CrossMark

CARDIOVASCULAR GENOMICS (TL ASSIMES, SECTION EDITOR)

The Genetics of Physical Activity

Xiaochen Lin[1,2] · Charles B. Eaton[1,3] · JoAnn E. Manson[4,5] · Simin Liu[1,2,6,7,8]

Abstract

Purpose of Review Physical activity (PA) is a well-established modifiable lifestyle determinant for multiple cardio-metabolic outcomes. While many psychosocial and environmental correlates of PA have been identified, current understanding of the genetic architecture that contributes to PA is still very limited, especially when compared to other phenotypes such as obesity and diabetes.

Recent Findings This review systematically and comprehensively assesses available evidence from animal experiments, family studies, population-based candidate gene analyses, and genome-wide association studies (GWAS) studying the genetics of physical activity patterns. It discusses the scientific evolution in the field of PA genetics, including the recognition of increased sample sizes, the shift from early family-based approaches to association-based design, and the rapidly advancement of enabling genotyping and sequencing technologies. In addition, this review points to the gaps in the current knowledge base, including the general lack of GWAS and whole-genome sequence analyses particularly understudied populations, and the need for large-scale collaborative effort in both observational and experimental settings. In this review, we also call for research utilizing systems biology strategies for PA genetic research and accounting for complex gene-environment interactions that may vary by race/ethnicity.

Summary The epidemic of physical inactivity has been a public health nemesis, encompassing a large burden of diseases and high societal costs. A better understanding of the genetic basis of PA can inform public health policies for the prevention, control, and treatment of many chronic diseases related to physical inactivity.

Keywords Physical activity · Genetics · Gene-environment interaction · Lifestyle intervention · Cardio-metabolic health

✉ Simin Liu
Simin_liu@brown.edu

1   Department of Epidemiology, Brown University, Providence, RI, USA

2   Center for Global Cardio-metabolic Health, Brown University, Providence, RI, USA

3   Department of Family Medicine, The Warren Alpert Medical School of Brown University, Providence, RI, USA

4   Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

5   Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

6   Division of Endocrinology, Department of Medicine, The Warren Alpert Medical School of Brown University, Providence, RI, USA

7   Department of Endocrinology, Guangdong General Hospital, Guangzhou, China

8   Department of Epidemiology and Medicine, Brown University, 121 South Main St, Providence, RI 02903, USA

Introduction

Physical activity (PA) is a well-established and modifiable lifestyle determinant for multiple cardio-metabolic outcomes, including obesity, type 2 diabetes (T2D), and cardiovascular diseases (CVD) [1]. The vital role of physical activity in regulating cardio-metabolic health is well-documented [2].

Globalization of unhealthy lifestyles, particularly increasingly sedentary behavior, has undoubtedly contributed to the global pandemic of cardio-metabolic disorders. It is estimated that approximately 3.2 million deaths worldwide each year are attributable to insufficient physical activity. In the United State

(US) alone, about 50.8% of all adults do not meet the Physical Activity Guidelines for aerobic activity, and only 20.8% meet the guidelines for both aerobic and muscle-strengthening activity in 2015 [3]. Within this context, approximately 35% of coronary heart disease (CHD) mortality that has been attributed to the lack of physical activity [4], leading to the estimated $117 billion health care expenditures annually in the US [5].

Interventions to enhance PA have met with modest results, at least partially attributable to physiological and metabolic constraints, facilities, built-in environment, economics, and policies. These challenges underscore the need for well-designed preventive measures, informed by personalized precision medicine to understand the etiology of physical inactivity as it relates to cardio-metabolic diseases. While many psychological, biological, social, and environmental correlates have been identified, current understanding of the genetic architecture contributing to PA is very limited, especially when compared to other phenotypes such as obesity and diabetes. Moreover, relative to populations of European ancestry, other ethnic groups are disproportionately affected by physical inactivity and yet underrepresented in almost all genetic studies.

Driven by the promises of genetic epidemiology to unravel the complex genetic basis of PA, various studies have attempted not only to identify the presence of a genetic component but also to target specific candidate loci. More recent studies on the genome-wide scale have provided additional insights into the genetic architecture underlying human PA. In spite of those efforts, three key questions about PA genetics are of particular interest: (1) whether there is a genetic component underlying human PA, (2) to what extent PA levels are affected by the genetic factors, and (3) exactly, what genetic factors and biological pathways are involved in determining PA levels. With the emerging techniques of genotyping, imputation, and next-generation sequencing, along with the increasing resources allocated to genetic studies, the ability to investigate PA genetics has substantially improved, and there is an urgent need for timely information from a comprehensive and in-depth review of the research on PA genetics.

Evidence from Animals

The role of genetic factors in regulating PA levels has long been postulated based on findings from animal models [6–8].

Knab et al. found that the expression of the dopamine 1 receptor gene (Drd1) was suppressed in the brains of the highly active inbred mice (C57L/J), compared to the less active inbred mice (C3H/ HeJ) [9]. The same research team conducted a follow-up study and noted that the administration of a specific D1-receptor agonist (SKF-81297) caused substantial reduction in PA levels of highly active mice, which confirmed

their previous findings and suggested that regulating the D1 receptor pathway has direct impact on PA levels [9].

Nescient helix-loop-helix-2 (Nhlh2) is a gene shown to transcriptionally regulate prohormone convertases in addition to numerous other neuropeptide genes [10]. It exhibits partial haplotype differences between mice models with varying PA levels and localizes with a previously known PA-related quantitative trait locus (QTL) on chromosome 3 (46 cM) [11]. These findings have been corroborated by DJ Good et al. who found that Nhlh2 knockout mice (N2KO), a line of obese mouse model, have substantially reduced PA levels with no overt differences in food intake [12]. Subsequently, a mechanistic model was proposed to illustrate how the acetylation status of NHLH2 influences the expression of the monoamine oxidase A (MAO-A) gene, and then the levels of MAO-A enzyme, which regulate PA levels through the dopaminergic pathway [12].

The glucose transporter-4 gene (GLUT-4), also known as the solute carrier family 2 facilitated glucose transporter, member 4 gene (SLC2A4), is critical in the regulation of glucose uptake in skeletal muscle and the maintenance of glucose homeostasis. The elevated expression levels of GLUT-4 in the skeletal muscles of mice, due to the genetic selection history, is associated with glycogen depletion, which may limit the amount of carbohydrates locally available to sustain muscle contraction [13]. Transgenic mouse model overexpressing levels of GLUT-4 selectively in fast-twitch skeletal muscles had substantially higher levels of PA compared to control mice [14]. Another study comparing selectively bred highly active mice vs. to control mice reported about 2.5 times higher levels of the glucose transporter-4, along with higher glycogen and glycogen-synthase activity, among the highly active mice [13]. These findings were further supported by the fact that Slc2a4 is actually located in two previously established PA-related QTL on chromosome 11 (40 cM) and a mini-muscle phenotype-related QTL, which were identified from linkage studies in mice [15•].

Animal models can be viewed as a practical alternative resource when investigating PA. As noted above, animal research strongly suggests the presence of a heritable component in the determination of one's degree of PA. Nevertheless, we should exercise caution when generalizing findings from mice to humans. Undoubtedly, evidence from animal models requires further scrutiny and validation through research in humans.

Evidence from Parent-Offspring Studies, Twin Studies, and Linkage Studies

Both parent-offspring studies and twin studies have been widely used to decompose familial resemblance in PA into environmental influences and genetic influences. The

estimates of heritability reported by family studies have been in the range of 0.16 ~ 0.29 [16–19], although a few studies with objectively measured PA levels have reported higher levels of heritability [20, 21], suggesting that heritability may be underestimated when PA levels are self-reported. Twin studies have also been extensively used to decompose familial resemblance and to estimate the heritability of PA. Twin studies have the advantage of better distinguishing the impact of shared vs. unique environmental influences and have produced generally higher heritability estimates of PA compared to parent-offspring studies ranging from 35 to 83% [22]. Although family-based studies of shared phenotypes are invaluable in estimating the genetic component underlying human PA, they cannot by themselves provide insight on the underlying genetic mechanisms at play. In addition, environmental factors are harder to control in humans compared to animal models and sample sizes in many of the early family-based studies were small.

Early linkage studies extended the family-based study design. By demonstrating cosegregation of a trait with microsatellite markers spread out evenly across the genome, linkage studies map variability of a trait to a genomic location [23, 24]. An early genome-wide linkage scan based on the Quebec Family study showed promising evidence suggested linkage between chromosome 2p22-p16 and PA [25] (Table 1). Two other linkage studies, based on data from the VIVA LA FAMILIA Study and the Netherlands Twin Registry, also found suggestive evidence at 18q and 19p13.3 [26, 27] (Table 1). While these linkage studies have shed some light on the location of causal genetic variants for PA, they are hampered by poor resolution of the initial genetic signal and have limited power to detect modest effects [23–27]. Given most traits are complex, featuring hundreds or thousands of variants contributing modest effects, it is perhaps not surprising that linkage studies have not been that helpful for PA.

Evidence for Candidate Genes

More recent studies have employed association-based methods including candidate gene analyses and genome-wide association studies (GWAS). Several candidate genes have been proposed based on functional relevance and evidence from animal experiments, parent-offspring studies, and twin studies (Table 2).

LEPR, which has previously been related to obesity, was found associated with PA phenotypes in Pima Indians [28] and prepubertal boys [29]. Among Pima Indians, carriers of the Arg223-encoding allele of the Gln223Arg polymorphism (rs1137101) were found to have an approximately ~ 200 kcal lower 24-h energy expenditure (EE) per day compared to non-carriers ($P = 0.01$) [28]. The same polymorphism in *LEPR* was found associated with PA levels measured by PA-related energy expenditure in a group of growing prepubertal boys of European ancestry (GlnGln 240.5 ± 11.8 kcal/day vs. GlnArg 266.6 ± 14.3 kcal/day vs. ArgArg 180.6 ± 21.0 kcal/day; $P = 0.02$) [29]. In a subsequent replication study, De Moor et al. found that the P value for the rs12405556 SNP in the *LEPR* gene is 9.70×10^{-4} [30••].

The calcium-sensing receptor (*CASR*) gene is involved in calcium homeostasis and regulates bone mass density and serum calcium concentrations [31]. Lorentzon et al., observed ~1.4 h per week lower PA levels among adolescent girls carrying the serine genotype at the A986S polymorphism (rs1801725)($P = 0.01$) [32].

The 3′-phosphoadenosine 5′-phosphosulfate synthase (*PAPSS2*) gene is also a promising PA-determining gene. This gene is widely expressed in muscles and appears to be involved in initial skeletal development, which may influence the capacity to perform physical activity in later life [33, 34]. The functional relevance of *PAPSS2* is also corroborated by evidence from mouse models showing that the homologous mouse gene exhibits partial differences in haplotype structure between a highly active inbred strain of mice (C57L/J) and the control mice (C3H-HeJ). Lastly, the chromosome region (10q23) harboring this gene has been linked to maximal exercise capacity in a previous genome-wide linkage study [35], and the only GWAS that has been conducted to date using logistic regression on PA status (PA-related energy expenditure ≤ 4 MET-h/week vs. > 4 MET-h/week) also supports a role of this gene in determining PA levels (rs10887741; effect allele, T; OR, 1.32; 95% CI, 1.17–1.49; pooled $P = 3.81 \times 10^{-6}$; pooled P adjusted for BMI $= 6.26 \times 10^{-6}$) [30••].

Another gene involved in the dopaminergic pathway, the dopamine 2 receptor (*DRD2*) gene was also found associated with multiple PA phenotypes [36]. This candidate gene study showed that a fragment length polymorphism in exon 6 (rs6275) of *DRD2* was associated with time spent in PA (C/C 1.24 ± 0.03 vs. C/T 1.33 ± 0.06 vs. T/T 0.99 ± 0.12; $P = 0.02$) among women from two family-based cohorts. Ethnicity-specific analysis suggested that this polymorphism was also associated with levels of sports (C/C 1.65 ± 0.06 vs. C/T 1.77 ± 0.10 vs. T/T 1.28 ± 0.15; $P = 0.02$) and occupational PA (C/C 2.28 ± 0.13 vs. C/T 2.38 ± 0.18 vs. T/T 1.79 ± 0.21; $P = 0.004$) among white women. The investigators speculated that the *DRD2*-risk allele carriers tend to have lower PA levels because of impaired or subtly reduced motor skills induced by dopamine receptor deficiencies or alterations, and thus are more prone to physical inactivity.

Evidence from a series of candidate gene studies suggests that the angiotensin-converting enzyme (*ACE*) may also contribute to PA levels in humans. An insertion/deletion polymorphism (rs1799752) in *ACE* determines the expression of *ACE* in human skeletal muscle which in turn regulates the levels of serum and tissue *ACE*, the production of angiotensin II, and the

Table 1   Evidence from linkage studies for human physical activity

Author	Year	Country	No.	Instrument	Phenotype	Locus	Genetic marker	P
Simonen et al.	2003	USA	767 white adults from 207 families	Questionnaire	Time spent on PA	11p15	C11P15_3	0.0089
						15q13.3	D15S165	0.009
				3-day PA diary	Degree of inactivity	2p22-p16	D2S2347	0.0012
						7p11.2	IGFBP1	0.0046
						20q13.1	PLC1	0.0074
					Total PA	13q22-q31	D13S317	0.0029
					Moderate to vigorous PA	4q28.2	UCP1	0.005
						7p11.2	IGFBP1	0.006
						9q31.1	D9S938	0.0028
						13a22-q31	D13S317	0.0067
Rico-Sanz et al.	2004	USA	351 white sibling pairs		VO2max in sedentary state	11p15.1	SUR	0.0014
					MPO in sedentary state	10q23	D10S677	0.0014
						13q33	D13S796	0.0098
						18q11.1	D18S1107	0.0095
						18q11.2	D18S866	0.0077
						18q12	D18S478	0.0064
			102 black sibling pairs		VO2max in sedentary state	1p31	LEPR	0.001
						7q32	LEPMSAT	0.0068
						7q36	D7S3070	0.0083
						7q36	NOS3	0.003
			329 white sibling pairs		VO2max in response to training	4q27	FABP2	0.0086
						7q34	D7S495	0.0089
						13q12	D13S787	0.0087
						13q12	D13S1243	0.0088
					MPO in response to training	1q21	S100STUI	0.0091
						4p15.1	D4S2397	0.0075
						4p13	D4S1627	0.0062
						5q23	D5S1505	0.002
			90 black sibling pairs		VO2max in response to training	1p31	LEPR	0.0017
						1p31	D1S198	0.01
						16q22	RAD1	0.0041
						16q22	D16S2624	0.01
						20q13.1	D20S857	0.0028
						20q13.1	D20S839	0.0031
					MPO in response to training	1q22	D1S398	0.0033
						13q11	D13S175	0.0055
Cai et al.	2006	USA	1030 siblings from 319 Hispanic families	Accelerometer	Total PA	18q	D18S64	7.00E-13
					Sedentary activity	18q	D18S1102-D18S474	4.00E-14
					Light PA	18q	D18S1102-D18S474	6.00E-10
					Moderate PA	18q	D18S64	4.00E-10
					Vigorous PA	18q	D18S64	0.01
De Moor et al.	2007	Netherlands	1738 MZ twins, DZ twins, and siblings from 622 families	Questionnaire	Exercise participation	19p13.3	D19S247	0.000766
De Moor et al.	2007	Great Britain	700 pairs of female DZ twins	Questionnaire	Athlete status	3q24	D3S1569	0.000501
						4q32.3	D4S1597	0.00167

PA physical activity, VO2max maximal oxygen uptake, MPO maximal power output, MZ monozygotic, DZ dizygotic

half-life of the bradykinin [37]. This same 287-base-pair insertion has also been linked to better endurance performance among elite distance runner [38], rowers [39], and mountaineers [40]. Thereafter, a study examining training-related changes in the mechanical efficiency of human skeletal muscles found that the presence of ACE insertion conferred an enhanced mechanical efficiency in trained muscles [41]. Lastly, a more recent candidate gene analysis reported that this polymorphism was also associated with the PA status in 355 untreated stage I hypertensives (inactive, 13.7% II genotype, 53.7% ID

genotype, 32.6% DD genotype; active, 26.6% II genotype, 55.4% ID genotype, 18% DD genotype; $P = 0.001$) [42].

The melanocortin-4 receptor (MC4R) gene has been reproducibly associated with obesity and other anthropometric traits [43, 44]. This gene is highly expressed in the hypothalamic area and is known to act within the central nervous system profoundly influencing energy expenditure, body weight, and other human behaviors [45]. Interestingly, it is located in a region on chromosome 18, which was also identified by a linkage study for PA [26]. In a previous candidate

Table 2 Candidate genes for human physical activity with functional relevancy and/or association evidence

Gene	Locus	Functional relevance	Association evidence	Author	Genetic marker	Phenotype	P
With association evidence							
LEPR	1p31.3	Energy homeostasis	Yes	Stefan et al.	rs1137101	PA-related energy expenditure	0.01
			Yes	Richert et al.	rs1137101	PA-related energy expenditure	0.02
			Suggestive	De Moor et al.[1]	rs12405556	PA-related energy expenditure ≤ 4 MET-h/week vs. > 4 MET-h/week	9.70E-04
CASR	3q13.33	Mineral ion homeostasis	Yes	Lorentzon et al.	rs1801725	Hours/week spent on PA	0.01
PAPSS2	10q23.2	Musculoskeletal development and function	Yes	De Moor et al.[1]	rs10887741	PA-related energy expenditure ≤ 4 MET-h/week vs. > 4 MET-h/week	3.81E-06
DRD2	11q23.2	Dopaminergic pathway	Yes	Simonen et al.	rs6275	Time spent on PA	0.02
						Sports score	0.02
						Occupational PA score	0.004
GABRG3	15q12	Dopaminergic pathway	Suggestive	De Moor et al.[1]	rs8036270	PA-related energy expenditure ≤ 4 MET-h/week vs. > 4 MET-h/week	4.61E-05
ACE	17q23.3	Musculoskeletal performance	Yes	Winnicki et al.	rs1799752	PA status	0.001
MC4R	18q21.32	Dopaminergic pathway	Yes	Loo et al.	MC4R-C-2745T	Moderate to strenuous PA	0.005
						PA status	0.01
						Time spent on PA	0.005
Without association evidence							
NHLH2	1p13.1	Dopaminergic pathway	No				
ACTN2/ACTN3	1q43/11q13.2	Musculoskeletal development and function	No				
DRD1	5q35.2	Dopaminergic pathway	No				
CYP19A1	15q21.2	Sex hormone profiles	No				
SLC2A4	17p13.1	Glucose uptake in skeletal muscle and glucose homeostasis	No				

PA physical activity, *MET* metabolic equivalent

[1] This is the first and only GWAS to date

gene analysis, Loos et al. found that the *MC4R-C-2745T* polymorphism was associated with moderate to strenuous PA ($P = 0.005$), PA status ($P = 0.01$), and time spent on PA during the past year ($P = 0.005$) in a combined sample of women and men [46].

Several additional genes may have been implicated in the regulation of human PA. For example, the gamma-aminobutyric acid type A receptor-gamma 3 (*GABRG3*) gene is located in a genomic region in chromosome 15, which has been previously linked to PA levels in a linkage study [25]. In addition, higher expression levels of this gene were observed after exhaustive exercise [47]. The potential role of *GABRG3* in the regulation of human PA was also corroborated by the suggestive evidence from the GWAS [30••]. Genes that encode muscle structure proteins have also been considered PA-determining genes. The alpha-actin 2 (*ACTN2*) and the alpha-actin 3 (*ACTN3*) genes encodes two structural proteins that are associated with the size of muscles and the integrity of muscle fibers and have been further implicated in the regulation of contraction and degree of muscle strength [48–51]. An impact of sex hormones on

PA levels has been previously observed [52], possibly mediated through the modulation of the androgen to estrogen ratio by aromatase. Interestingly, the cytochrome P450 family 19 subfamily A member 1 (*CYP19A1*) gene, which encodes aromatase, co-localizes with an established mouse PA-related QTL [53]. Although no quantitative trait loci (QTL) have been identified for PA in humans to date for this locus [54], strong evidence from animal models supports *CYP19A1* as a PA-determining gene [52]. As discussed above, the functional relevance of the *NHLH2*, *DRD1*, and *SLC2A4* has also been supported by evidence from animal models. However, additional association-based investigations are needed to confirm their roles in the regulation of human PA and to identify the causal variants.

Candidate genes for PA proposed to date appear to primarily involve pathways related to homeostasis, dopamine, and development and function of the musculoskeletal system (Table 2). Overall, the findings point to an important role of the nervous system as well as structural and metabolic changes in the muscle in influencing both the degree and the capacity to actually perform PA (Fig. 1).

Fig. 1 Potential mechanisms that regulate human physical activity. The engagement of homeostasis (solid line) coupled with dopaminergic reward pathway (dotted line) may be responsible for the variation in human physical activity
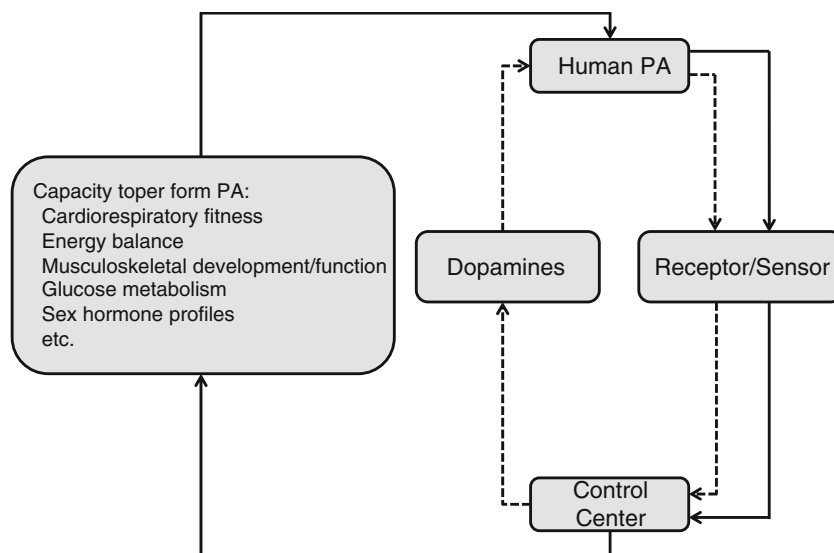
Evidence from GWAS

Candidate gene studies have led to limited findings given the biological pathways leading to human PA remain largely unknown. GWAS can overcome the challenge of candidate gene studies by embracing an unbiased approach to discovery through association.

Only one GWAS to date has been conducted for human PA (Table 2). In 2009, De Moor et al. performed the first GWAS of PA based on a HapMap Imputation dataset with $\sim 2.5 \times 10^6$ SNPs [30••]. By using a threshold of pooled P values $< 1.0 \times 10^{-5}$, the investigators identified a total of 37 novel SNPs in *PAPSS2* and two intergenic regions (2q33.1 and 18p11.32). The lead SNP in *PAPSS2*, rs10887741, appeared to be associated with the regular participation in leisure time exercise (60 min per week; OR 1.32, 95% CI 1.17–1.49; pooled $P = 3.81 \times 10^{-6}$; pooled P adjusted for BMI $= 6.26 \times 10^{-6}$). As aforementioned, evidence from animal models [11] and a previous genome-wide linkage study [35] support *PAPSS2*'s candidacy as a determinant of PA levels and physical fitness, through effects on the development of the musculoskeletal system [33, 34].

Despite these findings, we note that the first GWAS for PA has two substantial limitations. First, the study overall had limited power to detect causal variants with modest effects (Dutch: $n = 1644$; US: $n = 978$). Second, the study population was restricted to adults of European ancestry. Thus, the findings may not be generalizable to other race/ethnic group who are disproportionately affected by the consequence of physical inactivity and concurrently underrepresented in genetic studies. Nevertheless, the first GWAS for human PA represents a critical initial step to galvanize future efforts to understand the genetic basis of human PA.

Current GAPS and Future Directions

Multiple lines of evidence from animal models, parent-offspring studies, twin studies, candidate gene analyses, and GWAS suggest a genetic component underlying human PA. We hasten to note, however, that the other two key questions proposed in the beginning of this review (to what extent PA levels are affected by the genetic factors; exactly, what genetic factors and biological pathways are involved in determining PA levels) remain largely unanswered.

For the past few decades, genetic studies on human physical activity have mainly focused on testing a small group of markers from genomic regions of possible functional relevance or implicated in previous studies. We now know that a majority of these studies were grossly underpowered and restricted in discovery by low-resolution mapping techniques [15•, 22]. Furthermore, the initial explorations of PA genetics were conducted in family-based setting, which, despite its own advantages, heavily relies on the inheritance structure of the pedigrees, yields only positional candidates, and is not effective in detecting genetic markers with subtle effects [55].

As we shift to association-based methods and larger sample sizes, we are now beginning to see GWAS for human PA with large panels of SNPs [30••]. This shift has helped the field of PA genetics to form and evolve, but significant gaps in evidence of PA genetics remain.

In addition to identifying genetic variants across the genome associated with PA through GWAS, it is important for investigators to overlay advanced systems biology-based methodologies to findings from GWAS. Such methods will likely further enhance the statistical efficacy to detect the regulatory genetic modules for human PA clustered in pathways and networks with biologically plausible connections [22]. Advanced pathway and network analyses have proven to be

effective and efficient approaches to integrate data from GWAS and functional genomics in studying perturbations in biological pathways and gene networks in the etiology of other phenotypes and diseases, such as T2D and CVD [56]. Therefore, it is natural and logical to consider the possibility of applying such methodologies to the investigations of PA genetics.

Demonstrated in earlier parent-offspring studies and twin studies, the variation in human PA levels is most likely to be affected by the joint impact of both genetic and environmental factors (Fig. 2). While it is understandable that the field of PA genetics is currently focused on identifying causal variants with main effects that drive the variation in human PA, investigators should also not forget the potential of gene-environment interactions. Utilizing well-characterized environmental data from already established large-scale cohorts [57–59] provide a unique opportunity to identify mediating effects of environmental factors on the genetics-PA relation. Such gene-environment interaction studies will be essential to the further clarification of the mechanistic paradigm for the regulation of human PA. In addition, current evidence, without accounting for the complex gene-environment interaction, is primarily based on European-derived populations, which limits the generalizability and applicability of findings to minority populations.

Recent advances in GWAS and next generation sequencing have led to the identification of dozens of genetic variants that together could explain a significant amount of heritability for other phenotypes, such as body mass index (BMI) and T2D [60, 61]. Unfortunately, as pointed out earlier, there is generally a lack of evidence on PA from GWAS and GWAS-based meta-analyses compared to other phenotypes and diseases. Thus, it is not surprising that there has been a plea for large

GWAS and collaborative efforts to help unravel the genetic pathways that affect this important health-enhancing behavior [62].

Rapidly evolving portable technology has made it possible to further explore the genetic basis of PA in a more in-depth and comprehensive manner. Wearable devices for monitoring and tracking PA have become increasingly available and affordable, which provide a unique perspective and opportunity to investigate PA genetics using objective PA measures in large study populations [63, 64]. The availability of whole-genome sequencing (WGS) data for PA genetic studies will undoubtedly shed additional light on its genetic component, especially those rare variants that cannot be captured by standard GWAS. The Trans-Omics for Precision Medicine (TOPMed) program, as part of the Precision Medicine Initiative sponsored by the National Institute of Health (NIH), is an ongoing project that seeks to uncover disease risk factors and develop more targeted and personalized interventions by integrating WGS and other Trans-Omics data with molecular, behavioral, imaging, environmental, and clinical data [65]. In addition, as the field of PA genetics keeps evolving, it is essential to further improve our understanding of the molecular changes linking genetic factors to human PA behavior and then to the development of chronic diseases related from insufficient PA. The NIH sponsored Molecular Transducers of Physical Activity in Humans Consortium (MoTrPAC) [66], among other collaborative initiatives, aims to assemble a systematic and comprehensive molecular map detailing the molecular signals and mechanisms that transmit the health benefits of PA and construct a user-friendly database that can be used by clinicians and researchers to develop and test specific hypotheses. PA phenotyping utilizing wearable devices and resources like the TOPMed and MoTrPAC

Fig. 2 Mechanistic paradigm of the gene-environment interactions for human physical activity. Environmental factors may mediate and interact with genetic effects in the regulation of human physical activity
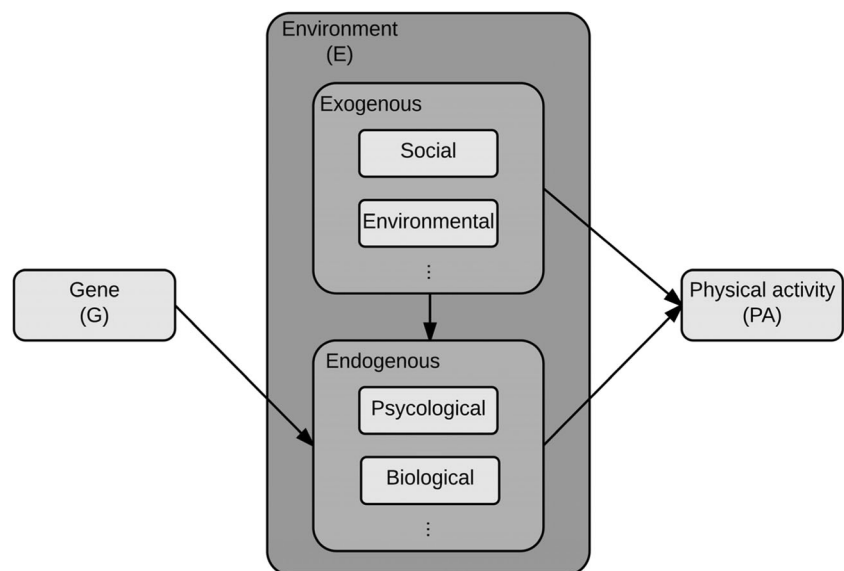
will allow the investigators to integrate genetic data on both common and rare variants, molecular signals and pathways, and multidimensional phenotypic and demographic data, to inform public health interventions, clinical practice, and the development of related policies and guidelines.

The global pandemics of physical inactivity and related chronic diseases are not only public health issues but also economic issues. In order to alleviate the tremendous health and economic burden associated with physical inactivity, a new generation of PA genetic research is urgently needed. The aforementioned advances are poised to revolutionize the field of PA genetic studies, and to provide pivotal insight into the genetic factors driving humans to perform PA. Future better-designed preventive strategies, informed by PA genetics, systems biology, bioinformatics, and ultimately "precision medicine," are expected to have major implications for the prevention, control, and treatment of chronic diseases related to physical inactivity.

Compliance with Ethical Standards

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1.   U.S. Department of Health and Human Services. Physical activity guidelines advisory committee report. Washington, DC: US Department of Health and Human Services; 2008. p. 2008.

2.   Lin X, Zhang X, Guo J, Roberts CK, McKenzie S, Wu WC et al. Effects of exercise training on cardiorespiratory fitness and biomarkers of cardiometabolic health: a systematic review and meta-analysis of randomized controlled trials. J Am Heart Assoc. 2015;4(7):e002014.

3.   Ma X, Ran Y, Gong J, Chen D. Sequential accelerated solvent extraction of polycyclic aromatic hydrocarbons with different solvents: performance and implication. J Environ Qual. 2010;39(6):2072–9.

4.   Ma PF, Xing H, Liao LJ, Chen B, Zhao QB, Quan Y, et al. Feasibility of using dried blood spots to detect HIV drug resistance genotyping. Zhonghua Yu Fang Yi Xue Za Zhi [Chinese J Prev Med]. 2010;44(11):993–8.

5.   Carlson SA, Fulton JE, Pratt M, Yang Z, Adams EK. Inadequate physical activity and health care expenditures in the United States. Prog Cardiovasc Dis. 2015;57(4):315–23. https://doi.org/10.1016/j.pcad.2014.08.002.

6.   Festing MF. Wheel activity in 26 strains of mouse. Lab Anim. 1977;11(4):257–8. https://doi.org/10.1258/002367777780936530.

7.   Lerman I, Harrison BC, Freeman K, Hewett TE, Allen DL, Robbins J, et al. Genetic variability in forced and voluntary endurance exercise performance in seven inbred mouse strains. J Appl Physiol (1985). 2002;92(6):2245–55. https://doi.org/10.1152/japplphysiol.01045.2001.

8.   Lightfoot JT, Turner MJ, Daves M, Vordermark A, Kleeberger SR. Genetic influence on daily wheel running activity level. Physiol Genomics. 2004;19(3):270–6. https://doi.org/10.1152/physiolgenomics.00125.2004.

9.   Knab AM, Bowen RS, Hamilton AT, Gulledge AA, Lightfoot JT. Altered dopaminergic profiles: implications for the regulation of voluntary physical activity. Behav Brain Res. 2009;204(1):147–52. https://doi.org/10.1016/j.bbr.2009.05.034.

10. Fox DL, Vella KR, Good DJ. Energy balance pathways converging on the Nhlh2 transcription factor. Front Biosci. 2007;12:3983–93.

11. Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, et al. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature. 2007;448(7157):1050–3. https://doi.org/10.1038/nature06067.

12. Good DJ, Li M, Deater-Deckard K. A genetic basis for motivated exercise. Exerc Sport Sci Rev. 2015;43(4):231–7. https://doi.org/10.1249/JES.0000000000000057.

13. Gomes FR, Rezende EL, Malisch JL, Lee SK, Rivas DA, Kelly SA, et al. Glycogen storage and muscle glucose transporters (GLUT-4) of mice selectively bred for high voluntary wheel running. J Exp Biol. 2009;212(Pt 2):238–48. https://doi.org/10.1242/jeb.025296.

14. Tsao TS, Li J, Chang KS, Stenbit AE, Galuska D, Anderson JE, et al. Metabolic adaptations in skeletal muscle overexpressing GLUT4: effects on muscle and physical activity. FASEB J. 2001;15(6):958–69.

15.• Roth SM, Thomis MA. Fundamental concepts in exercise genomics. Exercise Genomics. Berlin: Springer; 2011. p. 1–22. **This reference provides a thorough and comprehensive review of exercise genomics.**

16. Perusse L, Tremblay A, Leblanc C, Bouchard C. Genetic and environmental influences on level of habitual physical activity and exercise participation. Am J Epidemiol. 1989;129(5):1012–22.

17. Simonen RL, Perusse L, Rankinen T, Rice T, Rao DC, Bouchard C. Familial aggregation of physical activity levels in the Quebec Family Study. Med Sci Sports Exerc. 2002;34(7):1137–42.

18. Sallis JF, Patterson TL, Buono MJ, Atkins CJ, Nader PR. Aggregation of physical activity habits in Mexican-American and Anglo families. J Behav Med. 1988;11(1):31–41.

19. Choh AC, Demerath EW, Lee M, Williams KD, Towne B, Siervogel RM, et al. Genetic analysis of self-reported physical activity and adiposity: the Southwest Ohio Family Study. Public Health Nutr. 2009;12(8):1052–60. https://doi.org/10.1017/S1368980008003583.

20. Butte NF, Cai G, Cole SA, Comuzzie AG. Viva la Familia Study: genetic and environmental contributions to childhood obesity and its comorbidities in the Hispanic population. Am J Clin Nutr. 2006;84(3):646–54. quiz 73–4

21. Moore LL, Lombardi DA, White MJ, Campbell JL, Oliveria SA, Ellison RC. Influence of parents' physical activity levels on activity levels of young children. J Pediatr. 1991;118(2):215–9.

22. De Geus EJ, De Moor MH. Genes, exercise, and psychological factors. In: Genetic and molecular aspects of sport performance. Oxford: Blackwell Publishing; 2011. p. 294–305.

23. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996;273(5281):1516–7.

24. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genome-wide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet. 2001;69(5):936–50. https://doi.org/10.1086/324069.

25. Simonen RL, Rankinen T, Perusse L, Rice T, Rao DC, Chagnon Y, et al. Genome-wide linkage scan for physical activity levels in the Quebec Family Study. Med Sci Sports Exerc. 2003;35(8):1355–9. https://doi.org/10.1249/01.MSS.0000078937.22939.7E.

26. Cai G, Cole SA, Butte N, Bacino C, Diego V, Tan K, et al. A quantitative trait locus on chromosome 18q for physical activity and dietary intake in Hispanic children. Obesity (Silver Spring). 2006;14(9):1596–604. https://doi.org/10.1038/oby.2006.184.

27. De Moor MH, Posthuma D, Hottenga JJ, Willemsen G, Boomsma DI, De Geus EJ. Genome-wide linkage scan for exercise participation in Dutch sibling pairs. Eur J Hum Genet. 2007;15(12):1252–9. https://doi.org/10.1038/sj.ejhg.5201907.

28. Stefan N, Vozarova B, Del Parigi A, Ossowski V, Thompson DB, Hanson RL, et al. 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. 2002;26(12):1629–32. https://doi.org/10.1038/sj.ijo.0802161.

29. Richert L, Chevalley T, Manen D, Bonjour JP, Rizzoli R, Ferrari S. Bone mass in prepubertal boys is associated with a Gln223Arg amino acid substitution in the leptin receptor. J Clin Endocrinol Metab. 2007;92(11):4380–6. https://doi.org/10.1210/jc.2007-0932.

30.•• De Moor MH, Liu YJ, Boomsma DI, Li J, Hamilton JJ, Hottenga JJ, et al. Genome-wide association study of exercise behavior in Dutch and American adults. Med Sci Sports Exerc. 2009;41(10): 1887–95. https://doi.org/10.1249/MSS.0b013e3181a2f646. **This is the only GWAS for PA that has been published to date.**

31. Cole DE, Peltekova VD, Rubin LA, Hawker GA, Vieth R, Liew CC, et al. A986S polymorphism of the calcium-sensing receptor and circulating calcium concentrations. Lancet. 1999;353(9147): 112–5.

32. Lorentzon M, Lorentzon R, Lerner UH, Nordstrom P. Calcium sensing receptor gene polymorphism, circulating calcium concentrations and bone mineral density in healthy adolescent girls. Eur J Endocrinol. 2001;144(3):257–61.

33. Iida A, Simsek-Kiper PO, Mizumoto S, Hoshino T, Elcioglu N, Horemuzova E, et al. Clinical and radiographic features of the autosomal recessive form of brachyolmia caused by PAPSS2 mutations. Hum Mutat. 2013;34(10):1381–6. https://doi.org/10.1002/humu.22377.

34. Yerges-Armstrong LM, Chai S, O'Connell JR, Curran JE, Blangero J, Mitchell BD, et al. Gene expression differences between offspring of long-lived individuals and controls in candidate longevity regions: evidence for PAPSS2 as a longevity gene. J Gerontol A Biol Sci Med Sci. 2016;71(10):1295–9. https://doi.org/10.1093/gerona/glv212.

35. Rico-Sanz J, Rankinen T, Rice T, Leon AS, Skinner JS, Wilmore JH, et al. Quantitative trait loci for maximal exercise capacity phenotypes and their responses to training in the HERITAGE Family Study. Physiol Genomics. 2004;16(2):256–60. https://doi.org/10.1152/physiolgenomics.00035.2003.

36. Simonen RL, Rankinen T, Perusse L, Leon AS, Skinner JS, Wilmore JH, et al. A dopamine D2 receptor gene polymorphism and physical activity in two family studies. Physiol Behav. 2003;78(4–5):751–7.

37. Brown NJ, Blais C Jr, Gandhi SK, Adam A. ACE insertion/deletion genotype affects bradykinin metabolism. J Cardiovasc Pharmacol. 1998;32(3):373–7.

38. Myerson S, Hemingway H, Budget R, Martin J, Humphries S, Montgomery H. Human angiotensin I-converting enzyme gene and endurance performance. J Appl Physiol (1985). 1999;87(4): 1313–6.

39. Gayagay G, Yu B, Hambly B, Boston T, Hahn A, Celermajer DS, et al. Elite endurance athletes and the ACE I allele—the role of genes in athletic performance. Hum Genet. 1998;103(1):48–50.

40. Montgomery HE, Marshall R, Hemingway H, Myerson S, Clarkson P, Dollery C, et al. Human gene for physical performance. Nature. 1998;393(6682):221–2. https://doi.org/10.1038/30374.

41. Williams AG, Rayson MP, Jubb M, World M, Woods DR, Hayward M, et al. The ACE gene and muscle performance. Nature. 2000;403(6770):614. https://doi.org/10.1038/35001141.

42. Winnicki M, Accurso V, Hoffmann M, Pawlowski R, Dorigatti F, Santonastaso M, et al. Physical activity and angiotensin-converting enzyme gene polymorphism in mild hypertensives. Am J Med Genet A. 2004;125A(1):38–44. https://doi.org/10.1002/ajmg.a.20434.

43. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat Genet. 2008;40(6):768–75. https://doi.org/10.1038/ng.140.

44. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009;41(1):25–34. https://doi.org/10.1038/ng.287.

45. Berthoud HR, Morrison C. The brain, appetite, and obesity. Annu Rev Psychol. 2008;59:55–92. https://doi.org/10.1146/annurev.psych.59.103006.093551.

46. Loos RJ, Rankinen T, Tremblay A, Perusse L, Chagnon Y, Bouchard C. Melanocortin-4 receptor gene and physical activity in the Quebec Family Study. Int J Obes. 2005;29(4):420–8. https://doi.org/10.1038/sj.ijo.0802869.

47. Kawai T, Morita K, Masuda K, Nishida K, Sekiyama A, Teshima-Kondo S, et al. Physical exercise-associated gene expression signatures in peripheral blood. Clin J Sport Med. 2007;17(5):375–83. https://doi.org/10.1097/JSM.0b013e31814c3e4f.

48. Delmonico MJ, Kostek MC, Doldo NA, Hand BD, Walsh S, Conway JM, et al. Alpha-actinin-3 (ACTN3) R577X polymorphism influences knee extensor peak power response to strength training in older men and women. The journals of gerontology Series A. Biol Sci Med Sci. 2007;62(2):206–12.

49. MacArthur DG, North KN. ACTN3: a genetic influence on muscle function and athletic performance. Exerc Sport Sci Rev. 2007;35(1):30–4. https://doi.org/10.1097/JES.0b013e31802d8874.

50. Norman B, Esbjornsson M, Rundqvist H, Osterlund T, von Walden F, Tesch PA. Strength, power, fiber types, and mRNA expression in trained men and women with different ACTN3 R577X genotypes. J Appl Physiol (1985). 2009;106(3):959–65. https://doi.org/10.1152/japplphysiol.91435.2008.

51. Clarkson PM, Devaney JM, Gordish-Dressman H, Thompson PD, Hubal MJ, Urso M, et al. ACTN3 genotype is associated with increases in muscle strength in response to resistance training in women. J Appl Physiol (1985). 2005;99(1):154–63. https://doi.org/10.1152/japplphysiol.01139.2004.

52. Lightfoot JT. Sex hormones' regulation of rodent physical activity: a review. Int J Biol Sci. 2008;4(3):126–32.

53. Lightfoot JT, Turner MJ, Pomp D, Kleeberger SR, Leamy LJ. Quantitative trait loci for physical activity traits in mice. Physiol Genomics. 2008;32(3):401–8. https://doi.org/10.1152/physiolgenomics.00241.2007.

54. Salmen T, Heikkinen AM, Mahonen A, Kroger H, Komulainen M, Pallonen H, et al. Relation of aromatase gene polymorphism and hormone replacement therapy to serum estradiol levels, bone mineral density, and fracture risk in early postmenopausal women. Ann Med. 2003;35(4):282–8.

55. Ott J, Kamatani Y, Lathrop M. Family-based designs for genome-wide association studies. Nat Rev Genet. 2011;12(7):465–74. https://doi.org/10.1038/nrg2989.

56. Chan KH, Huang YT, Meng Q, Wu C, Reiner A, Sobel EM, et al. Shared molecular pathways and gene networks for cardiovascular disease and type 2 diabetes mellitus in women across diverse ethnicities. Circ Cardiovasc Genet. 2014;7(6):911–9. https://doi.org/10.1161/CIRCGENETICS.114.000676.

57. Women's Health Initiative. https://www.whi.org/. Accessed 26 Aug 2017.

58. Framingham Heart Study. https://www.framinghamheartstudy.org/. Accessed 26 Aug 2017.

59. Jackson Heart Study. https://www.jacksonheartstudy.org/. Accessed 26 Aug 2017.

60. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197–206. https://doi.org/10.1038/nature14177.

61. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, et al. The genetic architecture of type 2 diabetes. Nature. 2016;536(7614):41–7. https://doi.org/10.1038/nature18642.

62. Huppertz C, Bartels M, Groen-Blokhuis MM, Dolan CV, de Moor MH, Abdellaoui A, et al. The dopaminergic reward system and leisure time exercise behavior: a candidate allele study. Biomed Res Int. 2014;2014:591717. https://doi.org/10.1155/2014/591717.

63. Althoff T, Sosic R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. Nature. 2017;547(7663):336–9. https://doi.org/10.1038/nature23018.

64. McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts Cardiovascular Health Study. JAMA Cardiol. 2017;2(1):67–76. https://doi.org/10.1001/jamacardio.2016.4395.

65. NHLBI Trans-Omics for Prevision Medicine Whole Genome Sequencing Program (TOPMED). https://www.nhlbiwgs.org/. Accessed 26 Aug 2017.

66. Molecular Transducers of Physical Activity in Humans Consortium (MoTrPAC). https://commonfund.nih.gov/moleculartransducers/overview. Accessed 26 Aug 2017.