Journal Pre-proof

Sleep duration: A Review of Genome-wide Association Studies (GWAS) in Adults from 2007 to 2020

Victoria Garfield, PhD

PII: S1087-0792(20)30156-8

DOI: https://doi.org/10.1016/j.smrv.2020.101413

Reference: YSMRV 101413

To appear in: *Sleep Medicine Reviews*

Received Date: 8 June 2020

Revised Date: 14 August 2020

Accepted Date: 17 August 2020

Please cite this article as: Garfield V, Sleep duration: A Review of Genome-wide Association Studies (GWAS) in Adults from 2007 to 2020, *Sleep Medicine Reviews*, https://doi.org/10.1016/j.smrv.2020.101413.

Sleep duration: A Review of Genome-wide Association Studies (GWAS) in Adults from 2007 to 2020

Running head: Sleep duration genome-wide studies: 2007 to 2020

Victoria Garfield, PhD[1]

[1]MRC Unit for Lifelong Health and Ageing, Institute of Cardiovascular Science, University College London, 1-19 Torrington Place, WC1E 7HB, v.garfield@ucl.ac.uk, ORCID: https://orcid.org/0000-0002-5127-0890

Summary

A modest body of research exists in the area of human sleep genetics, which suggests that specific sleep phenotypes are, like many other complex traits, somewhat heritable. Until 2007 research into sleep genetics relied solely on twin studies, but in the last 13 years with the advent of huge biobanks and very large-scale genome-wide association studies, the field of molecular sleep genetics has seen important advances. To date, the majority have focused on self-reported sleep duration, but in recent years genome-wide association studies of objectively-measured sleep have emerged. These genetic studies have discovered multiple common genetic variants and as such, have provided insight into potential biological pathways, causal relationships between sleep duration and important disease outcomes using Mendelian randomisation. They have also shown that the heritability of these traits may not be as high as previously estimated. This article is the first to provide a detailed review of these recent advances in the genetic epidemiology of sleep duration. Studies were identified using both the GWAS Catalog and PubMed for completeness. Focus is on the genome-wide association studies published to date, including whether and how they have elucidated important biology and advanced knowledge in the area of sleep and health.

Keywords: sleep duration, GWAS, heritability, Mendelian randomisation, genome-wide association study, objective sleep duration, self-reported sleep duration.

Acknowledgements: VG is funded by a grant from Diabetes UK and the British Heart Foundation (15/0005250). The funders had no role in the review. The author declares that there are no conflicts of interest.

Abbreviations

ABCC9 – Adenosine triphosphate-binding cassette, sub-family C, member 9

ADRB1 – Beta-1 adrenergic receptor

AUTS2 – Activator of transcription and developmental regulator

BMI – Body mass index

CBWD2 – Cobalamin Synthase W Domain-Containing Protein 2 gene

CHARGE – Cohorts for Heart and Aging Research and Genomic Epidemiology

CPMC – Coriell Personalized Medicine Collaborative

DNA – Deoxyribonucleic acid

DRD2 – Dopamine Receptor D2

EAGLE Consortium – EArly Genetics and Lifecourse Epidemiology Consortium

EMR – Electronic Medical Records

GABA – Neurotransmitter gamma-aminobutyric acid

GABRB3 – Gamma-aminobutyric acid receptor subunit beta-3

GWA – Genome-wide association

GWAS – Genome-wide association study

HCRTR2 – Orexin receptor

HRS – Health and Retirement Study

MAPKAP1 – Target of rapamycin complex 2 subunit

MR – Mendelian randomisation

OSA – Obstructive sleep apnoea

PAX8 – paired box thyroid-specific transcription factor

PDE11A - Phosphodiesterase 11A

PRS – Polygenic risk score

SLC6A3 – Dopamine Transporter Solute Carrier Family C6, Member 4

SNP – Single nucleotide polymorphism

SPT-window – Sleep-period time window

TSHR – Thyroid stimulating hormone receptor

UKB – UK Biobank

VRK2 – Vaccinia Related Kinase 2

Introduction

Currently, a modest body of research exists in the area of human sleep genetics, which has shown that specific sleep phenotypes (duration, timing, quality) are somewhat heritable [1]. Twin studies have, for the most part, estimated that genetic factors account for between 30% to 65% of the variance in the duration, quality and patterns of sleep [2–5]. However, at least two adult twin studies found no evidence that sleep duration is heritable[6,7]. A recent meta-analysis of all sleep duration twin studies found marked heterogeneity across studies, which is likely to have contributed to the two above-mentioned studies' discrepant findings, in comparison to other twin studies[8]. The disparities in findings may be attributed to differences in sample characteristics (e.g. sample size, age or sex distribution); environmental factors pertaining to specific populations (e.g. cultural factors or differences in employment); and/or divergent methodological approaches (e.g. how sleep duration is measured, as some

studies might ask a single question about average sleep duration, whilst other studies calculate sleep duration from the time participants fall asleep and wake up, and some studies use a validated sleep instrument, which includes a question on duration)[8]. Notwithstanding the invaluable contribution that twin studies have made to the sleep genetics, due to some of these inconsistencies in results, novel and distinct approaches were required to further our understanding of sleep genetics. As such, the last 10 to 15 years have seen enormous advances in the field of Genetic Epidemiology and much of this is owed to the advent of genome-wide association studies (GWAS), including whether and how novel sleep duration genetic variants overlap with other important sleep phenotypes (Figure 1).

A GWAS is a hypothesis-free study in which the associations between millions of single nucleotide polymorphisms (SNPS – a type of common genetic variant present in at least 1% of the population) and a given phenotype (for example, sleep duration) are analysed using statistical analysis (usually linear/logistic regression, depending on whether the phenotype is continuous or categorical). The aim of a GWAS is to understand whether there are common causal genetic variants (SNPs) that contribute to common diseases and as such, this is based on the 'common disease-common variant' hypothesis [9]. GWAS have a stricter alpha level imposed on them than the majority of other studies and associations are deemed 'significant' at the genome-wide level if they are $<5x10^{-8}$ and in some cases now an even more strict threshold of $<5x10^{-9}$ is used, as sample sizes have become huge (such is the era of the mega-GWAS) with the advent of biobanks and huge consortia.

Multiple GWAS of sleep duration have now been published and data from these studies suggest that the heritability of distinct sleep phenotypes may not be as high as that observed in twin studies. At this point, it is important to explain the differences between SNP-based heritability and heritability estimates from twin studies. The twin study is a classical quantitative genetic approach that aims to unravel and quantify the contributions of genetic,

shared environmental (i.e. shared completely by twins and that contribute to them being similar, such as family socioeconomic status) and non-shared environmental factors (i.e. those that are unique to each twin and contribute to differences between them). Heritability in twin studies is usually estimated by comparing the resemblance between monozygotic (MZ, who share 100% of their genetic data and dizygotic (DZ, who share 50% of their genetic data) twins[10]. The twin method compares the similarity of MZ and DZ twins, who tend to grow up in very similar environments[11]. Similarity is assessed statistically with a correlation coefficient and if the correlation for MZ pairs is greater than that of DZ pairs, it is likely that individual differences in the phenotype of interest (e.g. sleep duration in this case) has some genetic basis[10]. On the other hand, SNP-based heritability refers to estimation of the variance explained in a phenotype (e.g. sleep duration) by all SNPs included in a GWAS in unrelated individuals[12]. Two of the most common methods are genomic relatedness matrix (GRM) restricted maximum likelihood (GREML)[12], which uses individual-level data and linkage disequilibrium score regression (LDSC) which exploits summary-level data GWAS data[12]. The former implements GREML using software such as genome-wide complex trait analysis (GCTA) and includes all available SNPs using a mixed model approach[10]. Fixed effects are usually factors such as age and sex and the total genetic effects for individual chromosomes or the whole genome constitute a random effect[10]. The latter involves regressions of GWAS summary statistics (i.e. from millions of SNPs) and assesses the extent to which each individual SNP tags other variants locally (also known as its 'LD score')[13]. The slope of the LDSC regression is then rescaled to provide heritability estimates explained by all of the SNPs incorporated into the estimation of the LD scores[13].

Importantly, though, some of the discrepancies between heritability estimates, known as the 'missing heritability'[14], from twin vs. SNP-based approaches may relate to method-specific issues. For example, SNP-based methods are unable to: include gene-environment

correlations (while the twin design is able to), capture a large amount of common genetic variants of small effect and variants that are rare but have large effect sizes[8]. However, it has also been proposed that twin studies may have overestimated heritability due to violations of the 'equal environments assumption' (that environmental factors are shared to an equal extent by MZ and DZ twins)[15], gene-gene interactions and/or gene-environment interactions[16].

This review is the first to bring together all of these findings from GWAS of sleep duration in one article, as it is both timely and important that sleep health professionals and researchers are up-to-date with these genetic findings. The GWAS Catalog database (https://www.ebi.ac.uk/gwas/home) was searched for published GWA studies up until 31st May, 2020, where the term 'sleep duration' was entered into the search box. For completeness, a further search was conducted in PubMed, using the following terms: ("GWAS"[Title/Abstract] OR "genome wide association"[Title/Abstract]) AND "sleep duration"[Title/Abstract]. Where possible, PRISMA guidelines were adhered to.

GWA studies of sleep duration: 2007 to 2020

The search described above in the GWAS Catalog yielded 33, while the search in PubMed provided 39, articles for screening. The PubMed search yielded more results because it also included some other study designs, such as Mendelian randomisation studies. After screening, a total of 13 studies (nine for self-reported and four for objective sleep duration) were retained for review (Table 1).

Table 1. *GWA studies included in the review*

Author, [reference], (year)	N	Number of novel genetic loci	Number of most strongly replicated genetic loci
Self-reported sleep duration GWA studies			
Gottlieb et al.[17] (2007)	749	1 suggestive SNP	N/A[a]
Allebrandt et al.[18] (2013)	4,251	1	0
Byrne et al.[1] (2013)	2,323	7 suggestive SNPs	0
Ollila et al.[19] (2014)	1,941	0	0
Gottlieb et al.[20] (2014)	47,180	7	0
Scheinfeldt et al. [21] (2015)	3,414	0	0
Jones et al.[22] (2016)	127,573	3	1 SNP near 2 SNPs discovered by Gottlieb et al. (2014) in the PAX8 gene
Lane et al.[23] (2017)	112,586	5	1 SNP replicated from Jones et al. (2016)[b]
Dashti et al.[24] (2019)	446,118	78	2 SNPs replicated from Jones et al. (2016) (1 in PAX8 and 1 in VRK2)
Jansen et al.[25](2019)	384,317	53	1 SNP in the SLC6A3 gene
Objective sleep duration GWA studies			
Spada et al.[26] (2016)	941	200 suggestive SNPs	0
Doherty et al.[27] (2018)	91,105	8	1 SNP in PAX8
Dashti et al.[24] (2019)	85,499	N/A	78 self-reported sleep duration SNPs replicated, with strongest signal in PAX8
Jones et al.[28] (2019)	85,670	10	1 SNP in PAX8

Note. [a]first GWA study to be published; [b]this was very likely to be due to substantial sample overlap, as the effect size was almost identical to that of Jones' (2016) study.

GWAS of self-reported sleep duration

GWAS have described SNPs that are associated with distinct measures of sleep, however they remain largely un-replicated and the effect sizes of the identified SNPs are very small (for example, the largest reported effect size to date is 4 minutes). However, the variant reported to have the largest effect size to date accounted for approximately 5% of the variation in sleep duration [18], yet this has not been replicated in subsequent studies. A more recent and much larger GWAS of sleep duration found that the maximum variance explained by a single variant was 0.07% [20].

The first GWAS of sleep duration was published in 2007. Seven hundred and forty nine participants were genotyped for 100,000 SNPs, and the analyses examined associations between these SNPs and self-reported usual sleep duration [17]. Only one intergenic SNP on chromosome 13 (rs6599077) was associated with sleep duration at $p=1.4x10^{-7}$. This means that no SNPs were associated with sleep duration at genome-wide significance, but this SNP was significant at a genome-wide suggestive level of significance. The GWAS to first report on self-reported sleep duration alone had one SNP (rs11046205) reach genome-wide significance [18], which is an intronic variant in the adenosine triphosphate-binding cassette, sub-family C, member 9 (ABCC9) gene. This gene is involved in encoding a potassium channel (K_{ATP}), which contributes to energy metabolism; it has also been associated with Cantú syndrome and dilated cardiomyopathy [29,30]. However, neither of these conditions are related to sleep. Another GWAS performed on 2,323 Australian individuals found no genome-wide significant SNPs for self-reported sleep duration, sleep time, latency, quality or depth [1]. Seven SNPs, however, were suggestive of associations and these seven variants are located on different chromosomes and nearby or within distinct genes.

In 2014 two GWA studies of sleep duration were published. The first used Finnish population-based data from 1941 adults who self-reported their sleep duration [19]. This

study yielded no genome-wide significant SNPs. The next GWAS was carried out in 47,180 individuals of European ancestry from the Cohorts for Heart and Aging Research and Genomic Epidemiology (CHARGE), and found seven loci associated with self-reported sleep duration, 4 of which are on chromosome 2 and 3 on chromosome 6[20]. A further 11 loci were suggestive of associations with sleep duration but did not reach the genome-wide significant threshold. The strongest is an intergenic variant, located on chromosome 2 near the paired box thyroid-specific transcription factor (PAX8). PAX8 encodes a nuclear protein, which is involved in the expression of thyroid-specific genes, as well as thyroid follicular cell development [31], whereas the Cobalamin Synthase W Domain-Containing Protein 2 gene (CBWD2) is highly expressed in the brain, but remains poorly characterised [20]. This association was found to be in the same direction in an African-American sample, although it was not genome-wide significant ($p=9.3*10^{-4}$). The authors did not however, assess SNP-based heritability in this study. Shortly after, in 2015, a GWA study in 3,414 individuals from the Coriell Personalized Medicine Collaborative (CPMC) emerged, but they did not identify any signals that were genome-wide significant [21]. This GWAS was very small in sample size, used slightly unconventional analytical approaches (i.e. adjusted for multiple covariates that are not usually included in GWA studies, included multiple ancestry groups but did not use wholly appropriate modelling techniques such as linear mixed modelling), which may have also contributed to their results.

In 2016 the first sleep GWAS that exploited the UK Biobank data was published. It included 127,573 UK Biobank participants (interim UKB genetic data release), from which three genome-wide significant variants associated with self-reported sleep duration emerged [22]. SNP-based heritability was estimated to be 7% in this study.

The main distinctions between this study, by Jones et al. [22] and that of Gottlieb and colleagues [20]were the following: the sample size was almost three times greater; SNP

heritability was estimated and, rather than combining several studies in a meta-analysis they were able to use a single, very large sample. Only three novel loci were found to be associated with self-reported sleep duration: rs62158211, rs17190618 and rs1380703 on chromosome 2. The effect alleles for each of these two SNPs were associated with a 2-minute decrease in sleep duration, whilst the effect allele for rs1380703 was associated with a 1.5-minute increase in sleep duration. rs62158211 is an intron in the *PAX8* gene and is in high LD with two variants previously reported by Gottlieb and colleagues [20]. Thus, Jones et al. 2016 [22] were the first to replicate an association in the same region as previously reported. rs17190618 and rs1380703 are both intronic variants within the Vaccinia Related Kinase 2 (VRK2) gene. GWAS have found this gene to be associated with schizophrenia [32] and epilepsy [33] [33] although not these specific sleep duration variants. The authors also used Mendelian randomisation (MR – a genetic epidemiology method commonly employed to try to understand causality) [34] to assess causal associations between body mass index (BMI), type-2 diabetes mellitus and sleep duration. They found no evidence of causal relationships with sleep duration, using 69 BMI and 55 diabetes variants, respectively. Bidirectional MR analyses were not deemed appropriate here, as only three SNPs associated with sleep duration emerged from this GWAS, which would likely pose a problem of weak instruments [35] for an MR study of sleep duration on BMI and diabetes.

In early 2017 a GWA study whose focus was not sleep duration (instead the focus was sleep disturbance phenotypes) identified five SNPs, one of which was the replicated PAX8 signal with qualitatively identical magnitude (2.34 minutes) to that of Jones and colleagues [23]. The authors estimated SNP heritability in their UKB subsample (N=112,586) at 10.3% and whilst this was marginally larger than Jones et al.'s estimate, it was comparable enough to suggest that some differences in phenotyping, exclusion criteria and analytical approach may explain this.

More recently, Dashti et al. [24] published the most comprehensive GWAS in 446,118 UKB participants of white European ancestry and found 78 independent variants associated with self-reported sleep duration, two of which were replicated from Jones et al. 2016 (one in VRK2 and one in PAX8). The authors observed that the largest effect was 2.44 minutes per allele for the PAX8 locus, whereas average effects were approximately one minute per allele. The 76 novel loci were identified across all autosomes, except 13, 21 and 22. Of these 76 loci, some particularly intriguing signals are in the dopaminergic (DRD2, SLC6A3) signalling pathway, orexin receptor (HCRTR2) and GABA (GABRR1) signalling system. Importantly, only a handful of the GWAS signals replicated in the CHARGE study (n=47,180), but none of these were below the conventional GWAS p-value threshold, while of the 78 loci, estimates for 55 had consistent signs. None of the sleep duration loci replicated in the paediatric EArly Genetics and Lifecourse Epidemiology (EAGLE) consortium (n=10,554), even at a 5% alpha threshold.

GWAS signals at four loci also overlapped with GWAS signals for other phenotypes. Specifically, the shorter sleep (<7 hours) allele was related to higher BMI, greater risk of Crohn's disease, febrile seizures and generalised epilepsy, cardiometabolic risk, but lower risk of interstitial lung disease. Of these phenotypes, genetic overlap with BMI has been explored in greater detail, but there appears to be only weak evidence for shared genetic aetiology, with a best-fit BMI polygenic risk score (PRS) explaining only 0.02% of the variance in sleep duration in a sample of 142,209 adults [36]. The authors also performed genetic correlation analyses that revealed shared pathways between sleep duration and psychiatric, anthropometric, cognitive and psychiatric phenotypes, while MR analyses showed bidirectional causal associations between sleep duration and schizophrenia, but not with diabetes or BMI. Estimated genome-wide SNP-based heritability in this study was 9.8%, making it only 2.8% higher than Jones' et al.'s estimate of 7%. The variance in sleep duration

explained by their 78 genome-wide significant SNPs was 0.69%, which is low, considering that for example, 97 BMI SNPs from a 2015 GWAS explain 3% of the variance in BMI [37]. This low variance explained in sleep duration points towards dozens of common genetic variants of particularly small effect size and the potential notion that sleep duration may not be as much due to common genetic variation as has previously been assumed.

Concurrently with the above study, Jansen and colleagues published a GWAS focused on insomnia using data from 1,331,010 individuals from 23andMe and UKB[25]. Alongside other sleep phenotypes (e.g. getting up, napping, snoring, daytime dozing, morningness) the authors also performed a GWAS of sleep duration in 384,317 UKB participants only, which yielded 53 genome-wide significant loci. Of these SNPs, 14 overlapped with insomnia, with a genetic correlation of -0.47, compared to fewer overlapping loci with the remaining sleep phenotypes. This was not surprising, given that fragmented rapid eye movement (REM) sleep is common in insomnia and relates to the fact that individuals suffering from this condition are likely to underestimate their sleep duration. Moreover, of the 53 sleep duration SNPs the locus that most strongly overlapped with Dashti et al's findings was an intron on chromosome 5 in the SLC6A3 gene, which had a similar effect size (Dashti: 0.9 vs. Jansen: 1.2 minutes' difference in sleep duration) in both studies.

GWAS of objective sleep duration

The first GWAS of objectively-measured sleep duration emerged in 2016 in 941 adults from the German LIFE study, who wore an actigraph for seven days [26]. It appears that the authors reported over 200 signals for objective sleep duration, but none were below the genome-wide significance threshold of $p<5x10^{-8}$. These variants have not been replicated.

In 2018, Doherty and colleagues published a large-scale GWAS of objectively-measured sleep duration in 91,105 UKB participants [27]. They identified eight variants (at a threshold of $p<5x10^{-9}$) associated with 7-day wrist-worn accelerometer-measured sleep duration explaining 0.39% of the variance in the phenotype, and of which two were novel. A previously-validated machine-learning model was used to distinguish sleep from other activity states (sedentary/walking/moderate intensity), as it was found to be valid for use in UKB participants [38]. The PAX8 signal from previous GWAS was replicated of the two novel SNPs one is near the MAPKAP1 gene and was associated with longer duration of sleep and the other is near the Activator of transcription and developmental regulator *(*AUTS2*)* gene and is associated with shorter sleep. MAPKAP1 encodes the target of rapamycin complex 2 subunit MAPKAP1 protein. In previous GWA studies, variants other than the one found here have been associated with a wide range of traits, including some that have been robustly related to sleep duration, such as haemodynamic phenotypes, resting heart rate, height and general cognitive ability. The activator of transcription and developmental regulator (AUTS2) is a protein coding gene, for which other variants, besides the sleep duration locus have been associated with some phenotypes relevant to sleep duration, including reaction time, mathematical ability, educational attainment, alcohol consumption, chronotype, BMI, haemodynamic measures, smoking behaviour, caffeine consumption and type-2 diabetes. The authors also observed genetic correlations between sleep duration and both anthropometric and cognitive phenotypes, specifically that increases in sleep duration were associated with lower fluid intelligence scores and poorer health status. They also performed MR analyses in 278,374 UKB participants who did not form part of the discovery sample, but these showed no evidence of causal relationships between sleep duration and a breadth of common traits/diseases, which is much in line with the work of Jones et al., 2016 [22] and Dashti et al., 2019 [24], reported earlier.

The study by Dashti and colleagues [24] reviewed earlier also performed GWAS analyses in 85,499 UKB participants who had accelerometry data, suggestive of substantial overlap with the objective sleep duration GWAS described above. However, the authors used an R package to infer accelerometer wear time. Sleep duration episodes were defined within the sleep-period time window (SPT-window) as periods of at least five minutes with no change greater than 5º related to the z-axis of the device. The authors then summed sleep duration across all sleep episodes to obtain a sleep duration phenotype. The 78 self-reported sleep duration SNPs were tested for associations with objective sleep duration in this UKB subsample and the most promising finding was that the PAX8 lead variant was associated with 2.7 minutes longer objective, as compared to 2.4 minutes with self-reported sleep duration. This was particularly reassuring, as this PAX8 variant is the most consistently-reported sleep duration GWAS signal. Moreover, the 78-SNP PRS was associated with longer objective duration of sleep.

One other GWAS of objective sleep duration published in 2019 in 85,670 UKB respondents found 11 common variants associated with sleep duration, 10 of which were novel [28]. The replication sample consisted of 5819 individuals with no loci reaching genome-wide significance, but the vast majority of signals were directionally concordant with that of the discovery sample. SNP heritability was estimated to be 19%, whilst the largest genome-wide heritability estimate for self-reported sleep is 9.8%, as reported earlier. This may reflect differences in measurement accuracy, where self-reported sleep duration is more likely to be the culprit, especially given that the genetic correlation the authors observed between self-reported and objective sleep duration was at best, modest (r_g=0.43). However, whilst we should be critical of self-report measures in general and sleep duration is no exception, a) the most robust self-reported sleep duration signal (in PAX8) was also associated with

accelerometer-measured sleep duration and b) mean self-reported and objective sleep durations were 7.2h and 7.3h in UKB, respectively, both of which represent healthy durations for adults [39]. As 10 novel loci emerged as associated with sleep duration but these were largely not replicated due to the small validation sample available, focus here is on what the authors identified as the potential lead causal variant, rather than the entire list of SNPs and genes that have not yet been replicated. rs17400325, a missense variant (p.Tyr727Cys) in the phosphodiesterase 11A (PDE11A) gene, that is highly expressed in the hippocampus was identified via fine-mapping as the lead causal variant. GWA studies have identified this SNP as associated with myopia, height, smoking initiation and educational attainment, while other variants in PDE11A have been linked to heart rate and haemodynamic traits. Mendelian randomisation analyses revealed no causal associations between accelerometer-measured sleep duration and any relevant traits identified in the genetic correlation analyses, which may be due to the limited number of genetic instruments and the fact that there were only 11 of them, which is not likely to make for a strong instrument in MR. In the opposite direction, MR identified a causal relationship between higher waist-hip-ratio and educational attainment, and shorter objective sleep duration.

Verdict on GWAS of sleep duration: what have these GWA studies taught us and why are they important (or not)?

What pathways, if any, have these GWA studies identified as biologically plausible?

In terms of how the above studies contributed to our understanding of the biology of sleep, the PAX8 gene (for which SNPs associated with sleep duration were identified and replicated in the GWASs by both the CHARGE and the UKB studies) encodes a protein which is involved in the expression of thyroid-related genes. PAX8 is associated with hypothyroidism and patients that do not receive treatment for this disease are more likely to have obstructive sleep apnoea (OSA) episodes [40]. Mutations in the PAX8 gene, amongst others, may result

in the thyroid stimulating hormone receptor (TSHR) gene being only partially activated. Whilst this is important, as the prevalence of hypothyroidism is approximately 2% (UK) [41], it is worth noting that the PAX8 gene is not highly expressed in the hypothalamus, for example, which is responsible for the regulation of sleep in the brain.

Variants in the VRK2 gene have (aside from duration of sleep) previously been associated with schizophrenia, a psychiatric illness which is known to have consequences for patients' sleep [42]. Evidence suggests that sleep disturbances may contribute to the onset of psychosis in young people [43]. However, similarly to the PAX8 gene pathways, this may not be as informative in terms of average sleep duration in the population, as these pathways relate to specific diseases, such as schizophrenia and hypothyroidism.

VRK2 and PAX8 are the only genes with variants that have been, what can now be referred to as robustly replicated, given that they emerged as 'hits' in a handful of large-scale GWA studies over the last six years and showed consistency both in terms of estimate size and direction. The remainder of the variants identified by these recent GWA studies still need to be replicated, possibly in even larger studies than UKB. Sleep duration changes as a function of age [44] and extreme sleep durations (short/long) are somewhat more prevalent than in previous generations [45]. Therefore, it is important to try to disentangle potential age-specific effects of these sleep duration SNPs and whether the magnitude of the association between for example, a sleep duration genetic risk score may differ by age stratum or whether distinct trajectories of the $PRS_{sleepduration}$ and sleep duration relationship are observed. Evidence from the Health and Retirement Study (HRS) suggests that this may at least be true for BMI, as it was observed that the associations between PRS_{BMI} and BMI differed by birth cohort, such that the magnitude of the relationship was larger among participants born after 1943, as compared to those born before 1924 [46]. The authors explained their findings by stating that perhaps BMI SNPs associate distinctly with actual BMI because of effect

modification due to the obesogenic environment that younger cohorts have been exposed to. While BMI and sleep duration are divergent phenotypes, they are both complex traits that have changed over time and are subject to life course alterations and thus, a similar study with sleep duration would be of interest.

How have these GWA studies contributed to knowledge in the area of sleep and health?

As mentioned earlier in this review, robustly replicated common genetic variants for complex traits do not simply end with publication of a GWA study, irrespective of how large it may be. Crucially, the advent of very large-scale GWA studies, alongside the now common practice of making summary statistics publicly available to the scientific community, has facilitated hundreds and hundreds of Mendelian randomisation studies. Whilst it would be wrong to try to persuade readers that MR is a panacea for causality, to deny that MR has allowed us to understand that associations which we have studied for decades using purely observational epidemiological designs are in fact causal, or that they are not in fact, causal, is perhaps equally wrong. Similarly, whilst the UK Biobank poses its own set of problems for researchers, particularly given its low response rate and now somewhat accepted selection bias issues, it would be unfair to deny its immense value in the advancement of multidisciplinary ageing science.

In the context of sleep duration, until a large and comprehensive GWAS was published in mid-2019 [24] we were largely incapacitated when considering how we might design and execute an MR study to understand whether sleep duration causally relates to a particular outcome of interest. This landmark study enabled the flurry of MR studies that emerged following its publication. The first of these was an MR study in a cohort of Chinese adolescents and they observed no unconfounded effect of sleep duration on diabetes, fasting glucose or glycated haemoglobin [47]. Since then, several other MR studies that attempt to assess the causal effect of sleep duration on important outcomes have emerged, including

sleep duration and: risk of breast cancer in women [48], myocardial infarction [49], glycaemic traits [50], cardiovascular disease and lipid profiles [51], haemoglobin and haematocrit [52], adiposity [53] and cognitive function and dementia [54]. The MR study of sleep duration and cognitive function/dementia was also the first to employ non-linear MR methods to try to understand the causal nature of the U-shaped association between sleep duration and these phenotypes.

Another recent study employed a neat design in which they investigated the relationship between sleep duration and disease prevalence, using the Partners Biobank which links electronic medical records (EMR) with genetic data [55]. They firstly examined the association between the 78-SNP PRS and self-reported sleep duration, followed by analyses of the PRS and 22 prevalent psychiatric, respiratory, cardiovascular, neurological, autoimmune and metabolic diseases, as well as breast cancer. The authors used these initial analyses as a type of screening, taking forward observed PRS-disease relationships into two-sample MR analyses. Thus, this study made valuable use of the novel sleep duration GWAS findings by demonstrating how we might integrate data on sleep, genomics and clinical endpoints using EMR.

It is however, also important to note that due to their inherent design (based on ‘common disease, common variant’ hypothesis) GWA studies are unable to provide any insight into monogenic sleep disorders. Sleep disorders that are likely to be monogenic (i.e. likely the result of a single gene mutation) and directly relate to duration of sleep are: Fatal Familial Insomnia (FFI), Advanced sleep-phase syndrome, Primary chronic insomnia and Narcolepsy with Cataplexy[56]. FFI (caused by a mutation in the Prion protein gene) appears to be the most serious of these disorders, as it is characterised by a clinical inability to sleep, dysautonomia and motor disturbances, which can rapidly lead to death[56]. However, whilst individuals with Advanced sleep-phase syndrome (linked to a mutation in the Period 2 gene)

do sleep, they sleep outside of normal bedtime hours; those with Primary chronic insomnia have a mutation in the Gamma-aminobutyric acid receptor subunit beta-3 (GABRB3) gene and suffer from chronic sleeplessness; and individuals who have Narcolepsy with Cataplexy (caused by a mutation in the Prepro-hyprocretin gene) experience excessive daytime sleepiness, cataplexy (strong emotions can trigger sudden loss of muscle tone), sleep paralysis, sleep onset rapid-eye movement (REM) and hypnagogic hallucinations[56].

Future directions and conclusions

In the last decade, GWAS has led to the discovery and characterisation of numerous common variants associated with both self-reported and objective sleep duration. The advent of large biobanks that possess a wealth of measurements, including deoxyribonucleic acid (DNA) have aided these discoveries enormously. The genetics of sleep duration is in a completely different place from where it was pre-2007 (the first sleep duration GWAS), as before this we relied largely on quantitative genetic studies (e.g. twin studies) and candidate gene studies were tasked with trying to understand the molecular basis of this incredibly complex phenotype. Downstream analyses assisted by these GWAS discoveries in this area now include various papers on causality (e.g. using methods such as MR) between sleep duration and important health outcomes, as well as other intriguing designs that integrate polygenic risk scoring, observational analyses, EMRs and MR analyses.

However, there are still open questions and certainly important limitations to discuss here. Firstly, the associations between sleep duration and some health outcomes may be non-linear and thus, it is important that when investigating causality this is also considered, as linear relationships, or lack thereof could be masking a different shape of association. This has only been done in one MR study to date, which examined causality between sleep duration and

cognitive function/dementia [54]. Secondly, in the context of both GWA and MR studies more ethnically diverse groups should be studied in larger numbers (some of the previous GWAS reported earlier have attempted replication in other ethnic groups, but the samples have mostly been too small). Whilst for example, the sleep duration SNP estimates, minor allele frequencies and heritability may well be similar across other ethnic groups, this is important to test empirically with well-powered studies. An example of this is a 2019 study in which neither the PAX8 nor the VRK2 loci replicated in a Japanese sample of 31,230 adults [57]. Thirdly, previous GWA studies' efforts to replicate the sleep duration signals in for example, the EAGLE cohorts have proved unsuccessful. Thus, we must be sure that these SNPs are also relevant to children and adolescents, as for some complex traits heritability may differ by age [58]. This is also important because some of the MR studies published to date have been in paediatric/adolescent cohorts [47,52,53] and have used the sleep duration SNPs discovered and replicated in adults. Fourthly, the divergence between SNP heritability estimates for objective (19%) vs. self-reported sleep duration (9.8%) raise questions and this is important to consider, given that the majority of cohort studies still only collect self-reported data, for obvious reasons. Of course, earlier it was noted that the genetic correlation for these two measures was only modest (r_g=0.43) and an objective measurement will usually trump a subjective report, yet this requires further investigation. For example, polysomnography (PSG) could be used to help disentangle this, but a particularly large sample is needed, which remains difficult to obtain due to practicalities and costs. Fifthly, as detailed above, MR studies using the sleep duration SNPs as instrumental variables have emerged over the last year or so. Whilst MR has proved to be a powerful tool for contributing to our understanding of causal relationships, as with any other aetiological epidemiology findings, triangulation is crucial here [59]. Sixthly, the field of sleep genetics could benefit from a move towards next generation sequencing (NGS) in the form of whole- exome and

genome sequencing (WES and WGS) studies[60]. In contrast to GWAS, WES studies are used to detect both rare and common genetic variants and focus on protein coding sequences [61]. A recent WES study, for example, found that a rare variant in the β1 adrenergic receptor (ADRB1) gene was associated with needing fewer hours of sleep (~6 hours) and heightened wakefulness[62]. WGS studies, however, have broader coverage of variants, as the name suggests, but are still expensive and challenges remain when it comes to sequencing particularly complicated areas of the genome[61].

In conclusion, GWAS and large biobanks have aided fundamental advances in the area of sleep (duration) genetics and have permitted long-awaited important downstream studies on sleep duration and health. Nevertheless, there are still important questions to be answered in the realm of sleep (duration) genetics and biology.

Practice points

1. This review provided a detailed account of GWA studies of sleep duration published between 2007 and 2020.
2. The advent of large-scale cohorts and biobanks permitted crucial advances in the area of sleep duration genetics.
3. These genetic association studies have provided important biological insights into both objective and subjective sleep duration.
4. These studies have also enabled, via designs such as Mendelian randomisation, the investigation of causal relationships between sleep duration and important health outcomes.

Research agenda

1. When using designs such as Mendelian randomisation (MR) to try to ascertain causality between sleep duration and important outcomes, future studies ought to not only perform linear analyses, but also focus on non-linear modelling, as U-shaped associations are common between sleep duration and multiple health outcomes.
2. For both GWA and MR studies in this area, focus should be shifted towards non-European ancestry samples and across diverse age-groups (e.g. children and adolescents), as these efforts, to date, have been lacking and largely underpowered.
3. Future research needs to conduct more in-depth investigations into why molecular (SNP-based) heritability estimates are doubled for objective vs. self-reported sleep duration.
4. Future studies should also consider how whole- genome and exome sequencing could support important advances in the field of sleep genetics.
5. MR is not a panacea for causality and as with any other study type, robust replication is required for these findings, exploiting triangulation methods detailed in the literature.

References

[1] Byrne EM, Gehrman PR, Medland SE, Nyholt DR, Heath AC, Madden PAF, et al. A genome-wide association study of sleep habits and insomnia. Am J Med Genet B Neuropsychiatr Genet 2013;162B:439–51. https://doi.org/10.1002/ajmg.b.32168.

[2] Partinen M, et al. Genetic and environmental determination of human sleep. Sleep 1990;6:179–85.

[3] Heath, A.C., Kendler. K.S., Eaves., L.J., & Martin. NG. Evidence for Genetic Influences on Sleep Disturbance and Sleep Pattern in Twins. Sleep 1990;13:318–35.

[4] Watson NF, Buchwald D, Vitiello M V., Noonan C, Goldberg J. A Twin Study of Sleep Duration and Body Mass Index. J Clin SLEEP Med 2010;6:11–7.

[5] Gehrman PR, Ghorai A, Goodman M, McCluskey R, Barilla H, Almasy L, et al. Twin-based heritability of actimetry traits. Genes, Brain Behav 2019;18:1–11. https://doi.org/10.1111/gbb.12569.

[6] Boomsma DI, Van Someren EJW, Beem AL, De Geus EJC, Willemsen G. Sleep during a regular week night: A twin-sibling study. Twin Res Hum Genet 2008;11:538–45. https://doi.org/10.1375/twin.11.5.538.

[7] Barclay NL, Eley TC, Maughan B, Rowe R, Gregory AM. Associations between diurnal preference, sleep quality and externalizing behaviours: a behavioural genetic analysis. Psychol Med 2011;41:1029–1040. https://doi.org/10.1017/S0033291710001741.

[8] Madrid-Valero JJ, Rubio-Aparicio M, Gregory AM, Sánchez-Meca J, Ordoñana JR. Twin studies of subjective sleep quality and sleep duration, and their behavioral correlates: Systematic review and meta-analysis of heritability estimates. Neurosci Biobehav Rev 2020;109:78–89. https://doi.org/10.1016/j.neubiorev.2019.12.028.

[9] *Ebrahim S, Davey Smith G. Mendelian randomization: Can genetic epidemiology help redress the failures of observational epidemiology? Hum Genet 2008;123:15–33. https://doi.org/10.1007/s00439-007-0448-6.

[10] Mayhew AJ, Meyre D. Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. Curr Genomics 2017;18:332–40. https://doi.org/10.2174/1389202918666170307161450.

[11] Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. Nat Rev

Genet 2002;3:872–82. https://doi.org/10.1038/nrg932.

[12] Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. Nat Genet 2017;49:1304–10. https://doi.org/10.1038/ng.3941.

[13] Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47:291–5. https://doi.org/10.1038/ng.3211.

[14] Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747–53. https://doi.org/10.1038/nature08494.

[15] Derks EM, Dolan C V., Boomsma DI. A test of the equal environment assumption (EEA) in multivariate twin studies. Twin Res Hum Genet 2006;9:403–11. https://doi.org/10.1375/183242706777591290.

[16] Young AI. Solving the missing heritability problem. PLoS Genet 2019;15:1–7. https://doi.org/10.1371/journal.pgen.1008222.

[17] *Gottlieb DJ, O'Connor GT, Wilk JB. Genome-wide association of sleep and circadian phenotypes. BMC Med Genet 2007;8 Suppl 1:S9. https://doi.org/10.1186/1471-2350-8-S1-S9.

[18] Allebrandt K V, Amin N, Müller-Myhsok B, Esko T, Teder-Laving M, Azevedo RVDM, et al. A K(ATP) channel gene effect on sleep duration: from genome-wide association studies to function in Drosophila. Mol Psychiatry 2013;18:122–32. https://doi.org/10.1038/mp.2011.142.

[19] Ollila HM, Kettunen J, Pietiläinen O, Aho V, Silander K, Kronholm E, et al. Genome-

wide association study of sleep duration in the Finnish population. J Sleep Res 2014;23:609–18. https://doi.org/10.1111/jsr.12175.

[20] *Gottlieb DJ, Hek K, Chen TH, Watson NF, Eiriksdottir G, Byrne EM, et al. Novel loci associated with usual sleep duration: The CHARGE Consortium Genome-Wide Association Study. Mol Psychiatry 2015;20:1232–9. https://doi.org/10.1038/mp.2014.133.

[21] Scheinfeldt LB, Gharani N, Kasper RS, Schmidlen TJ, Gordon ES, Jarvis JP, et al. Using the Coriell Personalized Medicine Collaborative Data to conduct a genome-wide association study of sleep duration. Am J Med Genet Part B Neuropsychiatr Genet 2015;168:697–705. https://doi.org/10.1002/ajmg.b.32362.

[22] *Jones SE, Tyrrell J, Wood AR, Beaumont RN, Ruth KS, Tuke MA, et al. Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. PLoS Genet 2016;12:e1006125. https://doi.org/10.1371/journal.pgen.1006125.

[23] Lane JM, Liang J, Vlasac I, Anderson SG, Bechtold DA, Bowden J, et al. Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. Nat Genet 2017;49:274–81. https://doi.org/10.1038/ng.3749.

[24] *Dashti HS, Jones SE, Wood AR, Lane JM, van Hees VT, Wang H, et al. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. Nat Commun 2019;10:1–12. https://doi.org/10.1038/s41467-019-08917-4.

[25] Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci

and functional pathways. Nat Genet 2019;51:394–403. https://doi.org/10.1038/s41588-018-0333-3.

[26] Spada J, Scholz M, Kirsten H, Hensch T, Horn K, Jawinski P, et al. Genome-wide association analysis of actigraphic sleep phenotypes in the LIFE Adult Study. J Sleep Res 2016. https://doi.org/10.1111/jsr.12421.

[27] *Doherty A, Smith-Byrne K, Ferreira T, Holmes M V., Holmes C, Pulit SL, et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. Nat Commun 2018;9:1–8. https://doi.org/10.1038/s41467-018-07743-4.

[28] Jones SE, van Hees VT, Mazzotti DR, Marques-Vidal P, Sabia S, van der Spek A, et al. Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. Nat Commun 2019;10:1–12. https://doi.org/10.1038/s41467-019-09576-1.

[29] Bienengraeber M, Olson TM, Selivanov VA, Kathmann EC, O'Cochlain F, Gao F, et al. ABCC9 mutations identified in human dilated cardiomyopathy disrupt catalytic KATP channel gating. Nat Genet 2004;36:382–7. https://doi.org/10.1038/ng1329.

[30] Harakalova M, van Harssel JJT, Terhal PA, van Lieshout S, Duran K, Renkens I, et al. Dominant missense mutations in ABCC9 cause Cantú syndrome. Nat Genet 2012;44:793–6. https://doi.org/10.1038/ng.2324.

[31] Mansouri A, Chowdhury K, Gruss P. Follicular cells of the thyroid gland require Pax8 gene function. Nat Genet 1998;19:87–90. https://doi.org/10.1038/ng0598-87.

[32] Steinberg S, de Jong S, Andreassen OA, Werge T, B??rglum AD, Mors O, et al. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. Hum Mol Genet 2011;20:4076–81. https://doi.org/10.1093/hmg/ddr325.

[33] International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address: epilepsy-austin@unimelb.edu.au. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. Lancet Neurol 2014;13:893–903. https://doi.org/10.1016/S1474-4422(14)70171-1.

[34] *Davey Smith G, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32:1–22. https://doi.org/10.1093/ije/dyg070.

[35] Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes M V, White J, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. Int J Epidemiol 2016:dyw088. https://doi.org/10.1093/ije/dyw088.

[36] Garfield V, Fatemifar G, Dale C, Smart M, Bao Y, Llewellyn CH, et al. Assessing potential shared genetic aetiology between body mass index and sleep duration in 142 , 209 individuals 2018:1–8. https://doi.org/10.1002/gepi.22174.

[37] 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;518:197–206.

[38] Willetts M, Hollowell S, Aslett L, Holmes C, Doherty A. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. Sci Rep 2018;8:7961. https://doi.org/10.1038/s41598-018-26174-1.

[39] Watson N, Badr M, Belenky G, Bliwise. Recommended Amount of Sleep for a Healthy Adult. J Clin Sleep Med J Clin Sleep Med J Clin Sleep Med 2015;1111:591–2. https://doi.org/10.5664/jcsm.4758.

[40] KR R, PH A, SS D, et al. Obstructive sleep apnea in hypothyroidism. Ann Intern Med 1984;101:491–4. https://doi.org/10.7326/0003-4819-101-4-491.

[41] Werhun A, Hamilton W. Are we overusing thyroid function tests? Br J Gen Pract 2013;63:404. https://doi.org/10.3399/bjgp13X670589.

[42] Wilson S, Argyropoulos S. Sleep in schizophrenia: time for closer attention. Br J Psychiatry 2012;200:273–4. https://doi.org/10.1192/bjp.bp.111.104091.

[43] Ruhrmann S, Schultze-Lutter F, Salokangas RKR, Heinimaa M, Linszen D, Dingemans P, et al. Prediction of psychosis in adolescents and young adults at high risk: results from the prospective European prediction of psychosis study. Arch Gen Psychiatry 2010;67:241–51. https://doi.org/10.1001/archgenpsychiatry.2009.206.

[44] Ancoli-Israel S. Sleep and its disorders in aging populations. Sleep Med 2009;10:S7–11. https://doi.org/https://doi.org/10.1016/j.sleep.2009.07.004.

[45] Chattu V, Manzar M, Kumary S, Burman D, Spence D, Pandi-Perumal S. The Global Problem of Insufficient Sleep and Its Serious Public Health Implications. Healthcare 2018;7:1. https://doi.org/10.3390/healthcare7010001.

[46] Walter S, Mejia-Guevara I, Estrada K, Liu SY, Glymour MM. Association of a genetic risk score with body mass index across different birth cohorts. JAMA - J Am Med Assoc 2016;316:63–9. https://doi.org/10.1001/jama.2016.8729.

[47] Wang J, Kwok MK, Au Yeung SL, Li AM, Lam HS, Leung JYY, et al. Sleep duration and risk of diabetes: Observational and Mendelian randomization studies. Prev Med (Baltim) 2019;119:24–30. https://doi.org/10.1016/j.ypmed.2018.11.019.

[48] Richmond RC, Anderson EL, Dashti HS, Jones SE, Lane JM, Strand LB, et al. Investigating causal relations between sleep traits and risk of breast cancer in women: mendelian randomisation study. BMJ 2019;365:l2327. https://doi.org/10.1136/bmj.l2327.

[49] Daghlas I, Dashti HS, Lane J, Aragam KG, Rutter MK, Saxena R, et al. Sleep Duration and Myocardial Infarction. J Am Coll Cardiol 2019;74:1304–14. https://doi.org/10.1016/j.jacc.2019.07.022.

[50] Bos MM, van Heemst D, Donga E, de Mutsert R, Rosendaal FR, Blauw GJ, et al. The Association between Habitual Sleep Duration and Sleep Quality with Glycemic Traits: Assessment by Cross-Sectional and Mendelian Randomization Analyses. J Clin Med 2019;8:682. https://doi.org/10.3390/jcm8050682.

[51] Zhuang Z, Gao M, Yang R, Li N, Liu Z, Cao W, et al. Association of physical activity, sedentary behaviours and sleep duration with cardiovascular diseases and lipid profiles: a Mendelian randomization analysis. Lipids Health Dis 2020;19:86. https://doi.org/10.1186/s12944-020-01257-z.

[52] Wang J, Kwok MK, Au Yeung SL, Li AM, Lam S, Leung GM, et al. The effect of sleep duration on hemoglobin and hematocrit: Observational and Mendelian randomization study. Sleep 2020. https://doi.org/10.1093/sleep/zsz325.

[53] Wang J, Li AM, Lam HSHS, Leung GM, Schooling CM. Sleep Duration and Adiposity in Children and Adults: Observational and Mendelian Randomization Studies. Obesity (Silver Spring) 2019;27:1013–22. https://doi.org/10.1002/oby.22469.

[54] *Henry A, Katsoulis M, Masi S, Fatemifar G, Denaxas S, Acosta D, et al. The relationship between sleep duration , cognition and dementia : a Mendelian randomization study 2019:1–12. https://doi.org/10.1093/ije/dyz071.

[55] Dashti HS, Redline S, Saxena R. Polygenic risk score identifies associations between sleep duration and diseases determined from an electronic medical record biobank. Sleep 2019;42:1–10. https://doi.org/10.1093/sleep/zsy247.

[56] Tafti YD and M. The genetic basis of sleep and sleep disorders. Curr Pharm Des 2008;14:3386–95. https://doi.org/10.1017/CBO9781139649469.

[57] Nishiyama T, Nakatochi M, Goto A, Iwasaki M, Hachiya T, Sutoh Y, et al. Genome-wide association meta-analysis and Mendelian randomization analysis confirm the influence of ALDH2 on sleep durationin the Japanese population. Sleep 2019;42:1–11. https://doi.org/10.1093/sleep/zsz046.

[58] Evans DM, Brion MJA, Paternoster L, Kemp JP, McMahon G, Munafò M, et al. Mining the Human Phenome Using Allelic Scores That Index Biological Intermediates. PLOS Genet 2013;9:1–15. https://doi.org/10.1371/journal.pgen.1003919.

[59] *Lawlor DA, Tilling K, Smith GD. Triangulation in aetiological epidemiology. Int J Epidemiol 2016;45:1866–86. https://doi.org/10.1093/ije/dyw314.

[60] Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci U S A 2015;112:5473–8. https://doi.org/10.1073/pnas.1418631112.

[61] Suwinski P, Ong CK, Ling MHT, Poh YM, Khan AM, Ong HS. Advancing personalized medicine through the application of whole exome sequencing and big data analytics. Front Genet 2019;10:1–16. https://doi.org/10.3389/fgene.2019.00049.

[62] Shi G, Xing L, Wu D, Bhattacharyya BJ, Jones CR, McMahon T, et al. A Rare Mutation of β$_1$-Adrenergic Receptor Affects Sleep/Wake Behaviors. Neuron 2019;103:1044-1055.e7. https://doi.org/10.1016/j.neuron.2019.07.026.

Figure 1. *Overlap between sleep duration SNPs and other sleep phenotypes*

Note. Information about overlapping SNPs and sleep phenotypes taken from the Sleep Disorder Knowledge Portal: http://sleepdisordergenetics.org/home/portalHome and the GWAS catalog: https://www.ebi.ac.uk/gwas/home