Physical activity, metabolic syndrome indicators and body composition interactions in Portuguese nuclear families

Daniel Monteiro de Vilhena e Santos

2013

U.PORTO

FACULDADE DE DESPORTO
UNIVERSIDADE DO PORTO

Physical activity, metabolic syndrome indicators and body composition interactions in Portuguese nuclear families

Dissertação apresentada ao Programa Doutoral em Ciências do Desporto (Decreto-Lei n.º 74/2006, de 24 de Março), com vista à obtenção do grau de Doutor em Ciências do Desporto, sob a orientação do Professor Doutor José António Ribeiro Maia, co-orientação do Professor Doutor Peter T. Katzmarzyk e suporte do Professor Doutor Vincent P. Diego.

Daniel Monteiro de Vilhena e Santos

Porto, Outubro 2013

FICHA DE CATALOGAÇÃO

Santos, D.M.V. (2013). **Physical activity, metabolic syndrome indicators and body composition interactions in Portuguese nuclear families.** Porto: Dissertação de Doutoramento apresentada à Faculdade de Desporto da Universidade do Porto.

Palavras-chave: ACTIVIDADE FÍSICA, SÍNDROME METABÓLICA, COMPOSIÇÃO CORPORAL, EPIDEMIOLOGIA GENÉTICA, INTERACÇÃO, FAMÍLIAS.

Numa casa portuguesa fica bem,
pão e vinho sobre a mesa.
E se à porta humildemente bate alguém,
senta-se à mesa com a gente.
Fica bem esta franqueza, fica bem,
que o povo nunca desmente.
A alegria da pobreza
está nesta grande riqueza
de dar, e ficar contente.

Quatro paredes caiadas,
um cheirinho à alecrim,
um cacho de uvas doiradas,
duas rosas num jardim,
um São José de azulejo,
mais o sol da primavera...
uma promessa de beijos...
dois braços à minha espera...

É uma casa portuguesa, com certeza!
É, com certeza, uma casa portuguesa!

No conforto pobrezinho do meu lar,
Há fartura de carinho.
E a cortina da janela é o luar,
mais o sol que bate nela...
Basta pouco, poucochinho p'ra alegrar
Uma existência singela...
É só amor, pão e vinho
e um caldo verde, verdinho
a fumegar na tigela.

Quatro paredes caiadas,
um cheirinho á alecrim,
um cacho de uvas doiradas,
duas rosas num jardim,
São José de azulejo
mais um sol de primavera...
uma promessa de beijos...
dois braços à minha espera...

É uma casa portuguesa, com certeza!
É, com certeza, uma casa portuguesa!
É uma casa portuguesa, com certeza!
É, com certeza, uma casa portuguesa!

Poema de Reinaldo Ferreira e Vasco Matos

DEDICATÓRIAS

In memoriam

"Saudade", de todas as palavras, é aquela que é mais nossa, mais portuguesa. Senti-la, vivê-la, estremece-nos o peito e espelha-nos os olhos. A minha avó Almerinda partiu cedo, cedo demais... Mas a sua voz ficou! A mais bela das vozes! Tantas vezes me amparou e guiou. Venho orgulhosamente de uma casa Portuguesa!

À minha mamã

Tudo o que sou... tudo o que faço... tudo o que tenho... tudo tem a tua marca... Praticamente 30 anos passaram e nada mudou! Tu, mamã... continuas a ser o farol da minha existência e essa é a maior das minhas fortunas!

À minha maninha

Como te amo maninha... és o meu maior orgulho, pela criança que foste, pela mulher que te tornaste!

À minha Barbita

A "estrelinha" mais brilhante da minha vida ilumina os meus acordares e aquece os meus descansos. O teu sorriso ultrapassa qualquer barreira e o teu abraço conforta qualquer desassossego... a nossa vida ainda agora começou e a maior prenda já está a caminho!!!☺

AGRADECIMENTO

Às famílias...

Eram umas seis e meia da manhã de apenas mais um sábado frio, preguiçoso, insosso, de Fevereiro... a persiana insistia em não subir e o conforto da flanela convidava a mais 5 minutos. Aqueles 5 minutos.

Estávamos em 2010 e aguardava-nos a primeira de muitas manhãs de sábado na companhia de famílias. Tantas famílias. De Santo Tirso a Vouzela, passando por Matosinhos, Leça da Palmeira e Porto. Famílias ricas, pobres e assim-assim. Umas mais completas que outras. Umas mais bem-dispostas que outras. Umas atrasadas. Outras adiantadas. Umas apressadas porque o sábado é apenas mais um dia da semana. Outras apressadas porque o pequeno-almoço estava já ali e o filho tinha um jogo importante. Outras apressadas porque sim... Porque o nosso tempo, o tempo de cada um de nós, é o nosso bem mais precioso.

A todas estas famílias, às famílias do Estudo Português de Famílias Saudáveis, endereço o meu mais sincero agradecimento pelo tempo que nos dedicaram.

Muito obrigado!!!

Aos professores de Educação Física e dos grupos PES...

“Daniel, é mesmo preciso fazer os re-testes?!”...

“Daniel, só conseguimos 6 famílias para este sábado...vou continuar a ligar-lhes!!”

“Daniel, já tem os relatórios prontos?”

“Sim Daniel, amanhã envio os testes de aptidão física”

“Sim Daniel, já recolhemos todos os aparelhos”

“DANIEL, MAIS QUESTIONÁRIOS????”

A lista continuaria sem qualquer esforço de memória. Estes quatro anos foram uma temporada de episódios. Uns mais fáceis que outros. Uns longos...

tão longos. Mas todos com protagonistas de elevadíssimo carácter, altruísmo e determinação – os professores de Educação Física e dos grupos PES.

Sempre tentaram que não vos tratasse por “Professor(a)”... Pelo menos nisto, não foram bem sucedidos!!

O meu sincero agradecimento acompanha o forte abraço amigo que vos envio: Alda Garrido, Ana pais de Sousa, Barros da Cunha, Cláudia Figueiredo, Adelino Passadiço, Miguel Branco, Rosa Pereira, Manuela Machado, José Soares, Pedro Guedes, Sérgio Lopes, Sérgio Silva, Duarte Baltar, Rui Rema, Ana Gomes, Carla Cruz, Clara Costa, Sofia Soares, Vitor Martins, Fernanda Torres, Nuno Correia, Ana Paula Carneiro, Diana Silva, Marta Costa, Joana França, Nuno Carvalho, Marco Costa, Pedro Leite, César Coelho, Duarte Dias, Ricardo Carneiro, António Campos, Diogo Guimarães, Edgar Ferreira, Daniel Rocha, Hélia Cruz, Tiago Ferreira, Fábio Ferreira, Andreia Ferreira, Cristina Sousa, Liliana Oliveira, Conceição Fernandes, Cristina Oliveira, Deolinda Viana, Helena Dias, Francisco Rodrigues, Madalena Silva, Rita Areias, Raquel Eira, Alfredo Calheiros, Aurélio Ribeiro, Joaquim Fontoura, José Manuel Ferreira, Luís Mortágua, Manuel Romão, Maria de Fátima Costa, Rosa Furriel, Pedro Pereira, Audrey Pais, Jorge Silva, Mariana Martins, Augusta Magalhães, Fátima Pinto, Manuela Guimarães, Vanda Teixeira, Teresa Assunção, Madalena Oliveira, António Gomes, Mário Pereira, Manuel Ribeiro

No meio do jardim fica sempre difícil, e talvez não seja importante, saber qual a flor com a côr mais viva, o cheiro mais intenso. Mas “arrisco” um agradecimento particular à Professora Paulina Ferreira, à Professora Amália Almeida, à Professora Ana Monteiro, às Professoras Cláudia Malafaya e Cristina Morêda e ao Professor Paulo Cálão.

Esta temporada terminou... Quem sabe, não haverá outra?

Aos funcionários...

Em todas as escolas ou centros de saúde houve sempre um sorriso à nossa espera. Um sorriso desprendido de interesse e apenas com vontade de ajudar. Nem que seja apenas a abrir um portão, a preparar uma sandes de queijo ou a aquecer um leite com chocolate. Alguém pronto a ir buscar um aquecedor para nos aconchegar e a oferecer-nos um café a meio da manhã.

Alguém que nos faz sentir bem. À D. Maria João em Santo Tirso, à D. Eva e ao Sr. Carlos em Matosinhos, entre muitos outros, o meu agradecimento pela resposta afirmativa a este nosso convite.

Às instituições...

Às direcções das Escolas Secundárias D. Dinis e João Gonçalves Zarco e às Escolas EB 2,3 de Paranhos, Leça da Palmeira e Vouzela, endereço o meu mais sincero agradecimento pela visão moderna e aberta que têm da actualidade e do papel fundamental que a Escola tem que assumir na formação de estilos de vida saudáveis.

Ao Centro de Saúde de Vouzela, na pessoa da Enfermeira-Chefe Maria Augusta Marques Almeida Costa, um agradecimento sentido pela dedicação a este projecto.

Um grande bem-hajam!

Ao karaté...

São muitos os ensinamentos que o meu mestre tentou (e continua a tentar) ensinar e que eu insisto em não aprender. Um dos mais vezes repetidos: "*não se pode estar em dois sítios ao mesmo tempo*"**.** Ao longo deste percurso, fui incapaz de estar apenas nele... Entre outros "desvarios", o karaté esteve sempre presente. Todos os dias da semana e muitos finais de semana. Ao final da tarde, os meus alunos, os meus atletas (como sou possessivo...) aguardam por mim. Ou melhor, aguardo eu por eles. Aguardo que me dediquem mais duas horas da vida deles. Que me ouçam e que me ensinem mais do que alguma vez lhes poderei ensinar. Como vos agradeço...

Aos amigos...

Daqueles amigos que vêm quando se lhes liga... Que nos dão o ombro para chorar... dos amigos que nos chamam à razão e que pensam em nós mesmo quando estamos longe... é desses amigos que estou a falar. São os únicos que cabem aqui! Não são amigos "virtuais" que "likam" de mim! São os que dizem presente na hora de festejar [a melhor despedida de solteiro de sempre!!! Bem, as melhores....:)] e na hora de consolar. É de vocês que quero falar. Quero que todos saibam como sou melhor pessoa por ter o privilégio de

partilhar a minha vida com voçês. Quero “berrar” como sou feliz ao sabor de uma tarde “hidratada” de conversas e parvoíces! Ou como a minha “família” é maior por vos ter perto! Quero que todos saibam que os meus amigos.. são os melhores amigos do Mundo, a todos vós, um grande beijo!!!

Zeca e Ariana; Mariana e Ricardo; Diogo e Joana; Tiago e Vânia; Pires e (.....); Ângelo e Albina; Capu; Polido e Ana; Mita; Vitor e Selminha, Rui, Joaninha e Francisca; Guto e Nice; Luís, Aninha, Maria e Santiago; Maria João.

Aos amigos da faculdade...

Apesar de agnóstico, apenas uma palavra move os meus lábios quando penso sobre as minhas amigas da faculdade: bênção!! De facto, foi uma bênção tê-las na minha companhia ao longo de 5 anos de trabalho. A Raquel, a Michele, a Fernanda e a Thayse vieram de longe, do outro lado do atlântico, para trabalhar! Para lutarem por um futuro melhor para si! São um exemplo que tantas vezes dou aos meus alunos – sair do conforto do lar e ter a coragem de lutar por algo melhor. Muito obrigado pelo exemplo que são e por tudo o que me deram. Não estaria aqui se não fossem vocês! E, não duvidem, tudo o que farei daqui para a frente, terá sempre um pouco de vós: o silêncio reflexivo da Tatá, o “valor do detalhe” da chefinha, a “descomplicação” da Mi e o carinho da Nandinha. Que o mundo vos sorria, como vocês me sorriram a mim!!!

Aos amigos Alcibíades Bustamante, Simonete Silva, Mafalda Roriz, Sara Pereira, Alessandra Borges, Pedro Silva, Rojapon Buranarugsa e Emanuel Passos, bem como, às colegas Sónia, Raquel, Sofia e Renata agradeço o forte empenho e disponibilidade para todas as recolhas de dados.

Aos professores da faculdade...

Estes cinco anos foram repletos de momentos muito importantes na minha formação.

O Prof. Doutor Gaston Beunen (*in memmorian*) permitiu-me o prazer de o ouvir algumas vezes e fiquei sempre com a mesma sensação: como sou pequeno...

À Prof. Martine Thomis agradeço os conhecimentos na área da genética quantitativa que fez o favor de me transmitir.

Um agradecimento ao Prof. Doutor Duarte Freitas pela partilha sincera dos seus conhecimentos e dos estudos que desenvolveu com perseverança na Região Autónoma da Madeira.

Às Professoras Doutoras Cláudia Forjaz e Carol Leandro e ao Prof. Doutor Luciano Basso agradeço o trato fácil e a partilha despreocupada.

Ao Prof. Doutor André Seabra agradeço a oportunidade de trabalhar no Laboratório de Cineantropometria e Estatística Aplicada.

Um agradecimento muito particular e sentido à Profª. Doutora Maria Manuel Jorge, pelo seu saber, pela forma inteligente, sagaz, com que me despertou para os temas da responsabilidade científica e do lugar da ciência no mundo actual.

Uma palavra de apreço e muita consideração para a Professora Doutora Denisa Mendonça pela forma afável como sempre me tratou e por ter aceite o convite para fazer parte da comissão de acompanhamento do meu Doutoramento.

E por último, mas decerto não somenos, um agradecimento muito especial ao Prof. Doutor Rui Garganta. Um verdadeiro Professor. Alguém que nos ensina pelos seus actos. Alguém que no mesmo laboratório nos consegue fazer pensar de forma distinta. Alguém que pensa “fora do baralho” mas que “não salta fora”. O Professor dedicou-me muito do seu tempo e do seu conhecimento... Gostava de o tratar por AMIGO, mas o respeito não me deixa.

À minha família....

Como todas as outras, é só mais uma... Mas esta, é a minha! E como gosto dela!!

Um agradecimento aos membros mais recentes: os meus sogros e cunhado. A forma como me receberam na vossa vida é uma lição que guardo junto a mim e à qual espero conseguir retribuir na mesma medida.

Um beijo para a minha avó Dina – que exemplo de modernidade tu és vovó; aos meus tios Cerejas, um beijinho pela forma como me ampararam quando precisei!!!

Aos meus tios Beatriz e Artur e à minha priminha Mariana um beijinho com muito carinho.

Um beijo especial para o meu pai – na vida temos que andar, outras vezes correr. Muitas vezes deslizamos, escorregamos ou tropeçamos, e acabamos mesmo por cair! Resta-nos o conforto que, tenhamos nós a força para nos levantar, haverá sempre alguém para andar a nosso lado...

Às mulheres da minha vida...

Que belas mulheres... Apaixonadas... Dedicadas... Esforçadas... Lutadoras... e frágeis ao mesmo tempo. Frágeis apenas para que eu possa tentar ampará-las, embora sejam sempre elas a amparar-me. A minha mamã, a minha maninha e a minha Barbita fazem-me sentir todos os dias a beleza da vida. Agradecer-vos, seria apenas um pálido esforço de retribuição por tudo o que me deram, dão e vão continuar a dar.

Resta-me um pedido: que me deixem sempre amparar-vos!!!

To my Professors and friends...

It is not easy to find words (especially in English) to express all my gratitude to Professor Peter T. Katzmarzyk, Professor Vincent Diego and Professor David-Alexandre Trégouet.

Professor David-Alexandre Tregouet left France to teach us. This was not only a challenge but also a great honour for me and I thank you for your efforts to get our paper published.

Professor Vincent Diego or as he prefers “my pal Vincent” is a blessing in this thesis. His fantastic statistical knowledge and skills allowed us to do some really “fancy” analysis of our data. Further, our dinners, my trip to Texas and the way you received me will never be forgotten. I hope we can find the time to be together once again, in Texas or in Porto. My home is always open for you. All the best, men!!!

It was in a sunny morning of March 2009 that Professor José Maia challenged me to do my PhD. Among other things he told me that he would ask Professor Peter Katzmarzyk to be my co-promoter. At that moment, I had no idea of whom he was talking about. I went to Pubmed and I thought I got an

idea… But then I met Professor Peter, I talked with him, I visited him, at his home, with his family. A scientist is much more than just an author… Professor Peter was always present to help me. He always had new ideas to develop our work and his efforts were essential to the success of this project. My words are simply too small to express all my gratitude to you. Thank you very much!!

Ao meu mestre...

Agora, mais que nunca, é que as teclas se movem, os dedos atrapalham, e as palavras nos falham... Outro dos ensinamentos do meu mestre : "*as palavras, são muito mais do que palavras, e por isso devem ser usadas com parcimónia*"... Tentarei estar à altura deste ensinamento, na certeza de que mais uma vez, o desiludirei!

O Professor Doutor José Maia é o coração e a alma de tudo o que está nesta tese, de tudo o que fiz durante a minha vida académica, e por isso, onde quer que esteja e o que quer que faça, carregarei orgulhosamente uma enorme dívida de gratidão. Ainda assim, esta é apenas a menor das fortunas que o Professor me ofereceu.

O Professor mostrou-me mundos novos, mundos que não sabia existirem e mundos que julgava inalcançáveis...

O Professor abriu-me as portas de sua casa e mostrou-me como vive. Mas o Professor fez mais, muito mais. O Professor pintou o meu Mundo com cores novas, deu-lhe contornos que eu nunca havia reconhecido e aromas que eu nunca havia experimentado, e por isso sou hoje uma pessoa muito mais feliz do que era há 5 anos atrás...

O Professor só mudou a minha vida... E essa é a maior dádiva que um aluno pode ousar receber do seu Mestre!!! Muito obrigado!!

ÍNDICE GERAL

ÍNDICE DE TABELAS

CHAPTER 1 - GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

INTRODUCTION

CHPATER 2 - RESEARCH PAPERS

GENETICS OF PHYSICAL ACTIVITY AND PHYSICAL INACTIVITY IN HUMANS

FAMILIAL AGGREGATION OF METABOLIC SYNDROME INDICATORS IN PORTUGUESE FAMILIES

GENOTYPE BY ENERGY EXPENDITURE INTERACTION WITH METABOLIC SYNDROME TRAITS: THE PORTUGUESE HEALTHY FAMILY STUDY

GENOTYPE BY ENERGY EXPENDITURE INTERACTION WITH BODY COMPOSITION TRAITS: THE PORTUGUESE HEALTHY FAMILY STUDY

ÍNDICE DE FIGURAS

CHAPTER 1 - GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

CHAPTER 2 - RESEARCH PAPERS

CHAPTER 3 - GENERAL OVERVIEW AND CONCLUSIONS

RESUMO

O propósito central desta tese foi estudar, em famílias nucleares, as interacções entre actividade física (AF), indicadores de síndrome metabólica (ISM) e composição corporal (CC) a partir de abordagens de Genética Quantitativa.

A amostra foi constituída por 2411 sujeitos, compreendendo 416 pais, 686 mães e 1309 filhos(as), provenientes de 563 famílias nucleares. Procedimentos estandardizados foram utilizados para medir o crescimento físico, CC, AF e ISM. Recorreu-se a análises univariadas e multivariadas implementadas nos *softwares* SPSS, GESEE, PEDSTATS e SOLAR.

As estimativas de heritabilidade de todos os ISM (0.12 a 0.60) e CC (0.21 a 0.34) foram significativas. Os ISM exibiram uma forte semelhança familiar sendo que as correlações entre familiares biologicamente aparentados foram de magnitude similar às correlações entre esposos. Os valores das correlações entre-traços foram baixos com a exceção do par HDL-triglicerídeos. O modelo de interacção Genótipo x Dispêndio Energético apresentou melhor ajustamento aos dados do que o modelo poligénico ($p<0.001$) para o perímetro da cintura, pressão arterial sistólica, glucose, colesterol total, triglicerídeos, índice de massa corporal, percentagem de massa gorda e percentagem de massa gorda no tronco.

Das principais conclusões, destacamos: (i) Actividade física e inactividade física parecem ter arquitecturas genéticas distintas; (ii) os ISM agregam-se no seio das famílas do FAMS; (iii) o ambiente partilhado exerce forte influência na expressão dos ISM; (iv) os ISM e de CC apresentam estimativas de heritabilidade significativas; (v) os ISM e a CC são significativamente influenciados pela interacção entre níveis de actividade física e o genótipo; (vi) a actividade física pode ser considerada um constrangimento ambiental que promove diferenças metabólicas e de composição corporal entre indivíduos que apresentam níveis diferenciados de AF.

Palavras-chave: Actividade Física, Síndrome Metabólica, Composição Corporal, Genética Quantitativa, Interação, Famílias.

ABSTRACT

The main purpose of this study was to study, in nuclear families, the interactions between physical activity (PA), physical inactivity (PI), metabolic syndrome indicators (MetSI) and body composition (BC) from a statistical human genetics perspective.

The sample was comprised of 2411 subjects: 416 fathers, 686 mothers and 1309 offspring from 563 nuclear families. Standardized procedures were used to measure growth, BC, PA and MetSI. Univariate and multivariate data analyses were used in standard software packages including SPSS, GESEE, PEDSTATS and SOLAR.

Heritability estimates for MetSI (0.12 to 0.60) and BC traits (0.21 to 0.34) were all significant ($p<0.05$). MetSI showed strong familial resemblance with correlations between biological relatives being of similar magnitude to those observed between spouses. With cross-trait correlations familial resemblance was very weak except for the High Density Lipoprotein Cholesterol-Triglyceride pair. The Genotype x Energy Expenditure interaction model fitted the data better than the polygenic model ($p<0.001$) for waist circumference, systolic blood pressure, glucose, total cholesterol and triglycerides, body mass index, fat mass percentage and trunk fat mass percentage.

The main conclusions from these studies were: (i) there are potentially different genetic influences on PA versus PI ; (ii) MetSI variance aggregates within families; (iii) most of the phenotypic variance in MetSI could be explained by shared environment; (iv) MetSI and BC traits present significant heritability estimates; (v) MetSI and BC trait expression is significantly influenced by the interaction between total daily energy expenditure and genotypes; (vi) PA may be considered an environmental variable that promotes metabolic and fat mass differences between individuals that are distinctively active.

Keywords: Physical Activity, Metabolic syndrome, Body Composition, Statistical Genetics, Interaction, Families.

CHAPTER 1

GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

INTRODUCTION

Given that the main purposes of this thesis are addressed in the Research Papers´ chapter, this brief introduction will firstly report the concept of interaction and that of nuclear family design as major concepts and as tools to search for the underlying "causes" of human variation in a varied set of phenotypes. Secondly, concise definitions of key concepts used throughout the thesis will be presented, namely physical activity, metabolic syndrome and body composition. A detailed description of *The Portuguese Healthy Families Study*, upon which this research was built, will be addressed in the third section of this introduction. Lastly, research questions tackled in the present dissertation are presented as well as its general outline.

1. Interaction of Nature and Nurture

"The interaction of nature and nurture is one of the central problems of genetics. We can only determine the differences between two different genotypes by putting each of them into a number of different environments."

(Haldane, 1946, p. 197)

The very first sentence of Haldane´s "*The interaction of nature and nurture*" paper written in 1946 is at the heart of this thesis, although addressed with a cross-sectional study, not with an experimental one. Basically, what Haldane meant was that a refined understanding of an individual's genetic potential can only be accomplished if the individual is exposed to distinct environmental constraints, usually different experimental conditions/exposures. Of course this is not true for those characteristics (e.g. eye colour) that are

totally under genetic determination. But, for the vast majority of human characteristics, i.e., metric phenotypes, that are genetically and environmentally driven, an analysis that surpasses the apparently simple addition of genetic and environmental effects is needed. In fact, as will be presented and discussed throughout the thesis, genes and environment are tightly related and co-dependent and can be identified even in cross-sectional studies using a family design.

1.1 GENOTYPE BY ENVIRONMENT INTERACTION

Genotype by Environment interaction (GxE) is the "backbone" of the present PhD thesis as it attempts to link physical activity with both metabolic syndrome and body composition indicators. It is based on the simple assumption that if a genotype[1] has an effect in one environment but fails to produce the same effect in a distinct environment, then the interaction with the environment is thought to have a mediating effect (Haldane, 1946). In other words, the genetic determination of traits might be influenced by particularities of the environment in which it is evolving, resulting in distinct outcomes depending on the environmental grades or classes/types (Haldane, 1946).

An early effort to classify GxE was made by Haldane (1946), as shown in Figure 1. Briefly, considering environments *X* and *Y* and subject *A* and subject *B*, Haldane proposed the existence of 4 types of interactions: (i) subject *A* and subject *B* have better performances in environment *X*, but subject *A* performs better than subject *B* in both environments; (ii) subject *A* still presents the better performance but only under environment *X*, as subject *B* presents the second and third better performances, under environments *X* and *Y*, respectively; (iii) subject *A* and subject *B* perform better in environments *X* and *Y*, respectively, but subject *A* performs better than subject *B* in both environments; and (iv) subject *A* still presents the better performance but only under environment *X*, as subject *B* presents the second and fourth better performances under environment *Y* and *X*, respectively. At this time, type *A* interaction was mostly

[1] Genotype – the genetic makeup of a cell, an organism or group of organisms with reference to a single trait or set of traits.

accepted by eugenics [2] scholars and type *B* interaction was studied by environmentalists[3].

Figure 1. Haldane's 4 types of interactions (adapted from Haldane, 1946)

		X	Y
Type 1	A	1	2
	B	3	4
Type 2	A	1	4
	B	2	3
Type 3	A	1	2
	B	4	3
Type 4	A	1	3
	B	4	2

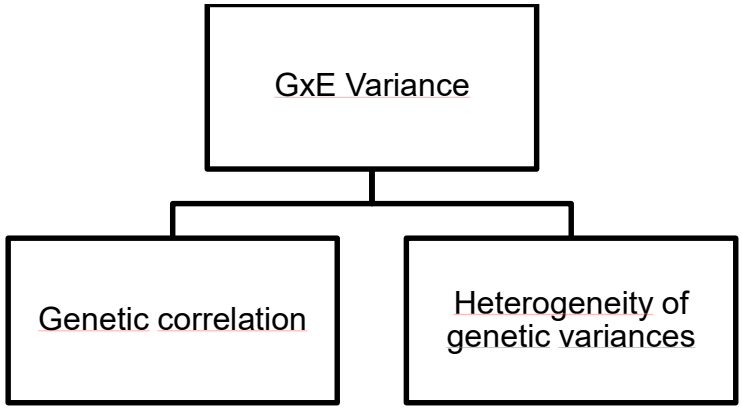
This idea was later statistically formalized by the well-known quantitative geneticist Douglas Scott Falconer (1952) in terms of a genetic correlation by assuming that a character measured in two environments represents two distinct traits. Falconer's suggestion was that a genetic correlation of 1 (or close to 1) means that no GxE interaction occurs. In other words, if there is a tight correlation between the same trait (here treated as distinct) in distinct environments, then the trait is under strong genetic influence and the environment has a negligible effect on trait expression. In contrast, if the genetic correlation is less than 1, then GxE interaction might have biological interest and would deserve further study.

[2] Eugenics – was the "so-science" that attempted to study the possibility of improving the qualities of the human species by selecting the genetically better adapted.
[3] Environmentalists – an individual who believes that differences between subjects or groups are explained, generally, by environmental factors.

The theoretical basis for Falconer's idea was provided by Alan Robertson (1959). Robertson provided a set of formulae with appropriate standard errors for the special cases in which two traits present the same heritability (i.e., the same intra-group correlation).

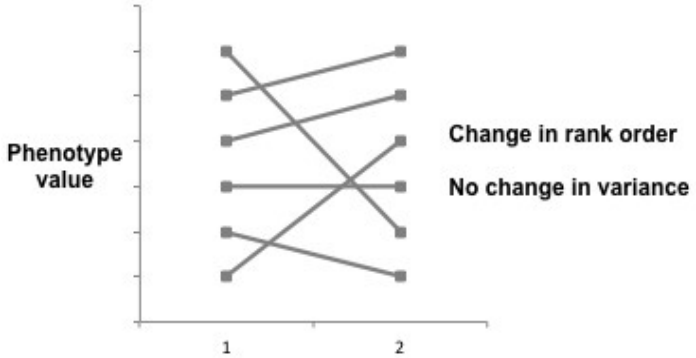
The work of the previous scientists yielded a standard foundation of GxE analysis in terms of a genetic correlation or genetic variance heterogeneity that was formalized in Robertson's (1959) seminal paper (Figure 2).

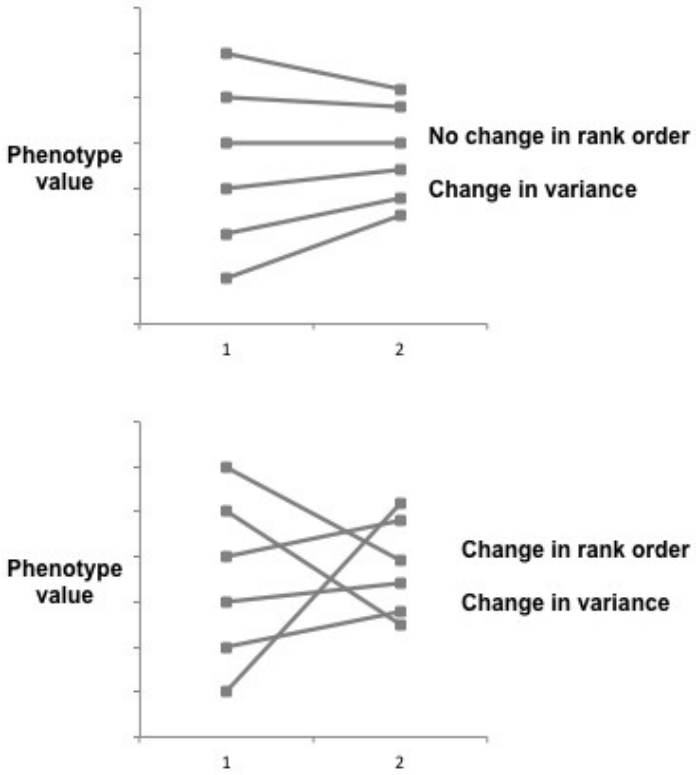
Figure 2. Genotype by Environment interaction components.



Further, Robertson (1959) suggested that interaction involves: (i) ranking changes between environments; (ii) changes in the magnitude of genetic and environmental effects between environments; and (iii) changes in phenotypic variance. Also, Bowman observed that this ranking and variance changes might occur at the same time or separately (Figure 3).

Figure 3. Robertson's Genotype by Environment interactions proposal [adapted from Bowman (1972)]





A similar method was later proposed by Dickerson (1962) when applying a two-way ANOVA approach to the analysis of GxE. The novelty of Dickerson's method is related to the possibility of testing the influence of many environments (i.e., classes), providing an estimate of the average degree of genetic correlation. This topic was subsequently addressed by Yukio Yamada (1962), who tried to elucidate which statistical model, random or mixed, should be used in the analysis and the potential differences between both approaches. Yamada also addressed the violation in the assumption of variance heterogeneity between environments.

More recently, GxE statistical analysis has become increasingly complex as incredible research instruments and tools are now available, namely those related to DNA analysis[4]. For instance, John Blangero (Blangero, 2009, p. 524) presented a "*model of quantitative trait variation that includes the effects of a*

[4] Some of these techniques, such as linkage studies, are family-data dependent, but some, such as genome wide association scans (GWAs) and association studies are not dependent upon biological relatives data.

single unknown major gene (MG), polygenes (PG), known (and therefore measured) environmental factors (E), and random environmental factors (e)".

1.2 NUCLEAR FAMILY DESIGN

Nuclear family studies are based on the relationships between biological relatives of distinct generations (Visscher et al., 2008). They are helpful in differentiating between genetic and environmental factors governing the variance found in many phenotypes namely the ones used in this thesis: physical activity/energy expenditure, metabolic syndrome and body composition indicators. This is feasible because of Ronald A. Fisher's seminal paper written in 1918 (Fisher, 1918), in which he resolved the apparent dispute between the Mendelians and Biometricians, and laid the fundamental ideas for the variance components approach. It is well known that in biological relatives genes will in most cases follow the Mendelian laws. This means that, on average, first-degree biological relatives, as in those relationships occurring in a nuclear family (except for the marital relationship), share 50% [5] of their genetic background (or 100% in the case of monozygotic twins) and share, generally, the same family environment.

1.2.1 Familial Correlations and Familial Aggregation

It is widely accepted that familial aggregation refers to the clustering of traits within a given family (Naj et al., 2012). A trait correlation between spouses and their siblings, or between siblings, provides a measure of shared inherited (biological and environment) information that influences the expression of a trait. A trait manifestation is influenced by both the genetic make-up of an individual and its environmental context. For example, a correlation of 0.35 between parent-offspring and/or siblings means that there is a trait similarity of 35% that is explained by the joint effects of shared genes [6] and shared environment (Falconer & Mackay, 1996). Despite its usefulness in describing familial similarities intra traits and/or between traits (Keen & Elston, 2003; Rao et al., 1987; Tregouet et al., 1999), familial correlations fail to inform about the

[5] This may not be totally true as DNA may suffer mutations during life.
[6] Gene – Very loosely defined, a gene is a DNA sequence unit that occupies a specific location on a chromosome and determines a particular characteristic in an organism.

individual sources of such similarity, i.e., do not allow a description of the magnitudes of genetic and environment influences on a trait.

1.2.2 Heritability Estimates

Heritability is a summary measure, always sample/population specific, of the genetic effects on a given trait (Bochud, 2012), and is estimated from twin and/or familial data. The computation of this quantity represents the second step in the genetic analysis of a trait, and its formulae are based on variance components or correlations. Heritability is defined as the amount of variation in any given trait that is due to genetic factors (Falconer & Mackay, 1996). This means that a non-zero heritability is mandatory for the "*detection of genes underlying a phenotype of interest*" (Bochud, 2012, p. 171), and that the correlation between a phenotype and genotype increases with increasing heritability estimates (Griffiths et al., 2000).

Heritability estimates provide relevant information in helping to disentangle the pathophysiology of human behaviours and/or diseases, as they offer a comparison between the relative contribution of genes and environment to the phenotypic expression of any trait (Bochud, 2012). However, heritability *per se* does not provide any information regarding the specific gene, or genes, involved in the regulation of a phenotype, which means that the genetic architecture of a trait cannot be depicted solely by heritability estimates.

1.2.3 Heritability Estimates in nuclear family designs

The nuclear or extended family designs are well-suited approaches to estimate heritability under the usual assumption that the total phenotypic variance of a quantitative trait (σ^2_T) is represented by the well-known decomposition $\sigma^2_T = \sigma^2_G + \sigma^2_E$, in which σ^2_G represents the genetic component and σ^2_E represents the environmental component. This corresponds to a crude interpretation of how phenotypic expression occurs, not accounting for gene by environment interactions, gene-environment correlations and gene-gene interactions. The genetic component comprises the additive genetic variance (σ^2_A), dominance variance (σ^2_D), and an epistatic variance (σ^2_I), such that $\sigma^2_G =$

$\sigma_A^2 + \sigma_D^2 + \sigma_I^2$. The environmental variance comprises the joint effects of the known and the unknown environmental factors associated with the trait, as well as measurements errors. These give rise to different heritability estimates, namely broad and narrow sense: broad sense heritability (H^2): $H^2 = \frac{\sigma^2_G}{\sigma^2_T}$; narrow sense heritability (h^2): $h^2 = \frac{\sigma^2_A}{\sigma^2_T}$. H^2 is the ratio of additive, dominant, and epistatic variance to the total phenotypic variance and refers to all the genetic contributions to a population's phenotypic variance; h^2 is the ratio of only the additive genetic variance to the total phenotypic variance, and represents the genetic information that is transmissible from one generation to the next. Each parent transmits a single allele[7] per locus[8] to each offspring. Thus, parent-offspring resemblance is the result of the average effect of single alleles.

Depending on the purpose of the research, variance components vary in their importance. For example, if the foremost question is related to transmissible information from one generation to the next, then the additive genetic variance is of utmost importance (Falconer & Mackay, 1996). On the other hand, if the question is related with the nonlinear interaction effects between alleles at the same locus then the dominance variance is of interest (Hill et al., 2008). The interaction effects between alleles at different loci[9] are estimated by the epistatic variance (Crow, 2010; Diego et al., 2013).

One key aspect in calculating heritability estimates from family designs is the assumption of random mating. Even though this is, generally observed, there are some cases that need careful attention given that interaction processes, social homogamy and phenotypic assortment may occur[10]. For instance, when studying height, it is important to acknowledge that, on average, people tend to feel attracted to people of the same height (Stulp et al., 2013),

[7] Alleles – is one of a number of alternative forms of the same gene or same group of genes.
[8] Locus – gene or genetic marker location in a chromosome.
[9] Loci – plural of locus.
[10] An interaction process, or influence process occurs if the correlation coefficient between spouses, in a given trait, increases with co-habitation increasing years; social homogamy refers to incidental resemblance between spouses given a known cultural or economic background; phenotypic assortment, or selection process occur when mating is linked to some known characteristic.

and that can lead to some bias[11] in the results. Another relevant aspect of heritability estimates is that they are population-specific (Diego et al., 2013; Hopper, 1993). Accordingly, it is possible to observe significant variations in the phenotypic variance across populations that are not only the result of the environment effects but also reflect additive genetic variance differences. Equally important is that heritability estimates are time limited and refer to a specific point in time (Visscher et al., 2008). This is why follow-up studies are of great assistance at unravelling the interrelationship between genes and the environment, as they allow for the monitoring of environmental changes that can lead to different heritability estimates in different points in time, even though no changes in the additive genetic variance is observed (de Andrade et al., 2002).

2. Key concepts

2.1 Physical Activity

The first key concept of this thesis is Physical Activity (PA). It is the most variable component of an individual's total daily energy expenditure (TDEE) – $TDEE = BMR + TEF + PA$ in which, BMR is the basal metabolic rate and refers to the energy needed to maintain body activities during rest; and TEF is the thermic effect of food and refers to the energy required to digest food (Dishman et al., 2004).

Physical activity is a broad term with a complex aetiology, both genetically and environment driven, and historically defined as "*any bodily movement produced by skeletal muscles that results in energy expenditure*" (Caspersen et al., 1985, p. 126). This definition allows for a wide-range of interpretations and uses, and it includes activity on the job [work physical activity (WPA)], activity in domestic affairs and transportation, and activity during free time [leisure time physical activity (LTPA) which may include sports activities and exercise]. Physical activity is often mistaken with two concepts that are closely linked:

[11] Bias – is a general statistical term meaning a systematic (not random) deviation from the true value.

1. Exercise – refers to a particular type of activity that has been defined as "*physical activity that is planned, structured, repetitive and purposive in the sense that improvement or maintenance of physical fitness or health is an objective*" (*Exercise, fitness and health a consensus of current knowledge*, 1990).
2. Physical Fitness (PF) – is a multi-dimensional concept that has been defined as "*the ability to carry out daily tasks with vigour and alertness, without undue fatigue and with ample energy to engage in leisure time pursuits and to meet the above average physical stress encountered in emergency situations*" (Clarke, 1979).

Physical activity is categorized by frequency, duration, intensity, and circumstances or purpose of the activity.

2.1.1 Measuring Physical Activity

Measuring PA has been a major challenge to researchers worldwide. It is not within the scope of this thesis to present a full description and explanation of all the methods available [for details see Dishman (2004)]. However, since PA is a major variable of the present thesis, here we present a summary of the pros and cons of some PA measurement methods that will, hopefully, justify our option for questionnaire data.

There are several available methods to measure PA, and are usually divided into direct calorimetry, indirect calorimetry, doubly labelled water, direct observation, objective and subjective methods (Dishman et al., 2004) :

- Direct calorimetry – quantifies energy expenditure by measuring the heat production of an individual in a sealed, insulated chamber.
- Indirect calorimetry – estimates the energy expenditure from oxygen consumption and carbon dioxide production.
- Doubly labelled water method – is the gold standard measure of total daily energy expenditure (Schoeller, 1999). It consists of the ingestion of a quantity of water labelled with a known concentration of naturally occurring, stable isotopes of hydrogen and oxygen. The body produces carbon dioxide and water as a result

of energy expenditure, and the differences between the isotope elimination rates are used to calculate total energy expenditure.

- Direct Observation – an investigator observes and records the activities of an individual during a period of time.
- Objective methods – includes monitoring devices such as pedometers and accelerometers that are capable of recording the frequency and intensity of human movements.
- Subjective methods – includes questionnaires, diaries, and 'parent proxy' questionnaires in which an individual is asked to recall and report participation in daily activities in a certain period or periods.

Despite being the most precise procedures for energy expenditure quantification, direct and indirect calorimetry, and doubly labelled water methods are not suited for epidemiological research as they involve expensive and/or unpractical materials, and have a high cost in financial terms. As such, we shall give our attention to the methods mostly used in nuclear family designs and, in particular, to the method used in this research.

Direct observation of an individual activity is the most reliable method to recall patterns and durations of PA. It has however major downfalls – data collection duration, and number of individuals assessed. In this method a researcher observes and reports the activities of an individual (or small group of individuals) over a period of time. This can be done directly or by image recording, but in both cases it is mandatory that all activities of interest be reported. As such, it is not possible to collect data from a substantial number of individuals at the same time, which makes this method an unfriendly approach to epidemiological research (Trost, 2007). For instance, direct observation is a valuable instrument if an investigator is interested in analysing patterns of activities during recess in school and its association with peer's acceptation (Huberty et al., 2013).

Direct measurement of PA by means of motion sensors such as pedometers and/or accelerometers has been increasingly used worldwide (Katzmarzyk et al., 2013; Mark & Janssen, 2009; Ness et al., 2007; Owen et al.,

2010; Rowlands, 2007; Steele et al., 2009). A pedometer is an instrument that records the number of steps of an individual. Recent models also allow for the quantification of steps in aerobic activity (e.g., Omron HJ-112), i.e., the number of steps at higher intensities. The main concerns related to the use of pedometers have to do with bias on daily activities, as individuals might feel encouraged to walk more if they are aware that they are not fulfilling current guidelines. Also, generally, pedometers only return information on the number of steps, not on the intensity or frequency, failing to contextualize the patterns of PA (Rowlands et al., 1997). Accelerometers are instruments that measure body acceleration and that indirectly estimate energy expenditure (Rowlands & Eston, 2007). Contrary to pedometers, accelerometers provide a measure of intensity and frequency of movement. However, there are still some limitations, namely their cost, their unsuitability for under-water activities and their incapability of recording information derived from static activity or activities in which the centre of gravity remains practically unmoved (e.g., cycling) (Rowlands & Eston, 2007).

Questionnaires are useful tools in large-scale studies as they are very cost- and time-efficient. They provide information on frequency (e.g., how many times did you played sports during last week?), intensity (e.g., were you exhausted at the end of those activities?), duration (e.g., how many minutes did you walk to go to school?) and circumstances (e.g., do you ride a bicycle to go to school?). Generally, questionnaires allow for the estimation of energy expenditure attributable to distinct types of PA (e.g., energy expenditure in sports activities). Questionnaires have, however, some limitations that need to be acknowledged. Firstly, they are highly vulnerable to individual's beliefs about PA. For instance, if a subject knows that it is important for his/her health to go to the gym, then he/she might overestimate the number of times that he/she has actually gone to the gym. Also, often it is quite difficult to quantify the amount of time involved in different activities, whether being related to memory lapses or to incapacity of recognizing the activity. For example, asking an individual how many times did he spent in moderate-to-vigorous activities (MVPA) might represent an issue as it can be hard to understand what the question is referring

to. Questionnaires also need to be population specific as there are cultural differences that need to be addressed when applying a questionnaire. For example, there is no need to ask 10 yr-old Portuguese children if they are involved in Physical Education classes since they are mandatory. However this is not true for other countries.

Diaries are a particular type of instrument that have a major merit: robust to memory lapses. As each individual completes them every day, the probability of misreporting is dramatically reduced. Still, they imply an effort from the participants that is not always easy to accomplish.

'Parent proxy' questionnaires refer to parental reports about their children's behaviours'. However, it has been shown that parents are often mistaken about their children's behaviours. For example, Robinson et al. (2006, p. 1) analysed the validity of parent proxy reporting of children's TV viewing and found that "*parents overestimate their child's television time compared to an objective measure when no television is present in the bedroom by 4 hours/week… in comparison to underestimating television time by over 3 hours/week… when the child has a television in their bedroom*".

2.1.2 Genetics of Physical Activity

The genetic foundations of physical activity have been explored mainly during the last three decades, and more specifically since the beginning of this century (Rankinen et al., 2010).

Familial aggregation studies have shown that PA h^2 estimates range from 0% (Perusse et al., 1989) to 60% (Butte et al., 2006). Twin studies yielded results even more disperse with additive genetic effects ranging from 0% (Joosen et al., 2005; Stubbe et al., 2005) to 85% (Stubbe et al., 2005). As for shared environment, the results are also quite different with 0% (De Moor et al., 2007b; De Moor et al., 2007c; Duncan et al., 2008; Eriksson et al., 2004; Joosen et al., 2005; Spinath et al., 2002) to 84% (Stubbe et al., 2005) of the variance explained. The same trend is observed for unique environment with and explained variance of 12% (Koopmans et al., 1994) to 72% (McCaffery et al., 2009). Suggestive linkages have been found with markers nearby different

activity-related genes: *EDNRB (Simonen et al., 2003b)*, *MC4R (Cai et al., 2006)*, *UCP1 (Simonen et al., 2003b)*, *FABP2 (De Moor et al., 2007b)*, *CASR (De Moor et al., 2007b)* and *SLC9A9 (De Moor et al., 2007b)*. Moreover, significant associations with PA phenotypes were found for *CASR* (Lorentzon et al., 2001), *ACE (Fuentes et al., 2002)*, *LEPR (Stefan et al., 2002)*, *MC4R* (Loos et al., 2005) and *DRD2 (Simonen et al., 2003a)* genes. Thus far, only one GWAs (De Moor et al., 2009) has studied PA and reported novel SNPs in the *PAPSS2* gene on chromosome *10q23.2* and in two intergenic regions on chromosomes *2q33.1* and *18p11.32*.

This topic will be further explored in the Research Papers´ chapter.

2.1.3 Epidemiology of Physical Activity

The major concern of this thesis is to investigate the relationship between PA and two health indicators: the Metabolic Syndrome (MetS) and Body Composition (BC). Here we present a short summary of the relationship between PA and health[12] in general, since MetS and BC associations with PA will be further explored in sections 2.2.3 and 2.3.3, respectively.

Reports about the importance of physical activity to health status tracks back to $2{,}500_{B.C.}$ in China, where physical exercise was recommended to promote health (Lyons & Petrucelli, 1978). This was also illustrated and eloquently postulated by the great Greek philosopher Hippocrates ($460\text{-}377_{A.C.}$) in his document *On Regimen in Acute Diseases* (Hipocrates): "*the sick will of course profit to a great extent from gymnastics with regard to the restoration of their health and the healthy will profit with regard to its maintenance, and those who exercise will profit with regard to the maintenance of their well-being and a lot more*" (Kritikos et al., 2009).

It has been, however, during the past 3 decades that PA epidemiological research has become a central focus of interest (Paffenbarger et al., 2001), with the advent of modern westernized societies, characterized by low levels of

[12] Health can be defined as a "*human condition with physical, social, and psychological dimensions, each characterized on a continuum with positive and negative poles. Positive health is associated with a capacity to enjoy life and to withstand challenges; it is not merely the absence of disease. Negative health is associated with morbidity and, in the extreme, with premature mortality*" (Exercise, fitness and health a consensus of current knowledge, 1990, p. 6).

energy expenditure and easy access to high caloric food that, together, are contributing to the obesity epidemic and associated morbidities ("Diet, nutrition and the prevention of chronic diseases", 2003).

The very first study in the modern era to address the topic of PA associations with health status (Morris et al., 1953), applying classic epidemiological approaches to investigate the prevalence of coronary heart disease (CHD) among London's bus conductors and drivers, was conducted by Dr. Jeremy Morris. Dr. Morris aimed to compare CHD prevalence between highly active conductors on London's double-decker buses and the drivers that spent most of the time at work seated. The results showed that conductors were at a lower risk of developing CHD than the drivers, and this was attributable to the PA differences between both occupations.

The London bus study was the "Big Bang" of Physical Activity Epidemiology and since then a large body of evidence has been emerging for the health benefits of having an active lifestyle (Bouchard, 2001). There is currently no agreed upon amount and intensity of PA that is necessary to have a protective effect on individual health. Moreover, the biological mechanisms that mediate the effects of PA on health have not been established (Gielen et al., 2010b).

Consensus efforts have focused primarily on the improvement and maintenance of physical fitness through PA ("American College of Sports Medicine position statement on the recommended quantity and quality of exercise for developing and maintaining fitness in healthy adults", 1978). These very first recommendations were of high intensity activities, very demanding and unaccommodating to sedentary people, who were consequently discouraged to become active (Wijndaele, 2007).

In the 1990's, and building on the increasing evidence supporting the beneficial effects of engaging in regular activities of moderate intensity, distinct European and US health related institutions recommended accumulating 30 minutes or more of moderate-intensity PA on most, if not all, week days for adults ("Nutrition and Diet for Healthy Lifestyles in Europe: Science and Policy

Implications. Proceedings of the European Conference. May 18-20, 2000. Crete, Greece", 2001). The recommendation for children and youth has been extended to 60 minutes per day (Strong et al., 2005).

These recommendations mainly target the involvement in regular physical activity of moderate intensity and increase the chance that these PA levels will have a preventive effect on developing metabolic and cardiovascular disorders. One key aspect noted by these institutions is the need to incorporate PA in daily routines, and to be active throughout the day[13]. It has also been suggested that greater amounts of PA, during long periods of time or of vigorous intensity, lead to additional health benefits.

2.2 METABOLIC SYNDROME

Metabolic syndrome (MetS) is a cluster of metabolic abnormalities characterized by glucose intolerance, hypertension, dyslipidaemia and obesity that co-occur in individuals more often than might be expected by chance, and has been associated with increased risk of type 2 diabetes (Bjerregaard-Andersen et al., 2013; Hsu, 2013), cardiovascular morbidity (Chen et al., 2006; Ivezic-Lalic et al., 2013), cardiovascular mortality (Ghaem Maralani et al., 2013; Lakka et al., 2002), and all-cause mortality (Baber et al., 2013; Jarrett et al., 2013).

The concept of MetS is the result of an ensemble of evidence suggesting that abnormalities in the metabolism of glucose, lipids and blood pressure are closely associated with insulin resistance[14]. It was firstly known as "*Syndrome X*" (Reaven, 1988), "*the deadly quartet*" (Kaplan, 1989), or the "*insulin resistance syndrome*" (DeFronzo & Ferrannini, 1991).

Although there is evidence that the MetS concept goes back to the 1920's (Cameron et al., 2004), its pathogenesis remains complex and vaguely understood thus far (Alberti et al., 2006). The core idea that is fundamental to

[13] The protective effect of having a non-sedentary daily routine has been linked to the NEAT – non exercise activity thermogenesis. For more details see Levine's work (Levine, 2002).
[14] Insulin resistance – Is a physiological condition in which cells fail to respond to the normal actions of the hormone insulin. Insulin plays an important regulatory function in the delivery of glucose into cells to provide them with energy. Insulin resistant cells cannot take in glucose, amino acids and fatty acids, which leads to a decrease in insulin/glucagon ratio. This inhibits glycolysis, leading to decreased energy production. The resulting hyperglycaemia may cause adverse health effects.

understanding MetS is the insulin resistance (IR) concept. This was first studied by Himsworth (1936) who observed differences in the insulin sensitivity of diabetes mellitus (DM) patients, which resulted in the recognition of two types of diabetes mellitus. IR has been associated with atherogenic dyslipidaemia (Ferrannini & Iozzo, 2006) and a proinflammatory state (Garg et al., 2003). However, the links between IR and most of the MetS indicators remain to be clarified and further research is needed to examine such associations (Alberti et al., 2006).

The other key idea that is relevant to the etiology of MetS is the recognition of two types of obesity: (i) android, in which fat accumulation occurs mainly in the abdominal area, and (ii) the gynoid, in which fat accumulation is well distributed in the whole body, particularly in the lower limbs (Vague, 1947). Obesity associations with hypertension, low high-density lipoprotein cholesterol (HDL-C), insulin resistance and hyperglycaemia are well studied (Schmidt et al., 1996). Vague's recognition is of utmost importance since it has been shown that among individuals with normal body mass index (BMI), but showing excess in central adiposity, the MetS features are present (Carr et al., 2004; Despres, 2006; Despres & Lemieux, 2006). This allows for the speculation that visceral adipose tissue is a foundation of numerous molecules that induce insulin resistance, such as nonesterified fatty acids (Iannello et al., 1998) and tumour necrosis factor-α (Aminzadeh et al., 2009). Present knowledge has identified other factors that influence MetS such as genetic background, sedentary behaviour, ageing, a proinflammatory state and hormonal deregulation (Potenza & Mechanick, 2009).

2.2.1 The definitions of the Metabolic Syndrome

A consensus definition of MetS has been a challenging endeavour (Alberti et al., 2009). The most common definitions were provided by the World Health Organization (WHO) (Alberti & Zimmet, 1998), the National Cholesterol Education Program Adult Treatment Panel III (NCEP-ATP) (Expert Panel on Detection & Treatment of High Blood Cholesterol in, 2001), and IDF (Alberti et al., 2006), as presented in Table 1.

Table 1. Metabolic syndrome most common definitions.

		Components										
		IR	BMI		TG		HDL-c		BP	Glu	WC		Alb
			♂	♀	♂	♀	♂	♀			♂	♀	
Definition	WHO[15]	4th quartile or DM2 or FGI or GI (normal glycaemia 110mg/dl or 6.1mmol/l)	> 30 and/or WHR > 0.9	> 30 and/or WHR >0 .85	> 150mg/dl (1.7mmol/L)		< 35mg/dl (0.9mmol/L)	< 39mg/dl (1.0mmol/L)	Under medication or SBP > 140mmHg or DBP > 90mmHg				Urinary excretion rate ≥ 20mg/min or Alb/Cre ≥ 30mg/g
	NCEP-ATP[16]				> 150mg/dl (1.7mmol/L)		< 40mg/dl (1.03mmol/L)	< 50mg/dl (1.3mmol/L)	≥ 130/85mmHg	≥ 110mg/dL (6.1mmol/L)	> 102cm	> 88cm	
	WHO (modified version)[17]	Hyperinsulinaemia: 4th quartile, DM2 or FGI	> 30 and/or WHR > 0.9	> 30 and/or WHR > 0.85	> 150mg/dl (1.7mmol/L)		< 35mg/dl (0.9mmol/L)	< 39mg/dl (1.0mmol/L)	Under medication or SBP > 140mmHg or DBP > 90mmHg		> 94cm	> 80cm	
	IDF[18]				> 150mg/dl (1.7mmol/L)		< 40mg/dl (1.03mmol/L)	< 50mg/dl (1.3mmol/L)	≥ 130/85mmHg	≥ 110mg/dL (6.1mmol/L)	Ethnic specific		

Legend: IR – insulin resistance; BMI – body mass index; TG – triglycerides; HDL-c – high density lipoprotein cholesterol; BP – blood pressure; Glu – glucose; WC – waist circumference; Alb – albumin; Cre – creatinine; DM2 – type 2 diabetes mellitus; FGI – fasting glucose intolerance; GI – glucose intolerance; WHR – waist-to-height ratio; SBP – systolic blood pressure; DBP – diastolic blood pressure

2.2.2 Genetics of Metabolic Syndrome

Metabolic Syndrome has been shown to be regulated to some extent by genetic effects. For instance, Vattikuti et al. (2012) found that 30 to 48% of the variance in MetS is due to genetic effects in The National Heart, Lung, and Blood Institute (NHLBI) Study; heritability estimates (h^2) varied from 0.33 (diastolic blood pressure) to 0.63 (HDL-cholesterol) (Tang et al., 2006). In this sample, significant familial intra-trait correlations were observed between parents and offspring for diastolic blood pressure (DBP), HDL-cholesterol (HDL) and fasting insulin (INS), supporting the hypotheses of shared effects (genetic and/or environmental) on these MetS indicators (Tang et al., 2006). Suggestive

[15] WHO definition – MetS is present if an individual is at risk for IR and for at least two of the other components.
[16] NCEP-ATP – MetS is present if an individual is at risk in three or more components.
[17] WHO definition (modified version) – MetS is present if an individual is at risk for IR and for at least two of the other components.
[18] IDF – MetS is present if an individual is at risk for WC and for at least two of the other components.

linkages have been found on chromosome *7q21* (Farook et al., 2012) and chromosome 1 (169.5–181.5 cM) (Ng et al., 2004) as well as candidate gene *CD36* and its flanking gene *GNAT3* seem to be associated with MetS.

This topic will be further explored in the Research Papers chapter and here it is only presented to provide a better understanding of our research project.

2.2.3 Metabolic Syndrome Epidemiology

The use of different MetS definitions and cut-off values have led to some differences in its observed prevalence across world regions (Boronat et al., 2005; Ford & Giles, 2003), which has then an influence in their associations with T2DM and cardiovascular disorders (Laaksonen et al., 2002). For instance, in the USA, in the *Third National Health and Nutrition Examination Survey* (Ford et al., 2002) conducted between 1988 and 1994, the prevalence of MetS in adults was 23.9% and 25.1% using NCEP and WHO definitions, respectively.

In a follow-up study in France (Guize et al., 2006), baseline MetS (IDF) prevalence in people over 40 years of age was 11% and 7.2% for men and women, respectively. After 3 years, the prevalence was 12.8% and 8.8% for men and women, respectively. Using NCEP-ATPIII criteria, a large Portuguese Hospital adult sample (age range: 18-96 years) showed a MetS prevalence of 27.5% (Fiuza et al., 2008), which is lower than the 34% found in the most recent USA data (Ervin, 2009), but higher than the French data reported above.

MetS prevalence increases with age (Ervin, 2009; Fiuza et al., 2008). For instance, in the Portuguese population, the MetS prevalence is 7.5% and 4.9% at 18-29 years and 36.5% and 39.2% at 50-59 years, for males and females, respectively (Fiuza et al., 2008). During adolescence, and using a specific IDF definition, the prevalence of MetS in the USA was of 4.5% (Franks et al., 2007). However, there were differences between boys (6.7%) and girls (2.1%), and between ethnic groups since the prevalence among Hispanics was 7.1%, Blacks was 3.0% and Whites was 4.5%. These differences have been attributed to cultural aspects (namely, nutritional habits and physical activity levels), and

socio-economic status (Katzmarzyk & Herman, 2007; Martinez-Gonzalez & Martin-Calvo, 2013; Zhan et al., 2012).

2.2.4 Metabolic Syndrome and Physical Activity

The relationship between individual indicators of the MetS and PA have been widely studied (Katzmarzyk & Herman, 2007; Strasser, 2012) and will be further addressed in this thesis. Here we present a very brief data summary, given that it will be thoroughly discussed in the Research Papers´ chapter.

Most of the studies conducted to analyse the association between these two phenotypes suggest that PA has a protective effect on the risk of developing MetS, or being at risk in any of the syndrome's components (Carroll & Dudfield, 2004; Kaino et al., 2013; Katzmarzyk & Herman, 2007) However, a clear delimitation of the type and amount of PA that produces significant effects in MetS is still needed (Laursen et al., 2012). This is to say that as moderate-to-high PA levels were found to have a preventive effect on the development of MetS (Katzmarzyk & Herman, 2007; Strasser, 2012), everyday low energy expenditure activities such as walking and cycling have been linked to negligible effects on MetS (Hahn et al., 2009).

2.3 BODY COMPOSITION

Body composition (BC) is the third key concept of this thesis and it refers to the body components that make up an individual's body mass. Diverse methods, instrumentation, and data analysis techniques enable the quantification of changes at the tissue level (most often fat mass and fat free mass) as a result of daily routines, such as physical activity and nutritional habits (Heymsfield, 2005).

The study of body composition[19] in epidemiological terms is a topic of great interest and a great challenge worldwide because of the obesity pandemic that is affecting both developed and most undeveloped countries ("Diet, nutrition and the prevention of chronic diseases", 2003). For epidemiologists, BC is an important screening tool of a population's health status, because of the known

[19] A thoroughly description of the history, methods, techniques, instrumentation and data analysis techniques related to body composition is provided in Heymsfield (2005).

associations between excess fat mass and cardiovascular disorders, in particular, and morbidity and mortality, in general (Heitmann et al., 1997; Lahmann et al., 2002).

As such, this section will briefly describe the definitions of obesity and overweight and its prevalence in different regions of the world. Some evidence of genetic regulation of BC and its association with physical activity will also be provided.

2.3.1 The definitions of Obesity and Overweight

Overweight and obesity are the result of a "caloric imbalance" as too few calories are being expended in relation to the amount of calories consumed (Daniels et al., 2005). Overweight is usually defined as having excess body weight (from fat, muscle, bone, water, or a combination of these factors) for a given height (National Institutes of Health, 2010), and obesity has been defined as having excess body fat (Krebs et al., 2007).

The vast majority of epidemiological studies rely on body mass index (BMI) to classify an individual as being overweight or obese. For adults, overweight is defined as BMI above 24.99 and lower than 30 kg/m^2. Obesity is defined as BMI at or above 30 kg/m^2. Obesity can be classified as class I (BMI >29.99 and <35 kg/m^2), II (BMI >34.99 and <40 kg/m^2) or III (BMI >39.99 kg/m^2). However, there are other indicators, such as percentiles for %fat mass or waist circumference that can be used to help define overweight and obesity.

As for childhood and adolescence, it has been difficult to find well-established criteria that relate whole-body fatness and its health outcomes (Krebs et al., 2007). Generally, cut-off points based on anthropometric measurement distributions (e.g. weight and BMI) have been used (Cole et al., 2000; Kuczmarski et al., 2000).

2.3.2 Genetics of Body Composition

Body composition has been found to be under significant genetic effects (Bellia et al., 2009; Butte et al., 2006; Jelenkovic et al., 2011; Luke et al., 2001; Mathias et al., 2009; Poveda et al., 2012). For instance, h^2 estimates of 38%

and 39% have been found for waist circumference in the Linosa Study (Bellia et al., 2009), and in a Spanish family study (Jelenkovic et al., 2011), respectively. Further, BMI heritability estimates yielded from twin studies range from 0.47 to 0.90 and from family studies range from 0.24–0.81 (Elks et al., 2012). As for body fat, different heritabilities have been found in distinct regions of the world. For instance, in Sweden, Wagner et al. (Wagner et al., 2013) found that the variance accounted by genes in body fat expression is 69%, whereas in Nigeria, Jamaica, and USA the percentages were 48%, 54%, 57%, respectively (Luke et al., 2001).

Moreover, complex DNA analysis, such as genome-wide association (GWA) studies identified several genetic factors associated with obesity both in childhood and adulthood, such as *FTO* and *MC4R* genes variants (Frayling et al., 2007; Loos et al., 2008). And the famous review paper by Rankinen et al. (2006, p. 1), states that "*176 human obesity cases due to single-gene mutations in 11 different genes have been reported, 50 loci related to Mendelian syndromes relevant to human obesity have been mapped to a genomic region, and causal genes or strong candidates have been identified for most of these syndromes*".

2.3.3 Overweight and Obesity Epidemiology

Since 1980, the worldwide obesity prevalence has doubled, and in 2008 more than 200 million men and nearly 300 million women were obese (Booth et al., 2008). This means that 35% of adults were overweight in 2008, and 11% were obese. Even though this was a major problem in developed countries, the truth is that this pandemic has been affecting developing countries as well and, in fact, 65% of humans live in countries where overweight and obesity are more responsible for the mortality rates than underweight (Booth et al., 2008).

In the United States, over 60% of the population is either overweight or obese (Wyatt et al., 2006). Moreover, in the United States, the obesity prevalence among children aged 6–11 years increased from 7% in 1980 to nearly 18% in 2010. Similarly, the obesity prevalence among adolescents aged 12–19 years increased from 5% to 18% over the same period (Ogden et al.,

2012). This is a major concern because of the known tracking of obesity from youth to adulthood (Freedman et al., 2005) and the associations between childhood obesity and health hazards during adulthood (Freedman et al., 2001).

In Portugal, a recent review (Carreira et al., 2012) has identified that between 1995 and 2005, the overweight prevalence increased by 3.2 % and 3.5 % and obesity prevalence by 7.4 % and 1.3 % among women and men, respectively. The overweight prevalence in Portugal is 38.6% and obesity prevalence is 13.8%, which leads to a total of 52.4% of individuals with excess weight (do Carmo et al., 2006).

2.3.4 Body Composition and Physical Activity

This topic has been extensively documented and here we present only a very brief contextualization [for extensive details please see (Heymsfield, 2005)].

Overweight and obesity have been highly associated with PA levels and patterns (Ballor & Keesey, 1991; Franz et al., 2007; Jakicic & Otto, 2005). Generally, %fat mass and overweight/obesity tend to decrease with increasing levels of PA and reduced sedentary time (DiPietro, 1999). Also, physical activity plays a role in the preservation of lean body mass (Poirier & Despres, 2001). Furthermore, there is strong evidence for a dose-response relationship between physical activity and weight loss as the rate of weight loss is dependent upon volume and intensity of physical activities (Ballor & Keesey, 1991; Jakicic & Otto, 2005; Poirier & Despres, 2001).

3. Problem analysis

This brief introduction aims only to help us in framing the main question we are trying to answer: **are the associations between PA and MetS and BC genetically "driven"?**

It is undoubtedly clear that both MetS and BC are complex traits, both are influenced by genetic and environmental effects, and both are influenced by PA levels. However, a clear understanding of the mechanisms that regulate PA

influence on MetS and BC is yet to be established. Although the present thesis will not solve this issue, we bring some new information and detail to this understanding using a nuclear family design. As such, we hypothesize that these mechanisms might be regulated to some extent by an interaction between an individual's genotype and its PA routines.

This topic is of wide importance since researchers and policy makers are trying to deal with the escalation of the obesity epidemic and the increases of sedentarism and their influence in MetS all over the world. For example, Cawley and Meyerhoefer (2012) estimate that obesity accounts for 21 percent of medical spending ($190 billion in 2005) and it is expected that if obesity trends keep its tendency, by 2030, obesity-related medical costs alone could rise by $48 to $66 billion per year in the USA (Wang et al., 2011). This means that, because of the association between obesity, cardiovascular disease and T2DM, MetS will probably have a large burden on world economics (Reynolds & He, 2005).

This scenario is a clear demonstration that a new paradigm in intervention programs is needed at the population level. However, such interventions are doomed to failure if a more sophisticated level of knowledge about the connecting paths between genes, physical activity, MetS and BC is not developed and established. In the present research, we used a family design in order to:

1. Quantify the similarity between relatives in PA, MetS and BC using familial correlations. If these traits are closely related within families this means that there are familial particularities that determine the differential expression of PA, MetS and BC.
2. Compute heritability estimates of MetS, BC and PA, which will provide a measure of the genetic relatedness on these traits' expression.
3. Analyse the importance of shared environment in the expression of PA and MetS. A correct understanding of the family dynamics that lead to MetS is necessary to develop and implement adequate

family-based intervention programs. This is why it is so important to verify the importance of shared-environment on determining PA and MetS.

4. Compute the correlations among the 5 components of MetS and identify possible sources of such correlations. If the MetS components are correlated then this would mean that its ”sources”, whether genetic or environmental, might be shared. This is a relevant concern when studying the MetS because shared genetic or environmental effects respectively imply common causal genotypes or environments.
5. Calculate the effect of the interaction between PA and genotype in the expression of MetS and BC. These will enhance our knowledge about a critical concern in PA epidemiology: why is there inter-individual variability in response to physical activity? In this case, if an interaction between genotype and PA (GxPA) that influences MetS and BC is observed, it may help us to better understand why some people become overweight or obese despite being active and some do not have MetS despite being inactive.

Some of the previous topics have already been addressed in well-known studies such as the Québec Family Study (Perusse et al., 1997) or the San Antonio Family Heart Study (Mitchell et al., 2003). Still, the main question is still unresolved: *why do people react differently to the same PA exposures?*

Present knowledge about the genetic influence on PA is not only limited, as it is often not in accordance. For example, PA heritability estimates ranges from 0 to 90% depending on the phenotype studied, the estimation method, the sample, the study design and the statistics applied. Furthermore, information about the genetic adaptation to PA exposures is scarce and non-existent in family studies.

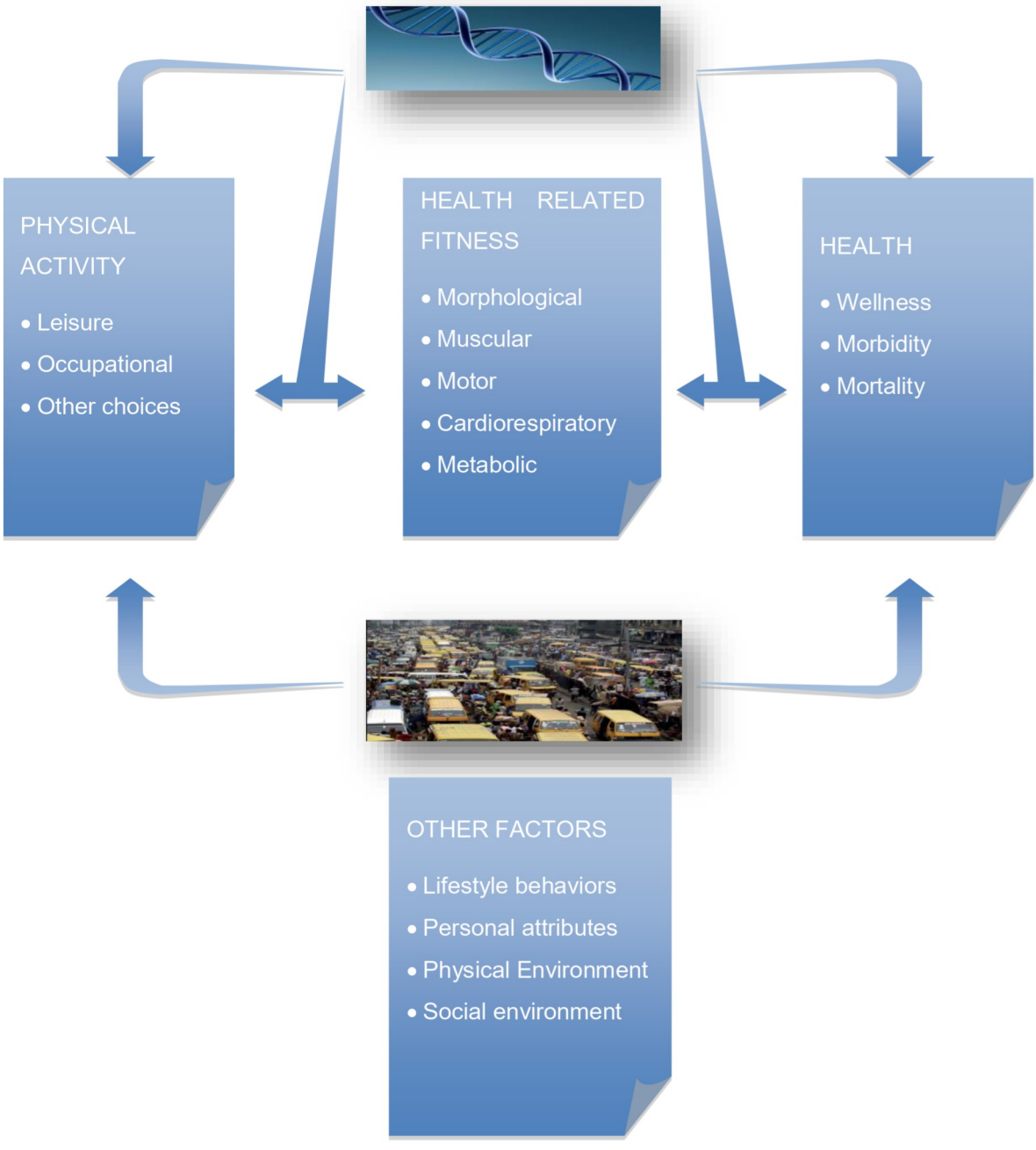
A study of these matters that explores the associations between these traits from a broad perspective, bringing together genetic and environmental factors, and their interactions must be conducted within a frame of reference. In

the next section we will briefly present the Bouchard and Shephard theoretical (heuristic) model aiming to link genetics, physical activity and health.

4. Frame of reference

It was in 1994 that Bouchard and Shephard, building on a consensus statement, presented a conceptual model to describe the interrelationships between physical activity, fitness and health (Fig. 4)

Figure 4. A consensus model describing the interrelationships among physical activity, health-related fitness (HRPF), and health status.



Basically, this broad model illustrates that physical activity, physical fitness and health influence each other in a complex and reciprocal manner. Very simply, we could postulate that lower levels of PA induce low fitness, which in turn produces poorer health status. Of course this sequence could be conceptualized the other way around. Another aspect that should be highlighted is the co-dependence of PA and health-related physical fitness (HRPF), as the two reciprocally influence each other. This has been shown previously in a study in which individuals with high fitness levels are more active than less fit individuals (Gielen et al., 2010a).

An important issue of this model, which is specifically associated with the scope of this thesis, is the importance of genetic factors and environment exposures in the variation of PA and health status. Moreover, the model clearly suggests that heredity not only exerts direct influence in both PA and health, but also influences the way by which PA promotes better fitness and subsequently leads to better health.

5. Outline of the Thesis and Research questions

This thesis follows the so-called Scandinavian model and comprises a collection of manuscripts that are published or submitted for publication. As all articles presented herein are in a series some repetition is unavoidable, especially in the methodology. Prior to the manuscripts presentation, a brief introduction is provided, followed by the research questions, sampling design and the applied methods. An overall summary and discussion is presented in the last chapter of this thesis as well as limitations and suggestions for future research.

The Research Papers´ chapter has four papers and comprises two parts. As exposed in section 3 of the present chapter, family studies present a number of advantages that can help us to improve health strategies thru physical activity. However, intervention programs will be better suited if (i) a more clear definition of PA is provided and (ii) the causes of PA phenotypic variance are

known. Hence, part one comprises a paper that aimed to review the literature, from 1981 to 2011, on the genetic determinants of physical activity and physical inactivity phenotypes in humans, considering them as distinct behaviours. Part two encompasses three original papers aiming to analyse the phenotypic variance of Metabolic Syndrome and Body Composition and the mediating role of Physical Activity in the expression of these traits. These studies aim to provide new light to the following research questions:

5.1 Research question 1: Are physical activity and physical inactivity genetically driven?

The mechanisms underlying physical activity regulation are yet to be fully understood. This has mainly been linked to its complexity, to difficulties related with the proper assessment of PA, which in turn has to do with the absence of a clear definition of what is being measured. Moreover, recent evidence highlight the notion that physical inactivity (PI) is not simply the nonexistence of activity, but is actually an individual construct that needs to be studied independently for its impacts on health status (Bray et al., 2009).

In this first paper an effort is made to summarize the available evidence on the genetic epidemiology of PA and PI, considering them as distinct behaviours. A review of the available literature on family, twin, linkage, association studies and genome-wide association scan (GWAs) on genetic determinants of PA and PI in humans, from 1981 to 2011, is presented.

We expect that this summary will provide strong evidence that not only PA is under genetic determination, but also that there is a huge variation in this determination depending upon the PA phenotype being assessed. Also, we anticipate observing a difference in the heritabilities of PA and PI that will enhance the point that these are distinct phenotypes with probably different underlying regulatory mechanisms.

5.2 Research question 2: Are Metabolic syndrome indicators correlated within families and are the Metabolic syndrome components genetically governed?

The first step towards an analysis of the genetic basis of a given trait is to calculate correlations among biological relatives. Previously, significant familial intra-trait correlations were observed between parents and offspring for diastolic blood pressure (DBP), HDL-cholesterol (HDL) and fasting insulin (INS) (Tang et al., 2006), which means that MetS physiology is shared (genetic and/or environmental) within families.

Here, we present intra-trait correlations for the five MetS traits in *The Portuguese Healthy Family Study* and we anticipate that these will be significant, which would, in the event that our anticipation holds true, support the point that, despite its origin, MetS is family-determined.

As previously stated, MetS aetiology is complex, both genetically and environmentally determined (Benyamin et al., 2007), and it has been shown that heritability estimates can vary from 0.30 (Vattikuti et al., 2012) to 0.63 (Tang et al., 2006). Understanding the extent to which MetS components are influenced by heredity in different parts of the world will add to the establishment of more appropriate intervention programmes, namely because risk behaviours associated with MetS may cluster within families (Vattikuti et al., 2012).

As such, in this analysis we present the heritability estimates of MetS components in The Portuguese Healthy Family Study, and we postulate that these estimates are significant and in accord with findings from other countries.

5.3 Research question 3: Are Metabolic syndrome components correlated with each other within families and between family members?

One aspect that has been on the heart of MetS is the co-dependence of its components. For instance, moderate to high correlations between total cholesterol (TC) and LDL-cholesterol (LDL) (r=0.84), LDL and HDL (r=0.62), TC and HDL (r=0.58), triacylglycerol (TRG) and TC (r=0.45), and TG and LDL (r=0.41) have been found (Pang et al., 2010). This evidence supports the fact

that there is additive genetic variation shared by MetS components, which simply means that these components are physiologically related.

In article 2 we present the cross-trait correlations between all 5 MetS components for each family member. We anticipate that waist circumference will be the trait more easily related with others, since it is an obesity marker. At the same time, we expect to find moderate correlations between spouses, since it is expected that obesity will be more prevalent in their age range.

Another interesting topic that we feel needs more careful attention is the possibility of the expression of one trait in a relative being correlated with a distinct trait in other relative. For instance, does father's blood pressure correlate with son's triglycerides?

Previously, body mass index/insulin correlations between biological relatives were found to be higher than between spouses, supporting the hypotheses of common transmissible factors for these two traits (Tregouet et al., 1999). This analysis is important as it allows speculating whether the shared environment is producing different effects in biological relatives. And, if so, the question is why?

In this thesis we present the cross-trait familial correlation for the 5 MetS components for each dyad within a family, and we expect that correlations will be higher among biological relatives than among spouses.

5.4 Research Question 4: To what extent are the Physical Activity effects on Metabolic syndrome components and Body Composition genetically determined?

The evidence of the preventive effect of PA on the risk of developing MetS and obesity has been widely reported and was earlier addressed in this thesis. However, a gap in the current knowledge has to do with the mechanisms that regulate such effects. One hypothesis that we explore here is the possibility that MetS and body composition expressions are the result of an interaction between an individual's genotype and the PA environment – Genotype by Environment Interaction (GxE).

We applied an elegant and sophisticated statistical technique capable of testing GxE interaction (DeYoung & Clark, 2012) and we hypothesize that genotype x energy expenditure (GxEE) interaction is a relevant determinant of variation in both MetS traits and body composition.

6. Study sample and methodology

6.1 STUDY SAMPLE

The Portuguese Healthy Family study, from the Portuguese Famílias Saudáveis (FAMS), investigates the relationship among MetS traits, physical activity, physical fitness and body composition in families. For prospective probands we focused on children and adolescents aged ≤ 18 years in schools from the Azores and Madeira archipelagos, and the north and central regions of mainland Portugal. School officers provided family lists based on the initial set of prospective probands, and families with at least two siblings were initially invited. However, given that families with 3 or more children are scarce in the Portuguese population (Rosa & Chitas, 2010), and to improve statistical power, we also invited one-offspring families through random eligibility. Children with chronic diseases, physical handicaps or psychological disorders were excluded as these conditions might impair their daily routines, namely their physical activities within schools and/or sports clubs. A total of 2411 subjects including 416 fathers, 686 mothers, and 1309 siblings, from 563 families were assessed. The ethics committee of the University of Porto approved the study, and written informed consent was obtained from all subjects.

6.2 PROCEDURES

6.2.1. Physical Growth

The standardized procedures of Lohman et al. (1988) were used to measure height (cm) and standing height (cm) with a Siber Hegner anthropometer [(GMP instruments) Holtain Ltd., England] of high precision (0.1 cm). Body mass was assessed with an electrical bioimpedance scale, TANITA BC-418 MA (Segmental Body Composition Analyser Tanita, Corporation, Tokyo, Japan Tanita scale®) with high precision (0.1 kg); Waist circumference

was measured with an anthropometric tape at the end of a normal expiration just above the iliac crest, using a non-elastic Holtain tape (Sanny, American Medical of Brazil, São Paulo, Brazil) with a precision of 0.1 cm.

Body mass index (BMI) was calculated as: $BMI = \frac{Weight\ (kg)}{height\ (m)^2}$. Overweight and obesity levels were defined as proposed by Cole et al. (2000) for children and adolescents.

6.2.2. Metabolic syndrome Indicators

Blood samples were collected after an overnight fast of at least 10 to 12 h. Glucose, TC, HDL, and TG were analysed with an LDX point of care analyser (LDX, 2003a). This method has been previously validated against a laboratory reference method (LDX, 2003b), and daily optical equipment checks were made according to manufacturer instructions.

Resting systolic (SBP) and diastolic (DBP) blood pressures were measured with an Omron Model M6 (HEM-7001-E) device according to the International Protocol of the European Society of Hypertension (Topouchian et al., 2006). Cuff sizes were modified depending on the size of the subject's arm. Subjects were seated in an upright position with the right arm resting on a table at the heart level. The first reading was performed after a 5 minute resting period. The other two readings were performed with three-minute breaks in between. The mean of the three blood pressure measurements was used for analysis.

6.2.3. Body composition

Body composition assessment followed the two-component model, fractioning body mass in fat mass and fat free mass. An electrical bioimpedance scale, TANITA BC-418 MA (Segmental Body Composition Analyser Tanita, Corporation, Tokyo, Japan Tanita scale®) with a precision of 0.1% for percentage measures and 0.1 kg for absolute measures, was used. Five distinct body segments were analysed: trunk, arms and legs. Fat mass and fat free mass was estimated and expressed in kg and % of body mass. Dual-energy X-ray Absorptiometry (DEXA) has previously validated this procedure (Pietrobelli

et al., 2004), with high correlation coefficients for all the measures. Further, algorithms used to estimate body fat were developed for European children, adolescents and adults.

6.2.4. Physical Activity

Using a 3-day physical activity diary (Bouchard et al., 1983), a trained technician interviewed each subject, recording the dominant activity for each 15-min period during 24 h by using a list of categorized activities. Categories from 1 to 9 refer to increasing levels of energy expenditure (METs) of each activity in which category 1 indicates very low energy expenditure such as sleeping or resting in bed, and category 9 refers to highly demanding physical work such as high-intensity sports. Approximate median energy cost for each of the nine categories in kcal/kg/15 min was used to compute the daily energy expenditure for each individual. The number of 15-min periods for each category was first summed over the 3-day period and weighted by its own median energy cost. Total energy expenditure (TEE) was then calculated by summing over the median energy cost of all nine categories and multiplying by subjects´ body weights. Total daily energy expenditure [TDEE (kcal/day)] was then calculated by dividing TEE by 3.

6.2.5. Data quality check

Data quality checks were performed in different stages: (i) assessment team was trained by experienced investigators; (ii) re-testing of a random sample; (iii) direct supervision of the author of this thesis of all the data collections; (iv) data entry checks and exploratory analysis for outliers identification.

6.2.6 Statistical Analysis

Exploratory, descriptive and inferential data analyses were performed in SPSS 18.0.

Family structure checks and descriptive data analyses were performed in PEDSTATS (Wigginton & Abecasis, 2005). Family correlations were calculated by means of Generalized Estimating Equations (Zhao et al., 1992) implemented

in the GESEE software developed by Tregouët et al. (1999). The remaining genetic analysis, heritability estimates and GxE, were performed in SOLAR 4.01 (Almasy & Blangero, 1998). All of these procedures are described in detail in the next chapter.

REFERENCES

Alberti, K. G., Eckel, R. H., Grundy, S. M., Zimmet, P. Z., Cleeman, J. I., Donato, K. A., Fruchart, J. C., James, W. P., Loria, C. M., & Smith, S. C., Jr. (2009). Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation, 120*(16), 1640-1645.

Alberti, K. G., Zimmet, P., & Shaw, J. (2006). Metabolic syndrome--a new world-wide definition. A Consensus Statement from the International Diabetes Federation. *Diabet Med, 23*(5), 469-480.

Alberti, K. G., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med, 15*(7), 539-553.

Almasy, L., & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet, 62*(5), 1198-1211.

American College of Sports Medicine position statement on the recommended quantity and quality of exercise for developing and maintaining fitness in healthy adults. (1978). *Med Sci Sports, 10*(3), vii-x.

Aminzadeh, M. A., Pahl, M. V., Barton, C. H., Doctor, N. S., & Vaziri, N. D. (2009). Human uraemic plasma stimulates release of leptin and uptake of tumour necrosis factor-alpha in visceral adipocytes. *Nephrol Dial Transplant, 24*(12), 3626-3631.

Baber, U., Gutierrez, O. M., Levitan, E. B., Warnock, D. G., Farkouh, M. E., Tonelli, M., Safford, M. M., & Muntner, P. (2013). Risk for recurrent coronary heart disease and all-cause mortality among individuals with chronic kidney disease compared with diabetes mellitus, metabolic syndrome, and cigarette smokers. *Am Heart J, 166*(2), 373-380 e372.

Ballor, D. L., & Keesey, R. E. (1991). A meta-analysis of the factors affecting exercise-induced changes in body mass, fat mass and fat-free mass in males and females. *Int J Obes, 15*(11), 717-726.

Benyamin, B., Sorensen, T. I., Schousboe, K., Fenger, M., Visscher, P. M., & Kyvik, K. O. (2007). Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? *Diabetologia, 50*(9), 1880-1888.

Bjerregaard-Andersen, M., Hansen, L., da Silva, L. I., Joaquim, L. C., Hennild, D. E., Christiansen, L., Aaby, P., Benn, C. S., Christensen, K., Sodemann, M., Jensen, D. M., & Beck-Nielsen, H. (2013). Risk of Metabolic Syndrome and Diabetes Mellitus Among Young Twins and Singletons in Guinea-Bissau. *Diabetes Care*.

Blangero, J. (2009). Statistical genetic approaches to human adaptability. 1993. *Hum Biol, 81*(5-6), 523-546.

Bochud, M. (2012). Estimating Heritability from Nuclear Family and Pedigree Data. In R. C. Elston, J. M. Satagopan & S. Sun (Eds.), *Statistical in Human Genetics: Methods and Protocols*. New York: Humana Press.

Booth, F. W., Laye, M. J., Lees, S. J., Rector, R. S., & Thyfault, J. P. (2008). Reduced physical activity and risk of chronic disease: the biology behind the consequences. *Eur J Appl Physiol, 102*(4), 381-390.

Boronat, M., Chirino, R., Varillas, V. F., Saavedra, P., Marrero, D., Fabregas, M., & Novoa, F. J. (2005). Prevalence of the metabolic syndrome in the island of Gran Canaria: comparison of three major diagnostic proposals. *Diabet Med, 22*(12), 1751-1756.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Theriault, G. (1983). A method to assess energy expenditure in children and adults. *Am J Clin Nutr, 37*(3), 461-467.

Bowman, J. (1972). Genotype x environment interactions. *Ann Genet Sel Anim, 4*(1), 117-123.

Bray, M. S., Hagberg, J. M., Perusse, 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. *Med Sci Sports Exerc, 41*(1), 35-73.

Butte, N. F., Cai, G., Cole, S. A., & Comuzzie, A. G. (2006). Viva la Familia Study: genetic and environmental contributions to childhood obesity and its comorbidities in the Hispanic population. *Am J Clin Nutr, 84*(3), 646-654; quiz 673-644.

Cai, G., Cole, S. A., Butte, N., Bacino, C., Diego, V., Tan, K., Goring, H. H., O'Rahilly, S., Farooqi, I. S., & Comuzzie, A. G. (2006). A quantitative trait locus on chromosome 18q for physical activity and dietary intake in Hispanic children. *Obesity (Silver Spring), 14*(9), 1596-1604.

Cameron, A. J., Shaw, J. E., & Zimmet, P. Z. (2004). The metabolic syndrome: prevalence in worldwide populations. *Endocrinol Metab Clin North Am, 33*(2), 351-375, table of contents.

Carr, D. B., Utzschneider, K. M., Hull, R. L., Kodama, K., Retzlaff, B. M., Brunzell, J. D., Shofer, J. B., Fish, B. E., Knopp, R. H., & Kahn, S. E. (2004). Intra-abdominal fat is a major determinant of the National Cholesterol Education Program Adult Treatment Panel III criteria for the metabolic syndrome. *Diabetes, 53*(8), 2087-2094.

Carreira, H., Pereira, M., Azevedo, A., & Lunet, N. (2012). Trends of BMI and prevalence of overweight and obesity in Portugal (1995-2005): a systematic review. *Public Health Nutr, 15*(6), 972-981.

Carroll, S., & Dudfield, M. (2004). What is the relationship between exercise and metabolic abnormalities? A review of the metabolic syndrome. *Sports Med, 34*(6), 371-418.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep, 100*(2), 126-131.

Cawley, J., & Meyerhoefer, C. (2012). The medical care costs of obesity: an instrumental variables approach. *J Health Econ, 31*(1), 219-230.

Chen, H. J., Bai, C. H., Yeh, W. T., Chiu, H. C., & Pan, W. H. (2006). Influence of metabolic syndrome and general obesity on the risk of ischemic stroke. *Stroke, 37*(4), 1060-1064.

Clarke, H. H. (1979). Academy approves physical fitness definition. *Physical Fitness Newsletter, 25*(1).

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ, 320*(7244), 1240-1243.

Crow, J. F. (2010). On epistasis: why it is unimportant in polygenic directional selection. *Philos Trans R Soc Lond B Biol Sci, 365*(1544), 1241-1244.

Daniels, S. R., Arnett, D. K., Eckel, R. H., Gidding, S. S., Hayman, L. L., Kumanyika, S., Robinson, T. N., Scott, B. J., St Jeor, S., & Williams, C. L. (2005). Overweight in children and adolescents: pathophysiology, consequences, prevention, and treatment. *Circulation, 111*(15), 1999-2012.

de Andrade, M., Gueguen, R., Visvikis, S., Sass, C., Siest, G., & Amos, C. I. (2002). Extension of variance components approach to incorporate temporal trends and longitudinal pedigree data analysis. *Genet Epidemiol, 22*(3), 221-232.

De Moor, M. H., Liu, Y. J., Boomsma, D. I., Li, J., Hamilton, J. J., Hottenga, J. J., Levy, S., Liu, X. G., Pei, Y. F., Posthuma, D., Recker, R. R., Sullivan, P. F., Wang, L., Willemsen, G., Yan, H., EJ, D. E. G., & Deng, H. W. (2009). Genome-Wide Association Study of Exercise Behavior in Dutch and American Adults. *Med Sci Sports Exerc, 41*(10), 1887-1895.

De Moor, M. H., Spector, T. D., Cherkas, L. F., Falchi, M., Hottenga, J. J., Boomsma, D. I., & De Geus, E. J. (2007b). Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. *Twin Res Hum Genet, 10*(6), 812-820.

De Moor, M. H., Stubbe, J. H., Boomsma, D. I., & De Geus, E. J. (2007c). Exercise participation and self-rated health: do common genes explain the association? *Eur J Epidemiol, 22*(1), 27-32.

DeFronzo, R. A., & Ferrannini, E. (1991). Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. *Diabetes Care, 14*(3), 173-194.

Despres, J. P. (2006). Is visceral obesity the cause of the metabolic syndrome? *Ann Med, 38*(1), 52-63.

Despres, J. P., & Lemieux, I. (2006). Abdominal obesity and metabolic syndrome. *Nature, 444*(7121), 881-887.

DeYoung, C. G., & Clark, R. (2012). The gene in its natural habitat: the importance of gene-trait interactions. *Dev Psychopathol, 24*(4), 1307-1318.

Dickerson, G. E. (1962). Implications of genetic-environmental interaction in animal breeding. *Animal Production, 4*, 47-63.

Diego, V. P., Kent, J. W., & Blangero, J. (2013). Familial Studies: Genetic Inferences. *International Encyclopedia of the Social & Behavioral Sciences, In Press*.

Diet, nutrition and the prevention of chronic diseases. (2003). *World Health Organ Tech Rep Ser, 916*, i-viii, 1-149, backcover.

DiPietro, L. (1999). Physical activity in the prevention of obesity: current evidence and research issues. *Med Sci Sports Exerc, 31*(11 Suppl), S542-546.

Dishman, R. K., Washburn, R. A., & Heath, G. W. (2004). *Physical activity epidemiology*. Champaign: Human Kinetics.

do Carmo, I., dos Santos, O., Camolas, J., Vieira, J., Carreira, M., Medina, L., Reis, L., & Galvao-Teles, A. (2006). Prevalence of obesity in Portugal. *Obes Rev, 7*(3), 233-237.

Duncan, G. E., Goldberg, J., Noonan, C., Moudon, A. V., Hurvitz, P., & Buchwald, D. (2008). Unique Environmental Effects on Physical Activity Participation: A Twin Study. *PLoS One, 3*(4).

Elks, C. E., den Hoed, M., Zhao, J. H., Sharp, S. J., Wareham, N. J., Loos, R. J., & Ong, K. K. (2012). Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol (Lausanne), 3*, 29.

Eriksson, M., Rasmussen, F., & Tynelius, P. (2004). Genetic and environmental factors in physical activity - The Swedish young male twins study. *International Journal of Obesity, 28*, S207-S207.

Ervin, R. B. (2009). Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003-2006. *Natl Health Stat Report*(13), 1-7.

Exercise, fitness and health a consensus of current knowledge. (1990). Champaign, Il.: Human Kinetics.

Expert Panel on Detection, E., & Treatment of High Blood Cholesterol in, A. (2001). Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA, 285*(19), 2486-2497.

Falconer, D. S. (1952). The problem of Environment and Selection. In *The American Naturalist* (Vol. 86, pp. 293-298): The University of Chicago Press.

Falconer, D. S., & Mackay, T. F. (1996). *Introduction to quantitative genetics*. Harlow, Essex: Longman.

Farook, V. S., Puppala, S., Schneider, J., Fowler, S. P., Chittoor, G., Dyer, T. D., Allayee, H., Cole, S. A., Arya, R., Black, M. H., Curran, J. E., Almasy, L., Buchanan, T. A., Jenkinson, C. P., Lehman, D. M., Watanabe, R. M., Blangero, J., & Duggirala, R. (2012). Metabolic syndrome is linked to chromosome 7q21 and associated with genetic variants in CD36 and GNAT3 in Mexican Americans. *Obesity (Silver Spring), 20*(10), 2083-2092.

Ferrannini, E., & Iozzo, P. (2006). Is insulin resistance atherogenic? A review of the evidence. *Atheroscler Suppl, 7*(4), 5-10.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Philosophical Transactions of the Royal Society of Edinburgh, 52*, 399–433.

Fiuza, M., Cortez-Dias, N., Martins, S., & Belo, A. (2008). Metabolic syndrome in Portugal: prevalence and implications for cardiovascular risk--results from the VALSIM Study. *Rev Port Cardiol, 27*(12), 1495-1529.

Ford, E. S., & Giles, W. H. (2003). A comparison of the prevalence of the metabolic syndrome using two proposed definitions. *Diabetes Care, 26*(3), 575-581.

Ford, E. S., Giles, W. H., & Dietz, W. H. (2002). Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *JAMA, 287*(3), 356-359.

Franks, P. W., Hanson, R. L., Knowler, W. C., Moffett, C., Enos, G., Infante, A. M., Krakoff, J., & Looker, H. C. (2007). Childhood predictors of young-onset type 2 diabetes. *Diabetes, 56*(12), 2964-2972.

Franz, M. J., VanWormer, J. J., Crain, A. L., Boucher, J. L., Histon, T., Caplan, W., Bowman, J. D., & Pronk, N. P. (2007). Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. *J Am Diet Assoc, 107*(10), 1755-1767.

Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., Perry, J. R., Elliott, K. S., Lango, H., Rayner, N. W., Shields, B., Harries, L. W., Barrett, J. C., Ellard, S., Groves, C. J., Knight, B., Patch, A. M., Ness, A. R., Ebrahim, S., Lawlor, D. A., Ring, S. M., Ben-Shlomo, Y., Jarvelin, M. R., Sovio, U., Bennett, A. J., Melzer, D., Ferrucci, L., Loos, R. J., Barroso, I., Wareham, N. J., Karpe, F., Owen, K. R., Cardon, L. R., Walker, M., Hitman, G. A., Palmer, C. N., Doney, A. S., Morris, A. D., Smith, G. D., Hattersley, A. T., & McCarthy, M. I. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science, 316*(5826), 889-894.

Freedman, D. S., Khan, L. K., Dietz, W. H., Srinivasan, S. R., & Berenson, G. S. (2001). Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa Heart Study. *Pediatrics, 108*(3), 712-718.

Freedman, D. S., Khan, L. K., Serdula, M. K., Dietz, W. H., Srinivasan, S. R., & Berenson, G. S. (2005). The relation of childhood BMI to adult adiposity: the Bogalusa Heart Study. *Pediatrics, 115*(1), 22-27.

Fuentes, R. M., Perola, M., Nissinen, A., & Tuomilehto, J. (2002). ACE gene and physical activity, blood pressure, and hypertension: a population study in Finland. *J Appl Physiol, 92*(6), 2508-2512.

Garg, R., Tripathy, D., & Dandona, P. (2003). Insulin resistance as a proinflammatory state: mechanisms, mediators, and therapeutic interventions. *Curr Drug Targets, 4*(6), 487-492.

Ghaem Maralani, H., Tai, B. C., Wong, T. Y., Tai, E. S., Li, J., Wang, J. J., & Mitchell, P. (2013). Metabolic syndrome and mortality in the elderly: a time-dependent association. *Diabetes Res Clin Pract, 99*(2), 209-216.

Gielen, S., Schuler, G., & Adams, V. (2010a). Cardiovascular effects of exercise training: molecular mechanisms. In *Circulation* (Vol. 122, pp. 1221-1238). United States.

Gielen, S., Schuler, G., & Adams, V. (2010b). Cardiovascular effects of exercise training: molecular mechanisms. *Circulation, 122*(12), 1221-1238.

Griffiths, A. J. F., Miller, J. H., & Suzuki, D. T. (2000). An Introduction to Genetic Analysis. *Quantifying heritability* 7th. 2013, disponível em http://www.ncbi.nlm.nih.gov/books/NBK21866/

Guize, L., Thomas, F., Pannier, B., Bean, K., Danchin, N., & Benetos, A. (2006). [Metabolic syndrome: prevalence, risk factors and mortality in a French population of 62 000 subjects]. *Bull Acad Natl Med, 190*(3), 685-697; discussion 697-700.

Hahn, V., Halle, M., Schmidt-Trucksass, A., Rathmann, W., Meisinger, C., & Mielck, A. (2009). Physical activity and the metabolic syndrome in elderly

German men and women: results from the population-based KORA survey. *Diabetes Care, 32*(3), 511-513.

Haldane, J. B. (1946). The interaction of nature and nurture. *Ann Eugen, 13*(3), 197-205.

Heitmann, B. L., Kaprio, J., Harris, J. R., Rissanen, A., Korkeila, M., & Koskenvuo, M. (1997). Are genetic determinants of weight gain modified by leisure-time physical activity? A prospective study of Finnish twins. *Am J Clin Nutr, 66*(3), 672-678.

Heymsfield, S. (2005). *Human Body Composition*: Human Kinetics.

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet, 4*(2), e1000008.

Himsworth, H. P. (1936). Diabetes mellitus: its differentiation into insulin-sensitive and insulin-insensitive types. *Lancet, 1* 127-130

Hipocrates. *On Regimen in Acute Diseases*: Kessinger Publishing.

Hopper, J. L. (1993). Variance components for statistical genetics: applications in medical research to characteristics related to human diseases and health. *Stat Methods Med Res, 2*(3), 199-223.

Hsu, C. H. (2013). Different impacts of metabolic syndrome components on insulin resistance in type 2 diabetes. *Int J Endocrinol, 2013*, 740419.

Huberty, J. L., Beets, M. W., Beighle, A., Saint-Maurice, P. F., & Welk, G. (2013). Effects of Ready for Recess, an Environmental Intervention, on Physical Activity in 3rd - 6th Grade Children. *J Phys Act Health*.

Iannello, S., Campione, R., & Belfiore, F. (1998). Response of insulin, glucagon, lactate, and nonesterified fatty acids to glucose in visceral obesity with and without NIDDM: relationship to hypertension. *Mol Genet Metab, 63*(3), 214-223.

Ivezic-Lalic, D., Bergman Markovic, B., Kranjcevic, K., Kern, J., Vrdoljak, D., & Vucak, J. (2013). Diversity of metabolic syndrome criteria in association with

cardiovascular diseases--a family medicine-based investigation. *Med Sci Monit, 19*, 571-578.

Jakicic, J. M., & Otto, A. D. (2005). Physical activity considerations for the treatment and prevention of obesity. *Am J Clin Nutr, 82*(1 Suppl), 226S-229S.

Jarrett, O. D., Wanke, C. A., Ruthazer, R., Bica, I., Isaac, R., & Knox, T. A. (2013). Metabolic syndrome predicts all-cause mortality in persons with human immunodeficiency virus. *AIDS Patient Care STDS, 27*(5), 266-271.

Joosen, A. M., Gielen, M., Vlietinck, R., & Westerterp, K. R. (2005). Genetic analysis of physical activity in twins. *Am J Clin Nutr, 82*(6), 1253-1259.

Kaino, W., Daimon, M., Sasaki, S., Karasawa, S., Takase, K., Tada, K., Wada, K., Kameda, W., Susa, S., Oizumi, T., Fukao, A., Kubota, I., Kayama, T., & Kato, T. (2013). Lower physical activity is a risk factor for a clustering of metabolic risk factors in non-obese and obese Japanese subjects: the Takahata study. *Endocr J, 60*(5), 617-628.

Kaplan, N. M. (1989). The deadly quartet. Upper-body obesity, glucose intolerance, hypertriglyceridemia, and hypertension. *Arch Intern Med, 149*(7), 1514-1520.

Katzmarzyk, P., & Herman, K. (2007). The role of physical activity and fitness in the prevention and treatment of metabolic syndrome. *Current Cardiovascular Risk Reports, 1*(3), 228-236.

Katzmarzyk, P. T., Barreira, T. V., Broyles, S. T., Champagne, C. M., Chaput, J. P., Fogelholm, M., Hu, G., Johnson, W. D., Kuriyan, R., Kurpad, A., Lambert, E. V., Maher, C., Maia, J., Matsudo, V., Olds, T., Onywera, V., Sarmiento, O. L., Standage, M., Tremblay, M. S., Tudor-Locke, C., Zhao, P., & Church, T. S. (2013). The International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE): design and methods. *BMC Public Health, 13*, 900.

Keen, K. J., & Elston, R. C. (2003). Robust asymptotic sampling theory for correlations in pedigrees. *Stat Med, 22*(20), 3229-3247.

Koopmans, J. R., Lorenz, J. P. V. D., & Boomsma, D. I. (1994). Smoking and sports participation. In U. Goldbourt, U. Faire & K. Berg (Eds.), *Factors in coronary heart disease* (pp. 217-235). Dordrecht: Kluwer Academic.

Krebs, N. F., Himes, J. H., Jacobson, D., Nicklas, T. A., Guilday, P., & Styne, D. (2007). Assessment of child and adolescent overweight and obesity. *Pediatrics, 120 Suppl 4*, S193-228.

Kritikos, A., Bekiari, A., Nikitaras, N., Famissis, K., & Sakellariou, K. (2009). Hippocrates' counselling with regard to physical exercise, gymnastics, dietetics and health. *Ir J Med Sci, 178*(3), 377-383.

Kuczmarski, R. J., Ogden, C. L., Grummer-Strawn, L. M., Flegal, K. M., Guo, S. S., Wei, R., Mei, Z., Curtin, L. R., Roche, A. F., & Johnson, C. L. (2000). CDC growth charts: United States. *Adv Data*(314), 1-27.

Laaksonen, D. E., Lakka, H. M., Niskanen, L. K., Kaplan, G. A., Salonen, J. T., & Lakka, T. A. (2002). Metabolic syndrome and development of diabetes mellitus: application and validation of recently suggested definitions of the metabolic syndrome in a prospective cohort study. *Am J Epidemiol, 156*(11), 1070-1077.

Lahmann, P. H., Lissner, L., Gullberg, B., & Berglund, G. (2002). A prospective study of adiposity and all-cause mortality: the Malmo Diet and Cancer Study. *Obes Res, 10*(5), 361-369.

Lakka, H. M., Laaksonen, D. E., Lakka, T. A., Niskanen, L. K., Kumpusalo, E., Tuomilehto, J., & Salonen, J. T. (2002). The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA, 288*(21), 2709-2716.

Laursen, A. H., Kristiansen, O. P., Marott, J. L., Schnohr, P., & Prescott, E. (2012). Intensity versus duration of physical activity: implications for the metabolic syndrome. A prospective cohort study. *BMJ Open, 2*(5).

LDX, C. (2003a). The Accuracy and Reproducibility of a Rapid, Fingerstick Method for Measuring a Complete Lipid Profile Is Comparable to a Reference Laboratory Method. *Cholestech Corporation*.

LDX, C. (2003b). The Accuracy and Reproducibility of a Rapid, Fingerstick Method for Measuring a Complete Lipid Profile Is Comparable to a Reference Laboratory Method (b). *Cholestech Corporation*.

Lohman, T. G., Roche, A. F., & Martorel, R. (1988). *Anthropometric Standardization Reference Manual*. Illinois: Human Kinetics Books Champaign.

Loos, R. J., Lindgren, C. M., Li, S., Wheeler, E., Zhao, J. H., Prokopenko, I., Inouye, M., Freathy, R. M., Attwood, A. P., Beckmann, J. S., Berndt, S. I., Prostate, L. C., Ovarian Cancer Screening, T., Jacobs, K. B., Chanock, S. J., Hayes, R. B., Bergmann, S., Bennett, A. J., Bingham, S. A., Bochud, M., Brown, M., Cauchi, S., Connell, J. M., Cooper, C., Smith, G. D., Day, I., Dina, C., De, S., Dermitzakis, E. T., Doney, A. S., Elliott, K. S., Elliott, P., Evans, D. M., Sadaf Farooqi, I., Froguel, P., Ghori, J., Groves, C. J., Gwilliam, R., Hadley, D., Hall, A. S., Hattersley, A. T., Hebebrand, J., Heid, I. M., Kora, Lamina, C., Gieger, C., Illig, T., Meitinger, T., Wichmann, H. E., Herrera, B., Hinney, A., Hunt, S. E., Jarvelin, M. R., Johnson, T., Jolley, J. D., Karpe, F., Keniry, A., Khaw, K. T., Luben, R. N., Mangino, M., Marchini, J., McArdle, W. L., McGinnis, R., Meyre, D., Munroe, P. B., Morris, A. D., Ness, A. R., Neville, M. J., Nica, A. C., Ong, K. K., O'Rahilly, S., Owen, K. R., Palmer, C. N., Papadakis, K., Potter, S., Pouta, A., Qi, L., Nurses' Health, S., Randall, J. C., Rayner, N. W., Ring, S. M., Sandhu, M. S., Scherag, A., Sims, M. A., Song, K., Soranzo, N., Speliotes, E. K., Diabetes Genetics, I., Syddall, H. E., Teichmann, S. A., Timpson, N. J., Tobias, J. H., Uda, M., Sardi, N. I. A. S., Vogel, C. I., Wallace, C., Waterworth, D. M., Weedon, M. N., Wellcome Trust Case Control, C., Willer, C. J., Fusion, Wraight, Yuan, X., Zeggini, E., Hirschhorn, J. N., Strachan, D. P., Ouwehand, W. H., Caulfield, M. J., Samani, N. J., Frayling, T. M., Vollenweider, P., Waeber, G., Mooser, V., Deloukas, P., McCarthy, M. I., Wareham, N. J., Barroso, I., Jacobs, K. B., Chanock, S. J., Hayes, R. B., Lamina, C., Gieger, C., Illig, T., Meitinger, T., Wichmann, H. E., Kraft, P., Hankinson, S. E., Hunter, D. J., Hu, F. B., Lyon, H. N., Voight, B. F., Ridderstrale, M., Groop, L., Scheet, P., Sanna, S., Abecasis, G. R., Albai, G., Nagaraja, R., Schlessinger, D., Jackson, A. U., Tuomilehto, J., Collins, F. S., Boehnke, M., & Mohlke, K. L. (2008). Common

variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet, 40*(6), 768-775.

Loos, R. J., Rankinen, T., Tremblay, A., Perusse, L., Chagnon, Y., & Bouchard, C. (2005). Melanocortin-4 receptor gene and physical activity in the Quebec Family Study. *Int J Obes (Lond), 29*(4), 420-428.

Lorentzon, M., Lorentzon, R., Lerner, U. H., & Nordstrom, P. (2001). Calcium sensing receptor gene polymorphism, circulating calcium concentrations and bone mineral density in healthy adolescent girls. *Eur J Endocrinol, 144*(3), 257-261.

Luke, A., Guo, X., Adeyemo, A. A., Wilks, R., Forrester, T., Lowe, W., Jr., Comuzzie, A. G., Martin, L. J., Zhu, X., Rotimi, C. N., & Cooper, R. S. (2001). Heritability of obesity-related traits among Nigerians, Jamaicans and US black people. *Int J Obes Relat Metab Disord, 25*(7), 1034-1041.

Lyons, A. S., & Petrucelli, R. J. (1978). *Medicine: An illustrated history*. New York: Harry N. Abrams.

Mark, A. E., & Janssen, I. (2009). Influence of bouts of physical activity on overweight in youth. *Am J Prev Med, 36*(5), 416-421.

Martinez-Gonzalez, M. A., & Martin-Calvo, N. (2013). The major European dietary patterns and metabolic syndrome. *Rev Endocr Metab Disord, 14*(3), 265-271.

McCaffery, J. M., Papandonatos, G. D., Bond, D. S., Lyons, M. J., & Wing, R. R. (2009). Gene X environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *Am J Clin Nutr, 89*(4), 1011-1018.

Morris, J. N., Heady, J. A., Raffle, P. A., Roberts, C. G., & Parks, J. W. (1953). Coronary heart-disease and physical activity of work. *Lancet, 265*(6796), 1111-1120; concl.

Naj, A. C., Park, Y. S., & Beaty, T. H. (2012). Detecting familial aggregation. In R. C. Elston, J. M. Satagopan & S. Sun (Eds.), *Statistical in Human Genetics: Methods and Protocols*. New York: Humana Press.

National Institutes of Health, N. H., Lung, and Blood Institute. (2010). *Disease and Conditions Index: What Are Overweight and Obesity?* : National Institutes of Health.

Ness, A. R., Leary, S. D., Mattocks, C., Blair, S. N., Reilly, J. J., Wells, J., Ingle, S., Tilling, K., Smith, G. D., & Riddoch, C. (2007). Objectively measured physical activity and fat mass in a large cohort of children. *PLoS Med, 4*(3), e97.

Ng, M. C., So, W. Y., Lam, V. K., Cockram, C. S., Bell, G. I., Cox, N. J., & Chan, J. C. (2004). Genome-wide scan for metabolic syndrome and related quantitative traits in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21-q25. *Diabetes, 53*(10), 2676-2683.

Nutrition and Diet for Healthy Lifestyles in Europe: Science and Policy Implications. Proceedings of the European Conference. May 18-20, 2000. Crete, Greece. (2001). *Public Health Nutr, 4*(2A), 337-434.

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2012). Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010. *JAMA, 307*(5), 483-490.

Owen, C. G., Nightingale, C. M., Rudnicka, A. R., Sattar, N., Cook, D. G., Ekelund, U., & Whincup, P. H. (2010). Physical activity, obesity and cardiometabolic risk factors in 9- to 10-year-old UK children of white European, South Asian and black African-Caribbean origin: the Child Heart And health Study in England (CHASE). *Diabetologia, 53*(8), 1620-1630.

Paffenbarger, R. S., Jr., Blair, S. N., & Lee, I. M. (2001). A history of physical activity, cardiovascular health and longevity: the scientific contributions of Jeremy N Morris, DSc, DPH, FRCP. *Int J Epidemiol, 30*(5), 1184-1192.

Pang, Z., Zhang, D., Li, S., Duan, H., Hjelmborg, J., Kruse, T. A., Kyvik, K. O., Christensen, K., & Tan, Q. (2010). Multivariate modelling of endophenotypes associated with the metabolic syndrome in Chinese twins. *Diabetologia, 53*(12), 2554-2561.

Perusse, L., Tremblay, A., Leblanc, C., & Bouchard, C. (1989). Genetic and environmental influences on level of habitual physical activity and exercise participation. *Am J Epidemiol, 129*(5), 1012-1022.

Pietrobelli, A., Rubiano, F., St-Onge, M. P., & Heymsfield, S. B. (2004). New bioimpedance analysis system: improved phenotyping with whole-body analysis. *Eur J Clin Nutr, 58*(11), 1479-1484.

Poirier, P., & Despres, J. P. (2001). Exercise in weight management of obesity. *Cardiol Clin, 19*(3), 459-470.

Potenza, M. V., & Mechanick, J. I. (2009). The metabolic syndrome: definition, global impact, and pathophysiology. *Nutr Clin Pract, 24*(5), 560-577.

Rankinen, T., Roth, S. M., Bray, M. S., Loos, R., Perusse, L., Wolfarth, B., Hagberg, J. M., & Bouchard, C. (2010). Advances in exercise, fitness, and performance genomics. *Med Sci Sports Exerc, 42*(5), 835-846.

Rankinen, T., Zuberi, A., Chagnon, Y. C., Weisnagel, S. J., Argyropoulos, G., Walts, B., Perusse, L., & Bouchard, C. (2006). The human obesity gene map: the 2005 update. *Obesity (Silver Spring), 14*(4), 529-644.

Rao, D. C., Vogler, G. P., McGue, M., & Russell, J. M. (1987). Maximum-likelihood estimation of familial correlations from multivariate quantitative data on pedigrees: a general method and examples. *Am J Hum Genet, 41*(6), 1104-1116.

Reaven, G. M. (1988). Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes, 37*(12), 1595-1607.

Robertson, A. (1959). The sampling variance of the genetic correlation coefficient. *Biometrics, 15*(3), 469-485.

Robinson, J. L., Winiewicz, D. D., Fuerch, J. H., Roemmich, J. N., & Epstein, L. H. (2006). Relationship between parental estimate and an objective measure of child television watching. *Int J Behav Nutr Phys Act, 3*, 43.

Rosa, M. J. V., & Chitas, P. (2010). *Portugal: os números*. Lisboa.

Rowlands, A. V. (2007). Accelerometer assessment of physical activity in children: an update. *Pediatr Exerc Sci, 19*(3), 252-266.

Rowlands, A. V., & Eston, R. G. (2007). The Measurement and Interpretation of Children's Physical Activity. *J Sports Sci Med, 6*(3), 270-276.

Schmidt, M. I., Watson, R. L., Duncan, B. B., Metcalf, P., Brancati, F. L., Sharrett, A. R., Davis, C. E., & Heiss, G. (1996). Clustering of dyslipidemia, hyperuricemia, diabetes, and hypertension and its association with fasting insulin and central and overall obesity in a general population. Atherosclerosis Risk in Communities Study Investigators. *Metabolism, 45*(6), 699-706.

Schoeller, D. A. (1999). Recent advances from application of doubly labeled water to measurement of human energy expenditure. *J Nutr, 129*(10), 1765-1768.

Simonen, R. L., Rankinen, T., Perusse, L., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., & Bouchard, C. (2003a). A dopamine D2 receptor gene polymorphism and physical activity in two family studies. *Physiol Behav, 78*(4-5), 751-757.

Simonen, R. L., Rankinen, T., Perusse, L., Rice, T., Rao, D. C., Chagnon, Y., & Bouchard, C. (2003b). Genome-wide linkage scan for physical activity levels in the Quebec Family study. *Med Sci Sports Exerc, 35*(8), 1355-1359.

Spinath, F. M., Wolf, H., Angleitner, A., Borkenau, P., & Riemann, R. (2002). Genetic and environmental influences on objectively assessed activity in adults. *Personality and Individual Differences, 33*(4), 633-645.

Steele, R. M., van Sluijs, E. M., Cassidy, A., Griffin, S. J., & Ekelund, U. (2009). Targeting sedentary time or moderate- and vigorous-intensity activity: independent relations with adiposity in a population-based sample of 10-y-old British children. *Am J Clin Nutr, 90*(5), 1185-1192.

Stefan, N., Vozarova, B., Del Parigi, A., Ossowski, V., Thompson, D. B., Hanson, R. L., Ravussin, E., & Tataranni, P. A. (2002). The Gln223Arg polymorphism of the leptin receptor in Pima Indians: influence on energy

expenditure, physical activity and lipid metabolism. *Int J Obes Relat Metab Disord, 26*(12), 1629-1632.

Strasser, B. (2012). Physical activity in obesity and metabolic syndrome. *Ann N Y Acad Sci*.

Strong, W. B., Malina, R. M., Blimkie, C. J., Daniels, S. R., Dishman, R. K., Gutin, B., Hergenroeder, A. C., Must, A., Nixon, P. A., Pivarnik, J. M., Rowland, T., Trost, S., & Trudeau, F. (2005). Evidence based physical activity for school-age youth. *J Pediatr, 146*(6), 732-737.

Stubbe, J. H., Boomsma, D. I., & De Geus, E. J. (2005). Sports participation during adolescence: a shift from environmental to genetic factors. *Med Sci Sports Exerc, 37*(4), 563-570.

Stulp, G., Buunk, A. P., Pollet, T. V., Nettle, D., & Verhulst, S. (2013). Are human mating preferences with respect to height reflected in actual pairings? *PLoS One, 8*(1), e54186.

Tang, W., Hong, Y., Province, M. A., Rich, S. S., Hopkins, P. N., Arnett, D. K., Pankow, J. S., Miller, M. B., & Eckfeldt, J. H. (2006). Familial clustering for features of the metabolic syndrome: the National Heart, Lung, and Blood Institute (NHLBI) Family Heart Study. *Diabetes Care, 29*(3), 631-636.

Topouchian, J. A., El Assaad, M. A., Orobinskaia, L. V., El Feghali, R. N., & Asmar, R. G. (2006). Validation of two automatic devices for self-measurement of blood pressure according to the International Protocol of the European Society of Hypertension: the Omron M6 (HEM-7001-E) and the Omron R7 (HEM 637-IT). *Blood Press Monit, 11*(3), 165-171.

Tregouet, D. A., Herbeth, B., Juhan-Vague, I., Siest, G., Ducimetiere, P., & Tiret, L. (1999). Bivariate familial correlation analysis of quantitative traits by use of estimating equations: application to a familial analysis of the insulin resistance syndrome. *Genet Epidemiol, 16*(1), 69-83.

Trost, S. G. (2007). State of the Art Reviews: Measurement of Physical Activity in Children and Adolescents. *American Journal of Lifestyle Medicine, 1*(4), 299-314.

Vague, J. (1947). Les Obésités: études biométriques. *Presse Med, 55*(30), 339.

Vattikuti, S., Guo, J., & Chow, C. C. (2012). Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. *PLoS Genet, 8*(3), e1002637.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nat Rev Genet, 9*(4), 255-266.

Wagner, H., Melhus, H., Pedersen, N. L., & Michaelsson, K. (2013). Genetic influence on bone phenotypes and body composition: a Swedish twin study. *J Bone Miner Metab*.

Wang, Y. C., McPherson, K., Marsh, T., Gortmaker, S. L., & Brown, M. (2011). Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet, 378*(9793), 815-825.

Wigginton, J. E., & Abecasis, G. R. (2005). PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics, 21*(16), 3445-3447.

Wijndaele, K. (2007). *Metabolic syndrome risk and mental health: relationship with physical activity and physical fitness in flemish adults*. Gent: Universiteit Gent. Relatorio de Estagio apresentado a

Wyatt, S. B., Winters, K. P., & Dubbert, P. M. (2006). Overweight and obesity: prevalence, consequences, and causes of a growing public health problem. *Am J Med Sci, 331*(4), 166-174.

Yamada, Y. (1962). Genotype by environment interaction and genetic correlation of the same trait under different environments. *遺伝學雜誌, 37*(6), 498-509.

Zhan, Y., Yu, J., Chen, R., Gao, J., Ding, R., Fu, Y., Zhang, L., & Hu, D. (2012). Socioeconomic status and metabolic syndrome in the general population of China: a cross-sectional study. *BMC Public Health, 12*, 921.

Zhao, L. P., Grove, J., & Quiaoit, F. (1992). A method for assessing patterns of familial resemblance in complex human pedigrees, with an application to the nevus-count data in Utah kindreds. *Am J Hum Genet, 51*(1), 178-190.

CHAPTER 2

RESEARCH PAPERS

ARTICLE 1

GENETICS OF PHYSICAL ACTIVITY AND PHYSICAL INACTIVITY IN HUMANS

Daniel Monteiro de Vilhena e Santos[1]

Peter T. Katzmarzyk[2]

André Filipe Teixeira Seabra[1]

José António Ribeiro Maia[1]

Published in Behaviour Genetics

DOI 10.1007/s10519-012-9534-1

[1] CIFI[2]D, Faculty of Sports, University of Porto, Portugal

[2] Pennington Biomedical Research Center, Louisiana State University System, USA

Abstract

Introduction: Emerging evidence suggests that physical activity and sedentary behaviour (reflected in physical inactivity [PI]) might be two different phenotypes that may have distinct underlying physiological mechanisms. The purpose of this review is to summarize the existing literature on the genetic determinants of PA and PI phenotypes in humans, considering them as distinct behaviours. **Methods:** Completed in January 2011, this review includes family studies, twin studies, association studies, genome-wide linkage studies and genome-wide association scan (GWAs) reporting different physical activity/inactivity-related phenotypes. **Results:** In regards to PA, familial aggregation studies resulted in heritability estimates ranging from 0% to 57%, and twin studies yielded heritability estimates (a^2) and shared environment (c^2) scores for PA phenotypes ranging from 0.00 to 0.85 and 0.00 to 0.84, respectively. Unique environmental (e^2) results are well dispersed from 0.12 to 0.72. Suggestive linkages were found with markers nearby different activity-related genes: *EDNRB*, *MC4R*, *UCP1*, *FABP2*, *CASR*, *SLC9A9*. Significant associations with PA phenotypes were found for *ACE*, *Gln223ARrg*, *MC4R* and *DRD2* genes. We found one GWAs that reported novel SNPs in the *PAPSS2* gene on chromosome *10q23.2* and in two intergenic regions on chromosomes *2q33.1* and *18p11.32*. Heritability estimates for PI ranged from 25% to 60% and linkage studies recorded higher LOD scores for PI versus PA. The *ACE* genotype was strongly associated with PI. **Conclusion:** There are potentially different genetic influences on PA versus PI phenotypes. Future studies should focus on the different genetic influences on PA and PI to improve our understanding of underlying determinants of these behaviours.

Keywords: Physical Activity; Physical Inactivity; Heritability; Candidate Gene Association; Genetic Linkage.

Introduction

Physical activity (PA) is a multifactorial trait with a complex etiology historically defined as “any bodily movement produced by skeletal muscles that results in energy expenditure” (Caspersen et al., 1985, p. 126). This broad definition may, however, be too simplistic to capture the range of human movement patterns that are related to health.

Recently, it has been suggested that physical inactivity (PI) per se may not simply represent the low end of the physical activity continuum, but may represent a different behavioral paradigm (Bray et al., 2009; Perusse et al., 1989). Indeed, it has been proposed that the “*emerging evidence for the role of sedentary behaviour on health, which may be independent of physical activity per se, finds us at the crossroad with respect to prescribing optimal daily human movement patterns for health*” (Katzmarzyk, 2010, p. 2717). Further, the mechanisms underlying the health outcomes associated with PI and PA may be linked to different biological/physiological pathways (Hamilton et al., 2007). Such a possibility has been sporadically studied, namely in rats and mice. For example, the idea that PI may impair lipid metabolism due to reduced lipoprotein lipase [(LPL) protein associated with PA and risk for Coronary Heart Disease (CHD)] activity in weight-bearing skeletal muscles and that this could be easily prevented by just performing non exercise activities - active living - in rats has been pursued (Bey & Hamilton, 2003). Bey and colleagues (2003) found, under acute and chronic inactivity conditions, severe reductions in muscle LPL and concluded that there is a potent regulatory process at the lower end of the PA continuum (i.e., PI) controlling LPL activity (Bey & Hamilton, 2003). Zderic and colleagues (2006) provided further evidence that PI motivates plasma lipids to suppress muscle LPL activity, after observing a difference of nearly 90% on heparin-releasable LPL activity between an inactive group and a control group. This impels us to wonder whether PA and PI may be associated with different genetic and environmental factors.

To date, most PA epidemiology research has focused on the psychosocial and environmental factors of distinct PA and PI phenotypes, and different

models and theories have been proposed to grasp the intricacies of its variation at the population level (Dishman et al., 1985). However, a clear and unequivocal interpretation of the net dynamics of their correlates is far from being unanimously accomplished, as they are influenced by demographic, biological, psychological, behavioral, social, cultural and environmental factors (Sherwood & Jeffery, 2000). This failure to clearly understand PA and PI physiology frustrates efforts to develop improved intervention programs and should prompt epidemiologists to follow different approaches.

Building on the promises of genetic epidemiology to unravel the complexities of PA and PI, several studies have attempted not only to identify the presence of genetic factors but also to investigate the importance of several candidate genes (Bray et al., 2009). We think that the time is right to present a comprehensive summary of research on the genetic epidemiology of PA and PI, considering them as distinct behaviours, rather than as a continuum of the same trait. As a step toward this end, we will review the available literature on family, twin, linkage and association studies as well as genome-wide association scan (GWAs) on genetic determinants of PA and PI in humans.

Methods

This present review, completed in January 2011, was conducted through systematic searches of Pubmed, Scopus and Web of Science databases for various combinations of the key words physical activity, physical inactivity, twins, genetic linkage, familial aggregation, familial resemblance, candidate gene association study. In addition, further hand searches or reference lists were done for all selected papers used in this review. We included family studies, twin studies, association studies, genome-wide linkage studies and genome-wide association studies (GWAs), published since 1980 until January 2011 regardless of sample size, method of data analysis, age range or gender groups. Different physical activity/inactivity-related phenotypes are reported: total activity (TA), total physical activity (TPA), physical activity energy expenditure (PAEE), light activity (LA), moderate activity (MA) and vigorous activity (VA), total daily activity (TDA), sports participation (SP), work physical

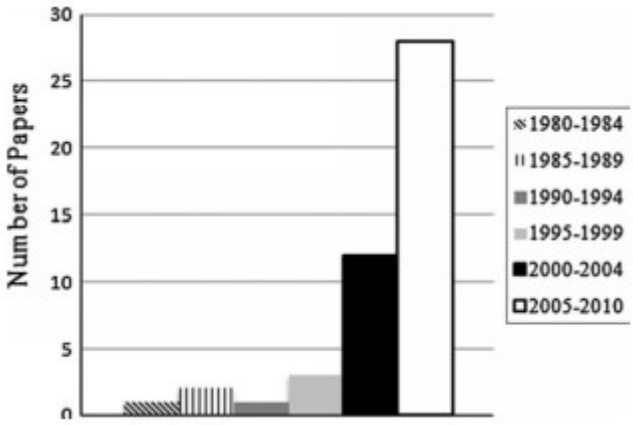
activity (WPA), vigorous exercise (VE), leisure-time physical activity (LTPA), physical activity during past year (PAPY), sedentary behaviour (SB), physical inactivity (PI), physical exercise (PE), daily physical activity (DPA), exercise participation (EP), moderate-to-vigorous physical activity (MVPA), moderate-to-vigorous activity (MVA), low physical activity (LPA), moderate physical activity (MPA), vigorous physical activity (VPA), sports physical activity (SPA), moderate-intensity leisure-time physical activity (MILTPA) and athlete status (AS). Due to the broad number of different phenotypes reported by the authors, and in order to provide a better understanding of the results, we decided to establish five distinct sets of phenotypes: (i) Physical Activity (PA)= TA, PAPY, TDA, TPA, LTPA, DPA, WPA, PAEE; (ii) Physical Inactivity (PI)= SB; (iii) Moderate-to-Vigorous Physical Activity (MVPA)= PE, MPA, MA, MILTPA, MVA; (iv) Vigorous Physical Activity (VPA)= VE, VA, EP, SPA, AS; (v) LPA= LA.

Results

The initial electronic search yielded an initial estimate of 1708 putative papers; after duplicate elimination 1176 studies were screened. After title and abstract analysis, 45 studies were selected according to the inclusion criteria: selected papers within the realm of genetic epidemiology had to provide information on heritability estimates, linkage results, identified genes or genetic markers, as well as recent analysis conducted by GWAs. Figure 1 shows the number of papers published per 5-year interval, showing an increase of interest in the genetics of PA since the start of the 21st century that is more obvious since 2005. From our research only 7 studies were published between 1980 and 1999.

Of the 45 studies, 7 estimated heritability coefficients based on information from family designs, 24 used twin samples to quantify additive genetic and environmental effects, 4 used the genome-wide linkage method to identify quantitative trait loci and 11 were based on association studies with putative genes. Only 1 GWAs study was retrieved.

Figure 1. Number of papers, per 5-year interval, on the genetics of PA.



PHYSICAL ACTIVITY

Family Studies

A very first step in genetic epidemiology applied to PA or PI research should start with the identification of familial aggregation estimated from odds ratios and/or intra-pair correlations in families. This approach is true for the most current trend across the frame of our search (Pérusse et al., 1988) and is followed by the computation of the effect size of genetic and environmental factors in each phenotype. Seven studies (Table I) were found focusing on PA and VPA.

Table I. Heritability estimates (h^2) of PA phenotypes from Family Studies.

Author / Country	Sample			Instrument	Phenotype	Heritability (h^2)
	N	Parents	Offspring			
Pérusse et al.1989 Canada	375 nuclear families (1610 subjects)	353 fathers (44.2±5.1) and 364 mothers (42.1±5.0).	477 sons (14.5±3.3) and 416 daughters (14.8±3.4).	B3DPAR	PA	h^2= 0.29
					EP	h^2= 0.00
Simonen et al. 2002 Canada	200 nuclear families from the *Quebec Family Study* (696 subjects)	140 fathers (54.5±7.3) and 172 mothers (52.2±7.6)	164 sons (26.3±9.2) and 220 daughters (28.1±10.4).	B3DPAR	TPA	h^2= 0.19±0.07
					MVPA	h^2= 0.16±0.03
					PYPA	h^2= 0.17±0.07
Mitchell et al. 2003 U.S.A.	42 pedigrees from *San Antonio Family Heart Study (1364)*			*Stanford 7-day physical activity recall questionnaire*	PA	h^2=0.09±0.05
Butte et al. 2006 U.S.A.	319 nuclear families from *Viva La Familia* study		1030 children (4-19 years)	Actiwatch accelerometer (Mini MItter Co. In. Bend. OR)	TPA	h^2= 0.57±0.09
					LPA	h^2= 0.39±0.11
					MPA	h^2= 0.50±0.08
					VPA	h^2= 0.32±0.11
Cai et al. 2006 U.S.A.	319 nuclear families from *Viva La Familia* study (1661 subjects)	631 parents	1030 children (4-19 years)	Actiwatch accelerometer (Mini MItter Co. In. Bend. OR)	TPA	h^2= 0.55±0.08
					LPA	h^2= 0.46±0.08
					MPA	h^2= 0.49±0.09
					VPA	h^2= 0.18±0.09
Seabra et al.2008 Portugal	2.375 nuclear families (9500 subjects).	2375 fathers (45.45±5.84) and 2375 mothers (42.92±5.47)	2425 sons (16.15±4.03) and 2325 daughters (16.01±3.98).	*Baecke Questionnaire of Habitual Physical Activity*	LTPA	h^2= 0.25±0.02
					SPA	h^2= 0.19±0.02
					TPA	h^2= 0.23±0.02
					WPA	h^2= 0.06±0.02
Choh et al. 2009 U.S.A.	219 males and 302 females (aged 18-86 yr) from 5 pedigrees - *Southwest Ohio Family Study*			*Baecke Questionnaire of Habitual Physical Activity*	LTPA	h^2= 0.17;c^2= 0.25
					SPA	h^2= 0.26; c^2= 0.13
					TPA	h^2= 0.29; c^2= 0.13
					WPA	h^2= 0.13; c^2= 0.13

Legend: PA - physical activity; EP – exercise participation; TPA – total physical activity; MVPA – moderate-to-vigorous physical activity; PYPA – past year's physical activity; LPA – low physical activity; MPA – moderate physical activity; VPA – vigorous physical activity; SPA – sport participation; WPA – work physical activity. B3DPAR – Bouchard 3-day physical activity record.; LTPA – leisure-time physical activity.

Using a sample of 375 nuclear families from the Quebec Family Study (QFS) and a complex path analysis model, Pérusse et al. (1989) estimated that 29% of PA variation was attributed to genetic predisposition. On the other hand the transmissible variance of VPA was only of cultural origin (b^2=12%) as no genetic influence was found. General PA referred to the sum of categorical scores 1 to 9 from the Bouchard three-day physical activity record (B3DPAR) (Bouchard et al., 1983), i.e., the total amount of daily activities, whereas VPA was defined as the sum of scores 6 to 9, i.e., exercise and sports activities. Generally, these results suggested that genetic effects are particularly meaningful for overall physical activity levels. The authors suggested the need for a better understanding of the two phenotypes in order to correctly interpret these results.

Simonen et al. (2002) assessed three different dimensions of PA in a sample from the QFS. Results suggested maximal heritability ranging from 16% to 19% of the phenotypic variation. Despite different methodological approaches – mainly on defining PA – these results were in accordance to those from the previous study.

Mitchell et al. (2003) used the Stanford 7-day physical activity recall questionnaire to verify the extent to which levels of PA aggregate in families. The results, in a sample of 42 extended pedigrees from the San Antonio Family Heart Study (SAFHS), revealed that modelling familial environmental effects as a shared household effect explained only 5% of PA variance and did not achieve statistical significance. However, when estimated as an upper-limit heritability, a significant 9% estimate was derived.

Viva la Familia Study (VLFS) was the first large-scale family study to use accelerometry to assess PA. Four different phenotypes were defined from PA to LPA and MVPA. Butte et al. (2006) and Cai et al. (2006) calculated genetic and environmental contributions to the phenotypic PA variation in 319 families. Maximal heritability ranged from 18% to 60%. Both studies suggested VPA as being less genetically determined whereas stronger genetic effects were found for PA.

In Portugal, Seabra et al. (2008) used four different PA phenotypes estimated by means of the Baecke questionnaire in a sample of 2,375 nuclear families. Genetic effects were computed using variance component models and ranged from 6% to 25%, after adjustment for the effects of multiple covariates. These results suggested a small, yet significant, influence of genetic factors on PA. LPA (25%) and PA (23%) were the phenotypes with the higher upper-limit heritability.

Choh et al. (2009) analysed 219 males and 302 females from 5 pedigrees of the Southwest Ohio Family Study (SOFS). Baecke questionnaire responses resulted in four different phenotypes related to LPA and PA. Using a maximum likelihood variance decomposition model resulted in 13% to 29% of genetic influence on PA levels that was significant in sport, leisure and total physical activity. No significant additive genetic effects were found on WPA.

In summary, maximal heritability estimates of PA ranged from 0 (Perusse et al., 1989) to 57% (Butte et al., 2006). The higher genetic contribution to PA was found when assessing PA with accelerometry. On the opposite side, the use of questionnaires resulted in a maximum value of 29% of additive genetic influence on PA variance. However, only two studies used accelerometry to assess PA, making it impossible to make any general inference. Across the studies, it is possible to observe a greater genetic contribution to PA than to VPA.

Twin Studies

Twin studies are another approach to determine the genetic and environmental effects on a phenotype. Monozygotic twins share the same genes identical by descent and dizygotic twins share on average 50% of their genes identical by descent; thus greater MZ resemblance is expected under assumptions of genetic contributions. Moreover, in contrast with family designs, twin designs allow us to jointly estimate shared environmental factors (c2), unique environmental factors (e^2) and additive genetic factors (a^2), which allows for a more precise evaluation of the relevance of genetics for a given trait.

Twin studies of PA have been published since 1981 and Table II summarizes the most relevant information of the 24 retrieved studies (Europe = 19; USA = 4; Across Countries = 1). Although the most relevant MZ-DZ twin comparisons are within-study rather than across studies, the complexity and diversity of models being tested within the realm of path analysis modelling, we nevertheless attempted a general overview instead of a study-by-study presentation.

Generally, MZ correlations [0.39 (Franks et al., 2005) to 0.98 (Beunen & Thomis, 1999) are greater than DZ correlations [-0.02 (Boomsma et al., 1989) to 0.72 (Koopmans et al., 1994)], meaning that genetic factors play an important role in individual differences in patterns and levels of PA. These correlations may be attributed either to a^2 and/or to c^2. Additive genetic effects and c^2 showed a similar pattern with a wide range of results for PA phenotypes from 0.00 (Joosen et al., 2005; Stubbe et al., 2005) to 0.85 (Stubbe et al., 2005) and 0.00 (De Moor et al., 2007b; De Moor et al., 2007c; Duncan et al., 2008; Eriksson et al., 2006; Joosen et al., 2005; Spinath et al., 2002) to 0.84 (Stubbe et al., 2005), respectively. Unique environmental (e^2) results are also well dispersed: 0.12 (Koopmans et al., 1994) to 0.72 (McCaffery et al., 2009).

In studies conducting gender specific analyses, h^2 are systematically higher in males [0.38 (Aaltonen et al., 2010) to 0.83 (Beunen & Thomis, 1999)] than in females [0.31 (Aaltonen et al., 2010) to 0.50 (Carlsson et al., 2006)]. On the other hand, c^2 seems to be more determinant on females PA [0.00 (De Moor et al., 2007a) to 0.54 (Beunen & Thomis, 1999)] than in males PA [0.00 (Beunen & Thomis, 1999; De Moor et al., 2007a; Maia et al., 2002) to 0.20 (Maia et al., 2002)]. Unique environmental factors (e^2) are fairly similar among genders, with a maximum of 0.63 in males and 0.67 in females both coming from the same study (Aaltonen et al., 2010).

With respect to the proportion of VPA explained by genetic factors, it ranged from 0.10 (McCaffery et al., 2009) to 0.85 (Stubbe et al., 2005). Furthermore, heritability estimates were much higher in males [0.83 (Beunen &

Thomis, 1999)] than females [0.44 (Beunen & Thomis, 1999)] for this particular level of PA.

The presence of age effects was investigated in two studies (Aaltonen et al., 2010; Stubbe et al., 2005; Stubbe et al., 2006), although following different methodologies and analysis, given their scope. In 2005, Stubbe and colleagues (Stubbe et al., 2005) verified a shift in the magnitude of genetic factors contributing to PA in their study using different age categories in their twin sample: 13-14 yrs, a^2=0; 15-16 yrs, a^2=0; 17-18 yrs, a^2=0.36; 19-20 yrs, a^2=0.85. In a follow-up study of a Finish twin cohort, Aaltonen et al. (2010) observed a decrease of the genetic influences from 44% of total variance at baseline to 34% at follow-up. Further, the genetic correlation (r_g = 0.72) between baseline and follow-up measures of PA seem to indicate that a considerable effect of the additive genetic influences at baseline remained present at follow-up.

MVPA and sedentary/inactivity phenotypes were jointly reported in one study (Fisher et al., 2010), and no multivariate analysis was conducted, e.g. genetic correlations and/or environmental correlations between them. All analysis were independently done in each phenotype.

Table II. Heritability estimates (h^2) and twins correlations of PA phenotypes from Twin Studies.

Author / Country	N	Twins correlations		Instrument	Phenotype	Heritability
		MZ	DZ			
Kaprio et al. 1981 Finland	1.537 MZ and 3.507 DZ twin pairs (aged over 18yr)	rMZ= 0.54	rDZ= 0.28	Questionnaire[20]	PA and SP	a^2= 0.62
Boomsma et al. 1989 Netherlands	44 MZ and 46 DZ twin pairs (aged 14-20 yr)	rMZM= 0.89 rMZF= 0.90	rDZM= 0.14 rDZF= 0.70 rOSDZ= -0.02	Questionnaire[21]	SP	a^2= 0.64; e^2= 0.36 $a^2_{♂}$= 0.77; $e^2_{♂}$= 0.23 $a^2_{♀}$= 0.35; $e^2_{♀}$= 0.65
Koopmans et al. 1994 Netherlands	1.587 twin pairs (aged 13-22 yr) and 1294 parents	rMZM= 0.89 rMZF= 0.85	rDZM= 0.60 rDZF= 0.72 rOSDZ= 0.35	Questionnaire[22]	SP	a^2= 0.48; c^2= 0.38; e^2= 0.12
Lauderdale et al. 1997 U.S.A.	3.344 male twin pairs (aged 33-51 yr)	rMZ=0.27-0.58	rDZ=0.07-0.44	Two sets of questions: 6 + 5.	MVPA	a^2= 0.08-0.58; c^2= 0.00-0.34
Aarnio et al. 1997 Finland	3.254 individual twins	rMZM= 0.72 rMZF= 0.64	rDZM= 0.45 rDZF= 0.41 rOSDZ= 0.22	Questionnaire[23]	PA	
Beunen et al. 1999 Belgium	91 twin pairs (aged 15 yr)	rMZM= 0.66 rMZF= 0.98	rDZM= 0.62 rDZF= 0.71 rOSDZ= 0.23	Number of hours spent in sport each week within the year preceding data collection.	SP	$a^2_{♂}$= 0.83; $c^2_{♂}$= 0.00; $e^2_{♂}$= 0.17 $a^2_{♀}$= 0.44; $c^2_{♀}$= 0.54; $e^2_{♀}$= 0.02
Spinath et al. 2002 Germany	168 MZ and 132 DZ twin pairs (aged 18-70 yr) from GOSAT	rMZ= 0.40	rDZ= 0.20	Kaulins and Willis M101 motion recorders	ACS	a^2= 0.42; e^2= 0.58
Maia et al. 2002 Portugal	411 twin pairs (aged 12-25 yr)	rMZM= 0.82 rMZF= 0.90	rDZM= 0.46 rDZF= 0.53 rOSDZ= 0.49	*Baecke Questionnaire of Habitual Physical Activity*(Baecke et al. 1982)	SP	$a^2_{♂}$= 0.68; $c^2_{♂}$= 0.20; $e^2_{♂}$= 0.12 $a^2_{♀}$= 0.40; $c^2_{♀}$= 0.28; $e^2_{♀}$= 0.32

[20] Leisure-time PA was measured by asking the opinion on the amount of PA currently engaged in, its intensity and duration, and number of years of PA engaged in the adult life.
[21] Dichotomous variable (Yes or No) - Subjects answered the question: "Have you been involved in sports activities during the past three months?"
[22] Dichotomous variable (Yes or No) - Subjects answered the question: "Do you participate in sports?"
[23] Two questions asked referring to frequency and intensity of PA, respectively.

		rMZM= 0.69 rMZF= 0.72	rDZM= 0.22 rDZF= 0.56 rOSDZ= 0.31		LTPA	$a^2_{♂}$= 0.63; $c^2_{♂}$= 0.00; $e^2_{♂}$= 0.37 $a^2_{♀}$= 0.32; $c^2_{♀}$= 0.38; $e^2_{♀}$= 0.30
Frederiksen et al. 2003 Denmark	616 MZ and 642 DZ twin pairs (aged45-68 yr)	rMZ= 0.54	rDZ= 0.13	Questionnaire[24]	SP	a^2= 0.49; c^2= 0.00; e^2= 0.51
Franks et al. 2005 U.S.A.	62 MZ and 38 DZ twin pairs	rMZ= 0.46	rDZ= 0.19	Questionnaire[25]	PA	a^2= 0.00; c^2= 0.65; e^2= 0.35
		rMZ= 0.39	rDZ= 0.18		PAEE	a^2= 0.00; c^2= 0.69; e^2= 0.31
Simonen et al. 2004 Finland	147 MZ and 153 DZ twin pairs (aged 35-70 yr)	Adolescent Exercise rMZ= 0.54	rDZ= 0.46	Questionnaire[26]	AE	A= 0.20; D= 0.23; C= 0.11; E= 0.46
		Adulthood Exercise rMZ= 0.54	rDZ= 0.13		CS	A= 0.13; D= 0.43; C= 0.00; E= 0.44
Joosen et al. 2005 Netherlands	12 MZ and 8 DZ twin pairs (aged 18-39 yr)				Respiration Chamber	
		rMZ=0.56	rDZ=0.43	Respiration Chamber	PA	a^2= 0.00; c^2= 0.41; e^2= 0.59
		rMZ=0.78	rDZ=0.60		AEE	a^2= 0.00; c^2= 0.68; e^2= 0.32
					Daily Life	
		rMZ= 0.88	rDZ= 0.42	Triaxial accelerometer (Tracmor; Philips Research, Eindhoven, Netherlands)	PA	a^2= 0.78; c^2= 0.00; e^2= 0.22
		rMZ= 0.82	rDZ= 0.64		AEE	a^2= 0.72; c^2= 0.00; e^2= 0.28
Stubbe et al. 2005	1.095 MZ. 811 DZ and				SP	13 – 14 yr age group

[24] Subjects answered the following question: "Do you in your leisure time participate in any of the following sports?: Jogging, gymnastics, swimming, tennis, badminton, football, handball, aerobics, rowing, table tennis, or volleyball." Subjects involved in any of these sports were considered active.

[25] Hours/week over the past year in which the child was typically engaged in sports and recreational activities requiring a greater expenditure of energy than normally needed for daily grooming, bathing, and eating – completed by the parent.

[26] Information for every event lasting at least 3 months was collected for exercise and other physical leisure time activities than exercise. Information included the time span of participation in years, months per year of participation, and mean frequency (times per week), duration (minutes per session), and intensity (light, moderate, strenuous), and whether participation was at a competitive level if sports related (yes/ no). The summary score was calculated by summing the weekly hours separately for exercise and other physical leisure time activities both for adolescence (age 12–18) and for adulthood (age 18 to age at interview).

Netherlands	722 opposite-sex twin pairs	rMZM= 0.88 rMZF= 0.87	rDZM= 0.82 rDZF= 0.84 rOSDZ= 0.47	Questionnaire[27]		a^2= 0.00; c^2= 0.84; e^2= 0.16
		rMZM= 0.80 rMZF= 0.83	rDZM= 0.82 rDZF= 0.84 rOSDZ= 0.47	Questionnaire		15 – 16 yr age group a^2= 0.00; c^2= 0.78; e^2= 0.22
		rMZM= 0.88 rMZF= 0.80	rDZM= 0.65 rDZF= 0.68 rOSDZ= 0.18	Questionnaire		17 – 18 yr age group a^2= 0.36; c^2= 0.47; e^2= 0.17
		rMZM= 0.86 rMZF= 0.83	rDZM= 0.35 rDZF= 0.53 rOSDZ= 0.48	Questionnaire		19 – 20 yr age group a^2= 0.85; c^2= 0.00; e^2= 0.15
Carlsson et al. 2006 Sweden	13.362 twin pairs (aged 14-46 yrs)	rMZM= 0.62 rMZF= 0.58	rDZM= 0.31 rDZF= 0.30	Past year questionnaire[28]	PA	$a^2_{♂}$= 0.57; $c^2_{♂}$= 0.03; $e^2_{♂}$= 0.40 $a^2_{♀}$= 0.50; $c^2_{♀}$= 0.06; $e^2_{♀}$= 0.44
Eriksson et al. 2006 Sweden	1.022 twin pairs	rMZ= 0.46	rDZ= 0.19	*Baecke Questionnaire of Habitual Physical Activity*	TPA	a^2= 0.49; e^2= 0.51
		rMZ= 0.39	rDZ= 0.18		PA	a^2= 0.40; e^2= 0.60
		rMZ= 0.58	rDZ= 0.23		OPA	a^2= 0.57; e^2= 0.43
		rMZ= 0.55	rDZ= 0.32		SPA	a^2= 0.56; e^2= 0.44
Stubbe et al. 2006 Australia, Denmark, Finland, the Netherlands, Norway, Sweden, and United Kingdom	37.051 Twin pairs	rMZM= 0.43-0.71 rMZF= 0.48-0.70	rDZM= 0.27-0.48 rDZF= 0.24-0.38 rOSDZ= 0.07-0.25	Questionnaire[29]	EP	a^2= 0.27-0.71; c^2= 0.37; e^2= 0.30-0.60

[27] Dichotomous variable (Yes or No) - Subjects answered the question: “Do you participate in sports regularly?”
[28] Average PA during leisure time in the past year.
[29] Different questions were asked in each of the countries involved.

De Moor et al. 2007a Netherlands	622 families (164 MZ and 236 DZ twin pairs and 582 siblings)			Questionnaire[30]	PE	$a^2_{♂}$= 0.68; $c^2_{♂}$= 0.00; $e^2_{♂}$= 0.32 $a^2_{♀}$= 0.46; $c^2_{♀}$= 0.00; $e^2_{♀}$= 0.54
De Moor et al. 2007c Netherlands	5.200 twins and siblings (aged 18-50 yr)	rMZM= 0.59 rMZF= 0.55	rDZM= 0.25 rDZF= 0.28 rOSDZ= 0.21	Questionnaire[10]	PE	a^2= 0.54; c^2= 0.00; e^2= 0.46
De Moor et al. 2007b Great Britain	793 MZ and 1000 DZ twin pairs (aged 51.9±12.8)	rMZ=0.66	rDZ=0.32	Questionnaire[31]	SP	a^2= 0.65; c^2= 0.00; e^2= 0.35
Duncan et al. 2008 USA	1.003 MZ (aged 29±13) and 386 DZ (aged 33±15) twin pairs.	rMZ= 0.43	rDZ= 0.30	Questionnaire[32]	PA 60 min cut-point	a^2= 0.45; e^2= 0.55
		rMZ= 0.30	rDZ= 0.25		PA 150 min cut-point	c^2= 0.28; e^2= 0.72
Wood et al. 2008 United Kingdom	325 MZ. 253 same-sex DZ and 258 opposite sex DZ twins (aged 8.51±0.42)	rMZ= 0.72	rDZ= 0.58	Actigraph	Actigraph Measurements	a^2= 0.92; c^2=0.00; e^2 = 0.08
McCaffery et al. 2009 USA	2710 MZ (aged 41.06±3.15) and 2327 (aged 41.07±2.80) DZ twin pairs.			Questionnaire[33]	VE	a^2= 0.10; c^2= 0.18; e^2= 0.72
Aaltonen et al. 2010 Finland	4280 MZ and 9276 DZ twins at baseline (1975) and 4383 MZ and 9439 DZ twins at follow-up (1981).	Baseline rMZ=0.54 Follow-up rMZ=0.43	Baseline rDZ=0.24 Follow-up rDZ=0.15	Questionnaire[34]	LTPA	Baseline a^2=0.44;e^2=0.56 $a^2_{♂}$=0.47;$e^2_{♂}$=0.54 $a^2_{♀}$=0.42;$e^2_{♀}$=0.56 Follow-up h^2=0,34;e^2=0,66 $h^2_{♂}$=0,38;$e^2_{♂}$=0.63 $h^2_{♀}$=0,31;$e^2_{♀}$=0,67

[30] Variable dichotomous (Yes or No) - Subjects answered the question: "Do you participate in exercise regularly?'. If the participants responded affirmative, further information on type, frequency and duration of exercise was gathered.
[31] Subjects were asked to indicate for a list of sports whether they had ever participated in each of these sports and what was the highest level at which they had ever competed in these sports.
[32] Subjects answered how many times per week they exercise moderately for at least 30 minutes and vigorously for at least 20 minutes. A physical activity measure was derived by summing the reported number of moderate and vigorous blocks of activity to estimate the total minutes per week of moderate-to-vigorous activity. Based on different cut points for total physical activity - at least 60 minutes per week, and at least 150 minutes per week – a dichotomous variable was created.
[33] VE was defined by self-reported regular participation in one or more of 5 common vigorous-intensity aerobic activities over the past 3 months. Subjects who endorsed in one or more activities were designated as vigorous exercisers, whereas participants who did not endorse any of these activities were designated as nonvigorous exercisers.
[34] The volume of leisure activity in metabolic equivalent units (MET index) was based on a series of structured questions on LTPA and physical activity during journeys to and from work.

Fisher et al. 2010 England	57 MZ (aged 11.06±0.59) and 60 DZ (aged 11.27±0.48) twin pairs.	rMZ= 0.76	rDZ= 0.71	Actigraph 7164	TPA	c^2= 0.73; e^2= 0.27
		rMZ= 0.69	rDZ= 0.52		MVPA	c^2= 0.61; e^2= 0.39

Legend: PA – physical activity; SP – sports participation; MVPA – moderate-to-vigorous physical activity; LTPA – leisure time physical activity; PE – physical exercise; TPA – total physical activity; SPA – sports physical activity; PAEE – physical activity energy expenditure; VE – vigorous exercise; AE – adolescent exercise; CS – competitive sports; EP – exercise participation; A – additive genetic factors; D – dominance genetic factors; C – common environmental factors; E – unique environmental factors; ACS – Actometer composite score; MZM - male monozygotic twins; MZF – female monozygotic twins; rDZM – male dizygotic twins; rDZF - female dizygotic twins; rOSDZ - opposite-sex dizygotic twins; MZ - monozygotic twins; DZ - dizygotic twins.

Linkage Studies

Linkage studies are the third step of genetic analysis and they aim to identify regions within the genome responsible for the variation of a particular phenotype. Only 4 (USA = 2; Europe = 2) linkage studies were found.

The first study was performed by Simonen et al. (2003a) using a total of 432 markers when performing this genome wide scan in 767 subjects from 207 nuclear families from the QFS. It should be emphasized that marker sets in linkage studies are highly polymorphic, meaning that a locus might exhibit moderate to high levels of intraspecific variation in the alleles available for that locus. Three different phenotypes related with PA and MVPA were assessed. Suggestive linkages were found in chromosomal regions *11p15*, *15q13.3* and *13q22-q31* for PA and *4q28.2*, *7p11.2*, *9q31.1* and *13q22-q31* for MVPA. The suggestive linkage at region *13q22-q31* may be associated with the gene encoding endothelin B receptor, found to mediate increases in spontaneous locomotor activity in mice (Nagasaka et al., 1999) and that has been mapped on chromosome *13q22*.

Cai et al. (2006) evaluated 1030 children and 631 parents from the VLFS. A panel of markers spaced an average of 10 centimorgans (cM) apart (range, 2.4 to 24.1 cM) was established for a total of 384 markers. Suggestive linkages with PA, LPA and MVPA were found on chromosomal region *18q21.1*. This region contains melanocortin 4-receptor gene (*MC4R*), which is known for its relevant function on energy expenditure regulation.

Table III. Promising and suggestive linkages with PA phenotypes.

Author / Country	N	Instrument	Phenotype	Locus	Genetic Marker	P
Simonen et al. 2003b U.S.A.	767 subjects from 207 nuclear family units of the Quebec Family Study	B3DPAR	TPA	*13q22-q31*	*D13S317*	0.029
			MVPA	*4q28.2*	*UCP1*	0.005
				7p11.2	*IGFBP1*	0.006
				9q31.1	*D9S938*	0.0028
				13q22-q31	*D13S317*	0.0067
		Past year questionnaire[35]	TSPA	*11p15*	*C11P15_3*	0.0089
				15q13.3	*D15S165*	0.009
Cai et al. 2006 U.S.A.	1030 children and 631 parents from the VIVA LA FAMILIA Study	Actiwatch accelerometer (Mini MItter, Bend, OR)	TPA	*18q*	*D18S64*	<0.001
			LPA	*18q*	*D18S1102-D18S474*	<0.001
			MPA	*18q*	*D18S64*	<0.001
			VPA	*18q*	*D18S64*	0.01
De Moor et al. 2007a Netherlands	622 families (164 MZ and 236 DZ twin pairs and 582 siblings)	Questionnaire[36]	EP	*19p13.3*	*D19S247*	<0.01
De Moor et al. 2007b Great Britain	700 DZ twin pairs (aged 54.6±11.5)	Questionnaire[37]	AS	*3q24*	*D3S1569*	<0.01
				4q32.3	*D4S1597*	<0.01

Legend: TPA – total physical activity; MVPA – moderate-to-vigorous physical activity; TSPA – time spent in physical activity; LPA – low physical activity; MPA – moderate physical activity; VPA – vigorous physical activity; EP – Exercise participation; AS – athlete status.

[35] Average PA during leisure time in the past year.
[36] Variable dichotomous (Yes or No) - Subjects answered the question: "Do you participate in exercise regularly?'. If the participants responded affirmative, further information on type, frequency and duration of exercise was gathered.
[37] Subjects were asked to indicate for a list of sports whether they had ever participated in each of these sports and what was the highest level at which they had ever competed in these sports.

In the Netherlands, De Moor et al. (2007a) performed a 400 marker 10cM genome scan in 622 families (164 MZ and 236 DZ twin pairs and 582 siblings). The phenotype VPA found suggestive linkage in all subjects in chromosome 19 nearby marker *D19S247*. Moreover, this was the first study reporting gender comparisons and females were found to contribute more strongly to the LOD score (maximum LOD in females=2.87 at 11 and 12cM, versus 0.83 in males at 9-12cM). The region does not harbour genes that have been related to exercise or physical activity phenotypes.

In a study with 700 British female DZ twin pairs, De Moor et al. (2007b) reported the first genome-wide linkage scan for VPA. Subjects were asked to indicate for a list of sports whether they had ever participated in each of those sports and what was the highest level at which they had ever competed in those sports. A total of 1946 markers were used and linkage signals for VPA were found on chromosomal region *3q22-q24* [including the sodium/hydrogen exchanger 9 (*SLC9A9*) and calcium-sensing receptor (*CASR*) gene] and *4q31-q34* [very close to uncoupling proten1 (*UCP1*) gene and to fatty-acid binding protein 2 (*FABP2*) gene], nearby markers *D3S1569* and *D4S1597*, respectively.

In summary, different phenotypes have been assessed by distinct instruments, in sample sizes ranging from 767 (Simonen et al., 2003a) to 4488 subjects (De Moor et al., 2007b), resulting in contrasting LOD scores from distinct genetic markers. No markers in common were discovered across studies, making it challenging to interpret the results. Suggestive linkages were found with markers nearby different activity-related genes: *EDNRB*, *MC4R*, *UCP1*, *FABP2*, *CASR*, *and SLC9A9*. The De Moor et al. (2007a), GWAs findings, however, were unable to overlap with suggested gene regions/genes from the linkage studies. Suggestions of association between the *CASR* gene and PA have been found in adolescent girls (Lorentzon et al., 2001). This gene is involved in muscle activity and is located just outside the confidence interval of the peak on *3q21* found by De Moor et al. (2007b). *FABP2* gene importance was firstly stressed by a linkage study carried out in 412 sibling pairs using 289 markers (Bouchard et al., 2000). Its involvement in the metabolism of long-chain fatty acids may be the reason for the suggestive linkage found for maximal

oxygen uptake in response to a 20-week exercise-training program on *4q26*. Again, *FABP2* gene is close to the suggestive linkage found on *4q31-q34* (De Moor et al., 2007b). The *UCP1* gene (involved in heat generation) was associated with MVPA in Simonen et al. (2003b) study. It is located in the 4q28-31 region nearby the suggestive linkage found by De Moor et al. (2007b).

Association Studies

Association studies usually follow a case-control design and report on a candidate gene associated with a particular phenotype. The ensemble of 10 studies presented in Table IVA analysed different genes: *CASR* (OMIM Number: 601199) - found to have an association between serum ionized calcium and a common polymorphism, *ala986* to *ser* (A986S), in the cytoplasmic tail of the calcium-sensing receptor (Scillitani et al., 2004), involved in calcium homeostasis; *LEPR* (OMIM Number: 601007) - a single-transmembrane-domain receptor of the cytokine receptor family, involved on the function of an adipocyte-specific hormone (LEP) that regulates adipose-tissue mass through hypothalamic effects on satiety and energy expenditure; *DRD2* (OMIM Number: 126450) – D2 dopamine receptor is a G protein-coupled receptor located on postsynaptic dopaminergic neurons whose signalling governs physiologic functions related to locomotion and hormone production; *MC4R* (OMIM Number: 155541) - melanocortin receptors are involved in the inhibition of feeding behaviour and the regulation of metabolism, thus are important on energy balance; *ACE* (OMIM Number: 106180) - Angiotensin I-converting enzyme is a dipeptidyl carboxypeptidase that plays an important role in blood pressure regulation and electrolyte balance by hydrolysing angiotensin I into angiotensin II, a potent vasopressor, and aldosterone-stimulating peptide. The enzyme is also able to inactivate bradykinin, a potent vasodilator. Further, angiotensin converter regulates exercise resistance and the increase of physical response to exercise; and FTO (OMIM Number: 610966) – associated with obesity.

Lorentzon et al. (2001) evaluated the association between the *A986S* polymorphism of the *CASR* and levels of PA in 97 female adolescents and

found that S allele carriers had significantly lower levels of PA (2,9 ± 2,6 vs. 4,3 ± 2,6 h/week; P< 0,01) than non-carriers.

Stefan et al. (2002) aimed to understand the influence of the *Gln223Arg* polymorphism of the *LEPR* on energy metabolism and adiposity in non-diabetic Pima Indians. PA was calculated by dividing the 24h energy expenditure by the sleeping metabolic rate (Snitker et al., 2001). The results showed that subjects homozygous for the *Arg223-encoding* allele had lower PA levels (1.36±0.09) than those homozygous (1.43±0.11) for the *Gln223-encoding* allele. This suggests the possibility of *Gln223Arg* polymorphism being associated with *LEPR* that has been connected with PA.

Fuentes et al. (2002) evaluated the association of the insertion/deletion polymorphism of the angiotensin-converting enzyme gene (*ACE I/D*) with MVPA, defined as the practice of at least 20-30 min of physical exercise during leisure time. A sample of 167 males and 288 females was assessed. Comparisons between genotypes showed no significant differences in *ACE DD*, *ID* and *II* distributions among the MVPA groups.

In Canada, Simonen et al. (2003a) used two different cohorts [(QFS and Heritage Family Study (HFS)] to study the association between a marker in *DRD2* and physical activity level. Two different questionnaires were used: QFS – B3DPAR (Bouchard et al., 1983); HFS – ARIC questionnaire. The authors designed a case-control study based on race and gender. For the QFS, a significant association (F=4.02, p=0.020) was found in the whole cohort between the time spent being physically active in the last year and *DRD2* genotype. However, no significant differences were observed between genotypes in PA phenotypes as assessed from the B3DPAR (Bouchard et al., 1983). Female *TT* homozygotes had significantly lower time spent in PA during past year (F=4.42, p=0.016) than *CC* homozygotes (25%) and *CT* heterozygotes (34%). Results in the HFS recorded differences between Blacks and Whites (p<0.001) for the *DRD2* allele frequencies. *T* allele frequencies were 63% among Blacks and 30% among Whites. Stratifying the cohort by sex and race resulted in significant associations between the sports index (F=5.93,

p=0.005) and work index (*F*=3.61, *p*=0.042) with *DRD2* genotype, among white females. Of those, *TT* homozygotes had 29% and 38% lower sports index than the *CC* homozygotes and *CT* heterozygotes, respectively. Moreover, White *TT* homozygote women also had 27% and 33% lower index values than *CC* homozygotes and *CT* heterozygotes, respectively. White women with a leisure time index above the median have significantly higher (χ^2=5.85, *p*=0.016) percentage of *T* allele carriers when comparing with those below the median.

Winnicki et al. (2004) looked into the association between the *ACE I/D* polymorphism and physical activity status in a sample of 355 never-treated, stage 1 hypertensives (265 men, 90 women) from the HARVEST study. The authors opted to divide the sample into two different groups based on PA level. *ACE* genotype was a strong predictor of PA level with subjects engaging in sports activities showing more than twice the frequency of the II genotype than their sedentary counterparts.

Loos et al. (2005) examined the contribution to physical activity levels of a polymorphism located nearby *MC4R* gene (*MC4R-C-2745T*). Significant associations were found between all PA phenotypes with *MC4R-C-2745T* polymorphism in this sample of 669 French-Canadian subjects from the QFS. Significantly lower score for MVPA was found in *T* allele homozygotes. A separate analysis of non-obese subjects resulted in a more pronounced association. Furthermore, the same trend was observed for the past year's PA with homozygotes for the *T*-allele spending on average 1 h/week less on PA compared to the other genotypes.

Richert et al. (2007) investigated the association of the *Gln223Arg* polymorphism in the leptin receptor gene (*LEPR*) with bone mass in 222 prepubertal boys from Switzerland. A significant difference (*p*=0.016) in PA among the three genotypes was found: homozygosity for the *Arg-encoding* allele was associated with lower PA compared with the other genotypes.

In Norway, Berentzen et al. (2008) hypothesized that *AA*-genotypes of *rs9939609 FTO* had lower levels of PA. Examinations of 234 obese and 323 controls were performed and no association between PA and *rs9939609 FTO*

was found. The same result was produced by Hakanen et al. (2009), who followed 438 children from age 7 months to 15 years and also found no association between *rs9939609 FTO* and PA index.

More recently, Liu et al. (2010) evaluated 1978 subjects from three cohorts – the Georgia Cardiovascular Twin study, the LACHY study and the APEX study. No significant association was found between *FTO rs9939609* and VPA, regardless the instrument used (questionnaire in the LACHY and APEX studies and accelerometry in the LACHY study).

In summary, different genes have been associated with PA phenotypes, although the results are quite dispersed. *FTO rs9939609* is known for its impact on fat mass regulation (Scuteri et al. 2007) but the two studies (Berentzen et al., 2008; Liu et al., 2010) investigating its influence on PA failed to demonstrate a significant and positive impact. *ACE* is linked to cardiovascular homeostasis through angiotensin II formation and bradykinin inactivation (Wang & Staessen, 2000) and Fuentes et al. (2002) hypothesized that it could exert some influence on PA but found no association. This was in opposition to the results of Winnicki et al. (2004) that found a strong predictor of PA in *ACE* genotype. *LEPR* regulates adipose-tissue mass through hypothalamus effects on fullness and energy use. Its administration in animals produces the activation of the sympathetic nervous system (SNS) (Tang-Christensen et al., 1999) that has been associated with measurements of spontaneous PA (Christin et al., 1993). Stefan et al. (2002) and Richert et al. (2007) results may reflect this evidence with a significant influence of *Gln223Arg* polymorphism on PA levels. The remaining studies focused their attention in different genes leading to less clear results in terms of the influence of each in PA. *MC4R* plays an important role in the physiology of obesity (Larsen et al., 2005) and has been linked to energy expenditure (Cole et al., 2010). A significant association with higher levels of PA was demonstrated by Loos et al. (2005). *DRD2* is involved in the amount of movement (in mice) and the rewarding system, as well as on motor control (Gingrich & Caron, 1993), making it an obvious candidate gene for Simonen et al. (2003a). However, the authors only found association between *DRD2* and PA as assessed by recall questionnaire.

Table IV. Association and GWAs studies.

Table IV A. Associations between polymorphisms and PA phenotypes

Author / Country	Locus	Gene	Polymorphism	N			Instrument	Phenotype	P
				D/D	D/R	R/R			
Lorentzon et al. 2001 USA	*3q21.1*	*CASR*	ALA986SER	71.1% A/A	25.8% A/S	3.1% S/S	Questionnaire	PYPA	0.01
Stefan et al. 2002 USA	*1p31.3*	*LEPR*	*Gln223ARG*	206 Gln/Gln	205 Gln/Arg	41 Arg/Arg	24h Energy Expenditure 24h Respiratory Quotient	24hEE PAL	0.008
Fuentes et al. 2002 Finland	*17q23.3*	*ACE*	*INS/DEL*	Frequent 36.5% DD	46% ID	17.5% II	Questionnaire[38]	MILTPA	0,269
				Non – Frequent 29.5% DD	50% ID	20.5% II			
Simonen et al. 2003a Canada	*11q23.2*	*DRD2*	*DRD2*	QFS – Women 165 C/C	131 C/T	23 T/T	[39]B3DPAR Past year questionnaire	TSPA TPA MVPA	0.333 0.722 0.868
				QFS – Men 197 C/C	180 C/T	25 T/T	B3DPAR Past year questionnaire	TSPA TPA MVPA	0.16 0.836 0.609
				10 C/C	43 C/T	38 T/T	HFS – Black Men		

[38] MILTPA was defined as the practice of at least 20–30 min of physical exercise during leisure time so that the person is at least a little out of breath and sweating. Subjects answered the following question: "How often do you practice physical exercise during leisure time for at least 20–30 min so that you are a little out of breath and sweating?" The possible answers to this question were: *1*) daily, *2*) 2–3 times a week, *3*) once a week, *4*) 2–3 times a month, *5*) few times a year or less, *6*) I can't do it because of disease or disability. Subjects who answered *1* or *2* were classified as having frequent MILTPA. Subjects who answered *3*, *4*, or *5* were classified as having nonfrequent MILTPA, and those who answered 6 were excluded from the study.

[39] Using a 3-d activity diary, which included one weekend day, subjects were instructed to record the dominant activity for each 15 min period during 24 h using a list of activities. Each period was given a score ranging from 1 to 9, with 1 corresponding to sleep and 9 to the activities characterized by the highest energy expenditure levels. The phenotype of participation in "moderate to strenuous physical activities" (categories 5–9), which included light manual work as well as strenuous exercise modes or intense manual work, was also used. A phenotype labelled as "total daily activity level" was defined as the sum of all categorical values (1–9).

							ARIC	SP	0.131
								LTPA	0.413
								WPA	0.617
				20 C/C	98 C/T	66 T/T	ARIC	SP	0.882
								LTPA	0.536
								WPA	0.938
				116 C/C	109 C/T	16 T/T	ARIC	SP	0.551
								LTPA	0.894
								WPA	0.314
				119 C/C	115 C/T	22 T/T	ARIC	SP	0.023
								LTPA	0.082
								WPA	0.004
Winnicki et al. 2004 Italy	*17q23.3*	*ACE*	*INS/DEL*	65 II	193 ID	97 DD	Questionnaire[40]	PA	0.002
Loos et al. 2005 Canada	*18q21.32*	*MC4R*	*MC4R-C-2745T*	360 CC	231 /T	51 TT	B3DPAR[41]	TPA	0.08
								MVPA	0.005
				331 CC	210 CT	45 TT	SAVT (h/week) on PA during the past year	PY	0.06
Richert et al. 2007 Switzerland	*1p31.3*	*LEPR*	*Gln223Arg*	49.8% Gln/Gln	40.4% Gln/Arg	9.8% Arg/Arg	Questionnaire	PAEE	
Berentzen et al. 2008 Norway	*16q12.2*	*FTO*	*rs9939609*	Obese			Questionnaire[42]	LTPA	
				26.4% TT	45% TA	28.6% AA			
				Controls					
				35.6% TT	47.2% TA	17.2% AA			

[40] Subjects were categorized as sedentary (class 0) if they did not regularly perform any sports activity, exercisers (class 1) if they performed regularly leisure physical activities, such as walking, gardening, yard working, etc., and/or performed sports like running, jogging, cycling, swimming, soccer, tennis at least once a week during the previous 2 months.

[41] Using a 3-d activity diary, which included one weekend day, subjects were instructed to record the dominant activity for each 15 min period during 24 h using a list of activities. Each period was given a score ranging from 1 to 9, with 1 corresponding to sleep and 9 to the activities characterized by the highest energy expenditure levels. The phenotype of participation in "moderate to strenuous physical activities" (categories 5–9), which included light manual work as well as strenuous exercise modes or intense manual work, was also used. A phenotype labelled as "total daily activity level" was defined as the sum of all categorical values (1–9).

[42] Self-administered questionnaire: sedentary, sedentary or light PA less than 2 h/wk; light PA, light PA 2–4 h/wk; moderate PA, more than 4 h/wk of light PA or 2–4 h/wk of more vigorous PA; high PA, more than 4 h/wk of moderate PA or regular heavy exercise/competitive sports several times per week.

Hakanen et al. 2009 Finland	*16q12.2*	*FTO*		31.9% TT	26.4% TA	28.5% AA	Questionnaire[43]	PAI	>0.99
Cole et al. 2010 USA	*1704*	*MC4R*	*MC4R*				Actiwatch accelerometer (Mini MItter, Bend, OR)	TPA	0.004
								ATMA	0.016
	602							ATVA	0.021
Liu et al. 2010 Europe and USA	*16q12.2*	*FTO*	*rs9939609*	181 TT	359 AT	204 AA	Questionnaire[44]	VPA	0.62
				123 TT	221 AT	129 AA	Actigraph 7164		0.26

Legend: D – dominant; R – recessive.

[43] Leisure-time physical activity index (PAI) was calculated as a multiple of the resting metabolic rate (hours per week) by multiplying the frequency, mean duration in minutes, and mean intensity of weekly leisure-time physical activity.
[44] PA was quantified using a modified version of the previous day physical activity recall, which recorded activities in 30-min time blocks for 24-h period.

Genome Wide Association scans (GWAs)

In 2009, De Moor and colleagues (2009) performed the only GWAs located by our search, aiming to identify genetic variants associated with adult LTPA in 1644 unrelated Dutch and 978 American adults of European ancestry (Table IV B). The analysis uncovered 37 novel SNPs for exercise participation that gather in three different genomic regions: in the *PAPSS2* gene [(OMIM Number: 603005) – encodes a sulfotransferase (*SULT*) enzyme that is involved in the sulfation of compounds such as lipids, carbohydrates, proteins, and exogenous drugs] on chromosome *10q23.2* and in two intergenic regions on chromosomes *2q33.1* and *18p11.32*. The *rs1240556* in the *LEPR* gene and the *rs8036270* in the *GABRG3* gene [(OMIM Number: 600233) - member of the *GABA-A* receptor gene family through which *GABA*, the major inhibitory neurotransmitter in the mammalian brain, acts, playing a role in regulating neuronal excitability throughout the nervous system] showed suggestive association. These results led the authors to suggest that PA is under the influence of many genetic variants with small effect sizes.

Table IV B. SNP for leisure time exercise behaviour (LTEB) that reach threshold for genome-wide significant association.

Author / Country	Phenotype	Instrument	Locus	Gene	Polymorphism	P
De Moor et al. 2009 Netherlands & USA	LTEB	Questionnaire [45]	*2q33.1*	*DNAPTP6*	*rs12612420*	<0.00001
			1023.2	*PAPSS2*	*rs10887741*	
			18p11.32	*C18orf2*	*rs8097348*	

Legend: TPA – total physical activity; 24hEE – twenty four hour energy expenditure; PAL – physical activity level; MVPA – moderate-to-vigorous physical activity; TSPA – time spent in physical activity; PY – past year physical inactivity; PAEE – physical activity energy expenditure; ATMA – awake time in moderate activity; ATVA – awake time in vigorous activity; MILTPA – moderate-intensity leisure time physical activity; LTEB – Leisure Time Exercise Behaviour; D – dominant; R - Recessive.

[45] Questions about type, frequency, and duration of exercise were asked. Subjects were classified into regular exercisers or nonexercisers using a threshold of 4 MET.h (metabolic equivalents.hours per week).

PHYSICAL INACTIVITY

Our research failed to find any study aimed exclusively at assessing genetic influences on PI phenotypes. Therefore, with the exception of the Fisher et al. (2010) twin study, all the subsequent studies have already been briefly addressed.

Family Studies

Of the seven family studies just three of them evaluated PI phenotypes (Table VA). Simonen et al. (2002) recorded maximal heritability estimates of 25% for physical inactivity as assessed by B3DPAR (Bouchard et al., 1983). The other two studies (Butte et al., 2006; Cai et al., 2006) were based in the same sample (VLFS) and assessed sedentary behaviour. Similar upper-limit heritability estimates were recorded: 0.60 and 0.57, respectively.

Twin Studies

Only one of the twenty four twin studies described above assessed PI phenotypes (Table VB). Fisher et al. (2010) studied a sample of 9-12 year-old same-sex twin pairs (234 individuals). Results from accelerometry failed to demonstrate significant differences between MZ (0.62) and DZ (0.48) twins correlations suggesting environmental influences were playing a role in the familial resemblance. In fact, the most parsimonious model for PI measures had no genetic contribution with common shared environment and non-shared environment explaining 55% and 45%, respectively, of the variance highlighting their importance on understanding PI patterns. It is possible that a power issue might be concomitant with the insufficiency in detecting small additive genetic effects.

Linkage Studies

Table VC summarizes information from PI phenotypes based on linkage studies.

The first study was performed by Simonen et al. (2003b), which found promising linkages of PI [as assessed by B3DPAR (Bouchard et al., 1983)] with

markers *D2S2347* and *D2S2305* on chromosome *2p22-p16* and suggestive linkages on chromosomes *7p11.2* and *20q13.1*. The second analysis was completed by Cai et al. (2006) using accelerometry data. The highest LOD score was 4.07 between markers *D18S1102* and *D18S474*. We should emphasize that both of these studies presented the highest LOD scores of all the phenotypes (PA and PI) presented so far, i.e., PI related phenotypes were found to be more genetically predisposed in comparison to all PA phenotypes.

Association Studies

Physical inactivity was evaluated in two association studies (Table VD). In the Winnicki et al. (2004) study *ACE* genotype was a strong predictor of PI level with *D* allele frequency significantly higher in the sedentary group. Close to 76% of *D/D* homozygotes were sedentary but only 48% of *I/I* homozygotes exhibited the same pattern. In Canada, Loos et al. (2005) reported a significant interaction between *MC4R-C-2745T* and generation. *T/T* homozygote offspring showed higher inactivity scores than heterozygotes or *C/C* genotypes. The same trend was not observed in parents.

Table V. Physical Inactivity (PI) genetic influences studies: family studies (VA), twin studies (VB), linkage studies (VC) and association studies (VD).

Table VA. Heritability estimates (h^2) of PI phenotypes from Family Studies.

Author / Country	Sample			Instrument	Phenotype	Heritability
	N	Parents	Offspring			
Simonen et al. 2002 Canada	200 nuclear families from the *Quebec Family Study* (696 subjects)	140 fathers (54.5±7.3) and 172 mothers (52.2±7.6)	164 sons (26.3±9.2) and 220 daughters (28.1±10.4).	B3DPAR[46]	PI	h^2= 0.25±0.07
Butte et al. 2006 U.S.A.	319 nuclear families from *Viva La Familia* study		1030 children (4-19 years)	Actiwatch accelerometer (Mini MItter Co, In, Bend, OR)	SB	h^2= 0.60±0.09
Cai et al. 2006 U.S.A.	319 nuclear families from *Viva La Familia* study (1661 subjects)	631 parents	1030 children (4-19 years)	Actiwatch accelerometer (Mini MItter, Bend, OR)	SB	h^2= 0.57±0.08

Table V B. Heritability estimates (h^2) and twin correlations of PI phenotypes from Twin Studies.

Author / Country	N	Twin correlations		Instrument	Phenotype	Heritability
		MZ	DZ			
Fisher et al. 2010 England 2010	57 MZ (aged 11.06±0.59) and 60 DZ (aged 11.27±0.48) twin pairs.	rMZ= 0.62	rDZ= 0.48	Actigraph 7164	SB	h^2= 0; c^2= 0.55; e^2= 0.45

[46] Using a 3-d activity diary, which included one weekend day, subjects were instructed to record the dominant activity for each 15 min period during 24 h using a list of activities. Each period was given a score ranging from 1 to 9, with 1 corresponding to sleep and 9 to the activities characterized by the highest energy expenditure levels. The phenotype "inactivity" was defined by the score on resting or very light activities (categories 1–4).

Table VC. Promising and suggestive linkages with PI phenotypes.

Author / Country	N	Genetic Marker	Locus	Instrument	Phenotype	P
Simonen et al. 2003b U.S.A.	767 subjects from 207 nuclear family units of the Quebec Family Study	*D2S2347*	*2p22-p16*	B3DPAR[47]	PI	0.0012
		D2S2305	*7p11.2*			0.0019
		IGFBP1	*20q13.1*			0.0046
		PLC1	*4q28-q31*			0.0074
Cai et al. 2006 U.S.A.	1030 children and 631 parents from the VIVA LA FAMILIA Study	*D18S1102* *D18S474*	*18q*	Actiwatch accelerometer (Mini MItter, Bend, OR)	SB	<0.001

Table VD. Associations between polymorphisms and PI phenotypes.

Author / Country	Locus	Gene	Polymorphism	N			Instrument	Phenotype	P
				D/D	D/R	R/R			
Winnicki et al. 2004 Italy	*17q23*	*ACE*	INS/DEL	65 I/I	193 I/D	97 D/D	Reaven Questionnaire	SB vs. PE	0.001
Loos et al. 2005 Canada	*18q22*	*MC4R*	*MC4R-C-2745T*	360 C/C	231 C/T	51 T/T	B3DPAR	PI	0.01

Legend: PI – physical inactivity; SB – sedentary behaviour. MZ - monozygotic twins; DZ - dizygotic twins.

[47] Using a 3-d activity diary, which included one weekend day, subjects were instructed to record the dominant activity for each 15 min period during 24 h using a list of activities. Each period was given a score ranging from 1 to 9, with 1 corresponding to sleep and 9 to the activities characterized by the highest energy expenditure levels. The phenotype “inactivity” was defined by the score on resting or very light activities (categories 1–4).

Summary and Conclusions

Over the years an increasing effort has been put forward to what is known about the genetic contribution to PA variation (Beunen & Thomis, 1999; Bray et al., 2009; Rankinen et al., 2001; Rankinen et al., 2010; Wolfarth et al., 2005). However, to the best of our knowledge, this is the first paper to aggregate results ranging from family studies, twin studies, linkage studies, association studies and GWAs on both PA and PI. This seemed like the appropriate time to do so as a new “Era” is approaching in Genetic Epidemiology related to PA and PI which is driven by technological advances and reduced costs in DNA sequencing.

The available results are quite ambiguous across studies despite the highly diversified nature of the available research. Familial aggregation studies resulted in maximal heritability estimates of PA ranging from 0% (Perusse et al., 1989) to 60% (Butte et al., 2006), and twin studies yielded MZ and DZ correlations of 0.39 (Franks et al., 2005) to 0.98 (Beunen & Thomis, 1999) and of -0.02 (Boomsma et al., 1989) to 0.72 (Koopmans et al., 1994), respectively. Both of these findings fail to clearly underline the systematic moderate-to-high importance of genes in the predisposition to be active.

The natural course of Genetic Epidemiology is moving from complex family structure analysis to an even more complex DNA analysis. Identifying genomic regions capable of carrying genes related to the trait has identified suggestive linkages with markers nearby different activity-related genes: *EDNRB*, *MC4R*, *UCP1*, *FABP2*, *CASR*, and *SLC9A9*. It should be emphasized, however, that no marker was present in more than one study. As pointed out earlier, different methods (namely on PA measuring or estimation), different density and marker spacing as well as diverse sample sizes and demographics might explain the discrepancy of the available results.

Another approach in DNA analysis is to look for associations between genotype frequencies of a candidate gene and the phenotype in association studies. Because of their physiological function, different genes have been studied - *FTO rs9939609*, *ACE*, *LEPR*, *Gln223Arg*, *MC4R* and *DRD2*.

Significant associations with PA phenotypes were found for *Gln223Arg*, *MC4R* and *DRD2* genes. *ACE* gene was also significantly associated with PA phenotypes in Winnicki et al. (2004) study but not in the Fuentes et al. (2002) research. Also, *MC4R* (Loos et al. 2005) and *DR2D* (Simonen et al., 2003a) associated positively only with higher levels of PA and PA during the past year, respectively.

In order to get a new breakthrough on PA understanding, genetic epidemiologists should strive to implement more powerful technologies and intensive computation techniques, such as GWAs. This holds the promise for, at least, replication of previous, more "modest" studies in addition to finding novel genes given the greater sample sizes, and the thousands of genetic markers involved. A first look with this new approach was published by De Moor and colleagues (2009) who failed to replicate all but the *LEPR* gene and *GABRG3* gene previous associations. However, novel SNPs in the *PAPSS2* gene on chromosome *10q23.2* and in two intergenic regions on chromosomes *2q33.1* and *18p11.32* were found. The absence of replication is not truly significant as different methodologies and PA phenotypes are being compared. Moreover, it should be emphasized that, to the best of our knowledge, only two studies (Good et al., 2008; Tsao et al., 2001), performed in mice, were able to clearly identify genes capable of altering PA. These studies identified NHLH2 (Good et al., 2008) as capable of promoting motivation or the ability to exercise and GLUT4 (Tsao et al., 2001) as capable to alter PA levels through increased glucose influx.

Almost 30 years passed since PA was formally defined (Caspersen et al., 1985) and it still remains unclear what researchers are really measuring and the best way to measure such a complex phenotype with its distinct modes, frequencies, intensities and duration. Indeed, one of the greatest challenges to a clearer understanding of genetic variation of PA is probably the measurement and quantification of PA itself. For example, the genetic contribution to PA if assessed by accelerometry is more obvious, but less applicable in real world situations to thousands of family members, than if estimated by questionnaires. This may imply that different instruments may measure and describe different

phenotypes. As such it is truly difficult to make a highly consistent statement on the magnitude of genetic predisposition to be active.

Also, as stated in the introduction, there has been a case for associating PA and PI with different genetic and environmental factors. In this review we found that, despite the smaller number of studies focusing in PI-related phenotypes, there is some ground to further pursue this suggestion. Although it seems that genetic influences are greater for PI than for PA phenotypes in some studies, care must be taken because of at least four issues: (i) the statistical power problem given the number of families, twin pairs, or cases and controls in each study sample, (ii) we still need to address the topic of whether or not PI represents distinct set of behaviours, or represents the lower end of the PA continuum; (iii) tests for the presence of pleiotropic effects and environmental covariation between PI and PA as two distinct phenotypes, or as ends in a continuum; (iv) a more precise conceptual description of the diverse facets of PI as well as an unequivocal operational definition measured with minimum amount of error is still lacking. The only study in opposition to this conclusion is the Fisher et al. (2010) twin study that reported no genetic influence on PI, although care must be taken in this conclusion given the sample size. Family studies recorded heritability estimates between 25% (Simonen et al., 2002) and 60% (Butte et al., 2006) and linkage studies (Cai et al., 2006) recorded the higher LOD scores for PI. *ACE* genotype was strongly associated with PI whereas *MC4R-C-2745T* failed to produce clear results.

It is quite intriguing that there may be distinct genetic influences on both PA (a health enhancing behaviour) and PI (a detrimental risk factor) if in fact PA and PI represent different, independent behavioral domains. Human anatomy and physiology has remained relatively unchanged over the past 10,000 years (Cordain et al., 1998), from an era of outdoor existence of hunting and foraging for foods. It's obvious that "*fighting for survival*" may have urged our ancestors to be more physically active than modern humans. For example, in the 1990s, Esparza and colleagues (2000) studied Mexican Pima Indians and their Arizonian Pima counterparts estimated to have been geographically separated

700-1000 years ago. The difference in energy expenditure was more than 500–600 more kcal/day in favour of Mexican Indians who still live a rural life.

But there are some suggestions that besides the time spent in MVPA for hunting or protection, our ancestors were quite inactive during the remainder of the day (~20 hr) (Lee et al., 2003; Tudge, 1999). Such a possibility would see us having to look at PA behaviour from another perspective: What if we were never supposed to be as active as we need to be today? Since the origin of life, all species struggle to adapt to an ever-changing environment and those who are successful are able to prevail. What if, by answering with ease to all human beings necessities, the contemporary environment is becoming a true challenge to us? In fact, it could be argued that PI is not really the reason for the current epidemiological constraints of obesity and cardiovascular diseases. There are plenty of other factors (e.g., nutrition, stress, etc.) that "compete" to decline health in our societies. As such, the need to increment PA levels might represent solely an adaptation to the environment to offset the expanded role of PI.

As we look forward to new results, building on the constant technological improvements, the time has probably come for multicentre research using all the possibilities of GWAs, Genome wide Haplotype Analysis, or combining transcriptomics in unravelling the genetics of these complex traits. The power of large cohort analyses and pooling data across cohorts has the potential to provide further insights on the genetics of physical activity and inactivity. To be successful in this huge enterprise, it is also very important that researchers focus on establishing more objective instruments and methodologies, allowing a more standardized assessment of PA and PI such as the new monitoring devices (combining information from accelerometry, heart rate, inclinometer and light sensor) and/or the multi-sensory armband that combines a two-axial accelerometry with galvanic skin response, near-body ambient temperature sensor, skin temperature, and heat flux sensor. Most PA and PI scores reported in this review relied on questionnaires whose validity and reliability have been reported (Montoye et al., 1996; Pereira et al., 1997). However it is important to acknowledge that errors of measurement may result in an overestimation of the

individual-specific environment as well as an underestimation of genetic factors. Moreover, it seems very clear that PA and PI are not only the result of combined effects of multiple genes (polygenic inheritance) but also the result of how these genes interact with the environment.

The results presented in this review suggest that a continued emphasis should be placed on developing interventions that are targeted at increasing physical activity and decreasing sedentary behaviours. These appear to be truly modifiable traits that are responsive to environmental perturbations.

Acknowledgments

Thanks are expressed to the three anonymous reviewers, as well as to the FCT – Fundação para a Ciência e a Tecnologia for granting this research (PTDC/DES/67569/2006 FCOMP-01-0124-FEDEB-09608 and SFRH / BD / 65290 / 2009).

References

Aaltonen, S., Ortega-Alonso, A., Kujala, U. M., & Kaprio, J. (2010). A Longitudinal Study on Genetic and Environmental Influences on Leisure Time Physical Activity in the Finnish Twin Cohort. Twin Res Hum Genet, 13(5), 475-481.

Berentzen, T., Kring, S. I., Holst, C., Zimmermann, E., Jess, T., Hansen, T., Pedersen, O., Toubro, S., Astrup, A., & Sorensen, T. I. (2008). Lack of association of fatness-related FTO gene variants with energy expenditure or physical activity. J Clin Endocrinol Metab, 93(7), 2904-2908.

Beunen, G., & Thomis, M. (1999). Genetic determinants of sports participation and daily physical activity. Int J Obes Relat Metab Disord, 23 Suppl 3, S55-63.

Bey, L., & Hamilton, M. T. (2003). Suppression of skeletal muscle lipoprotein lipase activity during physical inactivity: a molecular reason to maintain daily low-intensity activity. The Journal of physiology, 551(Pt 2), 673-682.

Boomsma, D. I., van den Bree, M. B., Orlebeke, J. F., & Molenaar, P. C. (1989). Resemblances of parents and twins in sports participation and heart rate. Behav Genet, 19(1), 123-141.

Bouchard, C., Rankinen, T., Chagnon, Y. C., Rice, T., Perusse, L., Gagnon, J., Borecki, I., 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. J Appl Physiol, 88(2), 551-559.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Theriault, G. (1983). A method to assess energy expenditure in children and adults. Am J Clin Nutr, 37(3), 461-467.

Bray, M. S., Hagberg, J. M., Perusse, 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. Med Sci Sports Exerc, 41(1), 35-73.

Butte, N. F., Cai, G., Cole, S. A., & Comuzzie, A. G. (2006). Viva la Familia Study: genetic and environmental contributions to childhood obesity and its comorbidities in the Hispanic population. Am J Clin Nutr, 84(3), 646-654; quiz 673-644.

Cai, G., Cole, S. A., Butte, N., Bacino, C., Diego, V., Tan, K., Goring, H. H., O'Rahilly, S., Farooqi, I. S., & Comuzzie, A. G. (2006). A quantitative trait locus on chromosome 18q for physical activity and dietary intake in Hispanic children. Obesity (Silver Spring), 14(9), 1596-1604.

Carlsson, S., Andersson, T., Lichtenstein, P., Michaelsson, K., & Ahlbom, A. (2006). Genetic effects on physical activity: results from the Swedish Twin Registry. Med Sci Sports Exerc, 38(8), 1396-1401.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep, 100(2), 126-131.

Choh, A. C., Demerath, E. W., Lee, M., Williams, K. D., Towne, B., Siervogel, R. M., Cole, S. A., & Czerwinski, S. A. (2009). Genetic analysis of self-reported physical activity and adiposity: the Southwest Ohio Family Study. Public Health Nutr, 12(8), 1052-1060.

Christin, L., O'Connell, M., Bogardus, C., Danforth, E., Jr., & Ravussin, E. (1993). Norepinephrine turnover and energy expenditure in Pima Indian and white men. Metabolism, 42(6), 723-729.

Cole, S. A., Butte, N. F., Voruganti, V. S., Cai, G., Haack, K., Kent, J. W., Jr., Blangero, J., Comuzzie, A. G., McPherson, J. D., & Gibbs, R. A. (2010). Evidence that multiple genetic variants of MC4R play a functional role in the regulation of energy expenditure and appetite in Hispanic children. American Journal of Clinical Nutrition, 91(1).

Cordain, L., Gotshall, R. W., Eaton, S. B., & Eaton, S. B., 3rd. (1998). Physical activity, energy expenditure and fitness: an evolutionary perspective. Int J Sports Med, 19(5), 328-335.

De Moor, M. H., Liu, Y. J., Boomsma, D. I., Li, J., Hamilton, J. J., Hottenga, J. J., Levy, S., Liu, X. G., Pei, Y. F., Posthuma, D., Recker, R. R., Sullivan, P. F., Wang, L., Willemsen, G., Yan, H., EJ, D. E. G., & Deng, H. W. (2009). Genome-Wide Association Study of Exercise Behaviour in Dutch and American Adults. Med Sci Sports Exerc, 41(10), 1887-1895.

De Moor, M. H., Posthuma, D., Hottenga, J. J., Willemsen, G., Boomsma, D. I., & De Geus, E. J. (2007a). Genome-wide linkage scan for exercise participation in Dutch sibling pairs. Eur J Hum Genet, 15(12), 1252-1259.

De Moor, M. H., Spector, T. D., Cherkas, L. F., Falchi, M., Hottenga, J. J., Boomsma, D. I., & De Geus, E. J. (2007b). Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. Twin Res Hum Genet, 10(6), 812-820.

De Moor, M. H., Stubbe, J. H., Boomsma, D. I., & De Geus, E. J. (2007c). Exercise participation and self-rated health: do common genes explain the association? Eur J Epidemiol, 22(1), 27-32.

Dishman, R. K., Sallis, J. F., & Orenstein, D. R. (1985). The determinants of physical activity and exercise. Public Health Rep, 100(2), 158-171.

Duncan, G. E., Goldberg, J., Noonan, C., Moudon, A. V., Hurvitz, P., & Buchwald, D. (2008). Unique Environmental Effects on Physical Activity Participation: A Twin Study. PLoS One, 3(4).

Eriksson, M., Rasmussen, F., & Tynelius, P. (2006). Genetic factors in physical activity and the equal environment assumption-- the Swedish young male twins study. Behav Genet, 36(2), 238-247.

Esparza, J., Fox, C., Harper, I. T., Bennett, P. H., Schulz, L. O., Valencia, M. E., & Ravussin, E. (2000). Daily energy expenditure in Mexican and USA Pima

indians: low physical activity as a possible cause of obesity. Int J Obes Relat Metab Disord, 24(1), 55-59.

Fisher, A., van Jaarsveld, C. H., Llewellyn, C. H., & Wardle, J. (2010). Environmental influences on children's physical activity: quantitative estimates using a twin design. PLoS One, 5(4), e10110.

Franks, P. W., Ravussin, E., Hanson, R. L., Harper, I. T., Allison, D. B., Knowler, W. C., Tataranni, P. A., & Salbe, A. D. (2005). Habitual physical activity in children: the role of genes and the environment. Am J Clin Nutr, 82(4), 901-908.

Fuentes, R. M., Perola, M., Nissinen, A., & Tuomilehto, J. (2002). ACE gene and physical activity, blood pressure, and hypertension: a population study in Finland. J Appl Physiol, 92(6), 2508-2512.

Gingrich, J. A., & Caron, M. G. (1993). Recent advances in the molecular biology of dopamine receptors. Annu Rev Neurosci, 16, 299-321.

Good, D. J., Coyle, C. A., & Fox, D. L. (2008). Nhlh2: a basic helix-loop-helix transcription factor controlling physical activity. Exerc Sport Sci Rev, 36(4), 187-192.

Hakanen, M., Raitakari, O. T., Lehtimaki, T., Peltonen, N., Pahkala, K., Sillanmaki, L., Lagstrom, H., Viikari, J., Simell, O., & Ronnemaa, T. (2009). FTO genotype is associated with body mass index after the age of seven years but not with energy intake or leisure-time physical activity. The Journal of clinical endocrinology and metabolism, 94(4), 1281-1287.

Hamilton, M. T., Hamilton, D. G., & Zderic, T. W. (2007). Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. Diabetes, 56(11), 2655-2667.

Joosen, A. M., Gielen, M., Vlietinck, R., & Westerterp, K. R. (2005). Genetic analysis of physical activity in twins. Am J Clin Nutr, 82(6), 1253-1259.

Katzmarzyk, P. T. (2010). Physical activity, sedentary behaviour, and health: paradigm paralysis or paradigm shift? Diabetes, 59(11), 2717-2725.

Koopmans, J. R., Lorenz, J. P. V. D., & Boomsma, D. I. (1994). Smoking and sports participation. In U. Goldbourt, U. Faire & K. Berg (Eds.), Factors in coronary heart disease (pp. 217-235). Dordrecht: Kluwer Academic.

Larsen, L. H., Echwald, S. M., Sorensen, T. I., Andersen, T., Wulff, B. S., & Pedersen, O. (2005). Prevalence of mutations and functional analyses of melanocortin 4 receptor variants identified among 750 men with juvenile-onset obesity. J Clin Endocrinol Metab, 90(1), 219-224.

Lee, K. E., Klein, B. E., & Klein, R. (2003). Familial aggregation of components of the multiple metabolic syndrome in the Framingham Heart and Offspring Cohorts: Genetic Analysis Workshop Problem 1. BMC Genet, 4 Suppl 1, S94.

Liu, G. F., Zhu, H. D., Lagou, V., Gutin, B., Stallmann-Jorgensen, I. S., Treiber, F. A., Dong, Y. B., & Snieder, H. (2010). FTO variant rs9939609 is associated with body mass index and waist circumference, but not with energy intake or physical activity in European- and African-American youth. Bmc Medical Genetics, 11.

Loos, R. J., Rankinen, T., Tremblay, A., Perusse, L., Chagnon, Y., & Bouchard, C. (2005). Melanocortin-4 receptor gene and physical activity in the Quebec Family Study. Int J Obes (Lond), 29(4), 420-428.

Lorentzon, M., Lorentzon, R., Lerner, U. H., & Nordstrom, P. (2001). Calcium sensing receptor gene polymorphism, circulating calcium concentrations and bone mineral density in healthy adolescent girls. Eur J Endocrinol, 144(3), 257-261.

Maia, J. A., Thomis, M., & Beunen, G. (2002). Genetic factors in physical activity levels: a twin study. Am J Prev Med, 23(2 Suppl), 87-91.

McCaffery, J. M., Papandonatos, G. D., Bond, D. S., Lyons, M. J., & Wing, R. R. (2009). Gene X environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. Am J Clin Nutr, 89(4), 1011-1018.

Mitchell, B. D., Rainwater, D. L., Hsueh, W. C., Kennedy, A. J., Stern, M. P., & Maccluer, J. W. (2003). Familial aggregation of nutrient intake and physical activity: results from the San Antonio Family Heart Study. Ann Epidemiol, 13(2), 128-135.

Montoye, H. J., Kemper, H. C. H., Washburn, R. A., & Saris, W. H. (1996). Measuring Physical Activity and Energy Expenditure Champaign: Human Kinetics Publishers.

Nagasaka, J., Tsuji, M., Takeda, H., & Matsumiya, T. (1999). Role of endothelin receptor subtypes in the behavioral effects of the intracerebroventricular administration of endothelin-1 in conscious rats. Pharmacol Biochem Behav, 64(1), 171-176.

Pereira, M. A., FitzerGerald, S. J., Gregg, E. W., Joswiak, M. L., Ryan, W. J., Suminski, R. R., Utter, A. C., & Zmuda, J. M. (1997). A collection of Physical Activity Questionnaires for health-related research. Med Sci Sports Exerc, 29(6 Suppl), S1-205.

Perusse, L., Tremblay, A., Leblanc, C., & Bouchard, C. (1989). Genetic and environmental influences on level of habitual physical activity and exercise participation. Am J Epidemiol, 129(5), 1012-1022.

Perusse, L., Tremblay, A., Leblanc, C., Cloninger, C. R., Reich, T., Rice, J., & Bouchard, C. (1988). Familial resemblance in energy intake: contribution of genetic and environmental factors. Am J Clin Nutr, 47(4), 629-635.

Rankinen, T., Perusse, 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(6), 855-867.

Rankinen, T., Roth, S. M., Bray, M. S., Loos, R., Perusse, L., Wolfarth, B., Hagberg, J. M., & Bouchard, C. (2010). Advances in exercise, fitness, and performance genomics. Med Sci Sports Exerc, 42(5), 835-846.

Richert, L., Chevalley, T., Manen, D., Bonjour, J. P., Rizzoli, R., & Ferrari, S. (2007). Bone mass in prepubertal boys is associated with a Gln223Arg amino acid substitution in the leptin receptor. J Clin Endocrinol Metab, 92(11), 4380-4386.

Scillitani, A., Guarnieri, V., De Geronimo, S., Muscarella, L. A., Battista, C., D'Agruma, L., Bertoldo, F., Florio, C., Minisola, S., Hendy, G. N., & Cole, D. E. (2004). Blood ionized calcium is associated with clustered polymorphisms in the carboxyl-terminal tail of the calcium-sensing receptor. J Clin Endocrinol Metab, 89(11), 5634-5638.

Seabra, A. F., Mendonca, D. M., Goring, H. H., Thomis, M. A., & Maia, J. A. (2008). Genetic and environmental factors in familial clustering in physical activity. Eur J Epidemiol, 23(3), 205-211.

Sherwood, N. E., & Jeffery, R. W. (2000). The behavioral determinants of exercise: implications for physical activity interventions. Annu Rev Nutr, 20, 21-44.

Simonen, R. L., Perusse, L., Rankinen, T., Rice, T., Rao, D. C., & Bouchard, C. (2002). Familial aggregation of physical activity levels in the Quebec Family Study. Med Sci Sports Exerc, 34(7), 1137-1142.

Simonen, R. L., Rankinen, T., Perusse, L., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., & Bouchard, C. (2003a). A dopamine D2 receptor gene polymorphism and physical activity in two family studies. Physiol Behav, 78(4-5), 751-757.

Simonen, R. L., Rankinen, T., Perusse, L., Rice, T., Rao, D. C., Chagnon, Y., & Bouchard, C. (2003b). Genome-wide linkage scan for physical activity levels in the Quebec Family study. Med Sci Sports Exerc, 35(8), 1355-1359.

Snitker, S., Tataranni, P. A., & Ravussin, E. (2001). Spontaneous physical activity in a respiratory chamber is correlated to habitual physical activity. Int J Obes Relat Metab Disord, 25(10), 1481-1486.

Spinath, F. M., Wolf, H., Angleitner, A., Borkenau, P., & Riemann, R. (2002). Genetic and environmental influences on objectively assessed activity in adults. Personality and Individual Differences, 33(4), 633-645.

Stefan, N., Vozarova, B., Del Parigi, A., Ossowski, V., Thompson, D. B., Hanson, R. L., Ravussin, E., & Tataranni, P. A. (2002). The Gln223Arg polymorphism of the leptin receptor in Pima Indians: influence on energy expenditure, physical activity and lipid metabolism. Int J Obes Relat Metab Disord, 26(12), 1629-1632.

Stubbe, J. H., Boomsma, D. I., & De Geus, E. J. (2005). Sports participation during adolescence: a shift from environmental to genetic factors. Med Sci Sports Exerc, 37(4), 563-570.

Stubbe, J. H., Boomsma, D. I., Vink, J. M., Cornes, B. K., Martin, N. G., Skytthe, A., Kyvik, K. O., Rose, R. J., Kujala, U. M., Kaprio, J., Harris, J. R., Pedersen, N. L., Hunkin, J., Spector, T. D., & de Geus, E. J. (2006). Genetic influences on exercise participation in 37,051 twin pairs from seven countries. PLoS One, 1, e22.

Tang-Christensen, M., Havel, P. J., Jacobs, R. R., Larsen, P. J., & Cameron, J. L. (1999). Central administration of leptin inhibits food intake and activates the sympathetic nervous system in rhesus macaques. J Clin Endocrinol Metab, 84(2), 711-717.

Tsao, T. S., Li, J., Chang, K. S., Stenbit, A. E., Galuska, D., Anderson, J. E., Zierath, J. R., McCarter, R. J., & Charron, M. J. (2001). Metabolic adaptations in skeletal muscle overexpressing GLUT4: effects on muscle and physical activity. FASEB J, 15(6), 958-969.

Tudge, C. (1999). Neanderthals, Bandits and Farmers: How Agriculture Really Began Yale University Press.

Wang, J. G., & Staessen, J. A. (2000). Genetic polymorphisms in the renin-angiotensin system: relevance for susceptibility to cardiovascular disease. Eur J Pharmacol, 410(2-3), 289-302.

Winnicki, M., Accurso, V., Hoffmann, M., Pawlowski, R., Dorigatti, F., Santonastaso, M., Longo, D., Krupa-Wojciechowska, B., Jeunemaitre, X., Pessina, A. C., Somers, V. K., & Palatini, P. (2004). Physical activity and angiotensin-converting enzyme gene polymorphism in mild hypertensives. Am J Med Genet A, 125A(1), 38-44.

Wolfarth, B., Bray, M. S., Hagberg, J. M., Perusse, 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. Med Sci Sports Exerc, 37(6), 881-903.

Zderic, T. W., & Hamilton, M. T. (2006). Physical inactivity amplifies the sensitivity of skeletal muscle to the lipid-induced downregulation of lipoprotein lipase activity. J Appl Physiol, 100(1), 249-257.

ARTICLE 2

FAMILIAL AGGREGATION OF METABOLIC SYNDROME INDICATORS IN PORTUGUESE FAMILIES

Daniel Monteiro de Vilhena e Santos[1]

Peter T. Katzmarzyk[2]

David-Alexandre Trégouet[3]

Thayse Natascha Gomes[1]

Fernanda Karina Santos[1]

José António Ribeiro Maia[1]

Published in BioMed Research International

Volume 2013 (2013), Article ID 314823, 7 pages

http://dx.doi.org/10.1155/2013/314823

[1] CIFI[2]D, Faculty of Sports, University of Porto, Portugal

[2] Pennington Biomedical Research Center, Louisiana State University System, USA

[3]INSERM, UMR_S 937, ICAN Institute for Cardiometabolism and Nutrition, Pierre and Marie Curie University (UPMC, Paris 6), Paris, France

ABSTRACT:

Background and Aims: Family studies are well suited to investigate the genetic architecture underlying the metabolic syndrome (MetS). The purposes of this paper were (i) to estimate heritabilities for each of the MetS indicators, and (ii) to test the significance of familial intra-trait and cross-trait correlations in MetS markers.

Methods and Results: This study included 1,363 individuals from 515 Portuguese families in which five MetS components, including waist circumference (WC), blood pressure (BP), HDL-cholesterol, triglycerides (TG), and glucose (GLU) were measured. Intra-trait and cross-trait familial correlations of these five components were estimated using Generalized Estimating Equations. Each MetS component was significantly heritable (h^2 ranged from 0.12 to 0.60) and exhibited strong familial resemblance with correlations between biological relatives of similar magnitude to those observed between spouses. With respect to cross-trait correlations, familial resemblance was very weak except for the HDL-TG pair.

Conclusions: The present findings confirm the idea of familial aggregation in MetS traits. Spousal correlations were, in general, of the same magnitude as the biological relatives' correlations suggesting that most of the phenotypic variance in MetS traits could be explained by shared environment.

Keywords: Familial, heritable, metabolic syndrome, trait, clustering.

Introduction

Metabolic syndrome (MetS) is a cluster of interrelated cardiovascular disease risk factors characterized by glucose intolerance, hypertension, dyslipidaemia and obesity and is associated with increased risk of cardiovascular disease mortality in adults (Isomaa et al., 2001; Lakka et al., 2002; Reaven, 1988). The mechanisms underlying the MetS physiopathogenesis are complex, both genetic and environmentally driven, and factors like physical inactivity and poor dietary habits are likely to condition the emergence and further development of cardiometabolic disorders (Alberti & Zimmet, 1998).

MetS prevalence has been increasing in the last decades, namely in the western, more modernized countries (Ford et al., 2004). However, available data is far from being concordant as different definitions and cut-points are being used to determine risk (Churilla et al., 2007). Using NCEP-ATPIII criteria, a large Hospital Portuguese sample of adults (age range: 18-96 yrs) showed a MetS prevalence of 27.5% (Fiuza et al., 2008), which compares to 34% in the USA (Ervin, 2009).

Studying families may be of great assistance in unravelling the importance of shared genes and environmental conditions since common behaviours within family lifestyles may trigger the appearance of MetS (Vattikuti et al., 2012). The first step to assess the genetic contribution to a trait is to calculate its familiality and to derive heritability estimates (h^2) (Statistical Human Genetics. Methods and Protocols, 2012). For example, in a recent report from Vattikuti et al. (2012), h^2 ranged from 0.30 to 0.48 in MetS traits. Also, in a study by Tang et al. (2006) h^2 were of the same magnitude, varying from 0.33 (DBP) to 0.63 (HDL). Moreover, in the NHLBI sample (Tang et al., 2006), significant familial intra-trait correlations were observed between parents and offspring for diastolic blood pressure (DBP), HDL-cholesterol (HDL) and fasting insulin (INS), supporting the hypotheses of shared effects (genetic and/or environmental) on these MetS indicators.

Bivariate familial correlation analysis can be of great help to explore the possibility of shared genetic and/or environmental effects among MetS traits. In this case, genetic correlation coefficients provide information on the amount of additive genetic variation shared between two MetS traits. This could be of great importance in understanding if two traits are more closely related than others. Potentially, this could influence the development of intervention programs. For instance, Tregouët et al. (1999) found that body mass index (BMI)/insulin (INS) correlations between biological relatives tend to be higher than between spouses, supporting the hypotheses of common transmissible factors for these two traits. Also, using a twin design with 405 MZ pairs and 290 DZ pairs, Pang et al. (2010) found moderate to high correlations between total cholesterol (TC) and LDL-cholesterol (LDL) (r=0.84), LDL and HDL (r=0.62), TC and HDL (r=0.58), triacylglycerol (TRG) and TC (r=0.45), and TG and LDL (r=0.41). In a recent review, Povel et al. (2011) verified that genetic correlation coefficients were strongest for waist circumference and HOMA-IR (median r=0.59), and for HDL and TG (median r=0.46). On the opposite side, the genetic correlation coefficients of BP and HDL (median r=0.09), BP and TG (median r=0.10), BP and fasting glucose (median r=0.13), and fasting glucose and HDL (median r=0.07) were the lowest. Vattikuti et al. (2012) found significant genetic correlations between GLU-INS (ρ_G=0.69), INS-TG (ρ_G=0.76) and TG-HDL (ρ_G=-0.59) in the Aric population.

The present report pursues two specific aims: (i) to estimate heritabilities (h^2) for each of the MetS indicators, and (ii) to test the significance of familial intra-trait and cross-trait correlations in MetS markers.

Materials and Methods

STUDY POPULATION

The Healthy Family study, from the Portuguese Famílias Saudáveis (FAMS), investigates the relationship among MetS traits, physical activity, physical fitness and body composition in families. Children and adolescents aged ≤ 18 years were recruited in schools from Azores and Madeira archipelagos, and north and central regions of mainland Portugal, and were

approached to freely participate in the study with their siblings and parents. School officers provided family lists, and families with at least two siblings were initially invited. However, given that families with 3 or more children are scarce in the Portuguese population (Rosa & Chitas, 2010), and to improve statistical power, we also invited one-offspring families through random eligibility. Children with chronic diseases, physical handicaps or psychological disorders were excluded as these conditions might impair their daily routines, namely their physical activities within schools and/or sports clubs. As such, our sample comprises 515 families with one or two offspring (see Table 1). Some relatives were not able to participate in the study and the resulting sample was: 252 fathers, 464 mothers, 317 sons and 330 daughters. The ethics committee of the Faculty of Sport, University of Porto, approved the study, and written informed consent was obtained from all subjects.

DATA COLLECTION

The standardized procedures of Lohman et al. (1988) were used to measure height with a Siber Hegner anthropometer (GMP instruments), and weight with a Tanita scale® (model BC-418 MA); Waist circumference was measured at the end of a normal expiration just above the iliac crest, using a non-elastic Holtain tape.

Blood samples were collected after an overnight fast of at least 10 to 12 h. Glucose, TC, HDL, and TG were analysed with an LDX point of care analyser (LDX, 2003a). This method has been previously validated against a laboratory reference method (LDX, 2003b), and daily optical equipment checks were made according to manufacturer instructions.

Resting systolic (SBP) and diastolic (DBP) blood pressures were measured with an Omron Model M6 (HEM-7001-E) device according to the International Protocol of the European Society of Hypertension (Topouchian et al., 2006). Cuff sizes were modified depending on the size of the subject's arm. Subjects were seated in an upright position with the right arm resting on a table at the heart level. The first reading was performed after a 5 minute resting period. The other two readings were performed with three-minute breaks in

between. The mean of the three blood pressure measurements was used for analysis.

STATISTICAL ANALYSIS

Descriptive continuous data are presented as means ± SDs.

Heritability estimates were computed using a maximum likelihood approach implemented in SOLAR v.4.3.1 software (Blangero & Almasy, 1997). In order to handle the non-independence of the family observations, the intra-trait and cross-trait familial correlation analyses were conducted within the framework of Generalized Estimating Equations (GEEs), an elegant and efficient alternative to Maximum Likelihood methods that does not require any distributional assumptions. GEEs are highly flexible for studying covariate effects on means and correlations, and are asymptotically robust to a misspecification of the exact pattern of correlations between observations. GEEs are in particular well suited to family data analysis where observations between relatives may be correlated due to shared environment and genetic factors (Tregouët & Tiret, 2000). For the current application, we used the GESEE software developed by Tregouët et al. (1999) as previously applied by Plancoulaine et al. (Plancoulaine et al., 2004; Plancoulaine et al., 2000; Plancoulaine et al., 2008).

According to the given family structure, up to eight different types of family intra-trait correlations could be estimated. For this particular study we decided to focus on: ρFM=father-mother, ρFO=father-offspring, ρMO=mother-offspring and ρSS=sibling-sibling. The generalized Wald test statistic was then employed to test specific hypotheses on these correlations: For example, we were particularly interested in testing whether: (1) all four correlations were equal to zero, meaning that there is no familial aggregation of MetS traits; (2) all correlations were equal to each other, which would suggest a strong shared environmental component; (3) ρFO=ρMO=ρSS suggesting that no sex differences exist in terms of parental transmission of MetS traits.

For the cross-trait analyses, two different types of correlations were studied: within-individual cross-trait correlations per class of relatives (e.g.

correlation between glucose and SBP in fathers) and inter-individual cross-trait correlations (e.g. correlation between glucose levels in a father and SBP in his offspring). For the latter, we assumed a symmetric correlation pattern for each pair of traits, with four inter-individuals cross-trait correlations according to classes of relatives: ρFM=spouses, ρFO=father-offspring, ρMO=mother-offspring, ρSS=sibling-sibling. This implies for example that the correlation between glucose in a father and SBP in his offspring is the same of the correlation between the SBP in the father and the glucose in his offspring.

Strong cross-trait correlations between biological relatives but not between spouses would suggest the existence of shared genetic factors influencing both traits. Additionally, observing cross-trait correlations between spouses would suggest shared environmental factors.

In all analyses, phenotypic means were adjusted for age, age^2, sex, age x sex, and age^2 x sex. The level of statistical significance was set at 0.05.

Results

Table 1 presents the descriptive data for the sample. A total of 1,363 subjects from 515 families were included. The average family size was 2.6 subjects. Correlations were derived from the following family combinations: 240 spousal, 216 father-offspring, 283 mother-offspring, and 208 sibling-sibling pairs.

All traits were found to be highly heritable (Table 2) with estimates ranging from 0.12 (GLU) to 0.60 (WC). This means that there are strong additive genetic effects on the expression of these MetS trait that justifies further specific analysis of their genetic architecture.

Table 3 shows the familial correlations and their 95% confidence intervals for each MetS trait. The absence of familial resemblance was rejected for each phenotype (all $p<10^{-4}$) and different general family correlations were observed.

Table I. Sample descriptive characteristics (means ± standard deviations).

	Fathers (n=252)	Mothers (n=464)	Sons (n=317)	Daughters (n=330)
Age (yrs)	46.8 ±8.9	43.8±8.6	14.0 ± 6.2	14.9 ±8.3
Height (cm)	169.9 ± 6.6	158.1 ±6.3	157.2 ±16.3	152.8 ± 12.5
Weight (kg)	81.8 ±13.2	68.6 ±12.4	55.9 ±17.8	53.1 ±15.1
BMI (kg/m^2)	28.3 ±4.1	27.5±4.7	22.1 ±4.3	22.4±4.6
WC (cm)	95.8 ±10.8	86.2±11.6	74.7 ±12.1	73.3 ± 13.5
SBP (mmHg)	133.7±17.3	124.9 ±16.5	116.0 ± 13.9	113.9 ±12.7
DBP (mmHg)	81.64 ± 10.62	77.37 ± 10.93	65.38 ± 9.51	67.62 ± 9.00
HDL (mg/dl)	46.4 ±15.1	55.8 ± 14.6	49.5 ± 14.9	52.8 ±13.6
TG (mmol/l)	1.74 ±1.11	1.28 ±0.68	0.78 ± 0.41	0.92±0.55
Glucose (mg/dl)	97.5 ±19.1	86.2 ±11.7	83.8 ± 8.7	82.7 ±8.6

Legend: BMI – body mass index; WC – waist circumference; SBP – systolic blood pressure; DBP – diastolic blood pressure; HDL – high density lipoprotein; TG – triglycerides.

Table II. Heritability (h^2) estimates of the different phenotypes in the Healthy Families Study

Trait	h^2	Std. Error	p-value	$CI_{95\%}$
WC	0.60	0.05	<0.001	0.52-0.67
SBP	0.50	0.05	<0.001	0.42-0.59
GLU	0.12	0.05	<0.001	0.03-0.20
HDL	0.44	0.05	<0.001	0.36-0.52
TG	0.29	0.06	<0.001	0.20-0.38

Legend: WC – waist circumference; SBP – systolic blood pressure; GLU – glucose; HDL – high density lipoprotein; TG – triglycerides.

For GLU, SBP and TG, the test for equality of all correlations was not rejected, p=0.24, p=0.21 and p=0.23, respectively. The resulting common familial correlations were GLU, ρ = 0.30 (95% confidence interval: [0.19 - 0.40]), SBP, ρ = 0.26 [0.19 - 0.32]), and TG, ρ = 0.16 [0.08 - 0.23]. For HDL (p<10-4) and WC (p<10-4), the test for equality of all correlations was significant. Furthermore, for HDL the biological relatives correlations were significantly different (p=0.01) from each other, with the lowest correlation of 0.28 for father-offspring and the highest of 0.49 for sib-sib. For WC no significant differences were observed between genetic related relatives (p=0.327), but mother-offspring correlation (ρ = 0.45 [0.34 – 0.55]) was higher than both father-offspring (ρ = 0.30 [0.25 – 0.50]) and sib-sib (ρ = 0.30 [0.18 – 0.42])

correlations. HDL and TG did not exhibit significant spouse correlation suggesting a low environment component for these phenotypes. However, TG correlations were also non-significant for father-offspring and sib-sib correlations.

Table III. Intra-trait familial correlations of MetS traits, and corresponding 95% confidence intervals.

Trait	ρ_{FM}	ρ_{FO}	ρ_{MO}	ρ_{SIBS}	All correlations are equal to zero	All correlations are equal	Equality between biological relatives
GLU	0.39 (0.18-0.57)	0.29 (0.15-0.43)	0.29 (0.17-0.41)	0.27 (0.13-0.39)	p<0.0001	p=0.236	
SBP	0.23 (0.09-0.37)	0.20 (0.09-0.31)	0.27 (0.18-0.36)	0.31 (0.17-0.44)	p<0.0001	p=0.211	
HDL	0.10 (-0.03-0.24)	0.28 (0.14-0.40)	0.39 (0.25-0.52)	0.49 (0.27-0.66)	p<0.0001	p<0.001	p=0.015
TG	0.15 (-0.01-0.31)	0.13 (-0.01-0.26)	0.22 (0.08-0.34)	0.12 (0.00-0.23)	p<0.0001	p=0.234	
WC	0.38 (0.25-0.50)	0.30 (0.19-0.41)	0.45 (0.34-0.55)	0.30 (0.18-0.42)	p<0.001	p=0.021	p=0.327

Legend: HDL – High Density Lipoprotein; GLU - Glucose; SBP – Systolic Blood Pressure; TG – Triglycerides; WC – Waist circumference.

Cross-trait within-individuals correlations are shown in Table 4. The pairs of traits that demonstrated significant family correlations were GLU-WC (p<0.001), HDL-SBP (p=0.039) and HDL-TG (p<0.01). Moreover it can be observed that glucose was poorly correlated with other MetS traits in a given individual, except with TG and WC in fathers only. Second, HDL was negatively correlated with TG in all class of relatives, and with SBP in daughters. Furthermore, SBP was positively correlated with sons' TG and parents' WC. Lastly, TG positively correlated with WC in male family members.

As for the cross-trait familial correlations computed for the four familial dyads (Table 5) results showed that within GLU-SBP (p<0.001), GLU-WC (p=0.027), GLU-TG (p=0.049), HDL-WC (p=0.010), SBP-WC (p=0.017) and SBP-TG (p=0.046), the specific familial correlations are significantly different from each other. From these, sibling correlations were only significant for GLU-WC (ρ=-0.10) and HDL-WC (ρ=0.14). On the other hand, spouses' correlations were significant for SBP-GLU (ρ=-0.12) and SBP-WC (ρ=0.13). Lastly, father-

offspring correlations were significant for GLU-TG (ρ=0.08) and mother-offspring correlations were significant for HDL-WC (ρ=0.08).

Table IV. Cross-trait correlations of MetS components within individuals according to each class of relatives

Trait	HDL	SBP	TG	WC
GLU	ρ_F=-0.01	0.08	0.13*	0.20*
	ρ_M=-0.02	-0.01	0.03	0.08
	ρ_S=-0.01	-0.07	0.02	-0.12
	ρ_D=-0.09	0.02	0.10	-0.08
HDL		-0.06	-0.17*	-0.04
		0.03	-0.11*	-0.06
		-0.15*	-0.12*	0.05
		-0.05	-0.13*	0.04
SBP			0.07	0.12*
			0.04	0.05
			-0.00	-0.00
			0.10*	-0.07
TG				0.11*
				0.04
				0.13*
				0.08

Correlation that are significantly different from 0 at $p < 0.05$ are labelled with an *.
Legend: HDL – High Density Lipoprotein; GLU - Glucose; SBP – Systolic Blood Pressure; TG – Triglycerides; WC – Waist circumference;
Each pair of traits is characterized by 4 types of within-individual correlations: in fathers ($\rho_{F.}$); in mothers (ρ_M); in sons (ρ_S); and in daughters (ρ_D).

Table V. Cross-trait inter-individuals correlations of MetS components according to each class of relatives

Trait	HDL	SBP	TG	WC
GLU	ρ_{FM}=0.07	-0.12*	0.04	0.08
	ρ_{FO}=0.06	0.01	0.08*	0.04
	ρ_{MO}=0.04	-0.12*	0.06	-0.02
	ρ_{SIBS}=0.04	-0.08	0.05	-0.10*
HDL		-0.08	-0.05	0.09
		0.08	-0.07	0.07
		-0.08	-0.08*	0.08*
		-0.09	-0.08	0.14*
SBP			0.07	0.13*
			0.01	0.05
			0.01	-0.00
			0.01	-0.07
TG				0.03
				0.08*
				0.07*
				0.07

*significantly different from 0 at $p < 0.05$.
Legend: HDL – High Density Lipoprotein; GLU - Glucose; SBP – Systolic Blood Pressure; TG – Triglycerides; WC – Waist circumference;
Each pair of traits is characterized by 4 types of inter-individual correlations: between spouses ($\rho_{FM.}$); father-offspring (ρ_{FO}); mother-offspring (ρ_{MO}) and sib-sib (ρ_{SS}).

Discussion

In this study detailed information about MetS in a sample of 515 Portuguese families is presented. We analysed the complex network of interrelationships between each trait within a family.

The present findings confirm that MetS traits are highly heritable in agreement with previous results (Tang et al., 2006; Vattikuti et al., 2012). For instance, in the present study, the heritability estimate for HDL was 0.44, which is close to previous estimates given by Vattikuti et al. (2012) and by the NHLBI (Tang et al., 2006) study, 0.48 and 0.63, respectively. Furthermore, physical exercise, as well as statins, has been shown to induce positive changes in HDL concentrations, via direct or mediated influences (Barter et al., 2010; Kelley et al., 2004). WC (h^2=0.60) and SBP (h^2=0.50) genetic factors were slightly higher in our study than in the NHLBI sample, which amounted to 0.42 and 0.33, respectively. Again, physical exercise and nutrition are well-known for their strong and positive impact on WC and SBP traits suggesting that shared environment will probably be responsible for a great amount of the phenotypic

variance in our sample. As for GLU and TG, the present results for heritability are lower when compared to other findings (Tang et al., 2006; Vattikuti et al., 2012) with h^2 ranging from 0.12 to 0.29. In any case, we have to emphasize that heritability estimates are also influenced by familial environment correlates, since no control was made for the possible effects of shared household characteristics during childhood and adulthood.

Familial intra-trait correlations of MetS components were estimated and familial similarity was found for each trait (all p<10-4) even though different magnitudes were established. Correlation equality was found in SBP, GLU and TG and between each biological family dyad in WC. Using the same statistical approach, Tregouët et al. (1999) found that spouse correlations were always different from zero, meaning that there is a possibility of a shared environmental factor influencing MetS traits. Furthermore, the Tregouët et al. (1999) results also highlighted that the correlations between biological relatives were always greater than between spouses for the five MetS traits. This was not the case in our sample where, besides HDL, spouse correlations were of the same magnitude as for the biological relatives' correlations. Spousal resemblance may be attributed to distinct processes, namely marital interaction (increasing mutual influence process across time of marriage), social homogamy (incidental resemblance due to some cultural and/or social background) and phenotypic assortment (selection process due to some characteristic) (Reynolds et al., 1996). It has been suggested that assortative mating by BMI was associated with increasing prevalence of obesity in offspring in several developed populations (Hebebrand et al., 2000; Katzmarzyk et al., 2002). If this is true, one could speculate that phenotypic correlations in MetS traits between spouses might be a function of assortative mating, although no studies have demonstrated, so far, this assumption. Also, it has been hypothesized that phenotypic variation is associated with the duration of the marriage as an indicator of social interaction (De Moor et al., 2011). Unfortunately, no adjustment was made for marriage years, but it is possible that the spousal correlations may reflect marital interaction processes since both spouses share many cultural assets, health beliefs, social background and experiences as

couples. Moreover, controlling for other lifestyle/behavioral variables, like impaired fetal growth that has been associated with insulin resistance and other MetS traits (Barker et al., 1993), would probably lead to a better understanding of the correlations. In this case, adjusting for life history variables during pregnancy would probably add information on siblings' correlations because of their exposure to possible different environments inutero. Other lifestyle variables have been used to limit to a minimum the influence of familial environment, but even when such adjustments were made, residual effects were still identified as capable of influencing the results (Tang et al., 2006). This ensemble of results supports a very strong argument for the existence of familial influence on MetS traits, due to additive and/or interactive genetic and shared environmental factors.

Lastly, cross-trait correlations for all MetS traits were computed. This is very relevant in helping disentangle the genetic and environmental mechanisms that may rule the association between traits. We took two different approaches: (i) analyse each family relative and (ii) analyse four different family dyads. Correlation in each relative was always significant and negative for HDL/TG. This result is in agreement with previous findings (Povel et al., 2011) in which low HDL is a result of inefficient catabolism of TG rich lipoproteins and reduced transfer of surface components to nascent HDL particles (Miller et al., 2007). Moreover, depletion of cholesterol HDL particles is also favoured in the presence of TG-enriched HDL particles that are more rapidly catabolized by lipases (Lamarche et al., 1999). Recently, a set of genetic variants in the *LPL* gene and in the *APOA1/APOC3/APOA4/APOA5* gene cluster was found to mediate the significant genetic correlation between TG and HDL (Kathiresan et al., 2009). On the other hand, our results fail to replicate some previous reports (Benyamin et al., 2007; Butte et al., 2005; Hong et al., 1997; Tang et al., 2006), but it needs to be recognized that the vast majority of the studies not only used different statistical methods as well as did not present information on the four dyads. For instance, we could not reproduce the results of Butte et al. (2005) in which environmental correlations of MetS traits were all significant, with the exceptions of SBP/GLU and TG/GLU, and genetic correlations were

significantly different from 1 (i.e., partial pleiotropy), meaning that the environment also plays a role on determining the relationship between the traits. However, it is undisputable that pleiotropy among some MetS components is present and it has been proven by familial studies (Hong et al., 1997; Mitchell et al., 1996) as well as by genome scans (Arya et al., 2004; de Andrade et al., 2002). Such incongruities may also be linked to population characteristics, and/or covariates adjustment, but the hard fact is that actual knowledge is not conclusive about the biological mechanisms that regulate the association between some MetS traits (Tang et al., 2006). And, in such cases in which there is evidence for shared physiology pathways, like in GLU and TG (Frayn, 2010), genetic analysis as the present report fail to consubstantiate that datum.

The cross-trait correlations between dyads yielded a different insight into the relationship between relatives and traits. In the present sample, WC correlated significantly with all of the other traits, at least in one dyad, meaning that body fat of a relative could be associated with other traits in another relative. On the other hand, in the Trégouet et al. (1999) study, the marker for obesity was BMI and it was not significantly correlated with TG. As TG is clearly related with body fat, it seems that environmental factors may play a key role in this association within a family, as nutritional habits and a sedentary lifestyle have been found to aggregate in families. Interestingly we found a very similar number of significant correlations for each dyad, even though they involved different traits. It seems clear from these results that MetS is truly a cluster of nested traits that are dependent of each other, whatever the nature of dependency: genetic or environmental.

Some limitations of the present study need to be addressed. Firstly, because our sample depends on the bias associated with a free health check-up, it may not be representative of the general Portuguese population. Moreover, it is possible that these families are in a healthier condition than the overall population, making it harder to detect clusters of cardiovascular risk factors in comparison to high-risk samples. Also, the possibility of false positive results due to multiple statistical testing should not be ruled out. Furthermore, we were not able to account for biological maturation in adolescents, which

could influence our results. Finally, the absence of data on insulin levels may limit the understanding of some traits' relationships.

This study has several strengths such as the large sample size and the reliance on continuous data for MetS traits, as well as the use of statistical methods such as GEE which allowed us to tackle the dependence among family members.

Conclusions

In summary, the present findings confirm the idea of familial aggregation in MetS traits. However, the spousal correlations were of the same magnitude as the biological relatives correlations (with the exception of HDL) impelling us to suggest that shared environment is of utmost importance on the phenotypic expression of MetS traits. Genetic pleiotropy might exist between HDL and TG. Future research on genetic variants mediating MetS traits' correlations would lead to a better understanding of MetS etiology.

Acknowledgments

This work was supported by the FCT – Fundação para a Ciência e a Tecnologia for granting this research (PTDC/DES/67569/2006, FCOMP-01-0124-FEDEB-09608 and SFRH/BD/65290/2009). PTK is partially supported by the Louisiana Public Facilities Authority Endowed Chair in Nutrition. We thank the participants in The Healthy Family Study. The authors state to have no conflict of interest. The guarantor is Daniel Santos.

References

Alberti, K. G., & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med, 15(7), 539-553.

Arya, R., Duggirala, R., Jenkinson, C. P., Almasy, L., Blangero, J., O'Connell, P., & Stern, M. P. (2004). Evidence of a novel quantitative-trait locus for obesity on chromosome 4p in Mexican Americans. Am J Hum Genet, 74(2), 272-282.

Barker, D. J., Hales, C. N., Fall, C. H., Osmond, C., Phipps, K., & Clark, P. M. (1993). Type 2 (non-insulin-dependent) diabetes mellitus, hypertension and hyperlipidaemia (syndrome X): relation to reduced fetal growth. Diabetologia, 36(1), 62-67.

Barter, P. J., Brandrup-Wognsen, G., Palmer, M. K., & Nicholls, S. J. (2010). Effect of statins on HDL-C: a complex process unrelated to changes in LDL-C: analysis of the VOYAGER Database. J Lipid Res, 51(6), 1546-1553.

Benyamin, B., Sorensen, T. I., Schousboe, K., Fenger, M., Visscher, P. M., & Kyvik, K. O. (2007). Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? Diabetologia, 50(9), 1880-1888.

Blangero, J., & Almasy, L. (1997). Multipoint oligogenic linkage analysis of quantitative traits. Genet Epidemiol, 14(6), 959-964.

Butte, N. F., Comuzzie, A. G., Cole, S. A., Mehta, N. R., Cai, G., Tejero, M., Bastarrachea, R., & Smith, E. O. (2005). Quantitative genetic analysis of the metabolic syndrome in Hispanic children. Pediatr Res, 58(6), 1243-1248.

Churilla, J. R., Fitzhugh, E. C., & Thompson, D. L. (2007). The Metabolic syndrome: How Definition Impacts the Prevalence and Risk in U.S. Adults: 1999-2004 NHANES. Metab Syndr Relat Disord, 5(4), 331-342.

de Andrade, M., Gueguen, R., Visvikis, S., Sass, C., Siest, G., & Amos, C. I. (2002). Extension of variance components approach to incorporate temporal trends and longitudinal pedigree data analysis. Genet Epidemiol, 22(3), 221-232.

De Moor, M. H., Willemsen, G., Rebollo-Mesa, I., Stubbe, J. H., De Geus, E. J., & Boomsma, D. I. (2011). Exercise participation in adolescents and their parents: evidence for genetic and generation specific environmental effects. Behav Genet, 41(2), 211-222.

Ervin, R. B. (2009). Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003-2006. Natl Health Stat Report (13), 1-7.

Fiuza, M., Cortez-Dias, N., Martins, S., & Belo, A. (2008). Metabolic syndrome in Portugal: prevalence and implications for cardiovascular risk--results from the VALSIM Study. Rev Port Cardiol, 27(12), 1495-1529.

Ford, E. S., Giles, W. H., & Mokdad, A. H. (2004). Increasing prevalence of the metabolic syndrome among U.S. Adults. Diabetes Care, 27(10), 2444-2449.

Frayn, K. N. (2010). Metabolic Regulation: A Human Perspective: Wiley-Blackwell.

Hebebrand, J., Wulftange, H., Goerg, T., Ziegler, A., Hinney, A., Barth, N., Mayer, H., & Remschmidt, H. (2000). Epidemic obesity: are genetic factors involved via increased rates of assortative mating? Int J Obes Relat Metab Disord, 24(3), 345-353.

Hong, Y., Pedersen, N. L., Brismar, K., & de Faire, U. (1997). Genetic and environmental architecture of the features of the insulin-resistance syndrome. Am J Hum Genet, 60(1), 143-152.

Isomaa, B., Almgren, P., Tuomi, T., Forsen, B., Lahti, K., Nissen, M., Taskinen, M. R., & Groop, L. (2001). Cardiovascular morbidity and mortality associated with the metabolic syndrome. Diabetes Care, 24(4), 683-689.

Johnson, W. D., Kroon, J. J., Greenway, F. L., Bouchard, C., Ryan, D., & Katzmarzyk, P. T. (2009). Prevalence of risk factors for metabolic syndrome in adolescents: National Health and Nutrition Examination Survey (NHANES), 2001-2006. Arch Pediatr Adolesc Med, 163(4), 371-377.

Kathiresan, S., Willer, C. J., Peloso, G. M., Demissie, S., Musunuru, K., Schadt, E. E., Kaplan, L., Bennett, D., Li, Y., Tanaka, T., Voight, B. F., Bonnycastle, L. L., Jackson, A. U., Crawford, G., Surti, A., Guiducci, C., Burtt, N. P., Parish, S., Clarke, R., Zelenika, D., Kubalanza, K. A., Morken, M. A., Scott, L. J., Stringham, H. M., Galan, P., Swift, A. J., Kuusisto, J., Bergman, R. N., Sundvall, J., Laakso, M., Ferrucci, L., Scheet, P., Sanna, S., Uda, M., Yang, Q., Lunetta, K. L., Dupuis, J., de Bakker, P. I., O'Donnell, C. J., Chambers, J. C., Kooner, J. S., Hercberg, S., Meneton, P., Lakatta, E. G., Scuteri, A., Schlessinger, D., Tuomilehto, J., Collins, F. S., Groop, L., Altshuler, D., Collins, R., Lathrop, G. M., Melander, O., Salomaa, V., Peltonen, L., Orho-Melander, M., Ordovas, J. M., Boehnke, M., Abecasis, G. R., Mohlke, K. L., & Cupples, L. A. (2009). Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet, 41(1), 56-65.

Katzmarzyk, P. T., Hebebrand, J., & Bouchard, C. (2002). Spousal resemblance in the Canadian population: implications for the obesity epidemic. Int J Obes Relat Metab Disord, 26(2), 241-246.

Kelley, G. A., Kelley, K. S., & Tran, Z. V. (2004). Aerobic exercise and lipids and lipoproteins in women: a meta-analysis of randomized controlled trials. J Womens Health (Larchmt), 13(10), 1148-1164.

Lakka, H. M., Laaksonen, D. E., Lakka, T. A., Niskanen, L. K., Kumpusalo, E., Tuomilehto, J., & Salonen, J. T. (2002). The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. JAMA, 288(21), 2709-2716.

Lamarche, B., Rashid, S., & Lewis, G. F. (1999). HDL metabolism in hypertriglyceridemic states: an overview. Clin Chim Acta, 286(1-2), 145-161.

LDX, C. (2003a). The Accuracy and Reproducibility of a Rapid, Fingerstick Method for Measuring a Complete Lipid Profile Is Comparable to a Reference Laboratory Method. Cholestech Corporation.

LDX, C. (2003b). The Accuracy and Reproducibility of a Rapid, Fingerstick Method for Measuring a Complete Lipid Profile Is Comparable to a Reference Laboratory Method (b). Cholestech Corporation.

Lohman, T. G., Roche, A. F., & Martorel, R. (1988). Anthropometric Standardization Reference Manual. Illinois: Human Kinetics Books Champaign.

Miller, M., Langenberg, P., & Havas, S. (2007). Impact of lowering triglycerides on raising HDL-C in hypertriglyceridemic and non-hypertriglyceridemic subjects. Int J Cardiol, 119(2), 192-195.

Mitchell, B. D., Kammerer, C. M., Blangero, J., Mahaney, M. C., Rainwater, D. L., Dyke, B., Hixson, J. E., Henkel, R. D., Sharp, R. M., Comuzzie, A. G., VandeBerg, J. L., Stern, M. P., & MacCluer, J. W. (1996). Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. Circulation, 94(9), 2159-2170.

Pang, Z., Zhang, D., Li, S., Duan, H., Hjelmborg, J., Kruse, T. A., Kyvik, K. O., Christensen, K., & Tan, Q. (2010). Multivariate modelling of endophenotypes associated with the metabolic syndrome in Chinese twins. Diabetologia, 53(12), 2554-2561.

Plancoulaine, S., Abel, L., Tregouet, D., Duprez, R., van Beveren, M., Tortevoye, P., Froment, A., & Gessain, A. (2004). Respective roles of serological status and blood specific antihuman herpesvirus 8 antibody levels in human herpesvirus 8 intrafamilial transmission in a highly endemic area. Cancer Res, 64(23), 8782-8787.

Plancoulaine, S., Abel, L., van Beveren, M., Tregouet, D. A., Joubert, M., Tortevoye, P., de The, G., & Gessain, A. (2000). Human herpesvirus 8

transmission from mother to child and between siblings in an endemic population. Lancet, 356(9235), 1062-1065.

Plancoulaine, S., Mohamed, M. K., Arafa, N., Bakr, I., Rekacewicz, C., Tregouet, D. A., Obach, D., El Daly, M., Thiers, V., Feray, C., Abdel-Hamid, M., Abel, L., & Fontanet, A. (2008). Dissection of familial correlations in hepatitis C virus (HCV) seroprevalence suggests intrafamilial viral transmission and genetic predisposition to infection. Gut, 57(9), 1268-1274.

Povel, C. M., Boer, J. M., & Feskens, E. J. (2011). Shared genetic variance between the features of the metabolic syndrome: heritability studies. Mol Genet Metab, 104(4), 666-669.

Reaven, G. M. (1988). Banting lecture 1988. Role of insulin resistance in human disease. Diabetes, 37(12), 1595-1607.

Reynolds, C. A., Baker, L. A., & Pedersen, N. L. (1996). Models of spouse similarity: applications to fluid ability measured in twins and their spouses. Behav Genet, 26(2), 73-88.

Rosa, M. J. V., & Chitas, P. (2010). Portugal: os números. Lisboa.

Statistical Human Genetics. Methods and Protocols. (2012). New York: Humana Press.

Tang, W., Hong, Y., Province, M. A., Rich, S. S., Hopkins, P. N., Arnett, D. K., Pankow, J. S., Miller, M. B., & Eckfeldt, J. H. (2006). Familial clustering for features of the metabolic syndrome: the National Heart, Lung, and Blood Institute (NHLBI) Family Heart Study. Diabetes Care, 29(3), 631-636.

Topouchian, J. A., El Assaad, M. A., Orobinskaia, L. V., El Feghali, R. N., & Asmar, R. G. (2006). Validation of two automatic devices for self-measurement of blood pressure according to the International Protocol of the European Society of Hypertension: the Omron M6 (HEM-7001-E) and the Omron R7 (HEM 637-IT). Blood Press Monit, 11(3), 165-171.

Tregouet, D. A., Herbeth, B., Juhan-Vague, I., Siest, G., Ducimetiere, P., & Tiret, L. (1999). Bivariate familial correlation analysis of quantitative traits by use of estimating equations: application to a familial analysis of the insulin resistance syndrome. Genet Epidemiol, 16(1), 69-83.

Tregouet, D. A., & Tiret, L. (2000). Applications of the estimating equations theory to genetic epidemiology: a review. Ann Hum Genet, 64(Pt 1), 1-14.

Vattikuti, S., Guo, J., & Chow, C. C. (2012). Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. PLoS Genet, 8(3), e1002637.

ARTICLE 3

GENOTYPE BY ENERGY EXPENDITURE INTERACTION WITH METABOLIC SYNDROME TRAITS: THE PORTUGUESE HEALTHY FAMILY STUDY

Daniel Monteiro de Vilhena e Santos[1]

Peter T. Katzmarzyk[2]

Vincent P. Diego[3]

Michele C. Souza[1]

Raquel N. Chaves[1]

John Blangero[3]

José António Ribeiro Maia[1]

Published in PlosOne

Published: Nov 18, 2013

DOI: 10.1371/journal.pone.0080417

[1] CIFI[2]D, Faculty of Sports, University of Porto, Portugal

[2] Pennington Biomedical Research Center, Louisiana State University System, USA

[3] Texas Biomedical Research Institute, San Antonio, TX, USA

Abstract

Moderate-to-high levels of physical activity are established as preventive factors in metabolic syndrome development. However, there is variability in the phenotypic expression of metabolic syndrome under distinct physical activity conditions. In the present study we applied a Genotype X Environment interaction method to examine the presence of GxEE interaction in the phenotypic expression of metabolic syndrome. A total of 958 subjects, from 294 families of The Portuguese Healthy Family study, were included in the analysis. Total daily energy expenditure was assessed using a 3-day physical activity diary. Six metabolic syndrome related traits, including waist circumference, systolic blood pressure, glucose, HDL cholesterol, total cholesterol and triglycerides, were measured and adjusted for age and sex. GxEE examination was performed on SOLAR 4.3.1. All metabolic syndrome indicators were significantly heritable. The GxEE interaction model fitted the data better than the polygenic model (p<0.001) for waist circumference, systolic blood pressure, glucose, total cholesterol and triglycerides. For waist circumference, glucose, total cholesterol and triglycerides, the significant GxEE interaction was due to rejection of the variance homogeneity hypothesis. For waist circumference and glucose, GxEE was also significant by the rejection of the genetic correlation hypothesis. The results showed that metabolic syndrome traits expression is significantly influenced by the interaction established between total daily energy expenditure and genotypes. Physical activity may be considered an environmental variable that promotes metabolic differences between individuals that are distinctively active.

Keywords: Energy Expenditure, Physical Activity, Genotype X Environment, Interaction, Metabolic syndrome.

Introduction

The understanding of the wide range of physical activity (PA) and total daily energy expenditure (TDEE) levels in different populations has been of utmost concern in epidemiological research, because of its relationship with health in general (Bouchard, 2001), and the metabolic syndrome (MetS)—a cluster of interrelated cardiovascular disease risk factors characterized by glucose intolerance, hypertension, dyslipidemia and obesity—in particular (Carroll & Dudfield, 2004). The definition of MetS, and its associated empirical cut-off points of different markers are controversial issues. Moreover, the mechanisms underlying MetS pathophysiology are extremely complex, as it is both genetically and environmentally driven (Teran-Garcia & Bouchard, 2007). In the recent Takahata study (Kaino et al., 2013), the results showed that total energy expenditure was, on average, lower among those individuals with the clustering of MetS risk factors, regardless of body mass index status. Also, it has been suggested that moderate-to-high PA levels have a preventive effect on the development of MetS (Katzmarzyk & Herman, 2007; Strasser, 2012). Further, recent data suggest that distinct PA frequencies and intensities may produce different effects in MetS expression, as a study has shown that everyday activities such as walking and cycling yielded minor effects on MetS whereas high intensity activities performed for more than two hours per week were associated with a lower prevalence of MetS (Hahn et al., 2009).

Phenotypic variation in different MetS indicators at the population level has been studied using different approaches (Athyros et al., 2012; Ford et al., 2008), and new efforts have been employed to disentangle the highly complex architecture of their genetic foundations (Almasy & Blangero, 2000). One aspect that remains to be addressed is the understanding of the differential effects of TDEE and PA on MetS traits via a possible interaction with genetic factors. This possibility was eloquently postulated by Booth and Lees (2007) in a revision of Arthur Beaudet´s concept of environmental-gene interaction (Beaudet, 1995) in which each “individual has their own array of disease susceptibility genes that will interact with physical inactivity to produce maladaptive changes in gene-

expression that often passes a clinical threshold into a chronic disease phenotype" (pg. 148). Data supporting this evidence in humans is scarce. Recently, in a 2004 statement by NIH (Release, 2004) it was indicated that a previously identified type 2 diabetes (T2D) predisposing polymorphism (Silander et al., 2004) would only express itself concurrently with uncertain genetic susceptibility factors, alongside environmental factors such as PA. In a study on rats, the results showed that after 11 generations of controlled selection in which high aerobic capacity animals were mated, those with high aerobic capacity had 12% lower mean 24 h blood pressure, 16% lower fasting blood glucose, 56% lower fasting plasma insulin, and 63% lower plasma triglycerides than those with low aerobic capacity (Wisloff et al., 2005). The logical inference was that genetically determined aerobic capacity is related to MetS traits (Bernal-Mizrachi & Semenkovich, 2006).

Contemporary technological advances and new statistical models allow geneticists to study EE and PA and its impact on phenotypic variance in different MetS traits using an escalating wealth of complex genetic models (Athyros et al., 2012). These developments permit the estimation of complex MetS traits using family data exploiting the elegant flexibility of maximum likelihood estimation techniques [see Elston and colleagues (Statistical Human Genetics. Methods and Protocols, 2012)]. These statistical techniques, mostly based on variance components (VC) models, provide a robust framework for statistical inference, which includes parameter and model-likelihood estimation, and associated likelihood ratio testing procedures (Blangero, 2009).

Under the VC model approach, we can formally test for genotype x environment (GxE) interaction (DeYoung & Clark, 2012). Very briefly, this interaction arises when a genotype yields distinct phenotypic expressions under contrasting environments (Falconer, 1981). Here we hypothesize that genotype x EE (GxEE) interaction is a potent determinant of variation in MetS traits. This possibility deserves further consideration, because it can inform the discussion around the preventive and systematic protective effects of EE and PA on MetS traits. For example, in the Nurses Health Study, it was concluded that PA interacts with T2D susceptibility, because among the women whose parents

had T2D, a 65% greater risk of developing T2D was found for those who were on the lower quintile of PA level when compared to those in the higher quintile. On the contrary, among the daughters of non-diabetic parents, being in the lower quintile of PA was associated with twice the risk of developing diabetes than being in the most active quintile.

The main purposes of the present study, using a nuclear family design, are to estimate the magnitude of genetic factors responsible for the architecture of MetS traits, and to study potential GxEE interaction.

Materials and Methods

The *Portuguese Healthy Family Study*, from the Portuguese *Estudo de Famílias Saudáveis Portuguesas* (FAMS), investigates the relationship among MetS traits, PA, physical fitness and body composition in nuclear families. Children and adolescents aged 10 to 18 years were recruited in schools from the north and central regions of mainland Portugal, and were approached to freely participate in the study with their siblings and parents. The ethics committee of the Faculty of Sport, University of Porto, approved the study, and written informed consent, and assent, was obtained from the parents (or guardians). Given that families with 3 or more children are scarce in the Portuguese population (Rosa & Chitas, 2010), a total of 500 families with at least one child were invited to participate in this study. Of these, 294 families agreed to participate with at least two family members (see Table 1).

Physical Activity

Using a 3-day physical activity diary (Bouchard et al., 1983), a trained technician interviewed each subject, recording the dominant activity for each 15-min period during 24 h by using a list of categorized activities. Categories from 1 to 9 refer to increasing levels of energy expenditure (METs) of each activity in which category 1 indicates very low energy expenditure such as sleeping or resting in bed, and category 9 refers to highly demanding physical work such as high-intensity sports. Approximate median energy cost for each of the nine categories in kcal/kg/15 min was used to compute the daily energy

expenditure for each individual. The number of 15-min periods for each category was first summed over the 3-day period and weighted by its own median energy cost. Total energy expenditure (TEE) was then calculated by summing over the median energy cost of all nine categories and multiplying by subjects´ body weights. Total daily energy expenditure [TDEE (kcal/day)] was then calculated by dividing TEE by 3.

BLOOD SAMPLING AND MEASUREMENTS OF CARDIOVASCULAR RISK FACTORS

Blood samples were collected after an overnight fast of at least 10-12 h. Glucose (GLU), total cholesterol, HDL-cholesterol (HDL), and triglycerides (TG) were analysed with an LDX point of care analyser (LDX, 2003a). This method has been previously validated against a laboratory reference method (LDX, 2003b), and daily optical equipment checks were made according to manufacturer instructions.

Resting systolic blood pressure (SBP) was measured with an Omron Model M6 (HEM-7001-E) device according to The International Protocol of the European Society of Hypertension (Topouchian et al., 2006). Cuff sizes were modified depending on the size of the participant's arm. Subjects were seated in an upright position and the right arm sitting on a table at the heart level. The first reading was performed after a 5 minute resting period. The other two readings were performed with three-minute breaks in between. The mean of the three blood pressure measurements was used for further analysis. All blood samples and blood pressure analysis were performed between 7:30 am and 10:30 am.

Waist circumference (WC) was measured with a Holtain flexible tape at the level of the smallest waist perimeter, with the subject standing erect with relaxed abdominal muscles and at the end of normal expiration.

STATISTICAL ANALYSIS

Univariate quantitative genetic procedures as implemented in SOLAR (Almasy & Blangero, 1998) under a special class of the multivariate linear model, namely the variance components (VC) approach, were used to estimate additive genetic and environmental VCs for each of the MS traits. Prior to all

modelling, age, age^2, sex and their relevant interactions were used as covariates in a preliminary VC model. Residuals were thus derived for each trait and were normalized using an inverse normal transformation, as previously advocated (Blangero et al., 2013; Diego et al., 2007). Heritability estimates (h^2) were computed using a maximum likelihood approach to estimate variance components under the standard polygenic model as implemented in SOLAR v.4.3.1 software (Almasy & Blangero, 1998).

To test for GxEE interaction, basic initial hypotheses were formulated regarding the variance/covariance relationship of a MetS indicator between family members with different levels of TDEE. As regards GxEE interaction, the fundamental null hypothesis is that the expression of a polygenotype (i.e., aggregate of all genotypes related to the expression of a phenotype) is independent of TDEE level. It can be shown from first principles that if there is no GxEE interaction, the same MetS indicator measured in subjects with different levels of TDEE will have a genetic correlation of 1.0 and the genetic variance will be homogeneous across all levels of TDEE (Blangero, 2009; Diego et al., 2003). On the contrary, if GxEE interaction is present, the genetic correlation will be significantly less than 1.0 and/or the genetic variance will not be the same among all levels of TDEE.

The foregoing requires that we model the variance and correlation as functions of TDEE levels. For the genetic variance function (and similarly for the environmental variance function), we modelled the variance using an exponential function to ensure positivity, which is required since any variance is a squared term (Blangero, 2009; Diego et al., 2003): $\sigma_g^2 = \exp\left[\alpha_g + \gamma_g\left(EE\right)\right]$ (Eq. 1), where α_g and γ_g are parameters to be estimated. An additional justification for the exponential function is suggested by the alternative name of this approach, namely the log-linear model of the variance: $\ln\sigma_g^2 = \alpha_g + \gamma_g\left(EE\right)$ (Eq. 2). That is, on taking the natural logarithm of the variance modelled as an exponential function, we have the equation of a line. In this form, the variance homogeneity null hypothesis obviously holds for a slope-term equal to 0: $\gamma_g = 0$. For the genetic correlation function, we modelled the genetic correlation as an

exponential decay function of the pairwise differences in TDEE levels: $\rho_g = \exp\left(-\lambda\left|EE_i - EE_j\right|\right)$ (Eq. 3), where λ is a parameter to be estimated as a function of the difference in TDEE levels between any two individuals *i* and *j*. Here we also have an elegant reexpression of the interaction null hypothesis, in this case regarding the genetic correlation, in that a genetic correlation equal to 1 is equivalent to $\lambda = 0$. This is because for $\lambda = 0$, we have $\rho_g = \exp\left(-\lambda\left|EE_i - EE_j\right|\right) = e^0 = 1$ (Eq. 4).

For reasons detailed in Diego et al. (2003), the likelihood ratio test statistics (LRTs) to test $\gamma_g = 0$ and $\lambda = 0$ are respectively distributed as χ_1^2, a chi-square random variable with 1 degree of freedom (d.f.), and $\left(\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2\right)$, a 50:50 mixture of chi-square random variable with a point-mass at 0, denoted by χ_0^2, and a chi-square with 1 d.f. Prior to examination of these hypotheses, however, we first confirmed if the overall GxEE interaction model provided a better fit to the data than the standard so-called polygenic model. The LRT for this comparison can be shown to be distributed as $\left(\frac{1}{2}\chi_2^2 + \frac{1}{2}\chi_3^2\right)$ (Diego, 2005).

Results

Table 1 presents basic descriptive information. Some relatives were not able to fully engage in the data collection procedures. As such, a total of 958 subjects, comprising 180 fathers, 253 mothers, 265 sons and 260 daughters, from 294 families were included. The average family size was 3.3 subjects. Families are, on average, young and the results are as expected as the mean values for all the MetS traits were consistently higher in parents than in offspring. Also, with the exception of HDL cholesterol, MetS indicators were higher in fathers than in mothers. The MetS profiles of sons and daughters were similar with daughters showing higher mean levels of TC, HDL and TG.

All MetS indicators showed highly significant h^2 estimates ranging from 0.21 (TG) to 0.59 (HDL) (Table 2), meaning that there are strong additive

genetic factors affecting their expression in family members that may justify a further specific analysis of their genetic architecture.

Table I. Sample descriptive characteristics (means ± standard deviations).

	Fathers (n=180)	Mothers (n=253)	Sons (n=265)	Daughters (n=260)
Age (yrs)	45.36 ± 5.17	43.49 ± 4.47	14.68 ± 2.78	14.40 ± 2.80
TDEE (kcal/day)	3561.79 ± 962.71	2788.37 ± 527.58	2280.57 ± 774.43	2024.85 ± 568.43
WC (cm)	92.34 ± 10.64	80.98 ± 8.99	72.74 ± 10.41	68.39 ± 8.56
SBP (mmHg)	131.7 ± 14.27	122.12 ± 16.14	117.60 ± 13.06	113.13 ± 10.49
GLU (mg/dl)	97.17 ± 13.19	87.77 ± 13.27	85.39 ± 9.10	83.24 ± 9.01
HDL (mg/dl)	44.58 ± 13.76	55.94 ± 14.37	47.88 ± 13.71	53.30 ± 12.85
TC (mg/dl)	196.72 ± 41.32	181.17 ± 32.17	140.64 ± 24.49	150.44 ± 26.42
TG (mg/dl)	140.33 ± 103.85	109.37 ± 62.63	62.55 ± 30.70	79.57 ± 50.00

Legend: WC – waist circumference; SBP – systolic blood pressure; GLU – glucose; HDL – high density lipoprotein; TC – Total cholesterol; TG – triglycerides.

Table II. Heritability estimates (h^2) and corresponding 95% confidence intervals (95%CI) of the different phenotypes in the Portuguese Healthy Families Study

Trait	h^2	Std. Error	p-value	95%CI
WC (cm)	0.34	0.07	<0.001	0.22-0.45
SBP (mmHg)	0.40	0.07	<0.001	0.27-0.51
GLU (mg/dl)	0.29	0.07	<0.001	0.18-0.40
HDL (mg/dl)	0.59	0.06	<0.001	0.48-0.69
TC (mg/dl)	0.51	0.07	<0.001	0.39-0.62
TG (mg/dl)	0.21	0.08	0.002	0.09-0.33

Legend: WC – waist circumference; SBP – systolic blood pressure; HDL – high density lipoprotein; TC – Total cholesterol; TG – triglycerides.

To test for the influence of TDEE and PA on the expression of MS indicators, the polygenic model was compared to the GxEE model by means of a log-likelihood ratio test (see Table 3). The GxEE interaction model is significantly better than the polygenic model for WC, SBP, GLU, TC and TG, meaning that the GxEE interaction model fits the data better than the polygenic model for these five traits.

However, to verify if there is GxEE interaction, the full model was compared to its constrained alternatives (i.e. setting or) for WC, SBP, GLU, TC and TG.

Table III. Results of log-likelihood ratio tests (LRT) and respective p-values contrasting a polygenic model vs a GEE model for each of the MS indicators

Trait	Polygenic LnL	GxTDEE LnL	LRT	p-value
WC	-380.061	-319.731	120.660	<0.0001
SBP	-370.926	-364.625	12.601	0.004
GLU	-444.331	-384.913	118.835	<0.0001
HDL	-340.745	-340.542	0.408	0.877
TC	-357.813	-343.478	28.669	<0.0001
TG	-380.194	-331.080	98.228	<0.0001

Legend: WC – waist circumference; SBP – systolic blood pressure; HDL – high density lipoprotein; TC – Total cholesterol; TG – triglycerides; LnL – log-likelihoods; LRT – Likelihood ratio test.

In Figure 1, we display the results for those traits that showed significant variance heterogeneity and a correlation function that is significantly different from 1. For WC, GLU, TC and TG, significant GxEE interaction was due to the rejection of the variance homogeneity hypothesis; i.e., variance heterogeneity. For WC and GLU, the null hypothesis that the genetic correlation (ρG) equals 1 was also significantly rejected. Figure 1a highlights that, for GLU, TC, TG, and WC, genetic variance increases with increasing levels of TDEE. On the other hand, figure 1b demonstrates that, for GLU and WC, the genetic correlation decreases as the differences between TDEE level increases among family members. It may be noticed that genetic correlation for GLU decays to 0 almost instantaneously. However, this is simply an artefact of the two correlations being plotted on one scale and therefore the genetic correlation that decreases at a faster rate seems to immediately go to 0. This is clarified in the next figure.

Figure 1. Genotype X Energy Expenditure genetic variance (a) and Genotype X Energy Expenditure genetic correlation (b).

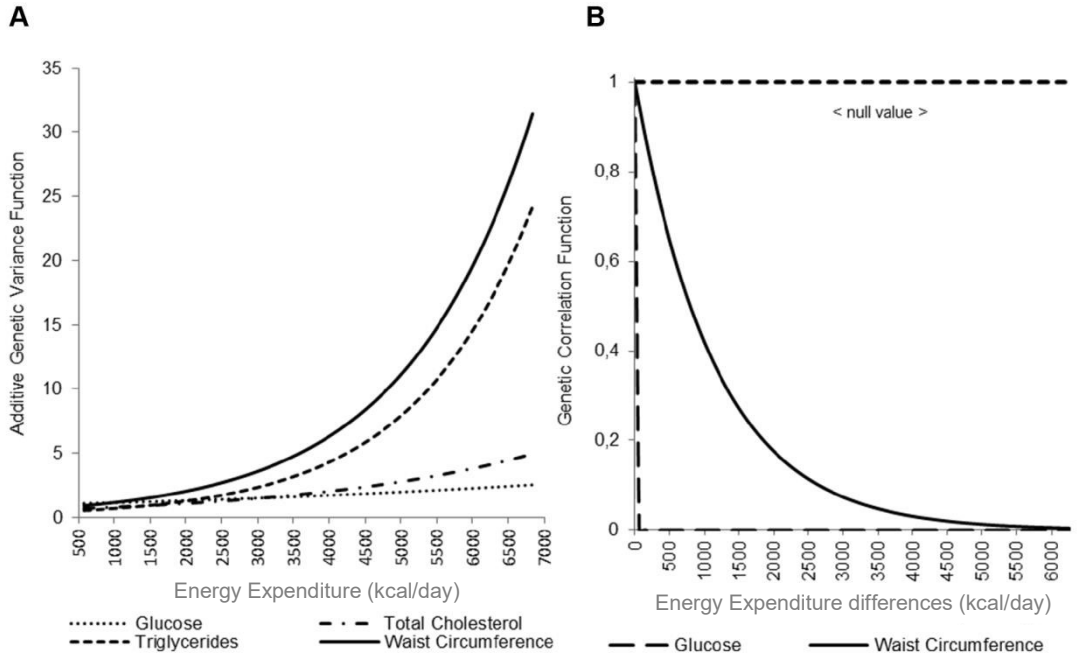
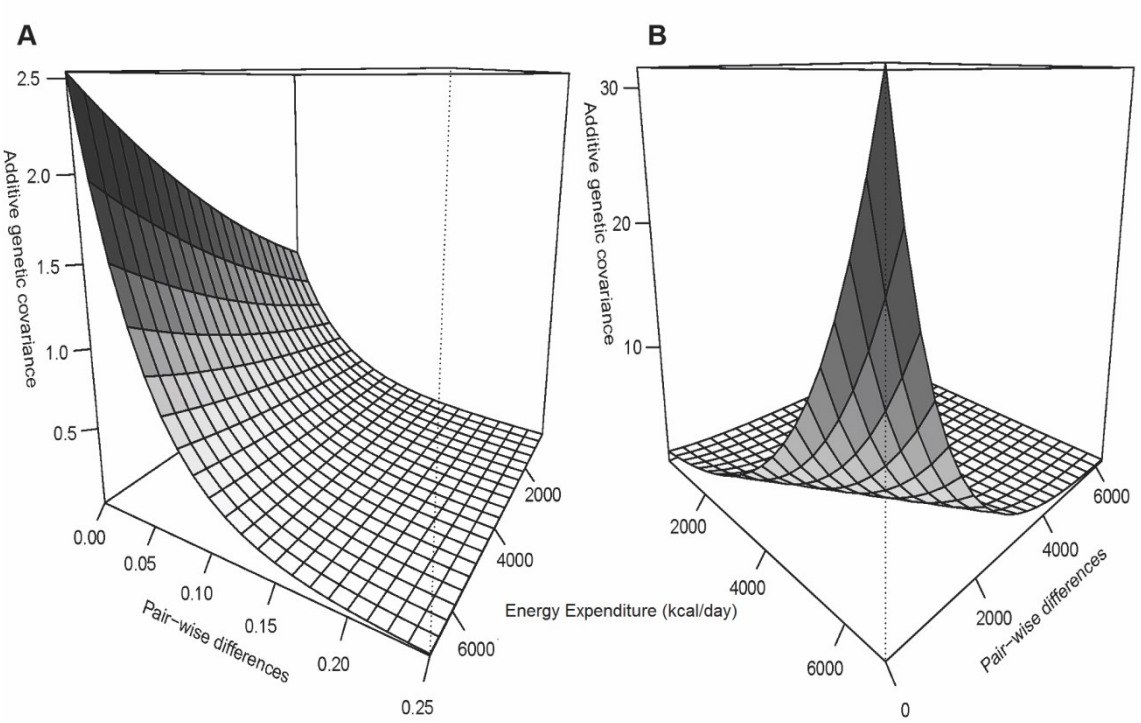


Figure 2 illustrates that, for GLU and WC, GxEE interaction is a joint function of genetic variance heterogeneity and a genetic correlation function not equal to one. Thus, we express them jointly as a covariance function in the vertical axis. Due to space restrictions on the 3-dimensional plot, we abbreviated the corresponding axes in figure 1 to TDEE and pair-wise differences. Moreover, because the traits change at different rates, and to show the correlation functions for both traits at the appropriate resolution, the pair-wise differences are shown on different scales. In fact, the GLU correlation function curve is decreasing at a much faster rate than WC, forcing us to change the scale.

Figure 2. Genetic covariance function for Fasting Glucose (a) and Waist Circumference (b). Total Daily Energy Expenditure units are in kcal/day.



Discussion

The present report aimed to assess the genetic variance present in MetS traits as well as to examine potential GxEE interaction that has an effect on MetS traits. Our results confirm the importance of genetics on MetS traits with all h^2 being significant and, more interestingly, highlight the importance of GxEE interaction in the phenotypic determination of MetS traits.

To the best of our knowledge this is the first attempt to apply a GxE interaction analysis to better understand the differential relationship between TDEE and PA with MetS risk factors. A GxE interaction effect is present when the phenotypic expressions of an environmental factor or behaviour is conditional to the genotype of an individual. This study provides evidence that there is a genetic basis for the variability in quantitative measures of the metabolic syndrome that is mediated by energy expenditure and/or physical activity.

The present findings confirm that MetS traits are highly heritable in agreement with previous results (Tang et al., 2006; Vattikuti et al., 2012). HDL

(h^2=0.59) was the most heritable of the MetS traits which is consistent with results from elsewhere ranging from 0.46 in Tehran Lipid and Glucose Study (TLGS) (Zarkesh et al., 2012), to 0.63 in the Family Heart Study (FHS) (Tang et al., 2006) population. On the other hand, TG (h^2=0.21), in disagreement with the data from TLGS (h^2=0.36) and FHS (h^2=0.48) (Tang et al., 2006; Zarkesh et al., 2012), was the least heritable trait. These differences could be attributable to distinct genetic architectures, namely gene frequencies and their value, as well as distinct environmental factors specific to each population. Also, distinct analysis strategies [e.g., Tang and colleagues (2006) made adjustments for a set of confounders that were not used in the present report], and different sample sizes might explain some of the variability in the TG heritability estimates.

It is well established that PA has a preventive effect on a variety of morbidities associated with cardiovascular diseases (Ford et al., 2008). For instance, studies on familial hypercholesterolemia (FH) in Utah families seem to demonstrate this very point. While FH, which leads to early-onset coronary heart disease, is caused by mutations in the LDL-receptor gene (*LDLR*), it was shown that heterozygous carriers of the *LDLR* mutation lived into their eighth and ninth decades in the 19th century but their descendants in the 20th century, who were also heterozygous carriers, lived only into their third and fourth decades (Hegele, 1992, 2002). These investigators documented that the carriers who lived in the 19th century were relatively more physically active and enjoyed a more nutritious diet than the carriers who lived in the 20th century. However, it has been shown that there is a considerable heterogeneity in the response to PA leading to distinct signals in terms of cardiovascular risk factors (Bouchard, 2001). For example, in the HERITAGE Family Study, after 20 weeks of supervised training sessions, a significant mean 3.6% increase in plasma HDL was observed together with a high inter-variability in responsiveness to training, ranging from a mean 9.3% decrease in Quartile 1 of HDL-C response to a mean 18% increase in Quartile 4. Moreover, the authors verified that only 15.5% of the variability was due to baseline variables and training adaptations, concluding that only a minor extent of the adaptation could be predictable by

nongenetic factors. This complex and controversial theme has been addressed in a recent paper by Bouchard et al. (2012) in which some participants, when exposed to regular exercise, ended up having worsened metabolic profiles. The preceding raises questions related to the importance of genetic susceptibility in explaining the variability in the response to similar levels of PA.

Our results showed that there is additive genetic variance heterogeneity for GLU, TC, TG, and WC across TDEE levels, and that for WC and GLU, the genetic correlation between distinct levels of TDEE was different from 1. These results may be related to work demonstrating that physical activity and inactivity have anti- and pro-inflammatory effects, respectively (Fenza & Fiorina, 2012; Hamer et al., 2012). It is widely believed that the MetS and MetS-associated diseases such as type 2 diabetes (T2D) and cardiovascular disease (CVD), are caused in a large part by chronic sub-clinical inflammation (Hansson, 2005; Hotamisligil, 2006). That we observed increasing additive genetic variance heterogeneity for GLU, TC, TG, and WC with increasing TDEE levels is consistent with upregulation in the genes involved in anti-inflammatory processes. Recent work (Richardson et al., 2013) showed significant SNP–moderate to vigorous PA interactions on BMI-for-age Z-scores in European-American at *GNPDA2* and *FTO* genes, and in Hispanic-American at *LZTR2/SEC16B*. In 2011 (Kilpelainen et al., 2011), a robust meta-analysis of 218,166 adults and 19,268 children found that, in adults, PA attenuated the effect (p-value for interaction = 0.001) of the minor (*A2*) allele of *rs9939609* on obesity. Even though this interaction failed to be statistically significant in children and adolescents, in adults the minor allele of *rs9939609* increased the odds of being obese less in the physically active group (odds ratio = 1.22/allele, 95% CI 1.19–1.25) than in the inactive group (odds ratio = 1.30/allele, 95% CI 1.24–1.36). These results seem to suggest that there are genes associated with increased MetS risk factors, such as obesity, that actually interact with different EE levels. Further, that we observed a genetic correlation function different from 1 across TDEE differences is consistent with the activation of genes involved in pro- and anti-inflammatory processes at relatively low and high PA levels, respectively (Whyte & Laughlin, 2010).

We speculate that potential epigenetic changes driven by EE might be linked to a major epigenetic modification - DNA methylation - that suppresses gene expression by modulating the access of the transcription machinery to the chromatin or by recruiting methyl-binding proteins (Barres et al., 2012). In a recent paper by Barres et al. (2009) global methylation values decreased after intense exercise, even after controlling for the effect of haemoglobin mRNA content. More specifically, the results highlighted that captured methylated promoters for metabolic genes, previously linked to type 2 diabetes (Simonen et al., 2003a), were lower after acute exercise, leading the authors to suggest that acute exercise induces gene-specific DNA hypo methylation in human skeletal muscle (Barres et al., 2009).

As mentioned earlier we were unable to find similar papers to which we could compare our results and assess the suitability of our interpretations. It is quite interesting that despite the great wealth of data on PA and MetS, be it from descriptive and/or interactions studies, results are still lacking on the complex and divergent effects of PA on MetS traits given individuals´ genotypes. Perhaps this may be due to the complex nature of MetS, the statistical challenges that GxE interaction poses, and the necessity of having large samples of families. Moreover, with the development of DNA sequence analysis, and the possibility of running genome wide associations scans (GWAs), many investigators have devoted their attention to finding specific loci associated with the phenotypic variability of MetS traits (Visscher et al., 2012). We feel as though it is extremely important to be able to merge the new evidence that are now being brought to life by GWAs with the results that are shown in this report and that consubstantiate that there is an underlying effect of genetic susceptibility in phenotypic expression of MetS traits.

Notwithstanding the importance of the present results, the fact that this study is based on a free health check-up, that may not be representative of the general Portuguese population, is a limitation as well as the relatively young and healthy sample of families may limit the generalizability of the results to older individuals who are more prone to develop MetS. Also, the lack of information about resting metabolic rate may influence the PAEE results,

although previous work (Simonen et al., 2003a) with this 3-day diary never considered this possibility. Another limitation is related to a lack of information on nutritional habits, although this information is never easy to evaluate. However, the large sample size, the reliance on continuous data for MetS traits, the use of state of the art statistical procedures and the novelty of the analysis in TDEE and PA genetic epidemiology research, are strengths of the present study.

In conclusion, the present results demonstrate that MetS trait expression is significantly influenced by the way in which the genotype “deals” with distinct TDEE levels. As such, PA may be considered an environmental variable that promotes metabolic differences between individuals that are distinctively active. This is valuable information for health practitioners. More efforts should be devoted to identify specific loci that control MetS traits and to test if those loci are regulated or not by TDEE and/or PA.

Acknowledgments

This work was supported by the FCT – Fundação para a Ciência e a Tecnologia: PTDC/DES/67569/2006 FCOMP-01-0124-FEDEB-09608 and SFRH/BD/65290/2009.

We thank the participants in The Portuguese Healthy Family Study.

References

Almasy, L., & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet, 62(5), 1198-1211.

Almasy, L., & Blangero, J. (2000). Challenges for genetic analysis in the 21st century: localizing and characterizing genes for common complex diseases and their quantitative risk factors. GeneScreen, 1(2), 113-116.

Athyros, V. G., Elisaf, M. S., Alexandrides, T., Achimastos, A., Ganotakis, E., Bilianou, E., Karagiannis, A., Liberopoulos, E. N., Tziomalos, K., & Mikhailidis, D. P. (2012). Long-term impact of multifactorial treatment on new-onset diabetes and related cardiovascular events in metabolic syndrome: a post hoc ATTEMPT analysis. Angiology, 63(5), 358-366.

Barres, R., Osler, M. E., Yan, J., Rune, A., Fritz, T., Caidahl, K., Krook, A., & Zierath, J. R. (2009). Non-CpG methylation of the PGC-1alpha promoter through DNMT3B controls mitochondrial density. Cell Metab, 10(3), 189-198.

Barres, R., Yan, J., Egan, B., Treebak, J. T., Rasmussen, M., Fritz, T., Caidahl, K., Krook, A., O'Gorman, D. J., & Zierath, J. R. (2012). Acute exercise remodels promoter methylation in human skeletal muscle. Cell Metab, 15(3), 405-411.

Beaudet, A. L. (1995). Genetics, biochemistry, and molecular basis of variant human phenotypes. In C. R. Scriver, A. L. Beaudet, W. S. Sly, D. Valle, J. B. Stanbury, J. B. Wyngaarden & D. S. Fredrickson (Eds.), The Metabolic and Molecular Bases of Inherited Disease (pp. 79). New York: McGraw-Hill.

Bernal-Mizrachi, C., & Semenkovich, C. F. (2006). Fast predators or fast food, the fit still survive. Nat Med, 12(1), 46-47; discussion 47.

Blangero, J. (2009). Statistical genetic approaches to human adaptability. 1993. Hum Biol, 81(5-6), 523-546.

Blangero, J., Diego, V. P., Dyer, T. D., Almeida, M., Peralta, J., Kent, J. W., Jr., Williams, J. T., Almasy, L., & Goring, H. H. (2013). A kernel of truth: statistical

advances in polygenic variance component models for complex human pedigrees. Adv Genet, 81, 1-31.

Booth, F. W., & Lees, S. J. (2007). Fundamental questions about genes, inactivity, and chronic diseases. Physiol Genomics, 28(2), 146-157.

Bouchard, C. (2001). Physical activity and health: introduction to the dose-response symposium. Med Sci Sports Exerc, 33(6 Suppl), S347-350.

Bouchard, C., Blair, S. N., Church, T. S., Earnest, C. P., Hagberg, J. M., Hakkinen, K., Jenkins, N. T., Karavirta, L., Kraus, W. E., Leon, A. S., Rao, D. C., Sarzynski, M. A., Skinner, J. S., Slentz, C. A., & Rankinen, T. (2012). Adverse metabolic response to regular exercise: is it a rare or common occurrence? PLoS One, 7(5), e37887.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Theriault, G. (1983). A method to assess energy expenditure in children and adults. Am J Clin Nutr, 37(3), 461-467.

Carroll, S., & Dudfield, M. (2004). What is the relationship between exercise and metabolic abnormalities? A review of the metabolic syndrome. Sports Med, 34(6), 371-418.

DeYoung, C. G., & Clark, R. (2012). The gene in its natural habitat: the importance of gene-trait interactions. Dev Psychopathol, 24(4), 1307-1318.

Diego, V. P. (2005). GENOTYPE × AGE INTERACTION, AND THE INSULIN-LIKE GROWTH FACTOR I AXIS IN THE SAN ANTONIO FAMILY HEART STUDY: A STUDY IN HUMAN SENESCENCE. New York: Graduate School of Binghamton University.

Diego, V. P., Almasy, L., Dyer, T. D., Soler, J. M., & Blangero, J. (2003). Strategy and model building in the fourth dimension: a null model for genotype x age interaction as a Gaussian stationary stochastic process. BMC Genet, 4 Suppl 1, S34.

Diego, V. P., Rainwater, D. L., Wang, X. L., Cole, S. A., Curran, J. E., Johnson, M. P., Jowett, J. B., Dyer, T. D., Williams, J. T., Moses, E. K., Comuzzie, A. G., Maccluer, J. W., Mahaney, M. C., & Blangero, J. (2007). Genotype x adiposity interaction linkage analyses reveal a locus on chromosome 1 for lipoprotein-associated phospholipase A2, a marker of inflammation and oxidative stress. Am J Hum Genet, 80(1), 168-177.

Falconer, D. S. (1981). Introduction to Quantitative Genetics. London, UK: Longman.

Fenza, R., & Fiorina, P. (2012). Physical Activity and Inflammation. In L. Luzi (Ed.), Cellular Physiology and Metabolism of Physical Exercise (pp. 99-108): Springer Milan.

Ford, E. S., Li, C., Zhao, G., Pearson, W. S., & Mokdad, A. H. (2008). Prevalence of the metabolic syndrome among U.S. adolescents using the definition from the International Diabetes Federation. Diabetes Care, 31(3), 587-589.

Hahn, V., Halle, M., Schmidt-Trucksass, A., Rathmann, W., Meisinger, C., & Mielck, A. (2009). Physical activity and the metabolic syndrome in elderly German men and women: results from the population-based KORA survey. Diabetes Care, 32(3), 511-513.

Hamer, M., Sabia, S., Batty, G. D., Shipley, M. J., Tabak, A. G., Singh-Manoux, A., & Kivimaki, M. (2012). Physical activity and inflammatory markers over 10 years: follow-up in men and women from the Whitehall II cohort study. Circulation, 126(8), 928-933.

Hansson, G. K. (2005). Inflammation, atherosclerosis, and coronary artery disease. N Engl J Med, 352(16), 1685-1695.

Hegele, R. A. (1992). Gene-environment interactions in atherosclerosis. Mol Cell Biochem, 113(2), 177-186.

Hegele, R. A. (2002). Environmental modulation of atherosclerosis end points in familial hypercholesterolemia. Atheroscler Suppl, 2(3), 5-7.

Hotamisligil, G. S. (2006). Inflammation and metabolic disorders. Nature, 444(7121), 860-867.

Kaino, W., Daimon, M., Sasaki, S., Karasawa, S., Takase, K., Tada, K., Wada, K., Kameda, W., Susa, S., Oizumi, T., Fukao, A., Kubota, I., Kayama, T., & Kato, T. (2013). Lower physical activity is a risk factor for a clustering of metabolic risk factors in non-obese and obese Japanese subjects: the Takahata study. Endocr J, 60(5), 617-628.

Katzmarzyk, P., & Herman, K. (2007). The role of physical activity and fitness in the prevention and treatment of metabolic syndrome. Current Cardiovascular Risk Reports, 1(3), 228-236.

Kilpelainen, T. O., Qi, L., Brage, S., Sharp, S. J., Sonestedt, E., Demerath, E., Ahmad, T., Mora, S., Kaakinen, M., Sandholt, C. H., Holzapfel, C., Autenrieth, C. S., Hypponen, E., Cauchi, S., He, M., Kutalik, Z., Kumari, M., Stancakova, A., Meidtner, K., Balkau, B., Tan, J. T., Mangino, M., Timpson, N. J., Song, Y., Zillikens, M. C., Jablonski, K. A., Garcia, M. E., Johansson, S., Bragg-Gresham, J. L., Wu, Y., van Vliet-Ostaptchouk, J. V., Onland-Moret, N. C., Zimmermann, E., Rivera, N. V., Tanaka, T., Stringham, H. M., Silbernagel, G., Kanoni, S., Feitosa, M. F., Snitker, S., Ruiz, J. R., Metter, J., Larrad, M. T., Atalay, M., Hakanen, M., Amin, N., Cavalcanti-Proenca, C., Grontved, A., Hallmans, G., Jansson, J. O., Kuusisto, J., Kahonen, M., Lutsey, P. L., Nolan, J. J., Palla, L., Pedersen, O., Perusse, L., Renstrom, F., Scott, R. A., Shungin, D., Sovio, U., Tammelin, T. H., Ronnemaa, T., Lakka, T. A., Uusitupa, M., Rios, M. S., Ferrucci, L., Bouchard, C., Meirhaeghe, A., Fu, M., Walker, M., Borecki, I. B., Dedoussis, G. V., Fritsche, A., Ohlsson, C., Boehnke, M., Bandinelli, S., van Duijn, C. M., Ebrahim, S., Lawlor, D. A., Gudnason, V., Harris, T. B., Sorensen, T. I., Mohlke, K. L., Hofman, A., Uitterlinden, A. G., Tuomilehto, J., Lehtimaki, T., Raitakari, O., Isomaa, B., Njolstad, P. R., Florez, J. C., Liu, S., Ness, A., Spector, T. D., Tai, E. S., Froguel, P., Boeing, H., Laakso, M., Marmot, M.,

Bergmann, S., Power, C., Khaw, K. T., Chasman, D., Ridker, P., Hansen, T., Monda, K. L., Illig, T., Jarvelin, M. R., Wareham, N. J., Hu, F. B., Groop, L. C., Orho-Melander, M., Ekelund, U., Franks, P. W., & Loos, R. J. (2011). Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. PLoS Med, 8(11), e1001116.

LDX, C. (2003a). The Accuracy and Reproducibility of a Rapid, Fingerstick Method for Measuring a Complete Lipid Profile Is Comparable to a Reference Laboratory Method. Cholestech Corporation.

LDX, C. (2003b). The Accuracy and Reproducibility of a Rapid, Fingerstick Method for Measuring a Complete Lipid Profile Is Comparable to a Reference Laboratory Method (b). Cholestech Corporation.

Release, N. N. (2004). Gene variants may increase susceptibility to type 2 diabetes.

Richardson, A. S., North, K. E., Graff, M., Young, K. M., Mohlke, K. L., Lange, L. A., Lange, E. M., Harris, K. M., & Gordon-Larsen, P. (2013). Moderate to vigorous physical activity interactions with genetic variants and body mass index in a large US ethnically diverse cohort. Pediatr Obes.

Rosa, M. J. V., & Chitas, P. (2010). Portugal: os números. Lisboa.

Silander, K., Mohlke, K. L., Scott, L. J., Peck, E. C., Hollstein, P., Skol, A. D., Jackson, A. U., Deloukas, P., Hunt, S., Stavrides, G., Chines, P. S., Erdos, M. R., Narisu, N., Conneely, K. N., Li, C., Fingerlin, T. E., Dhanjal, S. K., Valle, T. T., Bergman, R. N., Tuomilehto, J., Watanabe, R. M., Boehnke, M., & Collins, F. S. (2004). Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes. Diabetes, 53(4), 1141-1149.

Simonen, R. L., Rankinen, T., Perusse, L., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., & Bouchard, C. (2003a). A dopamine D2 receptor gene polymorphism and physical activity in two family studies. Physiol Behav, 78(4-5), 751-757.

Statistical Human Genetics. Methods and Protocols. (2012). New York: Humana Press.

Strasser, B. (2012). Physical activity in obesity and metabolic syndrome. Ann N Y Acad Sci.

Tang, W., Hong, Y., Province, M. A., Rich, S. S., Hopkins, P. N., Arnett, D. K., Pankow, J. S., Miller, M. B., & Eckfeldt, J. H. (2006). Familial clustering for features of the metabolic syndrome: the National Heart, Lung, and Blood Institute (NHLBI) Family Heart Study. Diabetes Care, 29(3), 631-636.

Teran-Garcia, M., & Bouchard, C. (2007). Genetics of the metabolic syndrome. Appl Physiol Nutr Metab, 32(1), 89-114.

Topouchian, J. A., El Assaad, M. A., Orobinskaia, L. V., El Feghali, R. N., & Asmar, R. G. (2006). Validation of two automatic devices for self-measurement of blood pressure according to the International Protocol of the European Society of Hypertension: the Omron M6 (HEM-7001-E) and the Omron R7 (HEM 637-IT). Blood Press Monit, 11(3), 165-171.

Vattikuti, S., Guo, J., & Chow, C. C. (2012). Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. PLoS Genet, 8(3), e1002637.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. Am J Hum Genet, 90(1), 7-24.

Whyte, J. J., & Laughlin, M. H. (2010). The effects of acute and chronic exercise on the vasculature. Acta Physiol (Oxf), 199(4), 441-450.

Wisloff, U., Najjar, S. M., Ellingsen, O., Haram, P. M., Swoap, S., Al-Share, Q., Fernstrom, M., Rezaei, K., Lee, S. J., Koch, L. G., & Britton, S. L. (2005). Cardiovascular risk factors emerge after artificial selection for low aerobic capacity. Science, 307(5708), 418-420.

Zarkesh, M., Daneshpour, M. S., Faam, B., Fallah, M. S., Hosseinzadeh, N., Guity, K., Hosseinpanah, F., Momenan, A. A., & Azizi, F. (2012). Heritability of the metabolic syndrome and its components in the Tehran Lipid and Glucose Study (TLGS). Genet Res (Camb), 94(6), 331-337.

ARTICLE 4

GENOTYPE BY ENERGY EXPENDITURE INTERACTION WITH BODY COMPOSITION TRAITS: THE PORTUGUESE HEALTHY FAMILY STUDY

Daniel Monteiro de Vilhena e Santos[1]

Peter T. Katzmarzyk[2]

Vincent P. Diego[3]

Thayse N. Gomes[1]

Fernanda K. Santos[1]

John Blangero[3]

José António Ribeiro Maia[1]

Under review in Biomed Research International

[1] CIFI[2]D, Faculty of Sports, University of Porto, Portugal

[2] Pennington Biomedical Research Center, Louisiana State University System, USA

[3] Texas Biomedical Research Institute, San Antonio, TX, USA

Abstract:

Objective: Levels of total daily energy expenditure (TDEE) have been negatively correlated with fat accumulation. However, evidence suggests that this association is highly variable. In the present study we applied a Genotype x Environment interaction method to examine the presence of GxTDEE interaction in the phenotypic expression of different body composition (BC) traits, namely body mass index (BMI), percent fat mass (%FAT), percent trunk fat mass (%TFAT), and waist circumference (WC).

Subjects: A total of 958 subjects from 294 families of The Portuguese Healthy Family Study were included in the analysis. TDEE was assessed using a 3-d physical activity recall. Body fat percentages were measured with a bioelectrical impedance scale. GxTDEE examination was performed using SOLAR 4.0 software.

Results: All BC traits were significantly heritable, with heritabilities ranging from 21% to 34%. The GxTDEE interaction model fitted the data better than the polygenic model for all traits. For all traits, a significant GxTDEE interaction was due to the rejection of the variance homogeneity hypothesis. For WC, GxTDEE was also significant by the rejection of the genetic correlation hypothesis.

Conclusion: These results suggest that fat accumulation is significantly influenced by the interaction between TDEE and genotype. TDEE is an environmental constraint that is related to the expression of individuals' BC genotypes, leading to variability in the phenotypic expression of BC traits.

Keywords: Energy Expenditure, Genotype X Environment, Interaction, Body Composition.

Introduction

The hypothesis that the development of many complex traits are the result of the interplay between genetic background and environmental influences has long been postulated (Haldane, 1946), and has been referred to as Genotype-by-Environment interaction (GxE) (Mackay, 2001). Under such a hypothesis it is expected that genetic effects are dynamically modulated by environmental exposures.

This concept has been used to study obesity for several decades (Bouchard, 2008), and there is a wealth of data confirming that environmental factors, whether related to nutritional habits and/or physical activity/exercise patterns, play key roles in the accumulation of body fat (Karnehed et al., 2006; Nelson et al., 2006). However, within a population sharing the same physical activity (PA) habits (in terms of levels and patterns), inter-individual variability in body composition is widely observed (King et al., 2008).

Genetic epidemiology research suggests that genetic factors accounts for 50% to 90% (Maes et al., 1997) of the total inter-individual variability in body fat accumulation. It remains, however, to be explained how environmental and behavioral factors, such as PA, affect the genetic influence on body composition. Twin-based studies have shown that genetic factors influence weight changes following different exercise patterns (Bouchard et al., 1990; Bouchard et al., 1994; MacDonald & Stunkard, 1990). For example, results from the Swedish Young Male Twins study (Karnehed et al., 2006) indicated that for those twins with genetic susceptibility for obesity, engaging in an active lifestyle had a preventive effect on accumulating fat. Accordingly, Mustelin et al. (2009) found an inverse additive genetic correlation between PA and BMI in both genders with correlations of -0.22 and -0.08 for females and males, respectively. More recently, an association study identified significant interactions between individual genes and self-reported PA, suggesting, for example, that the effect of the *FTO rs9939609* polymorphism on body fat accumulation is exacerbated by low levels of PA (Andreasen et al., 2008). Also, it has been shown that PA decreases the impact of *FTO* gene variants on

obesity (Kilpelainen et al., 2011). In a study with Danish and Finnish twin samples (Silventoinen et al., 2009), the results follow the same trend with an inverse association between PA and WC, BMI and % body fat as well as evidence that PA decreases both genetic and environmental variances of BMI and waist circumference. Moreover, using a GxE model, McCaffery et al. (2009) found that BMI is, on average, lower among those individuals that engage in vigorous activities, and that vigorous exercise significantly modified the additive genetic component of BMI, confirming the presence of an GxE interaction. Using an animal model, Noland et al. (2007) found that, even when exposed to a high fat diet, rats with inherited low oxidative capacity were heavier and hypertriglyceridemic when compared to high oxidative capacity rats. As such, it is highly likely that differences in PA patterns and levels may have different impacts on body composition changes within the same population. Accordingly, to better explain why some people become obese while others do not, it is important to understand how PA interacts with genotype and influences its association with body fat.

In the present study, using a nuclear family design, we bring together information on body composition and energy expenditure aiming (1) to estimate the magnitude of the genetic effects on body composition (BC) traits, and (2) to examine the Genotype x Total Daily Energy Expenditure (GxTDEE) interaction that may affect the impact of PA on BC traits. Our main hypothesis is that the genetic regulation of BC is affected by levels of PA.

Materials / subjects and Methods

The Portuguese Healthy Family Study, from the Portuguese *Estudo de Famílias Saudáveis Portuguesas* (FAMS), investigates the relationship among metabolic syndrome indicators, physical activity, physical fitness and body composition in nuclear Caucasian families. Children and adolescents aged 10 to 18 years were recruited in schools from the north and central regions of mainland Portugal, and were approached to freely participate in the study with their siblings and parents. Children with chronic diseases (such as asthma and diabetes), physical handicaps or psychological disorders that might impair their

daily routines and physical activities within schools and/or sports clubs were excluded. Given that families with 3 or more children are scarce in the Portuguese population (Rosa & Chitas, 2010), the sample comprises 294 families with only one or two siblings (see Table 1). The ethics committee of the Faculty of Sport, University of Porto, approved the study, and written informed consent, and assent, was obtained from all subjects.

PHYSICAL ACTIVITY

Using a 3-day physical activity diary (Bouchard et al., 1983), a trained technician interviewed each subject, recording the dominant activity for each 15-min period during 24 h by using a list of categorized activities. Categories from 1 to 9 refer to increasing levels of energy expenditure (METs) of each activity in which category 1 indicates very low energy expenditure such as sleeping or resting in bed, and category 9 refers to highly demanding physical work such as high-intensity sports. Approximate median energy cost for each of the nine categories in kcal/kg/15 min was used to compute the daily energy expenditure for each individual. The number of 15-min periods for each category was first summed over the 3-day period and weighted by its own median energy cost. Total energy expenditure (TEE) was then calculated by summing over the median energy cost of all nine categories and multiplying by subjects´ body weights. Total daily energy expenditure [TDEE (kcal/day)] was then calculated by dividing TEE by 3.

ANTHROPOMETRIC MEASUREMENTS

The standardized procedures of Lohman et al.(1988) were used to measure height with a Siber Hegner anthropometer (GMP instruments), and body composition was measured with a bioelectric impedance scale (TANITA BC-418 MA; Segmental Body Composition Analyser Tanita, Corporation, Tokyo, Japan). Two body composition traits were estimated - % body fat (%FAT) and % trunk fat (%TFAT). This impedance scale has been validated previously with Dual-energy X-ray Absorptiometry - DXA (Pietrobelli et al., 2004), a gold standard method for body composition measurement. Body mass index (BMI) was calculated as weight (kg)/height (m^2). Waist circumference was

measured at the end of a normal expiration just above the iliac crest, using a non-elastic Holtain tape.

STATISTICAL ANALYSIS

Univariate quantitative genetic procedures as implemented in SOLAR (Almasy & Blangero, 1998) under a special class of the multivariate linear model, namely the variance components (VC) approach, were used to estimate additive genetic and environmental VCs for each of the BC traits. Prior to all modeling, TDEE, age, age^2, sex, age-by-sex, and age^2-by-sex were used as covariates in a preliminary VC model. Residuals were thus derived for each trait and were normalized using an inverse normal transformation, as previously advocated (Blangero et al., 2013; Diego et al., 2007). Heritability estimates (h^2) were computed using a maximum likelihood approach to estimate variance components under the standard polygenic model as implemented in SOLAR v.4.3.1 software (Almasy & Blangero, 1998).

To test for GxTDEE interactions, basic initial hypotheses were formulated regarding the variance/covariance relationship of a BC indicator between family members with different levels of energy expenditure. With regards to GxTDEE interaction, the fundamental null hypothesis is that the expression of a polygenotype (i.e., aggregate of all genotypes related to the expression of a phenotype) is independent of TDEE levels. It can be shown from first principles that if there is no GxTDEE interaction, the same BC indicator measured in subjects with different levels of TDEE will have a genetic correlation of 1.0 and the genetic variance will be homogeneous across all levels of TDEE (Blangero, 2009; Diego et al., 2003). On the contrary, if GxTDEE interaction is present, the genetic correlation will be significantly less than 1.0 and/or the genetic variance will not be the same among all levels of TDEE.

The foregoing requires that we model the variance and correlation as functions of TDEE levels. For the genetic variance function (and similarly for the environmental variance function), we modeled the variance using an exponential function to ensure positivity, which is required since any variance is

a squared term (Blangero, 2009; Diego et al., 2003): $\sigma_g^2 = \exp\left[\alpha_g + \gamma_g\left(EE\right)\right]$, where α_g and γ_g are parameters to be estimated. An additional justification for the exponential function is suggested by the alternative name of this approach, namely the log-linear model of the variance: $\ln\sigma_g^2 = \alpha_g + \gamma_g\left(EE\right)$. That is, on taking the natural logarithm of the variance modeled as an exponential function, we have the equation of a straight line. In this form, the variance homogeneity null hypothesis obviously holds for a slope-term equal to 0: $\gamma_g = 0$. For the genetic correlation function, we modeled the genetic correlation as an exponential decay function of the pairwise differences in TDEE levels: $\rho_g = \exp\left(-\lambda\left|EE_i - EE_j\right|\right)$, where λ is a parameter to be estimated as a function of the difference in TDEE levels between any two individuals *i* and *j*. Here we also have an elegant reexpression of the interaction null hypothesis, in this case regarding the genetic correlation, in that a genetic correlation equal to 1 is equivalent to $\lambda = 0$. This is because for $\lambda = 0$, we have $\rho_g = \exp\left(-\lambda\left|EE_i - EE_j\right|\right) = e^0 = 1$. At the same time we employed a similar variance function for the residual environment variance but not an environmental correlation function due to the assumption of uncorrelated genetic and environmental effects. This allows us to guard against bias.

For reasons detailed in Diego et al. (2003), the likelihood ratio test statistics (LRTs) to test $\gamma_g = 0$ and $\lambda = 0$ are respectively distributed as χ_1^2, a chi-square random variable with 1 degree of freedom (d.f.), and $\left(\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2\right)$, a 50:50 mixture of chi-square random variable with a point-mass at 0, denoted by χ_0^2, and a chi-square with 1 d.f. Prior to examination of these hypotheses, however, we first confirmed if the overall GxTDEE interaction model provided a better fit to the data than the standard so-called polygenic model. The LRT for this comparison can be shown to be distributed as $\left(\frac{1}{2}\chi_2^2 + \frac{1}{2}\chi_3^2\right)$ (Diego, 2005).

Results

The basic descriptive data for TDEE and BC traits in fathers, mothers, sons, and daughters are presented in Table 1. Information from 294 families comprising 180 fathers, 253 mothers, 265 sons and 260 daughters was included. The average family size was 3.3 subjects. As expected, %FAT was higher in females than in males. Fathers and sons had consistently lower levels of %FAT than mothers and daughters, respectively. Sons and daughters´ average BMI were very similar.

Table I. Sample descriptive characteristics (means ± standard deviations).

	Fathers (n=180)	Mothers (n=253)	Sons (n=265)	Daughters (n=260)
Age (yrs)	45.4 ± 5.2	43.5 ± 4.5	14.7 ± 2.8	14.4 ± 2.8
Height (cm)	170.0 ± 6.7	158.6 ± 5.7	162.2 ± 12.9	156.3 ± 9.7
Weight (kg)	80.1 ± 13.2	66.9 ± 10.2	58.0 ± 16.2	53.6 ± 12.7
TDEE (kcal/day)	3561.8 ± 962.7	2788.4 ± 527.6	2280.6 ± 774.4	2024.9 ± 568.4
BMI (kg/m^2)	27.7 ± 4.1	26.6 ± 3.9	21.5 ± 4.1	21.7 ± 3.9
%FAT	23.0 ± 5.7	33.7 ± 5.9	20.0 ± 6.5	27.8 ± 6.2
% TFAT	24.6 ± 6.5	29.9 ± 7.0	16.9 ± 6.8	22.5 ± 7.6
WC (cm)	92.3 ± 10.6	81.0 ± 9.0	72.8 ± 10.4	68.4 ± 8.6

Legend: TDEE – total daily energy expenditure; BMI – body mass index; %FAT – fat percentage; %TFAT – trunk fat percentage; WC – waist circumference.

Heritability estimates (h^2) presented in table 2 were all highly significant ($p<0.001$), ranging from 0.21 ($CI_{95\%}$: 0.14,0.37) for %TFAT, to 0.34 ($CI_{95\%}$: 0.22,0.45) for WC meaning that the phenotypic expression of BC traits is in part due to moderate-to-strong additive genetic factors, which is a compelling argument to pursue further specific analysis of their genetic architecture.

Table II. Heritability estimates (h^2), standard errors, and corresponding 95% confidence intervals ($CI_{95\%}$) of the different phenotypes in the Portuguese Healthy Families Study.

Trait	h^2 ($CI_{95\%}$)	Std. Error	p-value
BMI	0.25 (0.14,0.37)	0.07	<0.001
%FAT	0.25 (0.14,0.37)	0.07	<0.001
%TFAT	0.21 (0.10,0.32)	0.07	<0.001
WC	0.34 (0.22,0.45)	0.07	<0.001

Legend: BMI – body mass index; %FAT – fat percentage; %TFAT – trunk fat percentage; WC – waist circumference.

The polygenic model was compared to the GxTDEE interaction model by means of a log-likelihood ratio test (see Table 3). The GxTDEE interaction model is significantly better than the polygenic model for all the BC traits implying that the GxTDEE model fits the data better than the polygenic model for each of these four traits. This means that inter-individual variability in the phenotypic expression of these body composition traits is to some degree explained by an interaction between genotype and total daily energy expenditure. As such, different genotype architectures lead to distinct expressions of body composition under the same energy expenditure levels. In table 4 we present the parameter estimates relevant to interpreting GxE interaction, namely the gamma and lambda parameters.

Table III. Results of log-likelihood ratio tests (LRT) and respective p-values contrasting a polygenic model vs a GxTDEE model for each of the body composition traits.

Trait	Polygenic LnL	GxTDEE LnL	LRT	p-value
BMI	-387.781	-338.660	98.243	<0.0001
% FM	-386.643	-370.572	32.144	<0.0001
% TFM	-387.543	-380.396	14.294	0.002
WC	-380.061	-319.731	120.660	<0.0001

Legend: BMI – body mass index; % FM – fat mass percentage; % TFM – trunk fat mass percentage; WC – waist circumference; LnL – log-likelihoods; LRT – Likelihood ratio test.

Table IV. Lambda and Gamma parameter estimates for each of the body composition traits.

Trait	Lambda	Gamma	Lambda LRT	Gamma LRT
BMI	0.0008 (0.0004,0.0017)	-	6.243	-
% FM	0.0014 (0.0008,0.0026)	-	11.597	-
% TFM	0.0012 (0.0006,0.0026)	-	7.031	-
WC	0.00009 (0.0005,0.0014)	0.0006 (0.0004,0.0007)	11.909	1.108

Legend: BMI – body mass index; % FM – fat mass percentage; % TFM – trunk fat mass percentage; WC – waist circumference; LRT – Likelihood ratio test.

Verification of GxTDEE interactions were made by comparing the full model to its constrained alternatives for BMI, % FAT, % TFAT, and WC.

The significant results for variance heterogeneity and genetic correlation are shown in Figure 1. For WC, the null hypotheses of a genetic correlation (ρ_g) equal to 1 and of genetic variance (σ^2_g) homogeneity are rejected. For the

remaining traits, a GxTDEE interaction was also observed but only by rejection of the genetic correlation hypothesis (Figure 1a). This means that despite variance homogeneity for BMI, %FAT, %TFAT, a significant interaction with TDEE was still present because the genetic correlation of these traits under distinct TDEE levels was not equal to 1. For example, if the genetic correlation between BMI under TDEE of 2500 kcal/day and a TDEE of 1500 kcal/day is 0.6, then we may speculate that if the TDEE environments differ then different genes are being activated and are being responsible for body composition expression. The null hypothesis of homogeneity in the genetic variance implies a straight line graph (i.e. slope equal to 0) at the level of the natural logarithm of the heritability given that the variances are modeled as exponential functions. Thus, Figure 1a shows that the genetic variance does vary as a function of the energy expenditure environment. Specifically, the genetic variance increases with increasing levels of energy expenditure, which means that the higher the TDEE values, the greater the differences in the set of genes activated that are responsible for WC expression. As for Figure 1b, the null hypothesis of a genetic correlation equal to 1 is graphically depicted by the horizontal line where the genetic correlation function equals 1. This means that under the null hypothesis the genetic correlation is not to be regarded as a function of differences in the environmental measure. Exponential curves that decay away from the null value simply indicate that the genetic correlation is in these cases a function of differences in the environmental measure.

Figure 1. Genotype X Energy Expenditure genetic variance (a) and correlation (b)

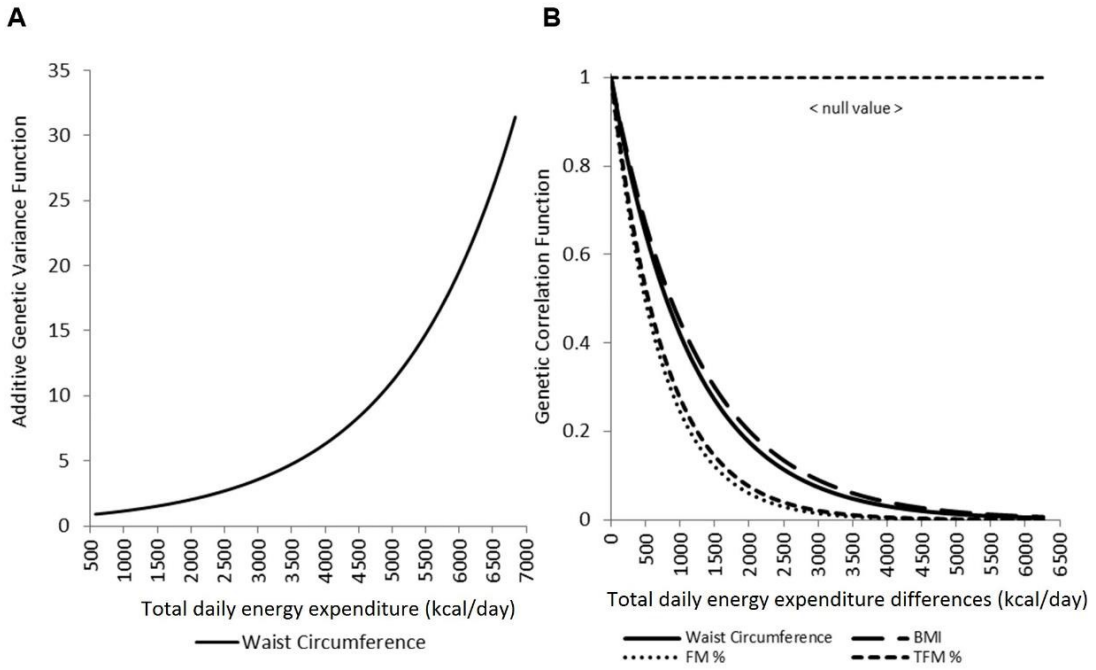
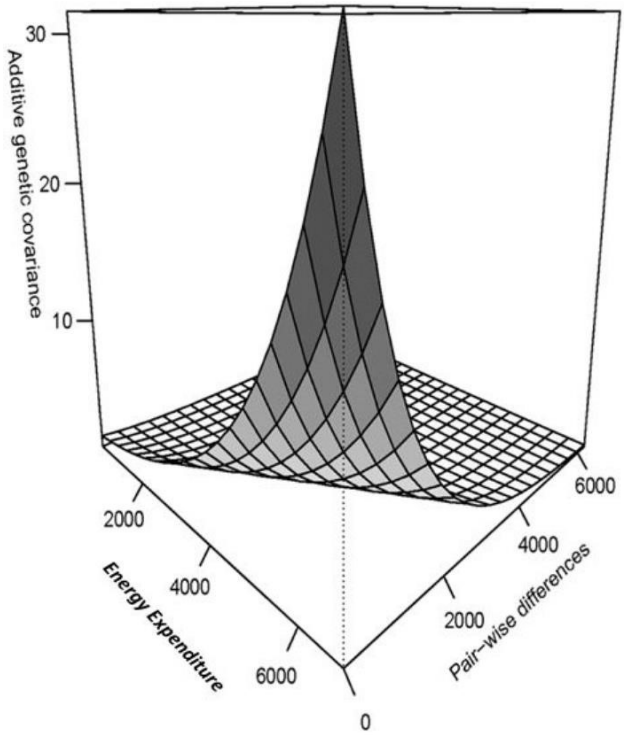


Figure 2 shows the simultaneous representation of the variance and correlation functions for WC, demonstrating that GxTDEE interaction for WC is a joint function of genetic variance heterogeneity and a genetic correlation different than one. In the figure, pair-wise differences refer to the differences between subjects in their TDEE levels.

Figure 2. Genetic covariance function for waist circumference.



Discussion

This study, based on a Portuguese sample of families, aimed to quantify the genetic variance of different BC traits as well as to examine the GxTDEE interaction in modulating the manifestation of these traits in family members. Our results confirm not only the importance of genetic factors in governing the expression of these BC traits, with all h^2 being significant, but most importantly showed the importance of GxTDEE interaction in fat accumulation. To the extent of our knowledge this is the first effort to apply a GxE interaction analysis, using a nuclear family-design study, to test the hypothesis that individual differences in phenotypic expression of BC traits are conditioned by their TDEE levels, i.e., inter-individual variability in different body composition traits is genetically driven and mediated by physical activity exposure.

Body composition heritability estimates reported here were all statistically significant which is in agreement with previous results (Bellia et al., 2009; Butte et al., 2006; Jelenkovic et al., 2011; Luke et al., 2001; Mathias et al., 2009; Poveda et al., 2012). Waist circumference was the most heritable of the four traits (h^2 = 0.34), and its value is comparable to the estimates of 0.38 found in the Linosa Study (Bellia et al., 2009) and 0.39 found in a study with 533 nuclear families from Spain (Jelenkovic et al., 2011). The heritability of BMI (h^2=0.25) is lower than those from Spain [h^2=0.44 (Jelenkovic et al., 2011)]. The same tendency was observed for body fat with our moderate heritability estimate of 0.25 contrasting with 0.69 in a Swedish sample (Wagner et al., 2013), 0.48 in Nigeria, 0.54 in Jamaica, 0.57 in USA (Luke et al., 2001), and 0.64 for males and 0.56 for females in a Chinese sample (Liu et al., 2012). These discrepancies are usually attributable to different sampling strategies and sample sizes, distinct statistical approaches used to estimate h^2, and use of distinct adjustments (different covariates). For instance, in our study, all of the h^2 estimations were controlled for the effect of TDEE which might explain this discrepancy of results. In summary, this wealth of data merely affirms the well-known dictum that heritability estimates are sample specific. Although our h^2 estimates are somewhat lower than the ones previously reported, we still have

from 1/5 to 1/3 of residual variance of BC traits explained by genetic factors, which is a compelling argument to further examine the underlying genetic architecture.

Over the years, researchers have been keen on studying the associations of different environmental exposures with BC (Perusse & Bouchard, 1999). This has mostly been done using a correlational approach between traits among family members (Liu et al., 2012; Mathias et al., 2009; Rose et al., 1998). Despite its usefulness in quantifying the degree and sign of the association between distinct BC phenotypes, correlations provide little information regarding the putative mechanisms that underlie such associations. GxE interaction analysis holds the promise of verifying if the association between an environmental factor (e.g., TDEE) and body fat accumulation is genetically driven, which may be of importance in understanding why people respond differently to physical exercise intervention programs (Bouchard & Rankinen, 2001).

In the present report, all BC traits were significantly influenced by a GxTDEE interaction through the rejection of the hypothesis of the genetic correlation being equal to 1 or/and the hypothesis of variance homogeneity. This means that the genotype effects are not exactly the same under different energy expenditure conditions, as they are not fully correlated between distinct TDEE environments. In fact, our results show that the greater the differences in the TDEE levels, the lower the genetic correlations, indicating that the genes influencing body composition traits differ under different TDEE levels. Similar approaches have been taken using DNA analysis (Li et al., 2010; Vimaleswaran et al., 2009). For example, Li et al. (2010) genotyped 12 SNPs in obesity-susceptibility loci of 20,430 individuals from the EPIC-Norfolk cohort, and reported that each additional BMI-increasing allele significantly increased the risk of obesity in the whole population, but significantly ($p_{interaction}$ = 0.015) more in inactive individuals [OR=1.158 ($CI_{95\%}$=1.118–1.199)] than in active individuals [OR=1.095 ($CI_{95\%}$=1.068–1.123)]. However, in the active group this increase was only 379 g, leading to the conclusion that being active may reduce the genetic predisposition to obesity by 40%. Also, the *FTO* gene was found, when

comparing active to nonactive individuals, to have a diminished influence on BMI (0.25 BMI increase per risk allele in active individuals vs 0.44 BMI increase per risk allele in nonactive individuals) and WC (0.64 cm increase per risk allele in active individuals vs 1.04 cm increase per risk allele in nonactive individuals) (Vimaleswaran et al., 2009). More recently, in a robust meta-analysis of 218,166 adults and 19,268 children the results showed that the association between *FTO* and obesity is diminished by 27% because of the effect of PA (Kilpelainen et al., 2011). In our sample, the genetic variance for WC increased with increasing levels of energy expenditure, which is not in line with previous results. This discrepancy may be related to the use of TDEE as a marker of PA, but still poses an argument for the necessity of continuing efforts to unravel the effects of PA at a genetic level that might influence different BC traits. For instance, the research by Lappalainen et al. (2009) also failed to find an association between exercise and the effect of *FTO* gene on weight changes, in a 4-year follow-up of 522 overweight or obese subjects, randomized to control and lifestyle intervention groups. We might speculate based on our results that WC changes' are influenced by a set of genes that is only "triggered" if an individual is exposed to high levels of energy expenditure.

This issue is highly challenging and important considering that in many countries researchers and policy makers are trying to deal with the obesity epidemic and associated morbidities not only from an health standpoint, but also from a financial view given the public burden in costs of obesity related morbidities (Bahia et al., 2012; Cawley & Meyerhoefer, 2012). This epidemic has been mostly connected to a fast changing environment (referred to as "obesogenic") characterized by inducing low levels of energy expenditure and persuasive ways of increasing caloric intake - that together constitute a difficult challenge to our genome (Li et al., 2010; Vimaleswaran et al., 2009) , but our results highlight that genetic adaptability to energy expenditure environments is probably more important than the environment itself. This has been proven previously in a highly cited experimental study with MZ twins (Bouchard et al., 1990) in which the variance in response to an overfeeding program of 100 days was three times greater between-pairs than within-pairs for BC traits. The same

trend was observed when MZ twins were subjected to an exercise protocol over a 93-day period. Once again, and under controlled nutrient intake, the differences in weight loss were more pronounced between-pairs than within-pairs (Bouchard et al., 1994). Both of these studies substantiate that the more genetically similar individuals are, the more similar they react to the same environment.

We think that our results add to the efforts in trying to disentangle these matters and help to substantiate the latter arguments by suggesting that the phenotypic expression of BC traits is the result of joint effects of genes, TDEE levels (environment) and their interactions.

Despite the relevance of the present results, some limitations should be acknowledged. Firstly, the sample used may not be representative of the general Portuguese population. Secondly the method chosen to estimate BC traits, in our case bioelectrical impedance analysis, even though having been previously validated with DXA (Pietrobelli et al., 2004), is not free from bias in its results although the precision of the equipment is ±1%. Nevertheless, this method has been widely used as a BC analyzer in many studies (Cummings et al., 2012; Du et al., 2013; Pausova et al., 2009). We also feel that the joint effects of the size of our sample, the use of state of the art statistical procedures, and the novelty of the analysis in PA genetic epidemiology research are strengths of the present study that warrant consideration.

Conclusions

In conclusion, the present results showed that the genetic expression of BC traits is significantly influenced by energy expenditure levels. Accordingly, physical activity may be considered an environmental variable that promotes inter-individual differences in BC traits through genetic mediation. This is valuable information for health practitioners. More efforts should be devoted to not only identify specific loci that control different BC traits but also to test if these loci are regulated or not by different PA levels.

Acknowledgments

This work was supported by the FCT – Fundação para a Ciência e a Tecnologia – (PTDC/DES/67569/2006 FCOMP-01-0124-FEDEB-09608 and SFRH/BD/65290/2009).

We thank the participants in The Healthy Family Study.

Conflict of interest

The authors declare no conflict of interest.

References

Almasy, L., & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet, 62*(5), 1198-1211.

Andreasen, C. H., Stender-Petersen, K. L., Mogensen, M. S., Torekov, S. S., Wegner, L., Andersen, G., Nielsen, A. L., Albrechtsen, A., Borch-Johnsen, K., Rasmussen, S. S., Clausen, J. O., Sandbaek, A., Lauritzen, T., Hansen, L., Jorgensen, T., Pedersen, O., & Hansen, T. (2008). Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes, 57*(1), 95-101.

Bahia, L., Coutinho, E. S., Barufaldi, L. A., Abreu Gde, A., Malhao, T. A., de Souza, C. P., & Araujo, D. V. (2012). The costs of overweight and obesity-related diseases in the Brazilian public health system: cross-sectional study. *BMC Public Health, 12*, 440.

Bellia, A., Giardina, E., Lauro, D., Tesauro, M., Di Fede, G., Cusumano, G., Federici, M., Rini, G. B., Novelli, G., Lauro, R., & Sbraccia, P. (2009). "The Linosa Study": epidemiological and heritability data of the metabolic syndrome in a Caucasian genetic isolate. *Nutr Metab Cardiovasc Dis, 19*(7), 455-461.

Blangero, J. (2009). Statistical genetic approaches to human adaptability. 1993. *Hum Biol, 81*(5-6), 523-546.

Blangero, J., Diego, V. P., Dyer, T. D., Almeida, M., Peralta, J., Kent, J. W., Jr., Williams, J. T., Almasy, L., & Goring, H. H. (2013). A kernel of truth: statistical advances in polygenic variance component models for complex human pedigrees. *Adv Genet, 81*, 1-31.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Theriault, G. (1983). A method to assess energy expenditure in children and adults. Am J Clin Nutr, 37(3), 461-467.

Bouchard, C. (2008). Gene-environment interactions in the etiology of obesity: defining the fundamentals. *Obesity (Silver Spring), 16 Suppl 3*, S5-S10.

Bouchard, C., & Rankinen, T. (2001). Individual differences in response to regular physical activity. *Med Sci Sports Exerc, 33*(6 Suppl), S446-451; discussion S452-443.

Bouchard, C., Tremblay, A., Despres, J. P., Nadeau, A., Lupien, P. J., Theriault, G., Dussault, J., Moorjani, S., Pinault, S., & Fournier, G. (1990). The response to long-term overfeeding in identical twins. *N Engl J Med, 322*(21), 1477-1482.

Bouchard, C., Tremblay, A., Despres, J. P., Theriault, G., Nadeau, A., Lupien, P. J., Moorjani, S., Prudhomme, D., & Fournier, G. (1994). The response to exercise with constant energy intake in identical twins. *Obes Res, 2*(5), 400-410.

Butte, N. F., Cai, G., Cole, S. A., & Comuzzie, A. G. (2006). Viva la Familia Study: genetic and environmental contributions to childhood obesity and its comorbidities in the Hispanic population. *Am J Clin Nutr, 84*(3), 646-654; quiz 673-644.

Cawley, J., & Meyerhoefer, C. (2012). The medical care costs of obesity: an instrumental variables approach. *J Health Econ, 31*(1), 219-230.

Cummings, N., Shields, K. A., Curran, J. E., Bozaoglu, K., Trevaskis, J., Gluschenko, K., Cai, G., Comuzzie, A. G., Dyer, T. D., Walder, K. R., Zimmet, P., Collier, G. R., Blangero, J., & Jowett, J. B. (2012). Genetic variation in SH3-domain GRB2-like (endophilin)-interacting protein 1 has a major impact on fat mass. *Int J Obes (Lond), 36*(2), 201-206.

Diego, V. P. (2005). *GENOTYPE × AGE INTERACTION, AND THE INSULIN-LIKE GROWTH FACTOR I AXIS IN THE SAN ANTONIO FAMILY HEART STUDY: A STUDY IN HUMAN SENESCENCE*. New York: Graduate School of Binghamton University. Relatorio de Estagio apresentado a

Diego, V. P., Almasy, L., Dyer, T. D., Soler, J. M., & Blangero, J. (2003). Strategy and model building in the fourth dimension: a null model for genotype x age interaction as a Gaussian stationary stochastic process. *BMC Genet, 4 Suppl 1*, S34.

Diego, V. P., Rainwater, D. L., Wang, X. L., Cole, S. A., Curran, J. E., Johnson, M. P., Jowett, J. B., Dyer, T. D., Williams, J. T., Moses, E. K., Comuzzie, A. G., Maccluer, J. W., Mahaney, M. C., & Blangero, J. (2007). Genotype x adiposity interaction linkage analyses reveal a locus on chromosome 1 for lipoprotein-associated phospholipase A2, a marker of inflammation and oxidative stress. *Am J Hum Genet, 80*(1), 168-177.

Du, H., Bennett, D., Li, L., Whitlock, G., Guo, Y., Collins, R., Chen, J., Bian, Z., Hong, L. S., Feng, S., Chen, X., Chen, L., Zhou, R., Mao, E., Peto, R., & Chen, Z. (2013). Physical activity and sedentary leisure time and their associations with BMI, waist circumference, and percentage body fat in 0.5 million adults: the China Kadoorie Biobank study. *Am J Clin Nutr, 97*(3), 487-496.

Haldane, J. B. (1946). The interaction of nature and nurture. *Ann Eugen, 13*(3), 197-205.

Jelenkovic, A., Poveda, A., & Rebato, E. (2011). Quantitative genetics of human morphology and obesity-related phenotypes in nuclear families from the Greater Bilbao (Spain): comparison with other populations. *Ann Hum Biol, 38*(4), 471-478.

Karnehed, N., Tynelius, P., Heitmann, B. L., & Rasmussen, F. (2006). Physical activity, diet and gene-environment interactions in relation to body mass index and waist circumference: the Swedish young male twins study. *Public Health Nutr, 9*(7), 851-858.

Kilpelainen, T. O., Qi, L., Brage, S., Sharp, S. J., Sonestedt, E., Demerath, E., Ahmad, T., Mora, S., Kaakinen, M., Sandholt, C. H., Holzapfel, C., Autenrieth, C. S., Hypponen, E., Cauchi, S., He, M., Kutalik, Z., Kumari, M., Stancakova, A., Meidtner, K., Balkau, B., Tan, J. T., Mangino, M., Timpson, N. J., Song, Y., Zillikens, M. C., Jablonski, K. A., Garcia, M. E., Johansson, S., Bragg-Gresham, J. L., Wu, Y., van Vliet-Ostaptchouk, J. V., Onland-Moret, N. C., Zimmermann, E., Rivera, N. V., Tanaka, T., Stringham, H. M., Silbernagel, G., Kanoni, S., Feitosa, M. F., Snitker, S., Ruiz, J. R., Metter, J., Larrad, M. T., Atalay, M., Hakanen, M., Amin, N., Cavalcanti-Proenca, C., Grontved, A., Hallmans, G., Jansson, J. O., Kuusisto, J., Kahonen, M., Lutsey, P. L., Nolan, J. J., Palla, L.,

Pedersen, O., Perusse, L., Renstrom, F., Scott, R. A., Shungin, D., Sovio, U., Tammelin, T. H., Ronnemaa, T., Lakka, T. A., Uusitupa, M., Rios, M. S., Ferrucci, L., Bouchard, C., Meirhaeghe, A., Fu, M., Walker, M., Borecki, I. B., Dedoussis, G. V., Fritsche, A., Ohlsson, C., Boehnke, M., Bandinelli, S., van Duijn, C. M., Ebrahim, S., Lawlor, D. A., Gudnason, V., Harris, T. B., Sorensen, T. I., Mohlke, K. L., Hofman, A., Uitterlinden, A. G., Tuomilehto, J., Lehtimaki, T., Raitakari, O., Isomaa, B., Njolstad, P. R., Florez, J. C., Liu, S., Ness, A., Spector, T. D., Tai, E. S., Froguel, P., Boeing, H., Laakso, M., Marmot, M., Bergmann, S., Power, C., Khaw, K. T., Chasman, D., Ridker, P., Hansen, T., Monda, K. L., Illig, T., Jarvelin, M. R., Wareham, N. J., Hu, F. B., Groop, L. C., Orho-Melander, M., Ekelund, U., Franks, P. W., & Loos, R. J. (2011). Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med, 8*(11), e1001116.

King, N. A., Hopkins, M., Caudwell, P., Stubbs, R. J., & Blundell, J. E. (2008). Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise-induced weight loss. *Int J Obes (Lond), 32*(1), 177-184.

Lappalainen, T. J., Tolppanen, A. M., Kolehmainen, M., Schwab, U., Lindstrom, J., Tuomilehto, J., Pulkkinen, L., Eriksson, J. G., Laakso, M., Gylling, H., & Uusitupa, M. (2009). The common variant in the FTO gene did not modify the effect of lifestyle changes on body weight: the Finnish Diabetes Prevention Study. *Obesity (Silver Spring), 17*(4), 832-836.

Li, S., Zhao, J. H., Luan, J., Ekelund, U., Luben, R. N., Khaw, K. T., Wareham, N. J., & Loos, R. J. (2010). Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Med, 7*(8).

Liu, R., Liu, X., Arguelles, L. M., Patwari, P. P., Zee, P. C., Chervin, R. D., Ouyang, F., Christoffel, K. K., Zhang, S., Hong, X., Wang, G., Xu, X., & Wang, X. (2012). A population-based twin study on sleep duration and body composition. *Obesity (Silver Spring), 20*(1), 192-199.

Lohman, T. G., Roche, A. F., & Martorel, R. (1988). *Anthropometric Standardization Reference Manual*. Illinois: Human Kinetics Books Champaign.

Luke, A., Guo, X., Adeyemo, A. A., Wilks, R., Forrester, T., Lowe, W., Jr., Comuzzie, A. G., Martin, L. J., Zhu, X., Rotimi, C. N., & Cooper, R. S. (2001). Heritability of obesity-related traits among Nigerians, Jamaicans and US black people. *Int J Obes Relat Metab Disord, 25*(7), 1034-1041.

MacDonald, A., & Stunkard, A. (1990). Body-mass indexes of British separated twins. *N Engl J Med, 322*(21), 1530.

Mackay, T. F. (2001). The genetic architecture of quantitative traits. *Annu Rev Genet, 35*, 303-339.

Maes, H. H., Neale, M. C., & Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet, 27*(4), 325-351.

Mathias, R. A., Deepa, M., Deepa, R., Wilson, A. F., & Mohan, V. (2009). Heritability of quantitative traits associated with type 2 diabetes mellitus in large multiplex families from South India. *Metabolism, 58*(10), 1439-1445.

McCaffery, J. M., Papandonatos, G. D., Bond, D. S., Lyons, M. J., & Wing, R. R. (2009). Gene X environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *Am J Clin Nutr, 89*(4), 1011-1018.

Mustelin, L., Silventoinen, K., Pietilainen, K., Rissanen, A., & Kaprio, J. (2009). Physical activity reduces the influence of genetic effects on BMI and waist circumference: a study in young adult twins. *Int J Obes (Lond), 33*(1), 29-36.

Nelson, M. C., Gordon-Larsen, P., North, K. E., & Adair, L. S. (2006). Body mass index gain, fast food, and physical activity: effects of shared environments over time. *Obesity (Silver Spring), 14*(4), 701-709.

Noland, R. C., Thyfault, J. P., Henes, S. T., Whitfield, B. R., Woodlief, T. L., Evans, J. R., Lust, J. A., Britton, S. L., Koch, L. G., Dudek, R. W., Dohm, G. L., Cortright, R. N., & Lust, R. M. (2007). Artificial selection for high-capacity endurance running is protective against high-fat diet-induced insulin resistance. *Am J Physiol Endocrinol Metab, 293*(1), E31-41.

Pausova, Z., Syme, C., Abrahamowicz, M., Xiao, Y., Leonard, G. T., Perron, M., Richer, L., Veillette, S., Smith, G. D., Seda, O., Tremblay, J., Hamet, P., Gaudet, D., & Paus, T. (2009). A common variant of the FTO gene is associated with not only increased adiposity but also elevated blood pressure in French Canadians. *Circ Cardiovasc Genet, 2*(3), 260-269.

Perusse, L., & Bouchard, C. (1999). Genotype-environment interaction in human obesity. *Nutr Rev, 57*(5 Pt 2), S31-37; discussion S37-38.

Pietrobelli, A., Rubiano, F., St-Onge, M. P., & Heymsfield, S. B. (2004). New bioimpedance analysis system: improved phenotyping with whole-body analysis. *Eur J Clin Nutr, 58*(11), 1479-1484.

Poveda, A., Jelenkovic, A., Salces, I., Ibanez, M. E., & Rebato, E. (2012). Heritability variations of body linearity and obesity indicators during growth. *Homo, 63*(4), 301-310.

Rosa, M. J. V., & Chitas, P. (2010). *Portugal: os números*. Lisboa.

Rose, K. M., Newman, B., Mayer-Davis, E. J., & Selby, J. V. (1998). Genetic and behavioral determinants of waist-hip ratio and waist circumference in women twins. *Obes Res, 6*(6), 383-392.

Silventoinen, K., Hasselbalch, A. L., Lallukka, T., Bogl, L., Pietilainen, K. H., Heitmann, B. L., Schousboe, K., Rissanen, A., Kyvik, K. O., Sorensen, T. I., & Kaprio, J. (2009). Modification effects of physical activity and protein intake on heritability of body size and composition. *Am J Clin Nutr, 90*(4), 1096-1103.

Vimaleswaran, K. S., Li, S., Zhao, J. H., Luan, J., Bingham, S. A., Khaw, K. T., Ekelund, U., Wareham, N. J., & Loos, R. J. (2009). Physical activity attenuates the body mass index-increasing influence of genetic variation in the FTO gene. *Am J Clin Nutr, 90*(2), 425-428.

Wagner, H., Melhus, H., Pedersen, N. L., & Michaelsson, K. (2013). Genetic influence on bone phenotypes and body composition: a Swedish twin study. *J Bone Miner Metab*.

CHAPTER 3

GENERAL OVERVIEW AND CONCLUSIONS

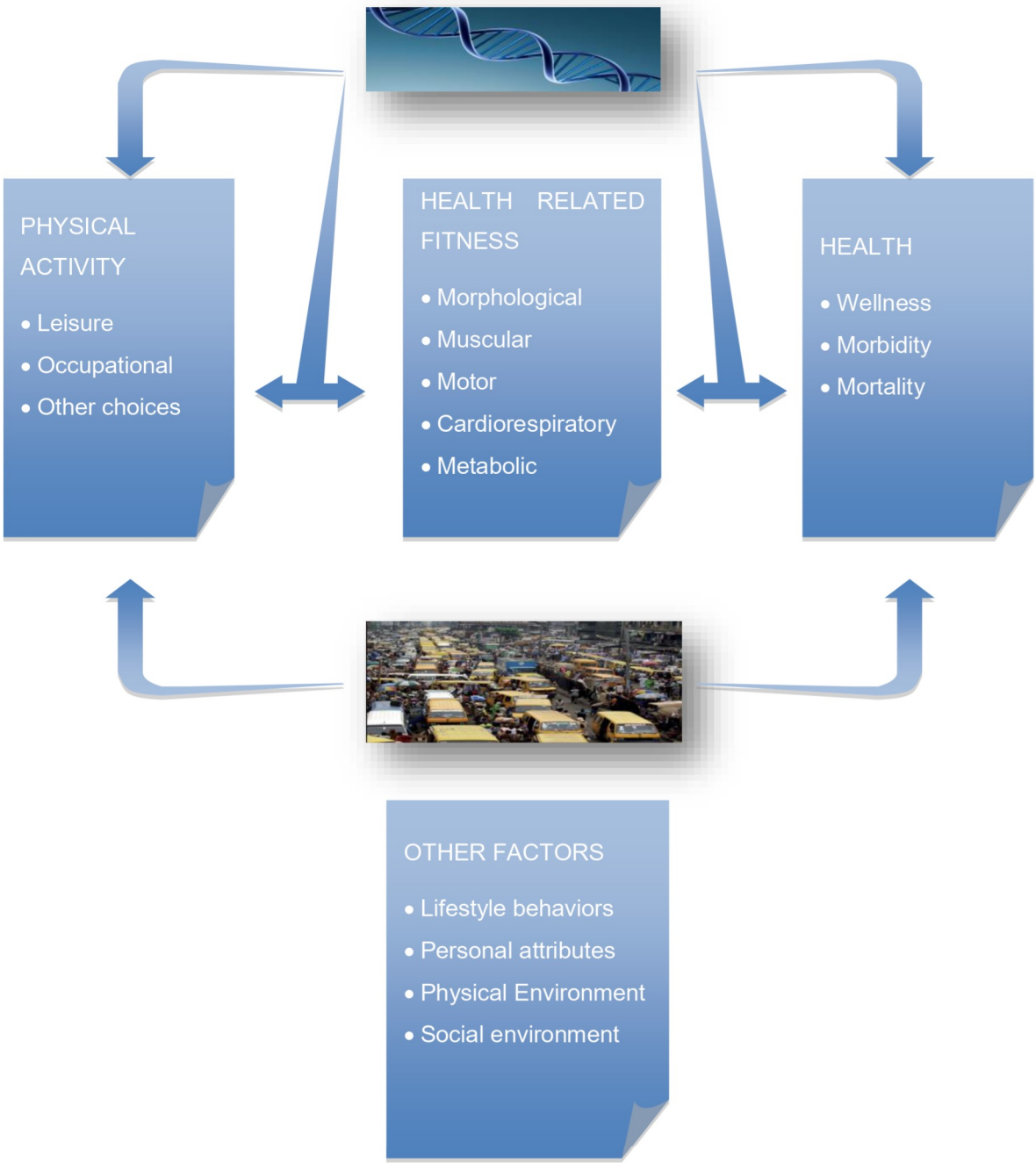
The major purpose of this thesis was to investigate the interactive roles of physical activity and metabolic syndrome and body composition variation in a sample of Portuguese nuclear families. This effort was presented in the previous chapters. Here we will address a general overview of the main findings and their implications, followed by the study limitations as well as opportunities for future research. Lastly, overall conclusions of this thesis will be presented.

GENERAL OVERVIEW

This research was based on the Portuguese Healthy Family Study, which uses a simultaneous cross-sectional and longitudinal design to investigate familial aggregation in cardiovascular risk factors, physical activity, body composition, nutritional behaviours and physical fitness. We began "*our journey*" by presenting an overall framing of the thesis and described the fundamental concepts that are at its core: (i) interaction, (ii) nuclear family, (iii) physical activity, (iv) metabolic syndrome and (iv) body composition (later expanded in the three empirical manuscripts and in the review paper), followed by a detailed description of the methodology that was used to address the research questions.

In order to present a clear and coherent path to this overview we will firstly deal with the four research questions of this research following Bouchard and Shephard´s model, as shown previously and repeated here (Figure 1).

Figure 1. A consensus model describing the interrelationships among physical activity, health-related fitness (HRPF), and health status.



The first issue to be addressed was the genetic "foundations" of physical activity (PA) and physical inactivity (PI) and the following research question was made: **are physical activity and physical inactivity genetically driven?**

We reviewed the available literature and the results were quite disperse as has been thoroughly discussed (but see paper nº 1). In summary, the available

knowledge on the genetics of PA showed that heritability estimates ranged from 0% to 85%, showing the substantial variability in the results so far. This dispersion extends to the influence of shared (0% to 84%) and unique environments (12% to 72%). This set of results based in families and twin studies supports the evidence that current knowledge is still inconclusive about the magnitude of genetic determination in PA.

More complex approaches have been applied in the past few years and suggestive linkages were found with markers nearby different activity-related genes: *EDNRB*, *MC4R*, *UCP1*, *FABP2*, *CASR*, *and SLC9A9*. However, it is not possible to contrast these results as none of these markers were studied more than once, i.e., have not been replicated. Also, significant associations with PA phenotypes were found for *ACE*, *Gln223ARrg*, *MC4R* and *DRD2* genes. In this case, however, it was possible to compare the results from the *ACE* gene and these were not concordant, as the Winnicki et al. (2004) study found a significant association with PA and the Fuentes et al. (2002) study was not able to. Moreover, *MC4R* (Loos et al., 2005) and *DR2D* (Simonen et al., 2003a) associated positively only with higher levels of PA and PA during the past year, respectively. The only GWAs analysis in the field of genetic epidemiology of PA found novel SNPs in the *PAPSS2* gene on chromosome *10q23.2* and in two intergenic regions on chromosomes *2q33.1* and *18p11.32* (De Moor et al., 2009). This set of results from linkage analysis, association studies and GWAs is not only a very valuable contribute, but also adds to the previous statement that regulatory mechanisms of PA are yet to be fully understood.

Furthermore, our review suggested that despite the lack of studies focusing on the genetic influences on PI, inactivity seems to be more genetically driven than PA. Briefly, heritability estimates for PI varied between 25% (Simonen et al., 2002) and 60% (Butte et al., 2006). Linkage studies recorded higher LOD scores in PI than in PA and *ACE* genotypes were strongly associated with PI (Cai et al., 2006). This set of results is a strong argument to further investigate the hypothesis of sedentary activities being independent of physical activity levels.

Reviewed papers made it clear, as postulated in Bouchard and Shephard´s model, that the PA "engine" is in part fuelled by individual genes. However, the degree to which genes fuel it and which genes are involved in its regulation are far from being extensively and clearly known. But the simple fact that information is available that confirms the genetic ruling of PA impels us to wonder whether the known positive effects of PA on MetS components and BC are related to genetic activation.

The second topic we addressed (in manuscripts 2, 3, and 4) was related with the genetic components of both the MetS and BC. The following research questions were asked: **(i) are metabolic syndrome components correlated within families, and are the metabolic syndrome components and body composition traits genetically governed?; and (ii) are metabolic syndrome components correlated with each other within families and between family members?**

In The Portuguese Healthy Family Study, familial similarity was found in all the MetS components (all $p<10^{-4}$), and correlation equality was found in SBP, GLU, TG and between each biological family dyad in WC. Using the same statistical approach, Trégouet et al. (1999) found that spouse correlations were always different from zero, meaning that there is a possibility of a shared environmental factor influencing MetS traits. In our study spouses' correlations were of the same magnitude as the biological relatives' correlations with the exception of HDL, which provides further support for the importance of shared environment. Further, results were in agreement with previous family studies that found significant heritability estimates for MetS (Bellia et al., 2009; Freeman et al., 2002) components and BC traits (Rebato et al., 2007).

Cross-trait correlations were always significant and negative for the HDL/TG pair. This is in agreement with previous findings (Povel et al., 2011) in which low HDL is a result of inefficient catabolism of TG rich lipoproteins and reduced transfer of surface components to nascent HDL particles (Miller et al., 2007). This implies the presence of a pleiotropic effect among MetS components as it has been proven by familial studies (Hong et al., 1997;

Mitchell et al., 1996) as well as by genome scans (Arya et al., 2004; de Andrade et al., 2002).

In The Portuguese Healthy Family Study waist circumference correlated significantly with all of the other traits at least in one dyad, meaning that body fat of a relative could be associated with other traits in another relative. In the Trégouet et al. (1999) study, BMI was not significantly correlated with triglycerides which means that environmental factors may play a key role in this association within a family, and be linked with previous evidence of family aggregation of nutritional habits and a sedentary lifestyle (Vauthier et al., 1996). It seems clear from these results that MetS is truly a cluster of nested traits that are dependent of each other, whatever the nature of dependency: genetic or environmental. This is in full agreement with our reference model that clearly illustrates the influence of environmental variables on the expression of health indicators.

The evidence for genetic influences on PA, MetS and BC constitute the foundation for the subsequent exploration of the relationship between PA and MetS and BC. So far, most prior research has focused on the physiological outcomes of exercise and how it triggers a chain of body reactions that are generally good to health (Gielen et al., 2010). For example, evidence shows that physical activity and inactivity have anti- and pro-inflammatory effects, respectively (Fenza & Fiorina, 2012; Hamer et al., 2012). Further, the positive effect of regular exercise on insulin and plasma lipids is well known (Duvivier et al., 2013).

Our questions, however, have not to do with physiological responses to physical activity but primarily deal with a “top-down” approach in which we infer about possible genetic “orders” that produces distinct physiological adaptations. That is, are genes sensitive to physical activity levels in such a way that they will act differently under distinct amounts or intensities of activities? The following research question was asked: **are the physical activity effects on**

metabolic syndrome components and body composition mediated by genetic factors?

To the best of our knowledge this is the first time that a genotype by energy expenditure model is applied to investigate the influence of PA on the expression of MetS and BC traits using a family study. Results showed that the GxEE interaction model fitted the data better than the polygenic model ($p<0.001$) for waist circumference, systolic blood pressure, glucose, total cholesterol, triglycerides, body mass index, percent fat mass and percent trunk fat mass. For waist circumference, glucose, total cholesterol, triglycerides, body mass index, percent fat mass and percent trunk fat mass the significant GxEE interaction was due to rejection of the variance homogeneity hypothesis. For waist circumference and glucose, GxEE was also significant by the rejection of the genetic correlation hypothesis. These results clearly show that an interaction occurs between genotype and physical activity that is responsible for some of the variation in the phenotypic expression of MetS and BC traits. We've shown that the greater differences in total daily energy expenditure, the greater variance in the genetic regulation of MetS and BC. This means that genes up-regulate the health enhancing qualities of PA, as regards MetS or BC. Once again, this body of results is in agreement with the model proposed by Bouchard and Shephard and help us to better understand the huge inter-variability in response physical activity (Bouchard & Rankinen, 2001), and why, for instance, some have difficulty in managing their body fat despite being physically active and why some do not have MetS regardless of having a sedentary life style[48].

Our results suggest that PA is to some extent genetically determined, and its expression may lead to modifications in the expression of genes responsible for BC and MetS.

This topic is highly complex and surpasses the boundaries of the present research. It seems clear that millions of years ago, human genes were selected to optimize aerobic metabolic pathways and conserve energy as an adaptation

[48] This statement serves just to illustrate the importance of our results. However, there are many other factors, namely nutritional habits that help to further explain these results.

to an environment with scarce food and ever-present dangers (Booth et al., 2008). Human activities were mostly of high intensity leading to adaptations in cardiac and vascular functions as a response to the metabolic demands of the working skeletal muscle under these conditions (Gielen et al., 2010).

Nowadays, physical activity levels are so much lower than they were supposed to be for our genetic background, that sedentary lifestyle in combination with high caloric nutritional habits has surpassed smoking as the major preventable cause of death in the United States (Mokdad et al., 2004). This is probably due to the absence of skeletal muscle work which influences the cardiac and vascular functions so that molecular cardiovascular effects of exercise changes in the present time are probably a return to normal values rather than an improvement in human biology (Gielen et al., 2010).

The mechanisms that are behind exercise-induced changes are not fully understood. For instance, when speaking about the beneficial effects of PA on cardiovascular health, there is some speculation about potential epigenetic changes driven by PA linked to a major epigenetic modification - DNA methylation - that suppresses gene expression by modulating the access of the transcription machinery to the chromatin or by recruiting methyl-binding proteins (Barres et al., 2012). Barres et al. (2009) have recently shown that global methylation values decrease after intense exercise, even after controlling for the effect of haemoglobin mRNA content. This means that captured methylated promoters for metabolic genes, previously linked to type 2 diabetes (Simonen et al., 2003a) were lower after acute exercise, leading the authors to suggest that acute exercise induces gene-specific DNA hypo-methylation in the human skeletal muscle (Barres et al., 2009).

As we've shown in our review paper, linkage, association and GWAs studies in PA have been leading to quite ambiguous results in terms of which genes are responsible for its expression at the population level. In addition, and due to its highly complex landscape, PA is a multifaceted behaviour that is undoubtedly under the influence of multiple genes that may influence each

other (epistasis) and we are yet to identify one single gene that is unquestionably associated with PA. All that we have are just candidate genes.

As for MetS and BC there is stronger evidence of genes that influence their expression (Chagnon et al., 2001; Hakanen et al., 2009). It has been shown that the *FTO* gene is associated with BMI (Hakanen et al., 2009) and *MC4R* and *MC5R* genes are related with obesity phenotypes (Chagnon et al., 2001). And, for example, blood pressure genetic analysis have identified 43 genetic variants associated with systolic, diastolic BP, and hypertension (Ehret & Caulfield, 2013) and a recent review (Boes et al., 2009) has identified 18 genes associated with HDL-C that together explain less than 10% of HDL-C variance, despite heritability estimates of near 80%, which means that there is a long road to travel until HDL-C genetic foundations are fully understood.

Despite these lapses in the current knowledge about the genetic regulation of PA, MetS and BC, our results substantiate what has been previously found. There is evidence that PA attenuates the effect of some of the obesity and MetS-related genes (Kilpelainen et al., 2011).

LIMITATIONS

Despite the relevance of the present results, various limitations should be acknowledged.

The first limitation is related to PA assessment. The methods used here are widely accepted in the field of physical activity epidemiology. As stated earlier (see Chapter 1), diaries are a good tool to assess physical activity levels and patterns and to estimate energy expenditure in “big” samples such as The Portuguese Healthy Family Study. However, the use of accelerometers and even pedometers would have added value to our efforts of more objective PA measuring, and thus provide more reliable information to the study of the association between physical activity and body composition or metabolic syndrome components. However, the cost-effectiveness of such methods makes it arguably difficult to implement. Furthermore, we would probably face a

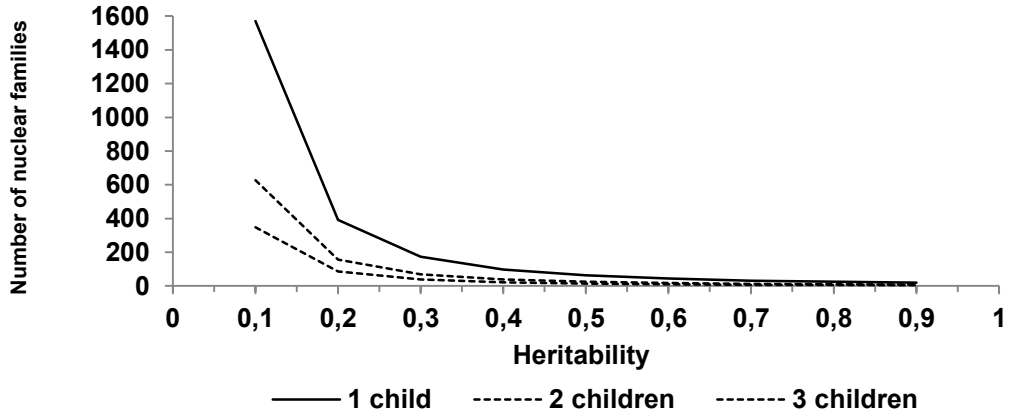
greater challenge than the financial constraint in asking parents to use accelerometers or pedometers during seven days. Even though the use of such devices does not limit daily routines, we wonder if parents would feel motivated enough to use them. Still, it is unquestionable that running a longitudinal study of objectively measured PA would have been a feat.

A point could be made about the phenotype used to analyse the relationship between PA and BC or MetS. As we've shown in the introduction, total daily energy expenditure is the sum of basal metabolic rate, thermic effect of food and PA. As such, it should not be confused with PA per se. The fact is that most of the research concerning the influence of PA in BC and MetS traits has focused on the positive impact of moderate-to-vigorous activities in preventing the hazards of MetS and excessive body fat (Katzmarzyk & Herman, 2007). However, we feel as though constraining our analysis to those few (or probably none) moments in contemporary daily routines in which people are involved in high intensity tasks, could be misleading in terms of what is the overall activity pattern of an individual during a day. Further, the 3-day activity diary used in the thesis allows for an adequate control over disparities between adults and youth routines, since each activity is independently assessed and weighted by its own metabolic value. This is something not possible with other questionnaires in which, for instance, parents' activities during their working hours are compared with offspring's' activities at school (Baecke et al., 1982). Also, and although PA and TDEE are distinct concepts, the reality is that the importance of PA in determining energy expenditure is only comparable to body weight's influence. This is determined by the simple fact that exercise increases basal metabolic rate (for instance, by promoting muscular protein synthesis) that leads to increasing levels of energy expenditure (Tsao et al., 2001). As such, the phenotype we used to investigate the interaction between energy expenditure and genotype that may influence the phenotypic expression of BC traits and MetS is securely associated with PA.

As regards sampling design, the use of nuclear families with two generations to estimate the genetic contribution to any given trait is somewhat limited. As described earlier (see Chapter 1), relatives can share, on average,

from 0 (non-biological relatives) up to 50% of their genes identical-by-descent (with the exception of monozygotic twins that share 100%), that when discussing the results needs to be accounted for. On the other hand, we used state-of-the-art statistical procedures to deal, specifically, with our problems whether GEE or variance components implemented in freely available software. Moreover, Schork and Schork (1993) performed a power analysis for detecting heritability under a range of nuclear family sizes. Based on the power curves presented here (Figure 2), we believe that we have adequate power to detect heritability in our traits just as long as the heritability is not too low (near 10%).

Figure 2. Number of families need to achieve 80% power to detect heritability.



A longitudinal analysis of our sample would allow us to verify how changes in physical activity/energy expenditure over time would associate with BC and MetS changes and would provide a better portrait of our results. Further, because our sample depends on the bias associated with a free health check-up, it may not be representative of the general Portuguese population. In addition, it is possible that these families are in a healthier condition than the overall population, making it harder to detect clusters of cardiovascular risk factors in comparison to high-risk samples. Also, the possibility of false positive results due to multiple statistical testing should not be ruled out. Unfortunately, we were not able to account for biological maturation in adolescents, which could influence our results. The absence of data on insulin levels may also limit the understanding of some traits' relationships. Moreover, the method chosen to estimate different BC traits, in our case bio-impedance analysis, even though

having been previously validated with DXA (Pietrobelli et al., 2004), is not free from bias in its results although the precision of the equipment is ±1%. Lastly, the lack of information from nutritional habits and extensive and precise socio-economic status is important and must be accounted for.

IMPLICATIONS AND OPPORTUNITIES FOR FUTURE RESEARCH

The present thesis suggests that there is plenty of work to be done in the field of genetic epidemiology using PA, MetS and BC as main research phenotypes. Firstly, we demonstrated that available information on the genetics of PA is scarce and ambiguous. This should lead investigators to pursue not only new loci and candidate genes associated with PA but also to use a multimode approach in PA assessment that would make the search more efficient; further a call is made to use similar definitions and instruments to measure PA so that comparisons between studies may be facilitated.

We presented information suggesting that physical inactivity is under genetic control that is independent of physical activity status. Research should now focus on determining the amount of variation in sedentarism that is explained by genes and then try to uncover which genes are responsible for such trait, as well as their mechanisms. This could be firstly done in a nuclear family design by assessing different indicators of sedentarism (for example, screen time, sitting time) over a period of time (longitudinal study of 3, 5 or 10 year lag would allow us to assess changes in family routines), controlling for the effect of PA, nutritional habits and other important behaviors.

We computed the heritability estimates of MetS components and its correlations within families. Familial aggregation in all MetS components in Portuguese families was observed. This is a clear indication that the step forward could be the implementation of intervention programs having families as their main targets. These families should be systematically educated on how to live a healthier life, using physical exercise programs as well as adequate nutritional habits (after the well-known motto – exercise is medicine). These intervention programs should be monitored in order to assess their suitability and efficiency to deal with cardiovascular risk factors in Portuguese families.

We verified that the preventive effects of PA on MetS and obesity might be mediated by genetic regulation. To further explore this possibility, the development of a study of objectively measured PA in nuclear families, during at least a decade, and analyzing its interaction with genotype that may influence not only MetS and BC but also physical fitness, motor coordination, social competencies and intellectual capacity, controlling for the effect of socio-economic status and nutritional habits would deeply enhance our knowledge about the importance of family in determining a variety of human behaviors.

CONCLUSIONS

In this final section we present the main conclusions of the present study.

1. There are potentially different genetic influences on PA versus PI phenotypes.
2. PA and PI appear to be truly modifiable traits that are sensitive to environmental challenges.
3. MetS components variance aggregates within families.
4. Spousal MetS components correlations were, in general, of the same magnitude as the biological relatives' correlations suggesting that most of the phenotypic variance in MetS traits could be explained by shared environment.
5. MetS components and BC traits present significant heritability estimates.
6. MetS components and BC traits expression is significantly influenced by the interaction established between total daily energy expenditure and genotypes.
7. PA may be considered an environmental variable that promotes metabolic and fat mass differences between individuals that are distinctively active.

REFERENCES

Arya, R., Duggirala, R., Jenkinson, C. P., Almasy, L., Blangero, J., O'Connell, P., & Stern, M. P. (2004). Evidence of a novel quantitative-trait locus for obesity on chromosome 4p in Mexican Americans. *Am J Hum Genet, 74*(2), 272-282.

Baecke, J. A., Burema, J., & Frijters, J. E. (1982). A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr, 36*(5), 936-942.

Barres, R., Osler, M. E., Yan, J., Rune, A., Fritz, T., Caidahl, K., Krook, A., & Zierath, J. R. (2009). Non-CpG methylation of the PGC-1alpha promoter through DNMT3B controls mitochondrial density. *Cell Metab, 10*(3), 189-198.

Barres, R., Yan, J., Egan, B., Treebak, J. T., Rasmussen, M., Fritz, T., Caidahl, K., Krook, A., O'Gorman, D. J., & Zierath, J. R. (2012). Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metab, 15*(3), 405-411.

Bellia, A., Giardina, E., Lauro, D., Tesauro, M., Di Fede, G., Cusumano, G., Federici, M., Rini, G. B., Novelli, G., Lauro, R., & Sbraccia, P. (2009). "The Linosa Study": epidemiological and heritability data of the metabolic syndrome in a Caucasian genetic isolate. *Nutr Metab Cardiovasc Dis, 19*(7), 455-461.

Boes, E., Coassin, S., Kollerits, B., Heid, I. M., & Kronenberg, F. (2009). Genetic-epidemiological evidence on genes associated with HDL cholesterol levels: a systematic in-depth review. In *Exp Gerontol* (Vol. 44, pp. 136-160). England.

Booth, F. W., Laye, M. J., Lees, S. J., Rector, R. S., & Thyfault, J. P. (2008). Reduced physical activity and risk of chronic disease: the biology behind the consequences. *Eur J Appl Physiol, 102*(4), 381-390.

Bouchard, C., & Rankinen, T. (2001). Individual differences in response to regular physical activity. *Medicine and Science in Sports and Exercise, 33*(6), S446-S451.

Butte, N. F., Cai, G., Cole, S. A., & Comuzzie, A. G. (2006). Viva la Familia Study: genetic and environmental contributions to childhood obesity and its comorbidities in the Hispanic population. *Am J Clin Nutr, 84*(3), 646-654; quiz 673-644.

Cai, G., Cole, S. A., Butte, N., Bacino, C., Diego, V., Tan, K., Goring, H. H., O'Rahilly, S., Farooqi, I. S., & Comuzzie, A. G. (2006). A quantitative trait locus on chromosome 18q for physical activity and dietary intake in Hispanic children. *Obesity (Silver Spring), 14*(9), 1596-1604.

Chagnon, Y. C., Rice, T., Perusse, L., Borecki, I. B., Ho-Kim, M. A., Lacaille, M., Pare, C., Bouchard, L., Gagnon, J., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., Bouchard, C., & Study, H. F. (2001). Genomic scan for genes affecting body composition before and after training in Caucasians from HERITAGE. *J Appl Physiol (1985), 90*(5), 1777-1787.

de Andrade, M., Gueguen, R., Visvikis, S., Sass, C., Siest, G., & Amos, C. I. (2002). Extension of variance components approach to incorporate temporal trends and longitudinal pedigree data analysis. *Genet Epidemiol, 22*(3), 221-232.

De Moor, M. H., Liu, Y. J., Boomsma, D. I., Li, J., Hamilton, J. J., Hottenga, J. J., Levy, S., Liu, X. G., Pei, Y. F., Posthuma, D., Recker, R. R., Sullivan, P. F., Wang, L., Willemsen, G., Yan, H., EJ, D. E. G., & Deng, H. W. (2009). Genome-Wide Association Study of Exercise Behavior in Dutch and American Adults. *Med Sci Sports Exerc, 41*(10), 1887-1895.

Duvivier, B. M., Schaper, N. C., Bremers, M. A., van Crombrugge, G., Menheere, P. P., Kars, M., & Savelberg, H. H. (2013). Minimal intensity physical activity (standing and walking) of longer duration improves insulin action and plasma lipids more than shorter periods of moderate to vigorous exercise (cycling) in sedentary subjects when energy expenditure is comparable. *PLoS One, 8*(2), e55542.

Ehret, G. B., & Caulfield, M. J. (2013). Genes for blood pressure: an opportunity to understand hypertension. In *Eur Heart J* (Vol. 34, pp. 951-961). England.

Fenza, R., & Fiorina, P. (2012). Physical Activity and Inflammation. In L. Luzi (Ed.), *Cellular Physiology and Metabolism of Physical Exercise* (pp. 99-108): Springer Milan.

Freeman, M. S., Mansfield, M. W., Barrett, J. H., & Grant, P. J. (2002). Heritability of features of the insulin resistance syndrome in a community-based study of healthy families. *Diabet Med, 19*(12), 994-999.

Fuentes, R. M., Perola, M., Nissinen, A., & Tuomilehto, J. (2002). ACE gene and physical activity, blood pressure, and hypertension: a population study in Finland. *J Appl Physiol, 92*(6), 2508-2512.

Gielen, S., Schuler, G., & Adams, V. (2010). Cardiovascular effects of exercise training: molecular mechanisms. *Circulation, 122*(12), 1221-1238.

Hakanen, M., Raitakari, O. T., Lehtimaki, T., Peltonen, N., Pahkala, K., Sillanmaki, L., Lagstrom, H., Viikari, J., Simell, O., & Ronnemaa, T. (2009). FTO genotype is associated with body mass index after the age of seven years but not with energy intake or leisure-time physical activity. *The Journal of clinical endocrinology and metabolism, 94*(4), 1281-1287.

Hamer, M., Sabia, S., Batty, G. D., Shipley, M. J., Tabak, A. G., Singh-Manoux, A., & Kivimaki, M. (2012). Physical activity and inflammatory markers over 10 years: follow-up in men and women from the Whitehall II cohort study. *Circulation, 126*(8), 928-933.

Hong, Y., Pedersen, N. L., Brismar, K., & de Faire, U. (1997). Genetic and environmental architecture of the features of the insulin-resistance syndrome. *Am J Hum Genet, 60*(1), 143-152.

Katzmarzyk, P., & Herman, K. (2007). The role of physical activity and fitness in the prevention and treatment of metabolic syndrome. *Current Cardiovascular Risk Reports, 1*(3), 228-236.

Kilpelainen, T. O., Qi, L., Brage, S., Sharp, S. J., Sonestedt, E., Demerath, E., Ahmad, T., Mora, S., Kaakinen, M., Sandholt, C. H., Holzapfel, C., Autenrieth, C. S., Hypponen, E., Cauchi, S., He, M., Kutalik, Z., Kumari, M., Stancakova, A., Meidtner, K., Balkau, B., Tan, J. T., Mangino, M., Timpson, N. J., Song, Y., Zillikens, M. C., Jablonski, K. A., Garcia, M. E., Johansson, S., Bragg-Gresham, J. L., Wu, Y., van Vliet-Ostaptchouk, J. V., Onland-Moret, N. C., Zimmermann, E., Rivera, N. V., Tanaka, T., Stringham, H. M., Silbernagel, G., Kanoni, S., Feitosa, M. F., Snitker, S., Ruiz, J. R., Metter, J., Larrad, M. T., Atalay, M., Hakanen, M., Amin, N., Cavalcanti-Proenca, C., Grontved, A., Hallmans, G., Jansson, J. O., Kuusisto, J., Kahonen, M., Lutsey, P. L., Nolan, J. J., Palla, L., Pedersen, O., Perusse, L., Renstrom, F., Scott, R. A., Shungin, D., Sovio, U., Tammelin, T. H., Ronnemaa, T., Lakka, T. A., Uusitupa, M., Rios, M. S., Ferrucci, L., Bouchard, C., Meirhaeghe, A., Fu, M., Walker, M., Borecki, I. B., Dedoussis, G. V., Fritsche, A., Ohlsson, C., Boehnke, M., Bandinelli, S., van Duijn, C. M., Ebrahim, S., Lawlor, D. A., Gudnason, V., Harris, T. B., Sorensen, T. I., Mohlke, K. L., Hofman, A., Uitterlinden, A. G., Tuomilehto, J., Lehtimaki, T., Raitakari, O., Isomaa, B., Njolstad, P. R., Florez, J. C., Liu, S., Ness, A., Spector, T. D., Tai, E. S., Froguel, P., Boeing, H., Laakso, M., Marmot, M., Bergmann, S., Power, C., Khaw, K. T., Chasman, D., Ridker, P., Hansen, T., Monda, K. L., Illig, T., Jarvelin, M. R., Wareham, N. J., Hu, F. B., Groop, L. C., Orho-Melander, M., Ekelund, U., Franks, P. W., & Loos, R. J. (2011). Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med, 8*(11), e1001116.

Loos, R. J., Rankinen, T., Tremblay, A., Perusse, L., Chagnon, Y., & Bouchard, C. (2005). Melanocortin-4 receptor gene and physical activity in the Quebec Family Study. *Int J Obes (Lond), 29*(4), 420-428.

Miller, M., Langenberg, P., & Havas, S. (2007). Impact of lowering triglycerides on raising HDL-C in hypertriglyceridemic and non-hypertriglyceridemic subjects. *Int J Cardiol, 119*(2), 192-195.

Mitchell, B. D., Kammerer, C. M., Mahaney, M. C., Blangero, J., Comuzzie, A. G., Atwood, L. D., Haffner, S. M., Stern, M. P., & MacCluer, J. W. (1996).

Genetic analysis of the IRS. Pleiotropic effects of genes influencing insulin levels on lipoprotein and obesity measures. *Arterioscler Thromb Vasc Biol, 16*(2), 281-288.

Mokdad, A. H., Giles, W. H., Bowman, B. A., Mensah, G. A., Ford, E. S., Smith, S. M., & Marks, J. S. (2004). Changes in health behaviors among older Americans, 1990 to 2000. In *Public Health Rep* (Vol. 119, pp. 356-361). United States.

Povel, C. M., Boer, J. M., & Feskens, E. J. (2011). Shared genetic variance between the features of the metabolic syndrome: heritability studies. *Mol Genet Metab, 104*(4), 666-669.

Rebato, E., Jelenkovic, A., & Salces, I. (2007). Heritability of the somatotype components in Biscay families. *Homo, 58*(3), 199-210.

Simonen, R. L., Perusse, L., Rankinen, T., Rice, T., Rao, D. C., & Bouchard, C. (2002). Familial aggregation of physical activity levels in the Quebec Family Study. *Med Sci Sports Exerc, 34*(7), 1137-1142.

Simonen, R. L., Rankinen, T., Perusse, L., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rao, D. C., & Bouchard, C. (2003a). A dopamine D2 receptor gene polymorphism and physical activity in two family studies. *Physiol Behav, 78*(4-5), 751-757.

Tregouet, D. A., Herbeth, B., Juhan-Vague, I., Siest, G., Ducimetiere, P., & Tiret, L. (1999). Bivariate familial correlation analysis of quantitative traits by use of estimating equations: application to a familial analysis of the insulin resistance syndrome. *Genet Epidemiol, 16*(1), 69-83.

Tsao, T. S., Li, J., Chang, K. S., Stenbit, A. E., Galuska, D., Anderson, J. E., Zierath, J. R., McCarter, R. J., & Charron, M. J. (2001). Metabolic adaptations in skeletal muscle overexpressing GLUT4: effects on muscle and physical activity. *FASEB J, 15*(6), 958-969.

Vauthier, J. M., Lluch, A., Lecomte, E., Artur, Y., & Herbeth, B. (1996). Family resemblance in energy and macronutrient intakes: the Stanislas Family Study. *Int J Epidemiol, 25*(5), 1030-1037.

Winnicki, M., Accurso, V., Hoffmann, M., Pawlowski, R., Dorigatti, F., Santonastaso, M., Longo, D., Krupa-Wojciechowska, B., Jeunemaitre, X., Pessina, A. C., Somers, V. K., & Palatini, P. (2004). Physical activity and angiotensin-converting enzyme gene polymorphism in mild hypertensives. *Am J Med Genet A, 125A*(1), 38-44.