Current Understanding of the Genetic Basis for Physical Activity¹–³

J. Timothy Lightfoot*  

Huffines’ Institute for Sports Medicine and Human Performance, Department of Health and Kinesiology, Texas A&M University, College Station, TX 77845-4243

Abstract

Although it is well known that physical activity prevents and ameliorates a large number of conditions and chronic diseases, it is also incontrovertible that physical inactivity is becoming more prevalent. This paradox has led some to suggest that genetic/biological factors influence activity levels as opposed to the classical notion that voluntary activity is solely regulated by environmental factors. There is a plethora of recent data showing that there is considerable genetic influence on activity levels in both humans and animals and emerging evidence suggesting potential genomic locations for those genetic factors. Several independent lines of evidence suggest that dopamine receptor 1 (Drd1) and nescient helix loop helix (Nhnh2) are excellent candidate genes for the regulation of physical activity, with several other potential candidate genes only partially supported. This foundation provides the basis for continuing work to identify additional candidate genes, to identify other genetic factors that are involved in the regulation of physical activity, and to investigate the mechanisms by which these genes and genetic factors regulate activity.  J. Nutr. 141: 526–530, 2011.

Overview of the problem

It is well known that regular moderate physical activity is linked to a variety of positive health outcomes. Operationally, regular moderate physical activity for health has been defined as the accumulation of 30 min/d of activity (1). With this amount of regular moderate activity, there are decreases of all-cause mortality rates (2), a decrease in rates of cardiovascular disease (3), a decrease in diabetes (4), a decrease in several forms of cancer (5), a decrease in risk of stroke (6), an increase in mitochondrial function (7), an association with an increased lifespan (8), and a decrease in obesity (9). Paradoxically, although this information has been well known for over 30 y, people in general have become less active. For example, in one of the largest studies of its kind, Troiano et al. (10) from the National Cancer Institute monitored daily activity in over 7000 individuals using accelerometers. Figure 1 shows the percentage of individuals who accumulated at least 30 min/d of activity. As is evident from the figure, although the initial percentages of children that were active at least 30 min/d is low at an average of 42%, as the groups aged, the amount of activity rapidly decreased to just 3.5% of the individuals in the 20- to 59-yr-old group who actually attained healthy physical activity levels. This level of inactivity is extremely detrimental to overall health, as demonstrated by Mokdad et al. (11), who noted that physical inactivity in combination with poor diet was the second actual leading cause of death in the United States. Further, it has been suggested that physical inactivity is the primary cause of death in $\sim 250,000$ cases/y (12) and costs the health care system at least $\sim 507$ billion/y (13). Thus, it can be agreed that physical inactivity, especially in conjunction with poor diet, has devastating consequences for our health and well-being.

It is recognized that most complex behaviors are regulated or governed primarily by items that fall into 1 of 3 groups: environmental aspects, genetic or biological aspects, and/or the interaction between the two. This is the case with obesity (9) and in the context of physical activity, it has been well established that environmental factors such as culture, the built environment, safety, and peer and parental support influence the rate of physical activity (14). However, in contrast with the amount of literature regarding environmental influence on activity, until the last 10 y there was little known about potential genetic and biological regulators of activity. This literature is currently starting to expand very quickly, but because the knowledge regarding the genetics of activity regulation is still slight, it is difficult to consider potential interaction between environmental factors and genetics in regulating activity. Thus, as a result of the state of the field, our overall objective has been to determine the biological regulatory mechanisms of physical activity that arise from genetic factors. In this sense, because the concern is about the biological control of activity and not how activity affects

---

¹ Published as a supplement to The Journal of Nutrition. Presented as part of the symposium entitled “Systems Genetics in Nutrition Research” given at the Experimental Biology 2010 meeting, April 27, 2010, in Anaheim, CA. The symposium was supported by an educational grant from USDA-NIFA, AFRI Award, and sponsored by the Nutrient Gene Interaction RIS and the University of Tennessee Obesity Research Center. The symposium was chaired by Naima Mousa-Stai-Moussa and cochaired by Brynn Voy. Guest Editor for this symposium publication was Kevin Schalinske. Guest Editor disclosure: Kevin Schalinske had no relationships to disclose.

² Supported by NIH NIAMS RO1AR050085.

³ Author disclosures: J. T. Lightfoot, no conflicts of interest.

* To whom correspondence should be addressed. E-mail: TLightfoot@hlkn.tamu.edu.
biology, physical activity is set as the dependent variable, with the independent variable being biological genetic underpinnings.

The model used
As the genetic regulation of activity is investigated, either human or rodent models can be used. In many cases, a mouse model is favored, because researchers can conduct breeding studies in mice to control the heterozygosity of the resulting population, the environmental factors can be controlled to a large extent, and, because of their short lifespans, mice can be measured across their lives. All of these factors are largely impossible using a human model. The mouse is a good model of the human genome, with recent papers suggesting ~75% homology between the human and mouse genome (15). This amount of genomic similarity means that the majority of the genetic results found in mice should be directly translatable to humans. An additional benefit of using mice as an experimental model is that unlike humans who have a heterozygous genome, multiple lines of homozygous inbred mouse strains are available (16). Thus, when using inbred strains, the amount and, in many cases, the location of heterozygosity induced can be controlled by rigorous breeding programs, which allow for the direct investigation of genetics on physical activity.

Unlike in humans, where the direct objective measurement of activity is complicated by a large variety of factors, it is relatively easy to directly measure activity in mice. The use of wheel-running as a measurement of activity levels in mice has a long history and is robust and repeatable (17). Although there are certainly methodological issues that have to be surmounted (e.g. wheel surface, size, length of measurement), wheel-running has been suggested as the most appropriate model of human voluntary exercise (18) and produces similar brain neurotransmitter changes as has been seen in humans (19,20), unlike forced treadmill exercise.

Investigating the genetic regulation of activity
A general systems genetics research design flow for determining the genetic basis of a phenotype usually consists of answering 4 questions sequentially (Fig. 2): 1) Is there a genetic influence on the trait of interest? 2) Where are the genes located that are associated with the trait of interest? 3) What are the identities of the involved genes? and 4) How do the genes work to regulate the phenotype? Regarding physical activity, the scientific literature is currently fairly clear regarding the first 2 questions, with data emerging regarding question 3 (the identity of the genes). It is anticipated that question 4 will continue to be a source of investigation and discovery for many years to come.

Are there genetic influences on physical activity?
From the available literature, it is clear that there is a significant genetic influence on physical activity. To date, there are at least 12 studies that have shown such (21–34). These studies have used both human and mouse models and the amount of heritability observed has ranged from 0.20 to 0.92 (i.e. genetics were responsible for 20–92% of the activity these subjects accomplished). Interestingly, although they showed significant genetic effects on activity, none of the human studies showed significant common environmental effects on activity.

In particular, 4 of these studies stand out. Joosen et al. (22), who reported the highest heritabilities of 0.92, conducted the only human study where activity was directly measured. Stubbe et al. (30) is actually the largest study to date, with this research group reporting data from almost 38,000 twin pairs with an
observed heritability of 0.48–0.71, depending on the nationality of the twins. Also, published almost simultaneously in 2005 in both a human model (29) and an animal model (33) was the suggestion that as individuals aged, the heritability of physical activity changed. Specifically, at ~18 y (~12 wk of age in the mouse), the genetic influence on physical activity markedly increased to ~80% in both studies. While both the Stubbe et al. (30) and Turner et al. (33) datasets did not cover the whole lifespan of their subjects, their data agree that at least at younger ages, heritability of activity is affected by age.

**Where are the genes located that are associated with physical activity?**

There are several designs that result in the production of genome maps indicating which parts of the genome are associated with a researcher’s trait of interest (Fig. 2). These genomic regions are known as quantitative trait loci (QTL) and are used to narrow down the number of possible genes that is related to physical activity. The identification of QTL through linkage studies is a common approach in beginning the identification of the genetic mechanisms associated with a trait of interest and has been employed with many traits in the past literature, including obesity (35,36). QTL can be categorized as either single-effect, where the genetic factors arising from the QTL act individually on activity, or as epistatic, where the genetic factors in the QTL must work interactively with genetic factors in other genomic locations before there is an effect on the phenotype. There are currently 7 studies, 3 studies in humans and 4 in mice, that have identified QTL associated with physical activity (27,37–42). Few of these QTL overlap because of differences in the activity indices used as phenotypes. However, there are some apparent chromosomal hot spots where several QTL overlap that provide good candidate regions for more investigation. For example, earlier work from our laboratory with 2 strains of mice showed a QTL on chromosome 13 with a large influence (6.4%) on distance, duration, and speed of running wheel activity (27). A similar QTL hotspot is on chromosome 11, where there appear to be several single-effect and epistatic QTL that colocalize near the mid-point of the chromosome and are associated with the mini-muscle phenotype, which is a unique characteristic conveying smaller muscle sizes but higher aerobic capacities in some mice bred for high activity (27,38,39). The majority of the discovered QTL to this point have been single-effect QTL. However, Leamy et al. (39) clearly showed that epistatic QTL are also associated with genetic control of physical activity and in fact provide as much, if not more, influence on the overall control as do the single-effect QTL. Thus, it appears that the current genomic maps that take into account both single-effect and epistatic QTL explain between 88 and 100% of the genetic variation in physical activity and provide excellent starting places to begin to identify the genes that are involved in the regulation of physical activity.

**What are the identities of the involved genes?**

It is easy to assume that it would be a relatively simple task to move from identifying QTL to identifying potential candidate genes within the QTL (the so-called quantitative trait genes or QTG). However, moving from QTL to QTG has proven to be a very difficult proposition, with very few genes identified through this method (43). In fact, most genes currently speculated to be involved in the regulation of physical activity have been named candidate genes based solely on their apparent functional relevance as opposed to their location in a known QTL. Therefore, to prevent the identification of a magnitude of false positive candidate genes, DiPetrillo et al. (44) noted that at least 3 independent lines of evidence are needed to declare a candidate gene. We suggest that this standard should be at least 4 independent lines of evidence due to the relative liberal nature of some of the suggested criteria. Specifically, some of the potential lines of evidence that might be generated are: 1) Is there functional relevance to the trait that we know about? 2) Does the gene localize within an identified QTL? 3) Is there a possible genomic structural variation in the gene that may give rise to a functional difference in the protein? 4) Is there a difference in gene expression with difference in trait? and 5) If a gene is manipulated, does this manipulation cause a change in the phenotype of interest? Although there are several speculated candidate genes, at this point, it appears that there are only 2 genes that meet the standard of having at least 4 independent lines of evidence: dopamine receptor 1 (Drd1) and nescent helix loop helix 2 (Nhlh2).

**Drd1.** It is relatively easy to make a case for the functional relevance of any dopamine gene to locomotion and physical activity, because it is well established that dopamine alterations are involved in attention deficit hyperactive disorder, Parkinson’s disease, and a variety of other locomotive diseases. When conducting a region-specific haplotype analysis, there are structural differences between high- and low-active animals on the Drd1 gene (personal communication, Tyrone Ceasar, University of Tennessee) and Drd1 falls into an activity-associated QTL that previously has been identified (27). Importantly, Knab et al. (45) showed that of 7 different dopamine genes, the brains of high-active animals (i.e. nucleus accumbens) presented with down-regulated Drd1 compared with low-active animals. Further, Rhodes et al. (46) showed that pharmacological manipulation of Drd1 altered physical activity. Thus, there exist at least 5 different lines of research that suggest that Drd1 is involved in the regulation of physical activity. The mechanism of how Drd1 regulates activity is not yet known.

**Nhlh2.** Another gene that appears to be a good candidate gene for the regulation of physical activity is nescent helix loop helix 2 (Nhlh2). Nhlh2 was a relatively recent discovery by Deborah Good et al. (47) and has a potential functional relationship with activity through its effect on B-endorphin production as well as an interaction with melanocortin-4 receptor. Additionally, Nhlh2 localizes into one of the epistatic QTL for activity discovered (chromosome 3, 39) and it shows partial haplotype structural differences in the gene between high- and low-active animals (personal communication, Tyrone Ceasar, University of Tennessee). Most importantly, Good, et al. (48) have shown that manipulation of Nhlh2 through knockout reduces wheel-running by ~50% compared with wild-type mice. Thus, Nhlh2 at this time presents potential functional relevance to physical activity, localizes into one of the identified QTLs, shows partial haplotype differences, and, with manipulation, alters activity levels, all evidence suggesting that Nhlh2 is an appropriate candidate gene for regulation of physical activity.

**Other genes.** Although there are only 2 genes that at this point have enough evidence to suggest that they are true candidate genes for the regulation of activity, there are several genes that have been suggested as potential candidate genes and that are currently the basis for ongoing study. Of those suggested genes, myostatin (Mstn), glucose transporter 4 (Slc2a4), and 3-phosphoadenosine 5-phosphosulfate synthase (Papss2) have the most evidence available at this time for their candidacy, with
all 3 showing functional relevance, localizing into identified QTL, and showing at least partial haplotype differences [personal communication, Dr. Trudy Moore-Harrison, University of North Carolina, Charlotte] (34).

Future areas of research
It is well accepted that ≥95% of the genome has unknown function, with long stretches of noncoding DNA interspersed between the traditional gene coding regions (49). Work in the past 10–12 y has shown that biomolecules that arise from these intergenic areas may have a regulating effect, primarily through post-transcriptional regulation of protein expression (e.g. short interfering RNA or micro RNA). Interestingly, there is preliminary evidence suggesting that 2 of the 3 significant QTL associated with activity in a large, human, genome-wide association study may actually be located in these intergenic areas (42). Evidence from a 41-strain mouse genome-wide association study from our laboratory also supports the location of several activity-associated QTL in intergenic areas (34). Additionally, although there is no preliminary data supporting it at this time, there is also potential that other, newly discovered genetic regulating mechanisms such as histone methylation, acetylation, and/or imprinting may be active in regulating activity. These potential areas certainly need a great deal more development and experimentation before anyone can state with confidence that there is a “nontraditional” type of genetic regulation associated with physical activity.

At this point, knowledge regarding the genetic control of physical activity is still in a very early state. It is clear that there is genetic influence on physical activity in both humans and animals strongly suggesting that one can be born with a predisposition toward inactivity or activity. Research has provided maps of the genetic locations of potential candidate genes and additional work in this area is in the pipeline. Additionally, potential candidate genes have been identified, with 


dnd1 and Nhlh2 having at least 4 independent lines of research supporting their candidacy. Interestingly, despite the ongoing work in identifying candidate genes, there is early evidence that we may have to consider other types of genetic regulation as being involved in determining levels of physical activity. Although it is unclear what impact increased knowledge of the genetic mechanisms would have on physical activity levels, it is certain that further understanding why an individual is either active or inactive can only prove helpful in the struggle against inactivity-induced diseases.

Acknowledgments
The sole author had responsibility for all parts of the manuscript.

Literature Cited


