SLEEP, 2022, 1–3

https://doi.org/10.1093/sleep/zsac128
Advance Access Publication Date: 4 June 2022
Editorial

EDITORIAL

Pinpointing the genetic and cellular links between sleep and metabolism

Susan T. Harbison*

Laboratory of Systems Genetics, Systems Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA

*Corresponding author. Susan T. Harbison, Laboratory of Systems Genetics, Systems Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, 10 Center Drive, Building 10, Room 7D13, Bethesda, MD 20892-1654, USA. Email: susan.harbison@nih.gov.

One great challenge of modern complex trait biology is to understand the causal role of variants identified in genome-wide association studies (GWAS). The reason for the difficulty is three-fold. First, the effector gene associated with a given GWAS variant is not usually known [1–3]. The usual approach is to assume that the closest gene(s) to the variant site is the effector gene. Moderate success in locating the effector gene has been demonstrated in flies with this approach, as linkage disequilibrium tends to decay rapidly, and role of the gene in a given complex trait is easy to verify using mutations [4, 5]. However, GWAS variants in mammals and humans often lie in long stretches of the genome that are in linkage disequilibrium. Many potential effector genes may occupy these large regions, so choosing the causal locus is not a trivial matter. Secondly, the GWAS variant may lie in a non-coding region; hence the variant may affect gene regulation, and the corresponding effector gene(s) could be located far upstream or downstream of the variant [2, 3]. Third, GWAS variants are present in all body tissues. Thus, even if the effector gene is known, it is not generally a straightforward affair to determine where in the body the gene may exert its effects. In this issue of *SLEEP*, Lasconi *et al*. demonstrate an approach that simultaneously addresses all of these issues to identify effector genes underlying the complex relationship between sleep and metabolism [6].

The relationship between sleep and metabolism has been appreciated for several decades. Observational studies of the 1990s and early 2000s revealed a small but persistent correlation between short sleep duration and obesity across different populations with differing demographic characteristics, including age, sex, and geographic location (reviewed in ref. [7]). Also, researchers noted a correlation between type 2 diabetes (T2D) and both short and long sleep duration (reviewed in ref. [8]). Around the same time, laboratory experiments in humans revealed that sleep restriction contributes to weight gain and misregulation of insulin signaling (reviewed in ref. [9]). This work inspired a search for potential genetic links between sleep and metabolic dysregulation using GWAS approaches.

Recent studies used the data from multiple GWAS to determine the relationship between sleep/sleep disorders and metabolic traits and disease. Two different analytical techniques were applied to genomic variant data: Mendelian Randomization (MR) and cross-trait Linkage Disequilibrium Score Regression (LDSR). MR leverages the genetic variants significantly associated with a trait in a GWAS to establish the causal relationship between the trait (referred to as an "exposure") and a target trait or disease (referred to as an "outcome") [10]. For example, using MR, insomnia was identified as a previously underappreciated risk factor for T2D [11, 12]. Not all MR studies verify the relationships noted in observational studies, however [13]. MR instrumental variables (i.e. the genetic variants associated with a trait) must meet certain conditions in order for the causal relationships to be valid [10]. The appropriate conditions are not always demonstrable and can be obviated by factors such as pleiotropy, linkage disequilibrium, selection bias, population stratification, maternal/paternal effects, and assortative mating, to name a few of the possibilities [10]. Like MR, cross-trait LDSR uses GWAS variant data to look for genetic correlations. However, LDSR uses the effects of all genetic variants tested, including those that do not reach statistical significance, assigning an LD score to each variant [14]. For example, LDSR demonstrated that insomnia and long sleep duration were genetically correlated with T2D [11, 15]. Significant genetic correlations indicate a shared genetic architecture between sleep and these metabolic traits, but which genes are shared remain unknown using this strategy.

Thus, both MR and LDSR have been instrumental in the identification of important relationships between sleep and other complex traits but as they rely on the genomic variant information, they carry over the limitations from that information, in particular the difficulties in localizing effector genes from non-coding variants and in discovering the relevant tissue substrates mentioned earlier.

Lasconi *et al*. use an assessment of the genetic correlations between sleep and metabolic traits from GWAS summary data as a starting point. They apply the published data for insomnia [11], long and short sleep duration [15], and chronotype [16] to establish the genetic correlation between sleep and metabolic traits including T2D [17], anthropometric traits [18–20], and traits related to insulin signaling [21] using LDSR. They confirm genetic correlations found by others and extend these data to find new genetic relationships among sleep and metabolic parameters.

In the next stage, Lasconi *et al*. bring in their innovative strategy for identifying effector genes and the tissues of the body where these effects might occur [6, 22–24]. Importantly, they hypothesize that genetic variants associated with sleep phenotypes may affect metabolism through the cells forming the pancreatic islets of Langerhans [6], small groups of cells in the pancreas that release metabolic hormones into the bloodstream [25]. Lasconi *et al*. focus on two of the five different types of cells found in the islets: the α cells, which release glucagon, and the β cells, which release insulin [25]. These cells are critical to the maintenance of glucose homeostasis in the body [25].

They then tested the idea that pancreatic islet cells may harbor the effector genes of sleep-associated loci. Using the genomic variants identified from the insomnia, sleep duration, and chronotype GWAS, they identified additional single nucleotide polymorphisms (SNPs) in the genome in high linkage disequilibrium with the sleep variants (i.e. $r^2 > 0.8$). This gave them a list of proxy SNPs. Next, they overlapped proxy SNPs with chromatin locations in the α and β cells identified using the Assay for Transposase-Accessible Chromatin with sequencing (ATAC-Seq). ATAC-Seq uses a hyperactive transposase to cut DNA and add adapter tags used for sequencing. Because the transposase does not cut DNA occupied by histones or transcription factors, ATAC-Seq identifies regions of open chromatin presumed to be transcriptionally active [26]. Thus, this strategy reduced the number of sleep-associated variants under consideration to those that mapped to open chromatin [24]. Partitioned LDSR revealed that the heritability enrichment of sleep traits in both the α and β cells was equivalent to that of human embryonic stem cells (hESC)-derived hypothalamic-like neuronal cells used as a positive control. The authors then examined the overlap of proxy SNPs within ATAC-Seq locations with putative interacting chromatin sites identified using Capture C. As Capture C maps regulatory element interactions to the promoter regions of known genes, the effector genes could then be mapped. The proxy SNPs mapped to open promoters corresponding to 76 effector genes in the α cells, and 63 effector genes in the β cells, which they characterized further with enrichment gene ontology analyses. Interestingly, some of the effector genes had previously known roles in metabolism. These putative effector genes link sleep GWAS loci to metabolism.

The Lasconi *et al*.'s study identifies putative effector genes underlying the relationship between sleep and metabolism that manifest in the α and β cells of the islets of Langerhans. As the authors note, their discovery merits further investigation of the role of these effector genes in α and β cells. Verification strategies might include some of the authors' previous approaches in human cell cultures [24], but additional strategies incorporating tools from model organisms are another possibility [1, 3]. The authors demonstrated that their approach is generalizable to other tissues and diseases, as they previously applied it to study bone mineral density [24] and inflammatory bowel disease [23]. This suggests that one only need look in other tissues to find additional effector genes that are the targets of sleep-associated loci. However, the formation of chromatin structures is specific to both cell type and developmental stage [27], and while surveys in model organisms have found many body tissues affected by sleep loss [28, 29], the full suite of cell types impacted by sleep loss and sleep disorders remains unknown. A future challenge is to determine which additional cell types may be the targets of sleep-associated loci. The Lasconi *et al*.'s study is an advance in the study of the connection between sleep and metabolism and signifies an important addition to the evidence that the effects of sleep and circadian-associated loci are not confined to the brain, but manifest in other parts of the body as well.

Funding

This work was supported by the Intramural Research Program of the National Institutes of Health (NIH), the National Heart, Lung, and Blood Institute.

Disclosure Statement

None declared.

Data Availability

No new data were generated or analyzed in support of this article.

References

1. Tran S, *et al*. Validation of candidate sleep disorder risk genes using zebrafish. *Front Mol Neurosci*. 2022;**15**:873520.
2. Tam V, *et al*. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;**20**(8):467–484.
3. Wangler MF, *et al*. *Drosophila* and genome-wide association studies: a review and resource for the functional dissection of human complex traits. *Dis Model Mech*. 2017;**10**(2):77–88.
4. Mackay TF, *et al*. The *Drosophila melanogaster* Genetic Reference Panel. *Nature*. 2012;**482**(7384):173–178.
5. Wu KJ, *et al*. Genotype influences day-to-day variability in sleep in *Drosophila melanogaster*. *Sleep*. 2018;**41**(2). doi:10.1093/sleep/zsx205
6. Lasconi C, *et al*. Variant-to-gene mapping analyses reveal a role for pancreatic islet cells in conferring genetic susceptibility to sleep-related traits. *Sleep*. 2022. doi:10.1093/sleep/zsac109
7. Cizza G, *et al*. A link between short sleep and obesity: building the evidence for causation. *Sleep*. 2005;**28**(10):1217–1220.
8. Ogilvie RP, *et al*. The epidemiology of sleep and diabetes. *Curr Diab Rep*. 2018;**18**(10):82.
9. Reutrakul S, *et al*. Sleep influences on obesity, insulin resistance, and risk of type 2 diabetes. *Metabolism*. 2018;**84**:56–66.
10. Sanderson E, *et al*. Mendelian randomization. *Nat Rev Methods Primers*. 2022;**2**:1–21.

11. Jansen PR, *et al*. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet*. 2019;**51**(3):394–403.
12. Yuan S, *et al*. An atlas on risk factors for type 2 diabetes: a wide-angled Mendelian randomisation study. *Diabetologia*. 2020;**63**(11):2359–2371.
13. Wang J, *et al*. Sleep duration and adiposity in children and adults: observational and Mendelian randomization studies. *Obesity (Silver Spring)*. 2019;**27**(6):1013–1022.
14. Bulik-Sullivan B, *et al*. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;**47**(11):1236–1241.
15. Dashti HS, *et al*. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nat Commun*. 2019;**10**(1):1100.
16. Jones SE, *et al*. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat Commun*. 2019;**10**(1):343.
17. Xue A, *et al*. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat Commun*. 2018;**9**(1):2941.
18. Locke AE, *et al*. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;**518**(7538):197–206.
19. Berndt SI, *et al*. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet*. 2013;**45**(5):501–512.
20. Shungin D, *et al*. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 2015;**518**(7538):187–196.
21. Manning AK, *et al*. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*. 2012;**44**(6):659–669.
22. Pahl MC, *et al*. Cis-regulatory architecture of human ESC-derived hypothalamic neuron differentiation aids in variant-to-gene mapping of relevant complex traits. *Nat Commun*. 2021;**12**(1):6749.
23. Lasconi C, *et al*. Variant-to-gene-mapping analyses reveal a role for the hypothalamus in genetic susceptibility to inflammatory bowel disease. *Cell Mol Gastroenterol Hepatol*. 2021;**11**(3):667–682.
24. Chesi A, *et al*. Genome-scale Capture C promoter interactions implicate effector genes at GWAS loci for bone mineral density. *Nat Commun*. 2019;**10**(1):1260.
25. Roder PV, *et al*. Pancreatic regulation of glucose homeostasis. *Exp Mol Med*. 2016;**48**:e219.
26. Klemm SL, *et al*. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet*. 2019;**20**(4):207–220.
27. Herrmann JC, *et al*. Making connections: enhancers in cellular differentiation. *Trends Genet*. 2022;**38**(4):395–408.
28. Everson CA, *et al*. Cell injury and repair resulting from sleep loss and sleep recovery in laboratory rats. *Sleep*. 2014;**37**(12):1929–1940.
29. Vaccaro A, *et al*. Sleep loss can cause death through accumulation of reactive oxygen species in the gut. *Cell*. 2020;**181**(6):1307–1328.e15.