in Sports & Exercise,* 11**,** 338-344.

FINCK, B. N. & KELLY, D. P. 2006. PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. *The Journal of Clinical Investigation,* 116**,** 615-622.

FLEURY, C., NEVEROVA, M., COLLINS, S., RAIMBAULT, S., CHAMPIGNY, O., LEVI-MEYRUEIS, C., BOUILLAUD, F., SELDIN, M. F., SURWIT, R. S., RICQUIER, D. & WARDEN, C. H. 1997. Uncoupling protein-2: a novel gene linked to obesity and hyperinsulinemia. *Nature Genetics,* 15**,** 269.

FOLLAND, J. P. & WILLIAMS, A. G. 2007. The adaptations to strength training: morphological and neurological contributions to increased strength. *Sports Medicine,* 37**,** 145-168.

FOSTER, C. 1983. $\dot{V}O_{2max}$ and training indices as determinants of competitive running performance. *Journal of Sports Sciences,* 1**,** 13-22.

FRAZER, K., BALLINGER, D., COX, D., HINDS, D., STUVE, L., GIBBS, R., BELMONT, J., BOUDREAU, A., HARDENBOL, P. & LEAL, S. 2007. A second generation human haplotype map of over 3.1 million SNPs. *Nature,* 449**,** 851-861.

GAILLARD, I., CLAUSER, E. & CORVOL, P. 1989. Structure of human angiotensinogen gene. *Dna,* 8**,** 87-99.

GAMBOA, J. L. & ANDRADE, F. H. 2012. Muscle endurance and mitochondrial function after chronic normobaric hypoxia: contrast of respiratory and limb muscles. *Pflugers Arch,* 463**,** 327-38.

GAYAGAY, G., YU, B., HAMBLY, B., BOSTON, T., HAHN, A., CELERMAJER, D. & TRENT, R. 1998. Elite endurance athletes and the ACE I allele–the role of genes in athletic performance. *Human Genetics,* 103**,** 48-50.

GIBBS, R., BELMONT, J., HARDENBOL, P., WILLIS, T., YU, F., YANG, H., CH'ANG, L., HUANG, W., LIU, B. & SHEN, Y. 2003. The International HapMap Project. *Nature,* 426**,** 789-796.

GIMENO, R. E., DEMBSKI, M., WENG, X., DENG, N., SHYJAN, A. W., GIMENO, C. J., IRIS, F., ELLIS, S. J., WOOLF, E. A. & TARTAGLIA, L. A. 1997. Cloning and Characterization of an Uncoupling Protein Homolog: A Potential Molecular Mediator of Human Thermogenesis. *Diabetes,* 46**,** 900-906.

GOLLNICK, P. D., ARMSTRONG, R., SALTIN, B., SAUBERT, C., SEMBROWICH, W. L. & SHEPHERD, R. E. 1973. Effect of training on enzyme activity and fiber composition of human skeletal muscle. *Journal of Applied Physiology,* 34**,** 107-111.

GOLLNICK, P. D., ARMSTRONG, R. B., SAUBERT IV, C. W., PIEHL, K. & SALTIN, B. 1972. Enzyme activity and fiber composition in skeletal muscle of untrained and trained men. *J Appl Physiol,* 33.

GOMEZ-GALLEGO, F., SANTIAGO, C., GONZÁLEZ-FREIRE, M., YVERT, T., MUNIESA, C. A., SERRATOSA, L., ALTMÄE, S., RUIZ, J. R. & LUCIA, A. 2009. The C allele of the AGT Met235Thr polymorphism is associated with power sports performance. *Applied Physiology, Nutrition, and Metabolism,* 34**,** 1108-1111.

GONZÁLEZ-ALONSO, J. 2007. Hyperthermia impairs brain, heart and muscle function in exercising humans. *Sports Medicine,* 37**,** 371-373.

GOTO, M., TERADA, S., KATO, M., KATOH, M., YOKOZEKI, T., TABATA, I. & SHIMOKAWA, T. 2000. cDNA cloning and mRNA analysis of PGC-1 in epitrochlearis muscle in swimming-exercised rats. *Biochemical and biophysical research communications,* 274**,** 350-354.

GREALY, R., HERRUER, J., SMITH, C. L., HILLER, D., HASELER, L. J. & GRIFFITHS, L. R. 2015. Evaluation of a 7-Gene Genetic Profile for Athletic Endurance Phenotype in Ironman Championship Triathletes. *PLoS One,* 10**,** e0145171.

GREALY, R., SMITH, C. L. E., CHEN, T., HILLER, D., HASELER, L. J. & GRIFFITHS, L. R. 2012. The genetics of endurance: Frequency of the ACTN3 R577X variant in Ironman World Championship athletes. *Journal of Science and Medicine in Sport,* 4**,** 365-371.

GULICK, T., CRESCI, S., CAIRA, T., MOORE, D. D. & KELLY, D. P. 1994. The peroxisome proliferator-activated receptor regulates mitochondrial fatty acid oxidative enzyme gene expression. *Proceedings of the National Academy of Sciences,* 91**,** 11012-11016.

HAGAN, R., SMITH, M. & GETTMAN, L. 1981. Marathon performance in relation to maximal aerobic power and training indices. *Medicine and Science in Sports and Exercise,* 13**,** 185-189.

HAGAN, R., UPTON, S., DUNCAN, J. & GETTMAN, L. 1987. Marathon performance in relation to maximal aerobic power and training indices in female distance runners. *British Journal of Sports Medicine,* 21**,** 3-7.

HAGBERG, J. M., FERRELL, R. E., MCCOLE, S. D., WILUND, K. R. & MOORE, G. E. 1998. $\dot{V}O_{2max}$ is associated with ACE genotype in postmenopausal women. *Journal of Applied Physiology,* 85**,** 1842-1846.

HAGBERG, J. M., MCCOLE, S. D., BROWN, M. D., FERRELL, R. E., WILUND, K. R., HUBERTY, A., DOUGLAS, L. W. & MOORE, G. E. 2002. ACE insertion/' deletion polymorphism and submaximal exercise hemodynamics in postmenopausal women. *Journal of Applied Physiology,* 92**,** 1083-1088.

HAGBERG, J. M., MOORE, G. E. & FERRELL, R. E. 2001. Specific genetic markers of endurance performance and $\dot{V}O_{2max}$. *Exercise and sport sciences reviews,* 29**,** 15-19.

HE, Z., HU, Y., FENG, L., BAO, D., WANG, L., LI, Y., WANG, J., LIU, G., XI, Y., WEN, L. & LUCIA, A. 2008. Is there an association between PPARGC1A genotypes and endurance capacity in Chinese men? *Scandinavian Journal of Medicine & Science in Sports,* 18**,** 195-204.

HOLLOSZY, J. O. 1967. Biochemical adaptations in muscle. Effects of exercise on mitochondrial oxygen uptake and respiratory enzyme activity in skeletal muscle. *J Biol Chem,* 242**,** 2278-82.

HRUSKOVICOVA, H., DZURENKOVA, D., SELINGEROVA, M., BOHUS, B., TIMKANICOVA, B. & KOVACS, L. 2006. The angiotensin converting enzyme I/D polymorphism in long distance runners. *Journal of Sports Medicine and Physical Fitness,* 46**,** 509-513.

HUBERT, C., HOUOT, A.-M., CORVOL, P. & SOUBRIER, F. 1991. Structure of the angiotensin I-converting enzyme gene. Two alternate promoters correspond to evolutionary steps of a duplicated gene. *Journal of Biological Chemistry,* 266**,** 15377-15383.

HUDSON, D., MOKONE, G., NOAKES, T. & COLLINS, M. 2004. The-55 C/T polymorphism within the UCP3 gene and performance during the South African Ironman Triathlon. *International Journal of Sports Medicine,* 25**,** 427-432.

JELAKOVIĆ, B., KUZMANIĆ, D., MILICIĆ, D., LAGANOVIĆ, M., RONCEVIĆ, T., SERTIĆ, J. & STAVLJENIĆ-RUKAVINA, A. 2000. Influence of angiotensin converting enzyme (ACE) gene polymorphism and circadian blood pressure (BP) changes on left ventricule (LV) mass in competitive oarsmen. *American Journal of Hypertension,* 13**,** 182A.

JEUNEMAITRE, X., SOUBRIER, F., KOTELEVTSEV, Y. V., LIFTON, R. P., WILLIAMS, C. S., CHARRU, A., HUNT, S. C., HOPKINS, P. N., WILLIAMS, R. R., LALOUEL, J.-M. & CORVOL, P. 1992. Molecular basis of human hypertension: Role of angiotensinogen. *Cell,* 71**,** 169-180.

JIANG, N., ZHANG, G., BO, H., QU, J., MA, G., CAO, D., WEN, L., LIU, S., JI, L. L. & ZHANG, Y. 2009. Upregulation of uncoupling protein-3 in skeletal muscle during exercise: a potential antioxidant function. *Free Radical Biology & Medicine,* 46**,** 138-45.

JONES, A. M. & CARTER, H. 2000. The effect of endurance training on parameters of aerobic fitness. *Sports Medicine,* 29**,** 373-86.

JONES, A. M. & DOUST, J. H. 1998. The validity of the lactate minimum test for determination of the maximal lactate steady state. *Med Sci Sports Exerc,* 30**,** 1304-13.

JOYNER, M. J. 1991. Modeling: optimal marathon performance on the basis of physiological factors. *Journal of Applied Physiology* 70**,** 683-687.

JOYNER, M. J. & COYLE, E. F. 2008. Endurance exercise performance: the physiology of champions. *The Journal of Physiology,* 586**,** 35-44.

JOYNER, M. J., RUIZ, J. R. & LUCIA, A. 2011. The two-hour marathon: who and when? *Journal of Applied Physiology,* 110**,** 275-277.

KARJALAINEN, J., KUJALA, U., STOLT, A., MÄNTYSAARI, M., VIITASALO, M., KAINULAINEN, K. & KONTULA, K. 1999. Angiotensinogen Gene M235T Polymorpism Predicts Left Ventricular Hypertrophy in Endurance Athletes. *Journal of the American College of Cardiology,* 34**,** 494-499.

KATCH, V., WELTMAN, A., SADY, S. & FREEDSON, P. 1978. Validity of the relative percent concept for equating training intensity. *Eur J Appl Physiol Occup Physiol,* 39**,** 219-27.

KOCH, W., LATZ, W., EICHINGER, M., GANSER, C., SCHÖMIG, A. & KASTRATI, A. 2005. Genotyping of the angiotensin I converting enzyme gene insertion deletion polymorphism by the Taqman method. *Clinical Chemistry,* 51**,** 1547-1549.

KRAUSS, S., ZHANG, C.-Y., SCORRANO, L., DALGAARD, L. T., ST-PIERRE, J., GREY, S. T. & LOWELL, B. B. 2003. Superoxide-mediated activation of uncoupling protein 2 causes pancreatic β cell dysfunction. *Journal of Clinical Investigation,* 112**,** 1831-1842.

LACOUR, J. R. & BOURDIN, M. 2015. Factors affecting the energy cost of level running at submaximal speed. *Eur J Appl Physiol,* 115**,** 651-73.

LANDER, E. & SCHORK, N. 1994. Genetic dissection of complex traits. *Science,* 265**,** 2037-2037.

LEHMAN, J. J., BARGER, P. M., KOVACS, A., SAFFITZ, J. E., MEDEIROS, D. M. & KELLY, D. P. 2000. Peroxisome proliferator–activated receptor γ coactivator-1 promotes cardiac mitochondrial biogenesis. *The Journal of clinical investigation,* 106**,** 847-856.

LEONE, T. C., LEHMAN, J. J., FINCK, B. N., SCHAEFFER, P. J., WENDE, A. R., BOUDINA, S., COURTOIS, M., WOZNIAK, D. F., SAMBANDAM, N. & BERNAL-MIZRACHI, C. 2005. PGC-1α deficiency causes multi-system energy metabolic derangements: muscle dysfunction, abnormal weight control and hepatic steatosis. *PLoS Biol,* 3**,** e101.

LEVICK, J. R. 2013. *An introduction to cardiovascular physiology*, Butterworth-Heinemann.

LIN, J., HANDSCHIN, C. & SPIEGELMAN, B. M. 2005. Metabolic control through the PGC-1 family of transcription coactivators. *Cell metabolism,* 1**,** 361-370.

LIN, J., WU, H., TARR, P. T., ZHANG, C.-Y., WU, Z., BOSS, O., MICHAEL, L. F., PUIGSERVER, P., ISOTANI, E. & OLSON, E. N. 2002. Transcriptional co-activator PGC-1α drives the formation of slow-twitch muscle fibres. *Nature,* 418**,** 797-801.

LING, C., POULSEN, P., CARLSSON, E., RIDDERSTRÅLE, M., ALMGREN, P., WOJTASZEWSKI, J., BECK-NIELSEN, H., GROOP, L. & VAAG, A. 2004. Multiple environmental and genetic factors influence skeletal muscle PGC-1α and PGC-1β gene expression in twins. *The Journal of clinical investigation,* 114**,** 1518-1526.

LIU, Y. J., LIU, P. Y., LONG, J., LU, Y., ELZE, L., RECKER, R. R. & DENG, H. W. 2005. Linkage and association analyses of the UCP3 gene with obesity phenotypes in Caucasian families. *Physiol Genomics,* 22**,** 197-203.

LUCIA, A., GÓMEZ-GALLEGO, F., BARROSO, I., RABADÁN, M., BANDRÉS, F., SAN JUAN, A. F., CHICHARRO, J. L., EKELUND, U., BRAGE, S., EARNEST, C. P., WAREHAM, N. J. & FRANKS, P.

W. 2005. PPARGC1A genotype (Gly482Ser) predicts exceptional endurance capacity in European men. *Journal of Applied Physiology,* 99**,** 344-348.

LUCIA, A., GÓMEZ-GALLEGO, F., SANTIAGO, C., BANDRES, F., EARNEST, C., RABADAN, M., ALONSO, J., HOYOS, J., CORDOVA, A. & VILLA, G. 2006. ACTN3 genotype in professional endurance cyclists. *International Journal of Sports Medicine,* 27**,** 880-884.

MA, F., YANG, Y., LI, X., ZHOU, F., GAO, C., LI, M. & GAO, L. 2013. The Association of Sport Performance with ACE and ACTN3 Genetic Polymorphisms: A Systematic Review and Meta-Analysis. *PLoS ONE,* 8**,** e54685.

MACARTHUR, D., SETO, J., RAFTERY, J., QUINLAN, K., HUTTLEY, G., HOOK, J., LEMCKERT, F., KEE, A., EDWARDS, M. & BERMAN, Y. 2007. Loss of *ACTN3* gene function alters mouse muscle metabolism and shows evidence of positive selection in humans. *Nature Genetics,* 39**,** 1261-1265.

MACARTHUR, D. G. & NORTH, K. N. 2004. A gene for speed? The evolution and function of α-actinin-3. *BioEssays,* 26**,** 786-795.

MACARTHUR, D. G. & NORTH, K. N. 2007. ACTN3: A genetic influence on muscle function and athletic performance. *Exercise and Sport Sciences Reviews,* 35**,** 30-344.

MACARTHUR, D. G., SETO, J. T., CHAN, S., QUINLAN, K. G. R., RAFTERY, J. M., TURNER, N., NICHOLSON, M. D., KEE, A. J., HARDEMAN, E. C., GUNNING, P. W., COONEY, G. J., HEAD, S. I., YANG, N. & NORTH, K. N. 2008. An Actn3 knockout mouse provides mechanistic insights into the association between alpha-actinin-3 deficiency and human athletic performance. *Human Molecular Genetics,* 17**,** 1076-86.

MACIEJEWSKA, A., SAWCZUK, M., CIESZCZYK, P., MOZHAYSKAYA, I. A. & AHMETOV, I. I. 2011. The PPARGC1Agene Gly482Ser in Polish and Russian athletes. *Journal of Sports Sciences,* 30**,** 101-113.

MALISOUX, L., FRANCAUX, M. & THEISEN, D. 2007. What do single-fiber studies tell us about exercise training? *Med Sci Sports Exerc,* 39**,** 1051-60.

MARC, A., SEDEAUD, A., GUILLAUME, M., RIZK, M., SCHIPMAN, J., ANTERO-JACQUEMIN, J., HAIDA, A., BERTHELOT, G. & TOUSSAINT, J. F. 2014. Marathon progress: demography, morphology and environment. *J Sports Sci,* 32**,** 524-32.

MARUSZAK, A., ADAMCZYK, J. G., SIEWIERSKI, M., SOZANSKI, H., GAJEWSKI, A. & ZEKANOWSKI, C. 2014. Mitochondrial DNA variation is associated with elite athletic status in the Polish population. *Scand J Med Sci Sports,* 24**,** 311-8.

MAUD, P. J. & FOSTER, C. 1995. *Physiological Assessment of Human Fitness,* Champaign, IL, Human Kinetics.

MCARDLE, W. D., KATCH, F. I. & KATCH, V. L. 2009. *Exercise Physiology: Energy, Nutrition, and Human Performance,* Baltimore, MA, Lippincott Williams & Wilkins.

MCPHEE, J. S., WILLIAMS, A. G., STEWART, C., BAAR, K., SCHINDLER, J. P., ALDRED, S., MAFFULLI, N., SARGEANT, A. J. & JONES, D. A. 2009. The training stimulus experienced by the leg muscles during cycling in humans. *Experimental Physiology,* 94**,** 684-94.

MILLS, M., YANG, N., WEINBERGER, R., VANDER WOUDE, D., BEGGS, A., EASTEAL, S. & NORTH, K. 2001. Differential expression of the actin-binding proteins, α-actinin-2 and-3, in different species: implications for the evolution of functional redundancy. *Human Molecular Genetics,* 10**,** 1335-1346.

MONTGOMERY, H., MARSHALL, R., HEMINGWAY, H., MYERSON, S., CLARKSON, P., DOLLERY, C., HAYWARD, M., HOLLIMAN, D., JUBB, M., WORLD, M., THOMAS, E., BRYNES, A., SAEED, N.,

BARNARD, M., BELL, J., PRASAD, K., RAYSON, M., TALMUD, P. & HUMPHRIES, S. 1998. Human gene for physical performance. *Nature,* 393**,** 221-222.

MONTGOMERY, H. E., CLARKSON, P., DOLLERY, C. M., PRASAD, K., LOSI, M.-A., HEMINGWAY, H., STATTERS, D., JUBB, M., GIRVAIN, M. & VARNAVA, A. 1997. Association of angiotensin-converting enzyme gene I/D polymorphism with change in left ventricular mass in response to physical training. *Circulation,* 96**,** 741-747.

MUNIESA, C. A., GONZÁLEZ-FREIRE, M., SANTIAGO, C., LAO, J., BUXENS, A., RUBIO, J. C., MARTÍN, M. A., ARENAS, J., GÓMEZ-GALLEGO, F. & LUCIA, A. 2010. World-class performance in lightweight rowing: is it genetically influenced? A comparison with cyclists, runners and non-athletes. *British Journal of Sports Medicine,* 44**,** 898-901.

MURPHEY, L. J., GAINER, J. V., VAUGHAN, D. E. & BROWN, N. J. 2000. Angiotensin-Converting Enzyme Insertion/Deletion Polymorphism Modulates the Human In Vivo Metabolism of Bradykinin. *Circulation,* 102**,** 829-832.

MUZZIN, P., BOSS, O. & GIACOBINO, J.-P. 1999. Uncoupling Protein 3: Its Possible Biological Role and Mode of Regulation in Rodents and Humans. *Journal of Bioenergetics and Biomembranes,* 31**,** 467-473.

MYERSON, S., HEMINGWAY, H., BUDGET, R., MARTIN, J., HUMPHRIES, S. & MONTGOMERY, H. E. 1999. Human angiotensin I-converting enzyme gene and endurance performance. *Journal of Applied Physiology,* 87**,** 1313-1316.

NAYA, F. J., MERCER, B., SHELTON, J., RICHARDSON, J. A., WILLIAMS, R. S. & OLSON, E. N. 2000. Stimulation of Slow Skeletal Muscle Fiber Gene Expression by Calcineurin in Vivo. *Journal of Biological Chemistry,* 275**,** 4545-4548.

NAZAROV, I., WOODS, D., MONTGOMERY, H., SHNEIDER, O., KAZAKOV, V., TOMILIN, N. & ROGOZKIN, V. 2001. The angiotensin converting enzyme I/D polymoprhsim in Russian athletes. *European Journal of Human Genetics,* 9**,** 797-801.

NIEMI, A. & MAJAMAA, K. 2005. Mitochondrial DNA and ACTN3 genotypes in Finnish elite endurance and sprint athletes. *European Journal of Human Genetics,* 13**,** 965-969.

NISHIDA, Y., IYADOMI, M., HIGAKI, Y., TANAKA, H., KONDO, Y., OTSUBO, H., HORITA, M., HARA, M. & TANAKA, K. 2015. Association between the PPARGC1A Polymorphism and Aerobic Capacity in Japanese Middle-aged Men. *Internal Medicine,* 54**,** 359-366.

NISR, R. B. & AFFOURTIT, C. 2014. Insulin acutely improves mitochondrial function of rat and human skeletal muscle by increasing coupling efficiency of oxidative phosphorylation. *Biochimica et Biophysica Acta,* 1837**,** 270-6.

NORRBOM, J., SUNDBERG, C. J., AMELN, H., KRAUS, W. E., JANSSON, E. & GUSTAFSSON, T. 2004. PGC-1α mRNA expression is influenced by metabolic perturbation in exercising human skeletal muscle. *Journal of Applied Physiology,* 96**,** 189-194.

NORTH, K. N., YANG, N., WATTANASIRICHAIGOON, D., MILLS, M. A., EASTEAL, S. & BEGGS, A. H. 1999. A common nonsense mutation results in α-actinin-3 deficiency in the general population. *Nature Genetics,* 21**,** 353-354.

OH, S. 2007. The distribution of I/D polymorphism in the ACE gene among Korean male elite athletes. *The Journal of Sports Medicine and Physical Fitness,* 47**,** 250-254.

PAPADIMITRIOU, I., PAPADOPOULOS, C., KOUVATSI, A. & TRIANTAPHYLLIDIS, C. 2008. The ACTN3 gene in elite Greek track and field athletes. *International Journal of Sports Medicine,* 29**,** 352-355.

PAPADIMITRIOU, I., PAPADOPOULOS, C., KOUVATSI, A. & TRIANTAPHYLLIDIS, C. 2009. The ACE I/D polymorphism in elite Greek track and field athletes. *The Journal of Sports Medicine and Physical Fitness,* 49**,** 459-463.

PAUL, M., MEHR, A. & KREUTZ, R. 2006. Physiology of local renin-angiotensin systems. *Physiological Reviews,* 86**,** 747-803.

PEDERSEN, S. B., LUND, S., BUHL, E. S. & RICHELSEN, B. 2001. Insulin and contraction directly stimulate UCP2 and UCP3 mRNA expression in rat skeletal muscle in vitro. *Biochemical and Biophysical Research Communications,* 283**,** 19-25.

PÉRUSSE, L., RANKINEN, T., RAURAMAA, R., RIVERA, M. A., WOLFARTH, B. & BOUCHARD, C. 2003. The Human Gene Map for Performance and Health- Related Fitness Phenotypes: The 2002 Update. *Medicine and Science in Sports and Exercise,* 35**,** 1248-1264.

PETTE, D. & STARON, R. S. 1997. Mammalian skeletal muscle fiber type transitions. *Int Rev Cytol,* 170**,** 143-223.

PILEGAARD, H., SALTIN, B. & NEUFER, P. D. 2003. Exercise induces transient transcriptional activation of the PGC-1α gene in human skeletal muscle. *The Journal of physiology,* 546**,** 851-858.

PUIGSERVER, P., WU, Z., PARK, C. W., GRAVES, R., WRIGHT, M. & SPIEGELMAN, B. M. 1998. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell,* 92**,** 829-839.

PUTHUCHEARY, Z., SKIPWORTH, J., RAWAL, J., LOOSEMORE, M., VAN SOMEREN, K. & MONTGOMERY, H. 2011. The ACE Gene and Human Performance: 12 Years On. *Sports Medicine,* 41**,** 433-448.

RANKINEN, T., BRAY, M. S., HAGBERG, J. M., PÉRUSSE, L., ROTH, S. M., WOLFARTH, B. & BOUCHARD, C. 2006. The human gene map for performance and health-related fitness phenotypes: the 2005 update. *Medicine and Science in Sports and Exercise,* 38**,** 1863.

RANKINEN, T., GAGNON, J., PÉRUSSE, L., CHAGNON, Y. C., RICE, T., LEON, A. S., SKINNER, J. S., WILMORE, J. H., RAO, D. C. & BOUCHARD, C. 2000a. AGT M235T and ACE ID polymorphisms and exercise blood pressure in the HERITAGE Family Study. *American Journal of Physiology - Heart and Circulatory Physiology,* 279**,** H368-H374.

RANKINEN, T., PÉRUSSE, L., GAGNON, J., CHAGNON, Y. C., LEON, A. S., SKINNER, J. S., WILMORE, J. H., RAO, D. C. & BOUCHARD, C. 2000b. Angiotensin-converting enzyme ID polymorphism and fitness phenotype in the HERITAGE Family Study. *Journal of Applied Physiology,* 88**,** 1029-1035.

RANKINEN, T., PÉRUSSE, L., RAURAMAA, R., RIVERA, M. A., WOLFARTH, B. & BOUCHARD, C. 2001. The human gene map for performance and health-related fitness phenotypes. *Medicine and Science in Sports and Exercise,* 33**,** 855-867.

RANKINEN, T., PÉRUSSE, L., RAURAMAA, R., RIVERA, M. A., WOLFARTH, B. & BOUCHARD, C. 2002. The human gene map for performance and health-related fitness phenotypes: the 2001 update. *Medicine and Science in Sports and Exercise,* 34**,** 1219-1233.

RANKINEN, T., PÉRUSSE, L., RAURAMAA, R., RIVERA, M. A., WOLFARTH, B. & BOUCHARD, C. 2004. The human gene map for performance and health-related fitness phenotypes: the 2003 update. *Medicine and science in sports and exercise,* 36**,** 1451-1469.

RANKINEN, T., WOLFARTH, B., SIMONEAU, J.-A., MAIER-LENZ, D., RAURAMAA, R., RIVERA, M. A., BOULAY, M. R., CHAGNON, Y. C., PÉRUSSE, L., KEUL, J. & BOUCHARD, C. 2000c. No association between the angiotensin-converting enzyme ID polymorphism and elite endurance athlete status. *Journal of Applied Physiology,* 88**,** 1571-1575.

RIGAT, B., HUBERT, C., ALHENC-GELAS, F., CAMBIEN, F., CORVOL, P. & SOUBRIER, F. 1990. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *Journal of Clinical Investigation,* 86**,** 1343.

RIGAT, B., HUBERT, C., CORVOL, P. & SOUBRIER, F. 1992. PCR detection of the insertion/deletion polymorphism of the human angiotensin converting enzyme gene (DCP1)(dipeptidyl carboxypeptidase 1). *Nucleic Acids Research,* 20**,** 1433.

ROTH, S., WALSH, S., LIU, D., METTER, E., FERRUCCI, L. & HURLEY, B. 2007. The ACTN3 R577X nonsense allele is under-represented in elite-level strength athletes. *European Journal of Human Genetics,* 16**,** 391-394.

RUIZ, J. R., ARTETA, D., BUXENS, A., ARTIEDA, M., GÓMEZ-GALLEGO, F., SANTIAGO, C., YVERT, T., MORÁN, M. & LUCIA, A. 2010. Can we identify a power-oriented polygenic profile? *Journal of Applied Physiology,* 108**,** 561-566.

RUIZ, J. R., GOMEZ-GALLEGO, F., SANTIAGO, C., GONZALEZ-FREIRE, M., VERDE, Z., FOSTER, C. & LUCIA, A. 2009. Is there an optimum endurance polygenic profile? *Journal of Physiology,* 587**,** 1527-34.

RUSSELL, A. P., FEILCHENFELDT, J., SCHREIBER, S., PRAZ, M., CRETTENAND, A., GOBELET, C., MEIER, C. A., BELL, D. R., KRALLI, A. & GIACOBINO, J.-P. 2003. Endurance training in humans leads to fiber type-specific increases in levels of peroxisome proliferator-activated receptor-γ coactivator-1 and peroxisome proliferator-activated receptor-α in skeletal muscle. *Diabetes,* 52**,** 2874-2881.

SADOSHIMA, J. & IZUMO, S. 1993. Molecular characterization of angiotensin II--induced hypertrophy of cardiac myocytes and hyperplasia of cardiac fibroblasts. Critical role of the AT1 receptor subtype. *Circulation Research,* 73**,** 413-423.

SAHLIN, K. 1992. Metabolic factors in fatigue. *Sports Med,* 13**,** 99-107.

SALTIN, B. & ASTRAND, P. 1967. Maximal oxygen uptake in athletes. *Journal of applied physiology,* 23**,** 353-358.

SAUNDERS, C. J., SEPTEMBER, A. V., XENOPHONTOS, S. L., CARIOLOU, M. A., ANASTASSIADES, L. C., NOAKES, T. D. & COLLINS, M. 2007. No association of the ACTN3 gene R577X polymorphism with endurance performance in Ironman Triathlons. *Annals of Human Genetics,* 71**,** 777-81.

SAUNDERS, P. U., PYNE, D. B., TELFORD, R. D. & HAWLEY, J. A. 2004. Factors affecting running economy in trained distance runners. *Sports Medicine,* 34**,** 465-485.

SCANAVINI, D., BERNARDI, F., CASTOLDI, E., CONCONI, F. & MAZZONI, G. 2002. Increased frequency of the homozygous II ACE genotype in Italian Olympic endurance athletes. *European Journal of Human Genetics,* 10**,** 576.

SCHMUTZ, J., WHEELER, J., GRIMWOOD, J., DICKSON, M., YANG, J., CAOILE, C., BAJOREK, E., BLACK, S., CHAN, Y. M., DENYS, M., ESCOBAR, J., FLOWERS, D., FOTOPULOS, D., GARCIA, C., GOMEZ, M., GONZALES, E., HAYDU, L., LOPEZ, F., RAMIREZ, L., RETTERER, J., RODRIGUEZ, A., ROGERS, S., SALAZAR, A., TSAI, M. & MYERS, R. M. 2004. Quality assessment of the human genome sequence. *Nature,* 429**,** 365-368.

SCHRAUWEN, P., XIA, J., WALDER, K., SNITKER, S. & RAVUSSIN, E. 1999. A novel polymorphism in the proximal UCP3 promoter region: effect on skeletal muscle UCP3 mRNA expression and obesity in male non-diabetic Pima Indians. *International Journal Of Obesity,* 23**,** 1242.

SCHUELKE, M., WAGNER, K. R., STOLZ, L. E., HUBNER, C., RIEBEL, T., KOMEN, W., BRAUN, T., TOBIN, J. F. & LEE, S. J. 2004. Myostatin mutation associated with gross muscle hypertrophy in a child. *N Engl J Med,* 350**,** 2682-8.

SCOTT, R., IRVING, R., IRWIN, L., MORRISON, E., CHARLTON, V., AUSTIN, K., TLADI, D., DEASON, M., HEADLEY, S. & KOLKHORST, F. 2010. ACTN3 and ACE genotypes in elite Jamaican and US sprinters. *Medicine and Science in Sports and Exercise,* 42**,** 107-12.

SCOTT, R. A., MORAN, C. N., WILSON, R., ONYWERA, V., BOIT, M., GOODWIN, W., GOHLKE, P., PAYNE, J., MONTGOMERY, H. E. & PITSILADIS, Y. P. 2005. No association between angiotensin converting enzyme (ACE) gene variation and endurance athlete status in Kenyans. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology,* 141**,** 169-175.

SETO, J. T., QUINLAN, K. G. R., LEK, M., ZHENG, X. F., GARTON, F., MACARTHUR, D. G., HOGARTH, M. W., HOUWELING, P. J., GREGOREVIC, P., TURNER, N., COONEY, G. J., YANG, N. & NORTH, K. N. 2013. ACTN3 genotype influences muscle performance through the regulation of calcineurin signaling. *Journal of Clinical Investigation,* 123**,** 4255-63.

SHADEL, G. S. & CLAYTON, D. A. 1993. Mitochondrial transcription initiation. Variation and conservation. *Journal of Biological Chemistry,* 268**,** 16083-16083.

SIMON, J., YOUNG, J. L., GUTIN, B., BLOOD, D. K. & CASE, R. B. 1983. Lactate accumulation relative to the anaerobic and respiratory compensation thresholds. *J Appl Physiol Respir Environ Exerc Physiol,* 54**,** 13-7.

ST-PIERRE, J., LIN, J., KRAUSS, S., TARR, P. T., YANG, R., NEWGARD, C. B. & SPIEGELMAN, B. M. 2003. Bioenergetic analysis of peroxisome proliferator-activated receptor γ coactivators 1α and 1β (PGC-1α and PGC-1β) in muscle cells. *Journal of Biological Chemistry,* 278**,** 26597-26603.

STEWART, C. E. & RITTWEGER, J. 2006. Adaptive processes in skeletal muscle: Molecular regulators and genetic influences. *Journal of Musculoskeletal and Neuronal Interactions,* 6**,** 73-86.

STUMVOLL, M., FRITSCHE, A., T'HART, L. M., MACHANN, J., THAMER, C., TSCHRITTER, O., VAN HAEFTEN, T. W., JACOB, S., DEKKER, J. M. & MAASSEN, J. A. 2004. The Gly482Ser variant in the peroxisome proliferator-activated receptor gamma coactivator-1 is not associated with diabetes-related traits in non-diabetic German and Dutch populations. *Experimental and clinical endocrinology & diabetes: official journal, German Society of Endocrinology [and] German Diabetes Association,* 112**,** 253-257.

SUMMERMATTER, S., SANTOS, G., PÉREZ-SCHINDLER, J. & HANDSCHIN, C. 2013. Skeletal muscle PGC-1α controls whole-body lactate homeostasis through estrogen-related receptor α-dependent activation of LDH B and repression of LDH A. *Proceedings of the National Academy of Sciences,* 110**,** 8738-8743.

SWANN, C., MORAN, A. & PIGGOTT, D. 2015. Defining elite athletes: Issues in the study of expert performance in sport psychology. *Psychology of Sport and Exercise,* 16**,** 3-14.

TANAKA, M., WANG, G. & PITSILADIS, Y. P. 2016. Advancing sports and exercise genomics: moving from hypothesis-driven single study approaches to large multi-omics collaborative science. *Physiological genomics***,** physiolgenomics. 00009.2016.

TERADA, S., GOTO, M., KATO, M., KAWANAKA, K., SHIMOKAWA, T. & TABATA, I. 2002. Effects of low-intensity prolonged exercise on PGC-1 mRNA expression in rat epitrochlearis muscle. *Biochemical and biophysical research communications,* 296**,** 350-354.

THE INTERNATIONAL HAPMAP CONSORTIUM 2005. A haplotype map of the human genome. *Nature,* 437**,** 1299-1320.

TONKONOGI, M., KROOK, A., WALSH, B. & SAHLIN, K. 2000. Endurance training increases stimulation of uncoupling of skeletal muscle mitichondria in humans by non-esterified fatty acids: an uncoupling- protein- mediated effect? *Biochemical Journal,* 351**,** 805-810.

TSIANOS, G., SANDERS, J., DHAMRAIT, S., HUMPHRIES, S., GRANT, S. & MONTGOMERY, H. 2004. The ACE gene insertion/deletion polymorphism and elite endurance swimming. *European Journal of Applied Physiology,* 92**,** 360-362.

TSUBOYAMA-KASAOKA, N., TSUNODA, N., MARUYAMA, K., TAKAHASHI, M., KIM, H., IKEMOTO, S. & EZAKI, O. 1998. Up-Regulation of Uncoupling Protein 3 (UCP3) mRNA by Exercise Training and Down-Regulation of UCP3 by Denervation in Skeletal Muscles. *Biochemical and Biophysical Research Communications,* 247**,** 498-503.

TREMBLAY, MS., AUBERT, S., BARNES, JD., SAUNDERS, TJ., CARSON, V., LATIMER-CHEUNG AE, CHASTIN, SFM., ALTENBERUG TM, & CHINAPAW, MJM., 2017. SBRN Terminology consensus project participants. Sedentary behaviour research network (SBRN) - terminoligy consensus project process outcome. *International Journal of Behavoiur and Natural Physiology*, 14:75

TUNSTALL, R. J., MEHAN, K. A., WADLEY, G. D., COLLIER, G. R., BONEN, A., HARGREAVES, M. & CAMERON-SMITH, D. 2002. Exercise training increases lipid metabolism gene expression in human skeletal muscle. *American Journal of Physiology-Endocrinology and Metabolism,* 283**,** E66-E72.

TURAL, E., KARA, N., AGAOGLU, S. A., ELBISTAN, M., TASMEKTEPLIGIL, M. Y. & IMAMOGLU, O. 2014. PPAR-alpha and PPARGC1A gene variants have strong effects on aerobic performance of Turkish elite endurance athletes. *Mol Biol Rep,* 41**,** 5799-804.

VANACKER, J. M., PETTERSSON, K., GUSTAFSSON, J. Å. & LAUDET, V. 1999. Transcriptional targets shared by estrogen receptor-related receptors (ERRs) and estrogen receptor (ER) α, but not by ERβ. *The EMBO journal,* 18**,** 4270-4279.

VIDAL-PUIG, A., SOLANES, G., GRUJIC, D., FLIER, J. S. & LOWELL, B. B. 1997. UCP3: An Uncoupling Protein Homologue Expressed Preferentially and Abundantly in Skeletal Muscle and Brown Adipose Tissue. *Biochemical and Biophysical Research Communications,* 235**,** 79-82.

VINCENT, B., DE BOCK, K., RAMAEKERS, M., VAN DEN EEDE, E., VAN LEEMPUTTE, M., HESPEL, P. & THOMIS, M. A. 2007. ACTN3 (R577X) genotype is associated with fiber type distribution. *Physiological Genomics,* 32**,** 58-63.

VISSCHER, P., BROWN, M., MCCARTHY, M. & YANG, J. 2012. Five Years of GWAS Discovery. *The American Journal of Human Genetics,* 90**,** 7-24.

VOGT, M., PUNTSCHART, A., GEISER, J., ZULEGER, C., BILLETER, R. & HOPPELER, H. 2001. Molecular adaptations in human skeletal muscle to endurance training under simulated hypoxic conditions. *Journal of Applied Physiology,* 91**,** 173-182.

WHIPP, B. J. & WASSERMAN, K. 1972. Oxygen uptake kinetics for various intensities of constant-load work. *J Appl Physiol,* 33**,** 351-6.

WICKLMAYR, M., DIETZE, G., BRUNNBAUER, H., RETT, K. & MEHNERT, H. 1983. Dose-dependent effect of bradykinin on muscular blood flow and glucose uptake in man. *Hoppe Seylers Z Physiol Chem,* 364**,** 831-3.

WILLIAMS, A. & FOLLAND, J. 2008. Similarity of polygenic profiles limits the potential for elite human physical performance. *The Journal of Physiology,* 586**,** 113-21.

WILLIAMS, A. G. 2011. Letter to the Editor: The two-hour Marathon: Genetics will become increasingly important. *Journal of Applied Physiology,* 110**,** 288.

WILLIAMS, A. G., RAYSON, M., JUBB, M., WORLD, M., WOODS, D., HAYWARD, M., MARTIN, J., HUMPHRIES, S. E. & MONTGOMERY, H. E. 2000. Physiology: The ACE gene and muscle performance. *Nature,* 403**,** 614.

WITTKE-THOMPSON, J. K., PLUZHNIKOV, A. & COX, N. J. 2005. Rational inferences about departures from Hardy-Weinberg equilibrium. *The American Journal of Human Genetics,* 76**,** 967-986.

WOLFARTH, B., BRAY, M. S., HAGBERG, J. M., PÉRUSSE, L., RAURAMAA, R., RIVERA, M. A., ROTH, S. M., RANKINEN, T. & BOUCHARD, C. 2005. The human gene map for performance and health-related fitness phenotypes: the 2004 update. *Medicine and Science in Sports and Exercise,* 37**,** 881-903.

WOODS, D., HUMPHRIES, S. & MONTGOMERY, H. 2000. The ACE I/D Polymorphism and Human Physical Performance. *Trends in Endocrinology & Metabolism,* 11**,** 416-420.

WU, Z., PUIGSERVER, P., ANDERSSON, U., ZHANG, C., ADELMANT, G., MOOTHA, V., TROY, A., CINTI, S., LOWELL, B. & SCARPULLA, R. C. 1999. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell,* 98**,** 115-124.

YANG, N., MACARTHUR, D. G., GULBIN, J. P., HAHN, A. G., BEGGS, A. H., EASTEAL, S. & NORTH, K. N. 2003. ACTN3 genotype is associated with human elite athletic performance. *American Journal of Human Genetics,* 73**,** 627-631.

YANG, N., MACARTHUR, D. G., WOLDE, B., ONYWERA, V. O., BOIT, M. K., LAU, S. Y. M.-A., WILSON, R. H., SCOTT, R. A., PITSILADIS, Y. P. & NORTH, K. N. 2007. The ACTN3 R577X Polymorphism in East and West African Athletes. *Medicine and Science in Sports and Exercise,* 39**,** 1985-1988.

ZAREBSKA, A., SAWCZYN, S., KACZMARCZYK, M., FICEK, K., MACIEJEWSKA-KARŁOWSKA, A., SAWCZUK, M., LEOSKA-DUNIEC, A., EIDER, J., GRENDA, A. & CIĘSZCZYK, P. 2013. Association of the rs699 (M235T) Polymorphism in the AGT Gene with Power but not Endurance Athlete Status. *Journal of Strength and Conditioning Research,* 26**,** 2898-2903.

ZHANG, B., TANAKA, H., SHONO, N., MIURA, S., KIYONAGA, A., SHINDO, M. & SAKU, K. 2003. The I allele of the angiotensin-converting enzyme gene is associated with an increased percentage of slow-twitch type I fibers in human skeletal muscle. *Clinical Genetics,* 63**,** 139-144.

ZHOU, M., LIN, B.-Z., COUGHLIN, S., VALLEGA, G. & PILCH, P. F. 2000. UCP-3 expression in skeletal muscle: effects of exercise, hypoxia, and AMP-activated protein kinase. *American Journal of Physiology - Endocrinology and Metabolism,* 279**,** E622-E629.

11 APPENDIX 1

The Genetic Profile of Elite Athletes
Questionnaire: Physical Activity & General Health

Manchester Metropolitan University

Thank you for your interest in this research study. Prior to participation, we would like you to answer a few questions concerning your ethnic origin, blood donation and athlete status. Please answer the following questions as honestly as you can.

Participant ID code: _____ Date of birth: _____

Gender (please circle): Male / Female Height: _____

Nationality: _____ Body weight: _____

What is your ethnic group? Please circle ONE section from **A** to **F**, then tick the appropriate box to indicate your background.

A) White: English ☐ Scottish ☐ Welsh ☐ N. Irish ☐ Other ☐

If other, please state here: _____

B) Mixed: White and Black Caribbean ☐ White and Black African ☐ White and Asian ☐ Other ☐

If other, please state here: _____

C) Asian: Indian ☐ Pakistani ☐ Chinese ☐ Japanese ☐ Other ☐

If other, please state here: _____

D) Black: Caribbean ☐ African ☐ Other ☐

If other, please state here: _____

E) Other ethnic background: ☐ Please state here: _____

F) I do not wish to state my ethnic origin ☐

Blood donation

To determine your genetic profile we would like to take a small (10 mL) blood sample from a superficial vein in your arm. Before doing so, please answer the following safety questions.

1. Have you ever been infected with any other blood-borne disease? __ Yes ☐ No ☐
2. Are you anaemic or receiving treatment for anaemia or iron deficiency? _____ Yes ☐ No ☐
3. Are you allergic to any of the following: Sterile alcohol pad Yes ☐ No ☐
   Plaster Yes ☐ No ☐

If you have answered YES to any of these questions and/or you would prefer not to provide a blood sample, a buccal (cheek cell) sample may be provided instead.