Is physical activity regulated by genetics?

Van Der Zee, Matthijs D.; De Geus, Eco

published in
Routledge Handbook of Sport and Exercise Systems Genetics
2019

DOI (link to publisher)
10.4324/9781315146287-9

document version
Publisher's PDF, also known as Version of record

6

IS PHYSICAL ACTIVITY REGULATED BY GENETICS? EVIDENCE FROM STUDIES IN HUMANS

Matthijs D. van der Zee and Eco de Geus

Introduction

Physical activity (PA) is a broad concept containing a variety of different human behaviors that share the common denominator of expected beneficial effects on mental and somatic health (1). Typically, daily PAs take up about 30% of total energy expenditure and reflect a mixture of obligatory activities related to transportation, work, or household chores and self-chosen PA behaviors including sports and exercise activities in leisure time. Objective measurements of daily PA in studies conducted in Europe and the US have confirmed the existence of large individual differences in activity levels (26, 27). These differences are observed in men and women alike, and persist after stratifying for age which is known to have a strong effect on average PA level (49).

To increase the success of intervention on this important health behavior, much research has been devoted to the causes of the individual differences seen in PA in the general population. In this chapter, we focus on the role played by genetic factors. Broadly, the genetic research on PA phenotypes can be classified into a) twin studies that partition the observed variance in PA phenotypes into environmental and genetic components based on well-established biometric models of human inheritance (48), and b) gene-finding studies that use either candidate gene-based approaches based on the known biological role of proteins (e.g., in neurotransmission) or agnostic whole-genome searches using genome-wide microsatellite (linkage) or single nucleotide polymorphism (SNP) markers to detect genomic loci harboring variants associated with PA. We will summarize the current evidence from each of these types of studies to address the question of whether PA is regulated by genetics.

Twin studies

PA behaviors appear to "run in the family," for example, the chance of one family member being a regular exerciser increases the chance of all other family members to be, or to become, an exerciser. Familial aggregation of PA can be investigated by computing correlations among relatives such as siblings, and parents and their offspring. However, siblings among each other, and parents and their offspring, share not just half of their genes, they also share a household, socioeconomic status, the neighborhood, and various other aspects of belonging to the

same family or living in the same neighborhood (the so-called shared environment), including parenting behaviors, family functioning, shared peers, etc.

Twin studies can separate the two mechanisms of familial aggregation by comparing the resemblance in monozygotic (MZ) or identical twins to the resemblance in dizygotic (DZ) or fraternal twins (48). When twins are reared together, the amount of this sharing of the family environment is the same for MZ and DZ twins. The important difference between MZ and DZ twins is that the former share most, if not all, of their genotypes, whereas the latter share on average only half of the genotypes segregating in that family. If the resemblance in PA within MZ pairs is larger than that in DZ pairs, this suggests that additive genetic factors (A) influence PA. If MZ resemblance is more than double as large, it suggests the influence of nonadditive (D) genetic factors. Additive genetic factors represent the sum of all linear effects of the genetic loci that influence the trait of interest. Nonadditive factors include dominance and epistatic interaction effects. If, however, the resemblance in PA in DZ twins is as large as it is in MZ twins, this points to shared environmental factors (C) as the cause of twin resemblance. Furthermore, the extent to which MZ twins do not resemble each other is ascribed to the unique environmental factors (E). These include all person-specific experiences such as differential jobs or lifestyles, accidents or other life events, and in childhood, differential treatment by the parents, going to different schools, and having nonshared friends and peers. Measurement error will also be subsumed by the unique environmental factor. For more details on the modeling of twin and family resemblance, see Chapter 3.

We have previously reviewed all twin and family studies on the heritability of total PA and regular sports and exercise behaviors in childhood or adolescent samples (50). In younger children, the shared environmental factors (C) explain the largest part of the variation in PA. However, the importance of these shared environmental factors decreases in adolescence and young adulthood, where genetic effects become the dominant factor explaining individual differences in both total PA and regular sports and exercise behaviors. A sample size-weighted meta-analysis showed PA heritability estimates of 20% (95% confidence interval (CI) 13–27%) in children, 35% (95% CI 17–52%) in early adolescents, and 53% (95% CI 47–59%) in late adolescents. This increase in heritability ran in parallel to a gradual decrease in the importance of shared environmental factors, which are the main cause of individual differences in PA and exercise behavior in young childhood, but then gradually become overwhelmed by the importance of genetic factors during adolescence. Using longitudinal follow-up of twins from childhood to adolescence, we showed that the increase in heritability, at least for voluntary exercise behaviors, was due to a strong increase in the genetic variance in the course of adolescence that was not paired to a similar increase in environmental variance (31). Whether the trend of increasing heritability is continued or curbed in young, middle, or late adulthood remains to be established.

Sufficient studies in adult twins (>18 years of age) have now accrued to allow a similar meta-analytic approach and to arrive at solid estimates of additive genetic and shared environmental variance for PA phenotypes in adulthood. We searched publications in the English language on human subjects in PubMed and Web-of-Science from January 1980 to December 2017 using the keywords ("Physical Activity" OR Exercise OR Sports OR Lifestyle) AND (Genetic OR Genes OR Linkage OR QTL OR Twin OR Family OR Familial OR Heritability) AND "Humans" [MeSH terms]. From the 850 putative papers, title and abstract analysis was used to select only those publications reporting MZ and DZ/sib correlations (in at least 30 complete pairs) and/or estimates of (non-)additive genetic, shared, and/or unique environmental variance components. Reference sections of selected papers were used to identify additional papers missed by these search terms. We then removed samples reporting on twins with a mean age

less than 18. Next we removed partly overlapping datasets from the same twin cohorts. When the same phenotype was used in largely the same age group, we only used the study reporting the largest dataset, which typically would be the most recent study. A final set of 27 adult twin studies (2, 3, 8, 11, 14–16, 20–24, 29–32, 34, 35, 40, 44, 45, 46, 47, 51, 53, 54, 58) were selected by the above search criteria.

Physical activity phenotypes

In the final set of 27 adult twin studies, we encountered a large variation in measurement instruments and PA measures used. By far the largest common denominator was the use of survey-based methods using subjective PA reporting on self or family members. In large population-based twin registries, surveys are often considered the only feasible strategy. However, subjective reporting of PA is vulnerable to distortion due to recall error and reporting biases (25). It is particularly difficult to estimate both the duration and frequency of PAs that are light to moderate in intensity, including common activities such as walking and standing, or household activities (5). For total PA and light to moderate PA, objective measurement strategies using indirect calorimetry, the doubly labeled water method, or movement sensors (60) are therefore superior.

When the focus is not on the detection of total PA, but when people are asked to report moderate to vigorous PA, specifically when confined to structured activities in leisure time, self-reporting seems to fare much better (10). The cognitive salience of such intensive PAs is higher than light to moderate intense activities occurring as part of daily routine. Reliability of self-reporting further increases when a restriction is made to the reporting of regular exercise and sports behaviors in leisure time. In our own research, for instance, we have shown a high short-term test–retest reliability (17, 55) as well as substantial tracking over longer periods of time for the weekly volume of voluntary exercise behavior (31).

Our meta-analysis was organized into the main domains of PA that we encountered in the literature: total physical activity (TPA), moderate to vigorous physical activity (MVPA, including separate measures for moderate or vigorous PA where applicable), leisure-time physical activity (LTPA), and voluntary exercise behavior (VEB). The first two categories (TPA and MVPA) cover activities that are only partly under the individual's control, whereas the latter two (LTPA and VEB) cover PA that is largely voluntary in nature. We excluded PA measures that deliberately excluded sports and exercise activities (e.g., Baecke's nonexercise-related PA index), measures exclusively reporting on PA in the occupational setting (e.g., Baecke's Work/School index), and measures of daily activities of low intensity (e.g., accelerometer derived time spent in low LPA).

To convert the measured PA behaviors into an actual summary metric for use in the genetic analyses, again a breadth of different strategies has been used. Most studies adopt an estimation of the total energy expenditure by PA as their focus. Estimation of energy expenditure is usually based on published compendia (4) that convert each activity into equivalents of the resting energy expenditure (1 MET, approximately 1 kcal/kg/hour). For VEB, most studies use a strategy akin to what we do in the Netherlands Twin Register, where we use an open format that allows reporting all sports and exercise activities that are performed for at least 3 months a year and then record duration, frequency, and intensity of each reported activity. Typically, activities are censored that do not reach a minimal threshold of intensity such as fishing or chess. Exercise related to swimming, sailing, or skiing that are restricted to the annual holidays is discarded. Physical education classes are also discarded as they are poorly standardized and the activities are not voluntary. Energy expenditure in all leisure-time VEB is then summed across

all valid exercise activities in a weekly MET-hour score by taking the sum of the products of their weekly frequency, average duration, and MET score.

Another major difference across studies is the type of scale used to quantify PA. Continuous interval scales seem most optimal in terms of statistical power but quite often either ordinal categories (e.g., tertiles of MET-hours weekly; frequency of LTPA as once per month, once per week, every day) or even nominal dichotomies (YES/NO regular exerciser; YES/NO adheres to PA guidelines) were used. This use of dichotomies and categories is mostly inspired by the frequency distribution of some of the LTPA and VEB phenotypes. For instance, moderate to vigorous PA in leisure time is very skewed. Likewise, many individuals do not engage in voluntary exercise behavior (zero score) and even in those who do the amount of exercise can still show a skewed distribution. Often no transformation is available that converts such mixed distributions to a normal distribution. For these phenotypes, therefore, the use of ordinal categories or a dichotomy can be meaningful. A liability threshold model can then be used in the genetic analysis to recapture the normal distribution of the latent "liability to be a (vigorous) exerciser."

Meta-analyses

To assess heritability and influence of common environmental factors in the four PA categories mentioned previously, meta-analysis on A and C estimates was conducted across twin studies. Within the 27-study database, the A and C estimates and their standard errors were often reported separately for males and females. However, if the A and C estimates were reported only for males and females combined, we assigned the same estimates to the male and female part of the sample, adjusting the *N* to reflect the number of male or female twins. Most studies used age and sex as covariates, but additional covariates were sometimes used too, including socioeconomic status, body mass index, or fitness. When multiple estimates were generated for different covariate compositions, we opted to use the estimates only correcting for age and sex effects on mean PA level. In these studies, structural equation-based variance decomposition modeling was by far the most used analytic strategy. Unless stated otherwise, the results from the most parsimonious structural equation model were used in the meta-analysis. This model was most often an AE or ACE model. If the AE model was used, C was set to zero. No studies reported a model with nonadditivity (D).

For each of the four PA phenotypes, the estimates for A and C were computed in a sample-size weighted meta-analysis across all available studies for males and females separately. Inverse variance weighting was not possible because not all studies provided either standard errors of the estimates or 95% CIs from which the standard errors could be approximated. Forest plots in Figures 6.1 and 6.2 present the characteristics and the A and C estimates per study, and the meta-analytic results per PA phenotype. Study characteristics are as follows: country, whether males and females were combined in A and C estimation, whether PA measurement was through surveys (SUBJ) or experimental (OBJ), type of scale used (DIchotomy, CATegorical, or CONtinuous), mean age, and sex-specific sample size. The estimates for A and C are listed in the forest plots as a function of mean sample age.

Results for total physical activity

For the meta-analysis of TPA (see Figures 6.1 and 6.2) a total of six studies using eight phenotypes was available, most using an objective measure to quantify PA, and all combining males and females. The meta-analytic heritability estimate was 48% for females and 51% for

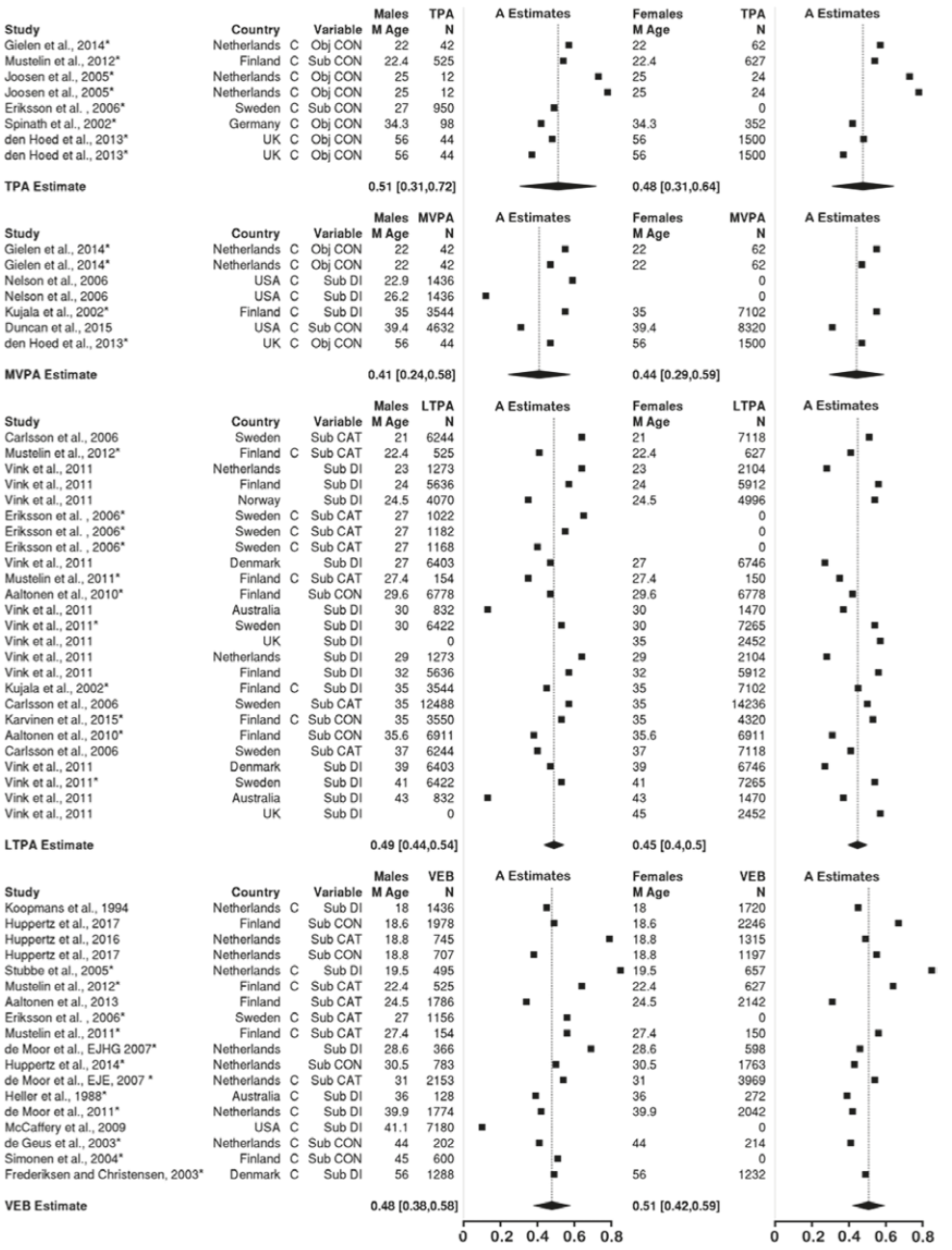

Figure 6.1 Forest plot containing study-specific and meta-analytic estimates of heritability (A) of total physical activity (TPA), moderate to vigorous activity (MVPA), leisure-time physical activity (LTPA), and voluntary exercise behavior (VEB), for males and females. Estimates include results from subjective (Sub), and objective (Obj), continuous (CON), categorical (CAT), and dichotomous (DI) variables. ⋆=studies where an AE model was fitted. C=estimates from a combined male and female sample.

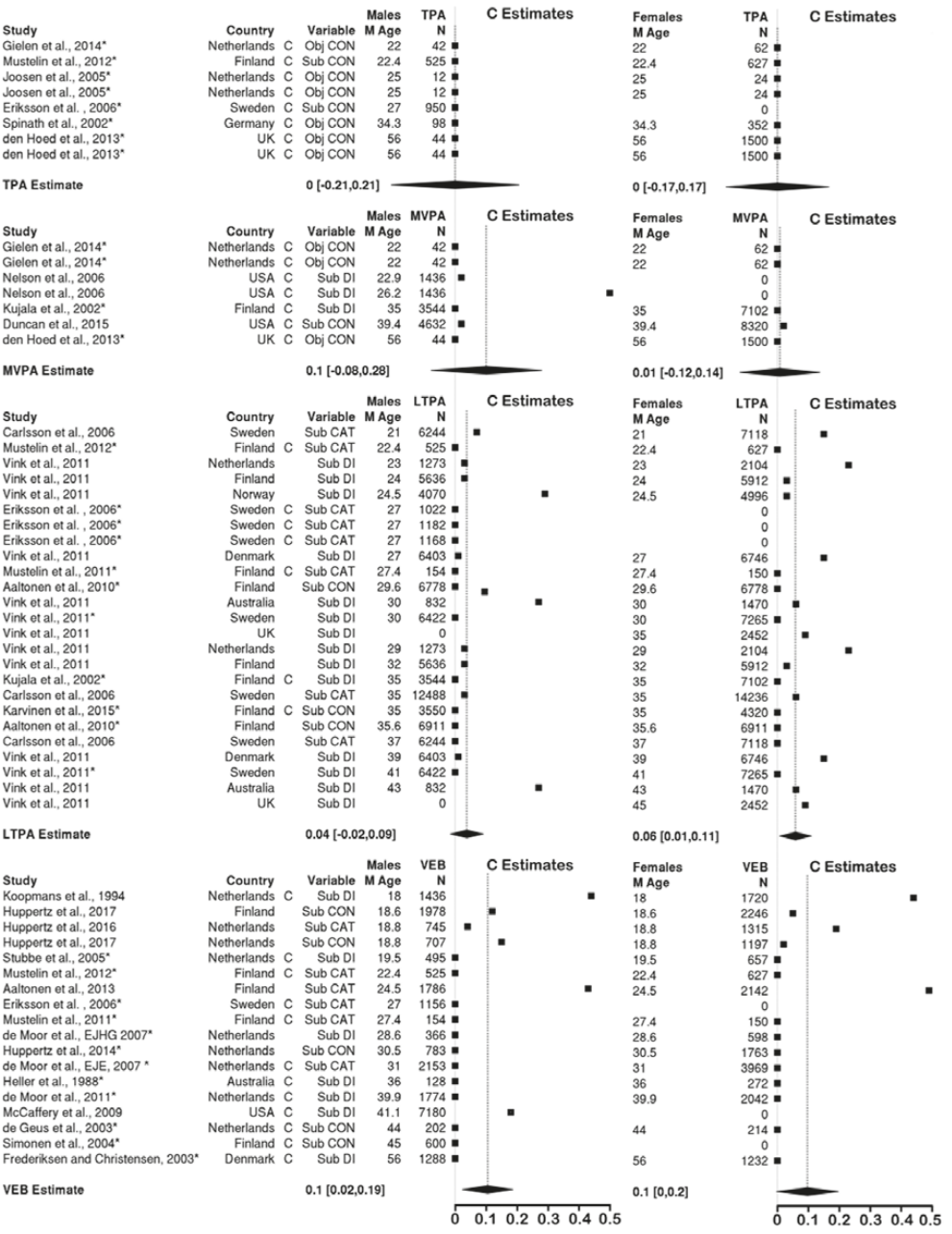

Figure 6.2 Forest plot containing study-specific and meta-analytic estimates of shared environmental contribution (C) to the variance in total physical activity (TPA), moderate to vigorous activity (MVPA), leisure-time physical activity (LTPA), and voluntary exercise behavior (VEB), for males and females. Estimates include results from subjective (Sub), and objective (Obj), continuous (CON), categorical (CAT), and dichotomous (DI) variables. *=studies where an AE model was fitted. C=estimates from a combined male and female sample.

males. Largest heritability estimates were found by Joosen et al. (34) who used the most reliable methods, namely the doubly labeled water method and more than 14 days of accelerometer recording, albeit in the smallest sample. The two survey methods using the Baecke questionnaire yielded results comparable to the objective methods. No evidence for common environmental effects was found.

Results for moderate to vigorous physical activity

For the meta-analysis of MVPA (see Figures 6.1 and 6.2) a total of five studies using seven phenotypes was available, all combining males and females. The meta-analytic heritability estimate was 44% for females and 41% for males. Comparable heritability was found for objective and subjective measures. Somewhat deviant results were found in the 26-year-old Add Health study participants (47). Using a 7-day recall questionnaire to obtain a dichotomy of MVPA frequency that indicated whether participants met national PA recommendations (>5 bouts/week), lower heritability was paired to the only evidence for a shared household effect on MVPA (50%).

Results for leisure-time physical activity

We used eight twin studies for the meta-analysis of LTPA (see Figures 6.1 and 6.2) reporting on more than 200,000 twins from seven countries. The meta-analytic heritability estimate was 49% for males and 45% for females with narrow confidence intervals (95% CI upper − lower=~10%) reflecting little heterogeneity. LTPA was only available from self-report measures in these very large epidemiological samples. In two countries, Norway and Australia, evidence for common environmental impact on LTPA was found (~28%). Comparing sample sizes, the absence of a large C in other twin studies cannot be attributed to low power.

Results for voluntary exercise behavior

VEB was also measured exclusively by survey or interview measures. A total of 15 different studies using 14 unique samples/measures yielded a meta-analytic heritability estimate of 48% for males and 51% for females (see Figure 6.1) with mild heterogeneity (95% CI upper − lower=~20%). Some evidence for common environmental effects were found, mostly driven by the younger samples (~10%, see Figure 6.2) with a particularly strong C effect in the young adult Finnish twins (43% males, 49% females).

Gene-finding studies

The above meta-analyses show that genetic variants play a key role in adult PA, whether it is total daily PA as measured with objective means (e.g., accelerometers), or self-reported voluntary sports and exercise activities in leisure time. We now turn to the gene-finding studies aiming to detect the actual "PA genes". For this, we have to use a narrative approach because not enough replication is currently available for any single locus to consider meta-analysis. The majority of gene-finding studies on TPA, MVPA, LTPA, and VEB have used a candidate gene approach. The use of this approach was predicated on the expectation that individual genetic variants would be associated with complex behavioral traits at a magnitude that would be small but detectable with hundreds, or a few thousand individuals. The advent of large international consortia performing meta-analyses on genome-wide association results for many complex behavioral

traits in samples as large as hundreds of thousands of participants has found this expectation to be untenable. The risk contributed by any single variant is tiny rather than just small, with only an increase of ~0.05 standard deviation per risk allele at best, necessitating much larger sample sizes to be able to detect them (59). Based on these types of concerns regarding candidate gene studies using small samples, we focus here mainly on the data-driven genome-wide approach in linkage and genome-wide association studies (GWAS).

Linkage

The three genome-wide linkage studies (7, 19, 52) performed to date have only produced suggestive hits, in keeping with their modest sample sizes ($767 < N < 1120$). In the Quebec family study, Simonen et al. (52) reported linkage for TPA (13q22–q31), MVPA (4q28.2, 7p11.2, 9q31.1, 13q22–q31), and VEB (11p15 and 15q13.3). Interestingly, the latter 15q linkage region contains the *GABRG3* gene in which a SNP (rs8036270) was found to be significantly associated with VEB in GWAS on a combined sample of 2622 Dutch and European American middle-aged adults (18). Significant association with *GABRG3* was replicated for LTPA in older adults in a GWAS in 10,684 European and 11,093 African Americans (42) although with different SNPs (rs72707657, rs12438610, rs12902711, rs12595253). The *GABRG3* gene may be involved in the aversive effects of exercise-induced fatigue because high expression levels of the *GABRG3* gene was found after a bout of exhaustive exercise (36). However, a GWAS in 13,980 Japanese older adults did not find significant association with any SNPs in the *GABGR3* gene region and the gene also did not surface in the other two GWAS.

In the Viva La Familia study (7), physical inactivity, recorded as the percentage time in sedentary activity, significantly mapped to markers D18S1102–D18S64 in the chromosome 18q21 region where the *MC4R* gene resides. An additional suggestive linkage signal for TPA was detected in the same region. *MCR4* had been implicated in smaller-scaled candidate gene studies (9, 43) but it failed to replicate in any of the five GWAS, two of which explicitly tested it as a candidate gene (i.e., at lower *P*-value). In the Viva La Familia study, a further linkage signal was found for MVPA in the region at 9q31.1, which harbors the *RN7SK* gene. A SNP (rs7023003) near this gene produced a suggestive ($P < 10^{-5}$) association with TPA in 8454 Korean participants (38). However, the same *RN7SK* SNP was explicitly but unsuccessfully tested for replication in Japanese participants (28) and the gene also did not surface in the other two GWAS.

Finally, in a sample from the Netherlands Twin Register (19), a suggestive linkage with exercise participation was found in all subjects on chromosome 19p13.3 (LOD 2.18). A SNP (rs12462609) in this region, located in the *CACNA1A* gene, was associated with vigorous PA in 261,055 older adult participants in the UK Biobank (39), and with TPA in 8454 older adult Koreans (38) whereas another SNP (rs111901094) in *GATAD2A* in the 19p13 linkage region was associated with VEB in the UK Biobank participants (39). The 19p13 region did not generate a significant or suggestive signal in the other two GWAS (28, 42). In summary, genes suggested from linkage regions (*GABRG3*, *MCR4*, *RN7SK*, *CACNA1A*) have not held up systematically when tested as candidate genes in GWAS.

Genome-wide association

Five GWAS (18, 28, 38, 39, 42) conducted for exercise behavior or related PA phenotypes were available at the time of writing. The three smallest GWAS ($2632 < N < 11,093$) did not detect significant associations after the required stringent correction for the multiple testing burden

that is inherent in the agnostic genome-wide approach (18, 38, 42). However, some genes listed as receiving suggestive evidence based on more lenient thresholds ($P < 10^{-5}$) showed interest-generating patterns of replication across multiple studies. For example, a suggestive association of VEB was found in Dutch and American participants of European descent (18) for an eQTL (rs12612420) that influences the expression of the *DNAPTP6* gene. This association was replicated (P=0.0092) in 16,026 Japanese participants (28).

Suggestive evidence for an association between VEB and the *PAPSS2* gene in Dutch and American participants (18) was replicated in 11,903 African American participants (42) for LTPA, although for different SNPs (rs10887741 vs. rs1819162). The *PAPSS2* gene has been linked to maximal exercise capacity (49) which may be a factor influencing the motivation to engage in voluntary exercise (12). Explicit testing of an association between VEB and *PAPSS2* in 13,980 Japanese participants, however, did not produce a significant replication (28). Suggestive evidence for an association with LTPA was found for rs116550874 at 1p36.23, and rs3792874, rs3792877, rs3792878, rs79173796 at 5q31.1 ($P < 8.61 \times 10^{-7}$) in 11093 African Americans and for rs28524846 at 14q23.3 ($P < 1.30 \times 10^{-6}$) in 10,684 European Americans (42). The authors point to the *ENO1*, *SLC22A5*, and *PDLIM4* genes as potential sources of the association signal at 1p36. *ENO1* encodes a glycolytic enzyme and the integral membrane protein SLC22A5 is associated with skeletal myopathy, whereas the PDLIM4 protein is involved in the pathway of actin cytoskeleton remodeling, and bone and skeletal muscle development. The rs28524846 SNP at 14q23 is an eQTL for *MPP5* and *APT6V1D* in nerve tissue, with the MPP5 protein known to regulate myelinating Schwann cells and *ATP6V1D* related to synaptic vesicle cycle and ATPase activity.

Encouragingly, the two largest GWAS to date did successfully detect genome-wide significant associations with PA. In a Japanese sample (N=16,016), Hara et al. (28) identified an association (Pmeta=2.2×10^{-9}) between LTPA and rs10252228, a SNP located in the intergenic region between the *NPSR1* and *DPY19L1* genes. Although functional links of intergenic SNPs to nearby genes should be interpreted with caution, Hara et al. (28) point out that the product of the *NPSR1* and that of the *DNAPTP6* genes previously found by the GWAS of de Moor et al. (18) are both involved in pulmonary function. Variants in *NPSR1* have been shown to be associated with asthma-related phenotypes, and *DNAPTP6* could be involved in bronchodilator response via the downregulation of ß2-adrenergic receptors. Impairments of pulmonary function could be a barrier to the adoption of (vigorous) PA.

In the UK Biobank (N=380,492), Klimentidis et al. (39) detected genome-wide significant genetic associations with touchscreen-survey based measures of MVPA. After applying corrections for work-related PA and an indicator of socioeconomic status, associations with MVPA were found at rs429358 (*APOE*), rs169504 (*PBX2*), rs4129572 (*EXOC4*), rs3094622 (*RPP21*), rs181220614 (*ARHGEF26-AS1*), rs149943 (*ZNF165*), and rs2988004 (*PAX5*). Intriguingly, the Alzheimer disease risk allele (E4) of the *APOE* gene was associated with higher levels of MVPA. As noted by Klimentidis et al. (39), *APOE* e4 carriers had a more favorable response to exercise, for instance, as suggested by a study showing larger aerobic fitness in response to training (56). This is in keeping with the theoretical notion that individuals who have higher exercise ability and/or trainability will find it easier to adopt regular exercise as a lifestyle (12, 13, 41).

From the UK Biobank self-reports (39), two dichotomous measures of more vigorous PA were defined: VPA by classifying participants as engaged in vigorous PA if they spend more than 25 minutes on activities "that make you sweat or breathe hard" for 3 or more days a week and VEB by classifying participants as regular vigorous exercisers if they spend more than 15 minutes on 2 or more days a week doing strenuous sports or other exercises. Significant genetic

association for VPA was found at rs1248860 (*CADM2*), rs2764261 (*FOXO3*), rs3781411 (*CTBP2*), rs12707131 (*EXOC4*), and rs328919 (*DPY19L1*) and for VEB at rs62253088 (*CADM2*), rs166840 (*AKAP10*), rs10946808 (*HIST1H1D*), rs75930676 (*SIPA1L1*), and rs4865656 (*LOC642366*). In addition, accelerometer data were available in just over 91,000 participants. From up to 7 days of accelerometer wear, overall acceleration was obtained as a measure of TPA and the fraction of accelerations >425 m*g* as a measure of VPA, yielding significant associations at rs55657917 (*CRHR1*) and rs185829646 (*ANKRD22*) for TPA and rs743580 (*PML*) and rs6433478 (*CIR1*) for VPA.

The *CADM2* gene, primarily expressed in the brain, surfaced in the UK Biobank study as a gene influencing vigorous sports and exercise-related PA. Previously this gene had been linked to risk-taking behavior and extraversion (6) as well as executive function (33). This shows a remarkable parallel to results reported in the largest candidate study to date by van der Mee et al. (57) where a polygenic dopaminergic risk score that summed the increaser alleles in *COMT* and *DAT1* for *both* executive function and reward sensitivity was associated with the volume of externally paced sports and exercise activities. Possibly, *CADM2*, like these dopaminergic genes operates through a "double whammy" of increased reward value of exercise and increased sports skills.

Conclusion

Although the gradual rise in heritability of PA seen in adolescence is not continued in adulthood, it is clear that genetic factors remain a major contributor to individual differences in adult PA behaviors. In the reports on over 283,904 adult twins, we find that about half of the variance in the four PA phenotypes can be explained by genetic factors (males 48%; 95% CI 44–52%; females 47%; 95% CI 44–50%). The substantial contribution of the shared environment (C) to childhood PA seems to have largely dissipated in adulthood (males and females 6%; 95% CI 3–10%). There is generally a good correspondence in estimates across individual studies even when they use samples from different countries or measures. For TPA and MVPA, objective measurement showed slightly higher genetic estimates than subjective measurement, but the differences are not striking. Major sex differences in genetic architecture of LTPA and VEB seemed to be absent; for TPA and MVPA estimates, the verdict is still out as analyses were mostly based on combined male/female samples.

The extant studies using a candidate gene, linkage, or whole-genome association approach have yielded a number of genes and variants worth careful monitoring in future gene-finding efforts. Previously, three biological themes have emerged as a potential source for "physical activity genes" from theory: (1) the brain circuitry related to motivational and affective aspects of PA; (2) the brain circuitry involved in the maintenance of energy intake/expenditure balance; and (3) the physiological determinants of the ability to perform (intense and/or prolonged) PA, ideally at an above average level (12, 13, 41). Many of the variants reviewed above do seem to fit these theoretical notions, but we issue a note of caution that this may partly reflect our deep desire (and uncanny ability) to make sense of data by "reasoning towards" our theoretical models. Rigorous replication followed by experimental validation (in animal models) is direly needed. However, the experience from many other complex behavioral traits allows us to end upbeat: once very large samples are amassed in international meta-analytic consortia, we can confidently expect large progress in our understanding of how PA is regulated by genetics.

References

1. **World Health Organization**. *Global Recommendations on Physical Activity for Health*. Switzerland: World Health Organization, 2010.
2. **Aaltonen S**, **Ortega-Alonso A**, **Kujala UM**, and **Kaprio J.** Genetic and environmental influences on longitudinal changes in leisure-time physical activity from adolescence to young adulthood. *Twin Res Hum Genet* 16: 535–543, 2013.
3. **Aaltonen S**, **Ortega-Alonso A**, **Kujala UM**, and **Kaprio J.** A longitudinal study on genetic and environmental influences on leisure time physical activity in the Finnish Twin Cohort. *Twin Res Hum Genet* 13: 475–481, 2010.
4. **Ainsworth BE**, **Haskell WL**, **Herrmann SD**, **Meckes N**, **Bassett DR**, Jr., **Tudor-Locke C**, **Greer JL**, **Vezina J**, **Whitt-Glover MC**, and **Leon AS.** 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc* 43: 1575–1581, 2011.
5. **Bassett DR**, Jr, **Cureton AL**, and **Ainsworth BE.** Measurement of daily walking distance-questionnaire versus pedometer. *Med Sci Sports Exerc* 32: 1018–1023, 2000.
6. **Boutwell B**, **Hinds D**, **23andMe Research Team**, **Tielbeek J**, **Ong KK**, **Day FR**, and **Perry JRB.** Replication and characterization of CADM2 and MSRA genes on human behavior. *Heliyon* 3: e00349, 2017.
7. **Cai GW**, **Cole SA**, **Butte N**, **Bacino C**, **Diego V**, **Tan K**, **Goring HH**, **O'Rahilly S**, **Farooqi IS**, and **Comuzzie AG.** A quantitative trait locus on chromosome 18q for physical activity and dietary intake in Hispanic children. *Obesity (Silver Spring)* 14: 1596–1604, 2006.
8. **Carlsson S**, **Andersson T**, **Lichtenstein P**, **Michaelsson K**, and **Ahlbom A.** Genetic effects on physical activity: Results from the Swedish twin registry. *Med Sci Sports Exerc* 38: 1396–1401, 2006.
9. **Cole SA**, **Butte NF**, **Voruganti VS**, **Cai G**, **Haack K**, **Kent JW** Jr, **Blangero J**, **Comuzzie AG**, **McPherson JD**, and **Gibbs RA.** Evidence that multiple genetic variants of MC4R play a functional role in the regulation of energy expenditure and appetite in Hispanic children. *Am J Clin Nutr* 91: 191–199, 2010.
10. **Craig CL**, **Marshall AL**, **Sjostrom M**, **Bauman AE**, **Booth ML**, **Ainsworth BE**, **Pratt M**, **Ekelund U**, **Yngve A**, **Sallis JF**, and **Oja P.** International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 35: 1381–1395, 2003.
11. **de Geus EJ**, **Boomsma DI**, and **Snieder H.** Genetic correlation of exercise with heart rate and respiratory sinus arrhythmia. *Med Sci Sports Exerc* 35: 1287–1295, 2003.
12. **de Geus EJ** and **de Moor MHM.** Genes, exercise, and psychological factors. In: *Genetic and Molecular Aspects of Sport Performance*, edited by Bouchard C and Hoffman EP. Chichester: Wiley, 2011, p. 294–305.
13. **de Geus EJ** and **de Moor MHM.** A genetic perspective on the association between exercise and mental health. *Ment Health Phys Act* 1: 53–61, 2008.
14. **de Moor MH**, **Posthuma D**, **Hottenga JJ**, **Willemsen G**, **Boomsma DI**, and **de Geus EJ.** Genome-wide linkage scan for exercise participation in Dutch sibling pairs. *Eur J Hum Genet* 15: 1252–1259, 2007.
15. **de Moor MH**, **Stubbe JH**, **Boomsma DI**, and **de Geus EJ.** Exercise participation and self-rated health: do common genes explain the association? *Eur J Epidemiol* 22: 27–32, 2007.
16. **de Moor MH**, **Willemsen G**, **Rebollo-Mesa I**, **Stubbe JH**, **de Geus EJ**, and **Boomsma DI.** Exercise participation in adolescents and their parents: Evidence for genetic and generation specific environmental effects. *Behav Genet* 41: 211–222, 2011.
17. **de Moor MH** and **de Geus, EJ**. Genetic influences on exercise behavior. In: *Lifestyle Medicine*, edited by Rippe JM. Boca Raton, FL: Taylor & Francis Group, 2012, p. 1367–1378.
18. **de Moor MH**, **Liu YJ**, **Boomsma DI**, **Li J**, **Hamilton JJ**, **Hottenga JJ**, **Levy S**, **Liu XG**, **Pei YF**, **Posthuma D**, **Recker RR**, **Sullivan PF**, **Wang L**, **Willemsen G**, **Yan H**, **de Geus EJ**, and **Deng HW.** Genome-wide association study of exercise behavior in Dutch and American adults. *Med Sci Sports Exerc* 41: 1887–1895, 2009.
19. **de Moor MH**, **Posthuma D**, **Hottenga JJ**, **Willemsen AHM**, **Boomsma DI**, and **de Geus EJC.** Genome-wide linkage scan for exercise participation in Dutch sibling pairs. *Eur J Hum Genet* 15: 1252–1259, 2007.
20. **den Hoed M**, **Brage S**, **Zhao JH**, **Westgate K**, **Nessa A**, **Ekelund U**, **Spector TD**, **Wareham NJ**, and **Loos RJ.** Heritability of objectively assessed daily physical activity and sedentary behavior. *Am J Clin Nutr* 98: 1317–1325, 2013.

21. **Duncan GE**, **Cash SW**, **Horn EE**, and **Turkheimer E.** Quasi-causal associations of physical activity and neighborhood walkability with body mass index: a twin study. *Prev Med* 70: 90–95, 2015.
22. **Eriksson M**, **Rasmussen F**, and **Tynelius P.** Genetic factors in physical activity and the equal environment assumption – the Swedish young male twins study. *Behav Genet* 36: 238–247, 2006.
23. **Frederiksen H** and **Christensen K.** The influence of genetic factors on physical functioning and exercise in second half of life. *Scand J Med Sci Sports* 13: 9–18, 2003.
24. **Gielen M**, **Westerterp-Plantenga MS**, **Bouwman FG**, **Joosen AM**, **Vlietinck R**, **Derom C**, **Zeegers MP**, **Mariman EC**, and **Westerterp KR.** Heritability and genetic etiology of habitual physical activity: A twin study with objective measures. *Genes Nutr* 9: 415, 2014.
25. **Hagstromer M**, **Ainsworth BE**, **Oja P**, and **Sjostrom M.** Comparison of a subjective and an objective measure of physical activity in a population sample. *J Phys Act Health* 7: 541–550, 2010.
26. **Hagstromer M**, **Troiano RP**, **Sjostrom M**, and **Berrigan D.** Levels and patterns of objectively assessed physical activity – a comparison between Sweden and the United States. *Am J Epidemiol* 171: 1055–1064, 2010.
27. **Hansen BH**, **Kolle E**, **Dyrstad SM**, **Holme I**, and **Anderssen SA.** Accelerometer-determined physical activity in adults and older people. *Med Sci Sports Exerc* 44: 266–272, 2012.
28. **Hara M**, **Hachiya T**, **Nishida Y**, **Shimanoe C**, **Tanaka K**, **Sutoh Y**, and **Shimizu A.** Genome-wide association study of leisure-1 time exercise behavior in Japanese adults. *Med Sci Sports Exerc*, 2018.
29. **Heller RF**, **O'Connell DL**, **Roberts DC**, **Allen JR**, **Knapp JC**, **Steele PL**, and **Silove D.** Lifestyle factors in monozygotic and dizygotic twins. *Genet Epidemiol* 5: 311–321, 1988.
30. **Huppertz C**, **Bartels M**, **de Geus EJ**, **van Beijsterveldt CEM**, **Rose RJ**, **Kaprio J**, and **Silventoinen K.** The effects of parental education on exercise behavior in childhood and youth: A study in Dutch and Finnish twins. *Scand J Med Sci Sports* 27: 1143–1156, 2017.
31. **Huppertz C**, **Bartels M**, **de Zeeuw EL**, **van Beijsterveldt CEM**, **Hudziak JJ**, **Willemsen G**, **Boomsma DI**, and **de Geus EJ.** Individual differences in exercise behavior: Stability and change in genetic and environmental determinants from age 7 to 18. *Behav Genet* 46: 665–679, 2016.
32. **Huppertz C**, **Bartels M**, **Jansen IE**, **Boomsma DI**, **Willemsen G**, **de Moor MH**, and **de Geus EJ.** A twin-sibling study on the relationship between exercise attitudes and exercise behavior. *Behav Genet* 44: 45–55, 2014.
33. **Ibrahim-Verbaas CA**, **Bressler J**, **Debette S**, **Schuur M**, **Smith AV**, **Bis JC**, **Davies G**, **Trompet S**, **Smith JA**, **Wolf C**, et al. GWAS for executive function and processing speed suggests involvement of the CADM2 gene. *Mol Psychiatry* 21: 189–197, 2016.
34. **Joosen AM**, **Gielen M**, **Vlietinck R**, and **Westerterp KR.** Genetic analysis of physical activity in twins. *Am J Clin Nutr* 82: 1253–1259, 2005.
35. **Karvinen S**, **Waller K**, **Silvennoinen M**, **Koch LG**, **Britton SL**, **Kaprio J**, **Kainulainen H**, and **Kujala UM.** Physical activity in adulthood: Genes and mortality. *Sci Rep* 5: 18259, 2015.
36. **Kawai T**, **Morita K**, **Masuda K**, **Nishida K**, **Sekiyama A**, **Teshima-Kondo S**, **Nakaya Y**, **Ohta M**, **Saito T**, and **Rokutan K.** Physical exercise-associated gene expression signatures in peripheral blood. *Clin J Sport Med* 17: 375–383, 2007.
37. **Kemper HC**, **Twisk JW**, and **van Mechelen W.** Changes in aerobic fitness in boys and girls over a period of 25 years: Data from the Amsterdam Growth And Health Longitudinal Study revisited and extended. *Pediatr Exerc Sci.* 25: 524–35, 2013.
38. **Kim J**, **Kim J**, **Min H**, **Oh S**, **Kim Y**, **Lee AH**, and **Park T.** Joint identification of genetic variants for physical activity in Korean population. *Int J Mol Sci* 15: 12407–12421, 2014.
39. **Klimentidis YC**, **Raichlen DA**, **Bea J**, **Garcia DO**, **Mandarino LJ**, **Alexander GE**, **Chen Z**, and **Going SB.** Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. *Int J Epidemiol*, 42: 1161–1176, 2018.
40. **Kujala UM**, **Kaprio J**, and **Koskenvuo M.** Modifiable risk factors as predictors of all-cause mortality: The roles of genetics and childhood environment. *Am J Epidemiol* 156: 985–993, 2002.
41. **Lightfoot JT**, **de Geus EJ**, **Booth FW**, **Bray MS**, **den Hoed M**, **Kaprio J**, **Kelly SA**, **Pomp D**, **Saul MC**, **Thomis MA**, **Garland T**, Jr, and **Bouchard C.** Biological/genetic regulation of physical activity level: Consensus from GenBioPAC. *Med Sci Sports Exerc* 50:863–873, 2017.
42. **Lin X**, **Chan KK**, **Huang Y-T**, **Luo X**, **Liang L**, **Wilson J**, **Correa A**, **Levy D**, and **Liu S.** Genetic determinants for leisure-time physical activity. *Med Sci Sports Exerc* 50: 1620–1628, 2018.
43. **Loos RJF**, **Rankinen T**, **Tremblay A**, **Pérusse L**, **Chagnon Y**, and **Bouchard C.** Melanocortin-4 receptor gene and physical activity in the Québec Family Study. *Int J Obes* 29: 420–428, 2004.

44. **McCaffery JM**, **Papandonatos GD**, **Bond DS**, **Lyons MJ**, and **Wing RR.** Gene X environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *Am J Clin Nutr* 89: 1011–1018, 2009.
45. **Mustelin L**, **Joutsi J**, **Latvala A**, **Pietilainen KH**, **Rissanen A**, and **Kaprio J.** Genetic influences on physical activity in young adults: A twin study. *Med Sci Sports Exerc* 44: 1293–1301, 2012.
46. **Mustelin L**, **Latvala A**, **Pietilainen KH**, **Piirila P**, **Sovijarvi AR**, **Kujala UM**, **Rissanen A**, and **Kaprio J.** Associations between sports participation, cardiorespiratory fitness, and adiposity in young adult twins. *J Appl Physiol* 110: 681–686, 2011.
47. **Nelson MC**, **Gordon-Larsen P**, **North KE**, and **Adair LS.** Body mass index gain, fast food, and physical activity: effects of shared environments over time. *Obesity (Silver Spring)* 14: 701–709, 2006.
48. **Plomin R**, **DeFries JC**, **Knopik V**, and **Niederheiser J.** *Behavioral Genetics*. New York: Worth, 2013.
49. **Rico-Sanz J**, **Rankinen T**, **Rice T**, **Leon AS**, **Skinner JS**, **Wilmore JH**, **Rao DC**, and **Bouchard C.** Quantitative trait loci for maximal exercise capacity phenotypes and their responses to training in the HERITAGE Family Study. *Physiol Genomics* 16: 256–260, 2004.
50. **Schutte NM**, **Bartels M**, and **de Geus EJ.** Genetics of physical activity and physical fitness. In: *Oxford Textbook of Children's Sport and Exercise Medicine* (3rd ed.), edited by Armstrong N and van Mechelen W. New York: Oxford University Press, 2017, p. 293–302.
51. **Simonen R**, **Levalahti E**, **Kaprio J**, **Videman T**, and **Battie MC.** Multivariate genetic analysis of lifetime exercise and environmental factors. *Med Sci Sports Exerc* 36: 1559–1566, 2004.
52. **Simonen RL**, **Rankinen T**, **Perusse L**, **Rice T**, **Rao DC**, **Chagnon Y**, and **Bouchard C.** Genome-wide linkage scan for physical activity levels in the Quebec family study. *Med Sci Sports Exerc* 35: 1355–1359, 2003.
53. **Spinath FM**, **Wolf H**, **Angleitner A**, **Borkenau P**, and **Riemann R.** Genetic and environmental influences on objectively assessed activity in adults. *Pers Individ Dif* 33: 633–645, 2002.
54. **Stubbe JH**, **Boomsma DI**, and **De Geus EJ.** Sports participation during adolescence: A shift from environmental to genetic factors. *Med Sci Sports Exerc* 37: 563–570, 2005.
55. **Stubbe JH**, **de Moor MH**, **Boomsma D**, and **de Geus EJ.** The association between exercise participation and well-being: A co-twin study. *Prev Med* 44: 148–152, 2006.
56. **Thompson PD**, **Tsongalis GJ, Seip RL**, **Bilbie C**, **Miles M**, **Zoeller R**, **Visich P**, **Gordon P**, **Angelopoulos TJ**, **Pescatello L**, **Bausserman L**, and **Moyna N.** Apolipoprotein E genotype and changes in serum lipids and maximal oxygen uptake with exercise training. *Metabolism* 53: 193–202, 2004.
57. **van der Mee DJ**, **Fedko IO**, **Hottenga JJ**, **Ehli EA**, **van der Zee MD**, **Ligthart L**, **van Beijsterveldt TCEM**, **Davies GE**, **Bartels M**, **Landers JG**, and **de Geus**, **EJC.** Dopaminergic genetic variants and voluntary externally paced exercise behavior. *Med Sci Sports Exerc* 50: 700–708, 2017.
58. **Vink JM**, **Boomsma DI**, **Medland SE**, **de Moor MH**, **Stubbe JH**, **Cornes BK**, **Martin NG**, **Skytthea A**, **Kyvik KO**, **Rose RJ**, **Kujala UM**, **Kaprio J**, **Harris JR**, **Pedersen NL**, **Cherkas L**, **Spector TD**, and **de Geus EJ.** Variance components models for physical activity with age as modifier: A comparative twin study in seven countries. *Twin Res Hum Genet* 14: 25–34, 2011.
59. **Visscher PM**, **Wray NR**, **Zhang Q**, **Sklar P**, **McCarthy MI**, **Brown MA**, and **Yang J.** 10 Years of GWAS discovery: Biology, function, and translation. *Am J Hum Genet* 101: 5–22, 2017.
60. **Westerterp KR.** Physical activity and physical activity induced energy expenditure in humans: Measurement, determinants, and effects. *Fron Psychol* 4: 90, 2013.