Journal of Child Psychology and Psychiatry
ACAMH for Child and Adolescent Mental Health
Check for updates
Journal of Child Psychology and Psychiatry **:* (2020), pp **–** doi:10.1111/jcpp.13252

Genome-wide DNA methylation patterns associated with sleep and mental health in children: a population-based study

Maria Elisabeth Koopman-Verhoeff,[1,2,3] Rosa H. Mulder,[1,2,4] Jared M. Saletin,[3] Irwin Reiss,[5] Gijsbertus T.J.van derHorst,[6] Janine F. Felix,[2,5] Mary A. Carskadon,[3] Henning Tiemeier,[1,7] and Charlotte A.M. Cecil[1,2,8]

[1]Department of Child and Adolescent Psychiatry, Erasmus University Medical Center, Sophia Children's Hospital, Rotterdam, The Netherlands; [2]The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [3]EP Bradley Hospital Sleep Laboratory, Alpert Medical School of Brown University, Providence, RI, USA; [4]Institute of Education and Child Studies, Leiden University, Leiden, The Netherlands; [5]Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [6]Department of Molecular Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands; [7]Department of Social and Behavioral Science, Harvard TH Chan School of Public Health, Boston, MA, USA; [8]Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Background: DNA methylation (DNAm) has been implicated in the biology of sleep. Yet, how DNAm patterns across the genome relate to different sleep outcomes, and whether these associations overlap with mental health is currently unknown. Here, we investigated associations of DNAm with sleep and mental health in a pediatric population. **Methods:** This cross-sectional study included 465 10-year-old children (51.3% female) from the *Generation R Study*. Genome-wide DNAm levels were measured using the Illumina 450K array (peripheral blood). Sleep problems were assessed from self-report and mental health outcomes from maternal questionnaires. Wrist actigraphy was used in 188 11-year-old children to calculate sleep duration and midpoint sleep. Weighted gene co-expression network analysis was used to identify highly comethylated DNAm 'modules', which were tested for associations with sleep and mental health outcomes. **Results:** We identified 64 DNAm modules, one of which associated with sleep duration after covariate and multiple testing adjustment. This module included CpG sites spanning 9 genes on chromosome 17, including *MAPT* – a key regulator of Tau proteins in the brain involved in neuronal function – as well as genes previously implicated in sleep duration. Follow-up analyses suggested that DNAm variation in this region is under considerable genetic control and shows strong blood–brain concordance. DNAm modules associated with sleep did not overlap with those associated with mental health. **Conclusions:** We identified one DNAm region associated with sleep duration, including genes previously reported by recent GWAS studies. Further research is warranted to examine the functional role of this region and its longitudinal association with sleep. **Keywords:** DNA methylation; sleep; psychopathology; accelerometer; epigenetics; epigenome.

Introduction

Sleep is increasingly recognized as an important factor in child mental health. Sleep disturbances, such as short sleep and shifted circadian rhythm, often develop in late childhood and have been implicated in mental health problems (Gregory & Sadeh, 2016; Wolfson & Carskadon, 1998). While poor sleep can exacerbate mental health difficulties (Lovato & Gradisar, 2014; Owens, 2005), mental health problems can also precede and worsen sleep (Verhoeff et al., 2018). Thus, the association between sleep and mental health is complex and likely bidirectional (Gregory & Sadeh, 2016). The mechanisms underlying this association, however, remain unknown.

Complex traits, including sleep, result from the interplay of genetic and environmental influences (Romens, McDonald, Svaren, & Pollak, 2015). How these factors jointly influence normative sleep, or the development of sleep problems, is currently unclear. Epigenetic processes such as DNA methylation (DNAm) have been proposed as a mechanism of interest (Massart et al., 2014; Morales-Lara, De-la-Pena, & Murillo-Rodriguez, 2018). Differential DNAm has been linked to a broad range of developmental outcomes, including sleep, as well as mental and physical health problems (Barker, Walton, & Cecil, 2018; Breton et al., 2017). Most research on this topic emerges from animal models (Gaine, Chatterjee, & Abel, 2018), with only a handful of studies examining DNAm and sleep in humans. These have typically relied on small samples of adults with dysregulated sleep (e.g. shift workers) and utilized a candidate gene approach focusing primarily on 'clock' genes: genes driving circadian rhythms in metabolism, physiology, and behavior (Cedernaes et al., 2015; Gaine et al., 2018; Lahtinen et al., 2019; Wong et al., 2015). In contrast, we are aware of only two epigenetic studies during development, both of which examined adolescence. One reported an association in 18- to 19 year-olds

Conflict of interest statement: No conflicts declared.

between sleep duration and DNAm of *DOCK1*, a gene influenced by circadian rhythmicity (Huang et al., 2017). The second found that higher DNAm in metabolic genes *PPARA* and *HSD11B2* was associated with shorter sleep, specifically in girls (Jansen et al., 2019).

Despite these promising preliminary findings, existing research has been limited in four key ways, namely (a) the use of small samples of adults or older adolescents; (b) a focus on a candidate gene approach; (c) the lack of multimodal assessments of sleep, making it unclear whether associations between sleep and DNAm differ between self-report and objective measures (e.g. actigraphy); and (d) despite evidence showing that mental health is related to DNAm alterations (Barker et al., 2018) and sleep (Gregory & Sadeh, 2016), no study has examined these factors jointly.

To address these gaps, we examined the relationship between genome-wide DNAm, sleep, and mental health in a general population sample of 10-year-old children – an important period for the development of sleep and mental health problems alike. The aims of our study were twofold: first, to characterize cross-sectional associations of DNAm with reported (i.e. dyssomnia symptoms) and actigraphy-assessed (i.e. sleep duration and midpoint) sleep using both a genome-wide approach and an targeted approach focusing on well-characterized clock genes to maximize comparability with existing studies; and second, to investigate whether sleep-associated DNAm patterns are also associated with common mental health problems. Findings were tested for consistency in a small independent sample.

Methods

Participants

This cross-sectional study included 10-year-old children of European ancestry (51.3% female) from the *Generation R Study*, a prospective population-based cohort from fetal life onward. Pregnant women (expected delivery date April 2002–January 2006) living in Rotterdam, the Netherlands, were invited to participate (Kooijman et al., 2016). The current analyses are based on children who had DNAm data and subjectively assessed sleep (n = 410). Of these, 188 also had actigraphy data. Written informed consent was obtained for all participants. The Medical Ethical Committee of the Erasmus MC, University Medical Center Rotterdam approved the study.

Measures

DNA Methylation. Five-hundred nanograms of DNA were extracted from peripheral blood at age 10 and underwent bisulfite conversion with the EZ-96 DNA Methylation kit (Shallow) (Zymo Research Corporation, Irvine, CA). Samples were plated onto 96-well plates in no specific order. DNAm was analyzed with the Illumina Infinium Human Methylation 450K BeadChip (Illumina Inc., San Diego, CA). Quality control of samples was performed using standardized criteria using the CPACOR workflow (Lehne et al., 2015). Probes with a detection p value above background ≥ 1E-16 were set to missing per array. Arrays with observed technical problems including failed bisulfite conversion, hybridization, or extension, and arrays with a mismatch between child sex and sex determined by the chr X and Y probe intensities were removed. Nonautosomal probes were excluded. Additionally, only arrays with a call rate >95% per sample were processed further. Methylation beta values outside a range of the 25th percentile minus 3*interquartile range to the 75th percentile plus 3*interquartile range were set to missing. The final dataset contained 425 samples, analyzing 458,563 CpG sites. For our targeted approach, we examined DNAm levels of CpG sites that were annotated to well-characterized clock-related genes (939 CpG sites across 39 genes (van den Berg et al., 2017, Table S1). For each CpG site, beta values represent the ratio of methylated signal divided by the sum of the methylated and unmethylated signals plus 100.

Child-reported dyssomnia symptoms. At age 10 years, children completed six questions of the Sleep Disturbance Scale for Children (Bruni et al., 1996) about perceived sleep, for example, 'Do you find it difficult to fall asleep?'; 'If you wake up at night, do you find it difficult to fall asleep again?'; 'Do you feel rested when you wake in the morning?' (previously described in Koopman-Verhoeff et al., 2019). The questions were rephrased for our pediatric population. Responses were scored on a three-point Likert scale ('No', 'Sometimes', or 'Yes'; α = .64). Items were summed; higher scores indicate greater sleep problems.

Actigraphy-estimated sleep. Sleep patterns were estimated with wrist tri-axial actigraphy (GENEActiv) on the nondominant wrist for five consecutive school nights in 188 children at age 11 (i.e. after DNA sampling) (Koopman-Verhoeff et al., 2018; Koopman-Verhoeff et al., 2019). The Geneactiv accelerometers were set a frequency of 50 Hz. The binary files were processed with the R-package GGIR (van Hees et al., 2014). Accompanying sleep diaries were collected and used to guide actigraphy analyses. Sleep duration was estimated as the total time scored sleep between falling asleep and final waking. Sleep midpoint was estimated as the halfway point between sleep onset and final waking. Sleep duration and midpoint were averaged across the week, excluding weekends to best approximate typical school-day sleep patterns and to minimize the influence of atypical weekend events.

Child psychopathology. The Child Behavior Checklist 6–18 (CBCL/6–18) was assessed using maternal reports at age 10 to derive broadband Internalizing and Externalizing problem-scales (Achenbach & Rescorla, 2001). The CBCL/6–18 is widely used internationally and has been found to be generalizable across 23 societies, including the Netherlands (Ivanova et al., 2010). Mothers rated various emotional and behavioral problems of the child in the previous six months on a three-point scale (0 = not true, 1 = somewhat true, 2 = very true).

Covariates

Sex of the child was obtained from medical records and maternal characteristics by questionnaires. Maternal education was defined by the highest attained educational level and classified into two categories (higher vocational education and university: yes or no). Correction for sample plate and cell type proportions was also applied. We used the Houseman method (Houseman et al., 2012) to estimate relative proportions of six white blood cell subtypes (CD4 + T-lymphocytes, CD8 + T-lymphocytes, NK (natural killer) cells, B-lymphocytes, monocytes, and granulocytes), based on a standard reference population (Reinius et al., 2012).

Statistical analysis

We had nearly complete cases, with four participants missing data on maternal education (defined as highest educational level). These participants were excluded from the analysis. Statistical analyses were performed in R (R Core Team, 2014), following three steps:

Step 1. Associations between DNA methylation and sleep. We applied weighted gene co-expression network analysis (WGCNA, Langfelder & Horvath, 2008) – a system-level data reduction approach – to reduce the dimensionality of the data and identify clusters (so-called 'modules') of highly comethylated DNAm sites across genome. As such, rather than focusing on individual sites or genes, WGCNA enables utilization of correlation patterns between sites to identify wider DNAm networks, which may also be functionally related (Botía et al., 2017). Block-wise network construction was run using default settings (power threshold of 6; minimal module size of 30 sites; merge cut height of 0.25). Each derived module was colored by size automatically and summarized by a 'module eigengene' (ME) value, the first principal component of the given module. We numbered the derived modules by significance with outcome for simplicity. CpG sites that do not comethylate were assigned to an 'unclassified' module. WGCNA analyses were performed twice: first based on the entire genome-wide data (i.e. hypothesis free; n = 458,563 sites) and second based on the subset of clock genes (i.e. targeted approach, n = 939 CpG sites).

Next, we tested bivariate correlations between the comethylated modules and the three sleep outcomes (i.e. child-reported dyssomnia symptoms, actigraphy-estimated sleep duration, and midpoint sleep). We selected modules that were associated with sleep outcomes after Bonferroni correction for multiple testing (0.05/n modules*3 sleep measures) (Chuang et al., 2017). These modules were further examined using linear regression models controlling for batch, cell types, child sex and age, and maternal education.

Modules that were significantly associated with sleep were examined further using publicly available resources to characterize (a) their genomic location; (b) potential genetic influences, by checking whether the CpG sites included in the modules are known to be *polymorphic* (i.e. overlapping with single nucleotide polymorphisms [SNPs]; Chen et al., 2013), linked to *methylation quantitative trait loci* (mQTLs; i.e. SNPs that associate with DNAm levels, either in *cis* or in *trans*; http://www.mqtldb.org/; GCTA set; Gaunt et al., 2016) or *heritable*, based on twin data (i.e. explained by additive genetic influences as opposed to shared and nonshared environmental influences; Hannon et al., 2018); and (c) blood–brain concordance, based on postmortem data from 122 individuals with DNAm from whole blood and four brain regions (the prefrontal cortex, entorhinal cortex, superior temporal gyrus, and cerebellum (https://epigenetics.essex.ac.uk/bloodbrain/; Hannon, Lunnon, Schalkwyk, & Mill, 2015).

Step 2: Testing the overlap of associations with mental health. Bivariate correlations between the comethylated modules, sleep, and mental health measures were examined to establish whether associations of DNAm and sleep are colocalized on the genome with associations of DNAm and internalizing and externalizing problems.

Step 3. Generalizability in independent sample. Associations identified in Steps 1 and 2 were estimated in an independent sample of 63 older adolescents (14.5 ± 0.3 years, 54% girls) of the Generation R Study to judge generalizability of results, with information on DNAm available at 10 years and actigraphy-assessed sleep at 14 years (i.e. prospective association). The children in this sample were recruited for a second actigraphy study at a later age than the first study described above due to logistic reasons (no repeated measurements).

Results

Characteristics of the study sample are presented in Table 1. For correlations across sleep and mental health variables, see Table S2. The average midpoint sleep was 2:49 (SD = 35 min), and the mean sleep duration was 7:36 (SD = 40 min).

Are DNAm patterns associated with sleep outcomes in children?

Genome-wide analyses. We identified 64 comethylated modules, containing between 30 and 65,804 CpG sites (Table S3). The majority of sites were unclassified (n = 261,374), suggesting they did not correlate strongly enough to form modules. Two modules correlated with sleep after Bonferroni correction for multiple testing (0.05/64 modules * 3 outcomes = 0.00026042) – both of which associated with sleep duration (module1 r = −.18, p = .00006, module2 r = −.18, p = .0001) (Table 2), but not with sleep midpoint or dyssomnia symptoms. Only the association between module1 and sleep duration remained significant in a regression model adjusting for covariates (ß = −.22, 95 CI%: −0.37 to −0.07, p = .004). As a sensitivity analysis, we replaced the missing values (n = 4) on maternal highest educational level attained, yielding highly consistent results. Additionally, as time of blood sampling corrected for the time of habitual awakening could be of influence, we re-ran analyses adjusting for these variables and found that results remained highly consistent (ß = −.19, 95 CI%: −0.34 to −0.05,

Table 1 Sample characteristics

Demographics	Reported dyssomnia symptoms (N = 410)	Actigraphic sleep (N = 188)
Sex, female, %	234 (50.3%)	93 (49.5%)
Age (years)	9.8 ± 0.3	11.7 ± 0.1
Maternal education, %		
Low and Intermediate	152 (32.7%)	64 (34.0%)
High	308 (66.2%)	121 (64.4%)
Dyssomnia symptoms, self-reported (score; range)	10.80 (8.00–18.00)	10.86 (6.00–17.00)
Sleep duration, actigraphy (hr:min)	–	7:35 ± 0:44
Midpoint sleep, actigraphy, time (hr:min)	–	02:48 ± 0:35
Internalizing problems, mother-reported, mean (SD)	4.16 (4.38)	4.03 (4.28)
Externalizing problems, mother-reported, mean (SD)	3.41 (4.25)	3.16 (3.82)

p = .008). Lastly, as cell proportions are estimated, rather than derived from actual cell counts, we reran analyses without cell type correction to test stability of associations and found that results were highly consistent (ß = −.22, 95 CI%: −0.36 to −0.07, p = .004).

Targeted circadian clock CpG site analyses. The targeted WGCNA approach containing exclusively clock-related genes identified five modules (ranging from 19–300 CpG sites over 10–39 genes), each including CpG sites spanning multiple genes, as opposed to clustering by gene. The majority of the CpG sites were unclassified (n = 540). No modules were associated with sleep outcomes after multiple testing correction.

Functional characterization of module1. Annotation to genes and genomic region: Module1 contained 32 sites spanning 9 genes (Table S4). The largest number of sites (n = 6) were annotated to the Microtubule-Associated Protein Tau (*MAPT*) gene. The CpGs of module1 were highly correlated with each other (Figure 1), as well as with sleep duration, and were all located in the chromosome 17q21.31 region, chr17:43502999-62843696, with the exception of one CpG site on chromosome 5.

Genetic influences: Six of the CpGs included in module1 were previously identified as polymorphic (three of which in *MAPT*), and twelve (37.5%) were found to be associated with mQTLs on chromosome 17, with a total of 71 associations (between 4 and 10 associations per CpG). The CpG site located on chromosome 5 (cg07870213) associated with both mQTLs on chromosome 5 in *cis* and chromosome 17 in *trans*, all of which were located in the module1 region (chr17:41993881-44852612). Finally, 10 of the 32 CpG sites in module1 had twin heritability estimates available, all of which showed moderate to strong genetic influences (r = .34–1.00).

Blood–brain concordance: For all but one of the CpG sites in module1, DNAm levels in blood correlated significantly with DNAm levels in at least one brain region. The three *MAPT* CpG sites that associated most strongly with sleep duration showed high blood–brain correlations (Figure S1). Of these, cg24801230 (one of the sites found to be polymorphic) showed an almost perfect correlation (r = .99) between blood and brain, with DNAm levels across tissues clustering into three alleles (Figure S1).

Are DNAm-sleep associations overlapping with child psychiatric symptoms?

No modules were associated with internalizing and externalizing problems after correction for multiple testing. Generally, we found weak associations between the DNAm modules and internalizing (strongest association: r = .15, p = .001) and externalizing problems (strongest association: r = .14, p = .002). Associated modules did not overlap with those identified for sleep duration (Figure S2).

Are results consistent in an independent sample?

The association between module1 and sleep duration was tested in an independent sample of older children, in order to test for consistency across developmental stage. Results from a regression analysis, controlling for covariates, yielded a highly comparable effect size (Discovery: ß = −.22, 95 CI%: −0.37 to −0.07, p = .004; Generalization sample**:** ß = −.23, 95 CI%: −0.50 to 0.04, p = .09), although the association was not statistically significant, likely due to the larger confidence intervals resulting from the use of a smaller sample (1/3 of discovery sample).

Discussion

The current study utilized a network-based approach to investigate associations between genome-wide DNA methylation, sleep, and mental health in a pediatric population. We highlight here two key findings. First, we found that DNAm patterns associated with sleep duration, but not with other sleep parameters. Specifically, our hypothesis-free analyses identified one DNAm module associated with actigraphy-assessed sleep duration. This module (a) contained 32 sites annotated to multiple genes previously linked to sleep duration in GWASes, including *MAPT*; (b) showed strong evidence of genetic influences based on molecular and twin data; and (c) showed cross-tissue concordance between blood and brain. In contrast, hypothesis-

Table 2 Associations between DNAm modules and actigraphy-derived sleep duration in children (N = 188)

	A. Correlations of the WGCNA modules with sleep duration				B. Standardized regression coefficients		
Module	r	p value	N cpgs	N genes	β	CI	p value
Module1	−.18	.00006	32	9	−.22	−.37 to −.07	.004
Module2	−.18	.0001	5845	3462	−.14	−.54 to 36	.07

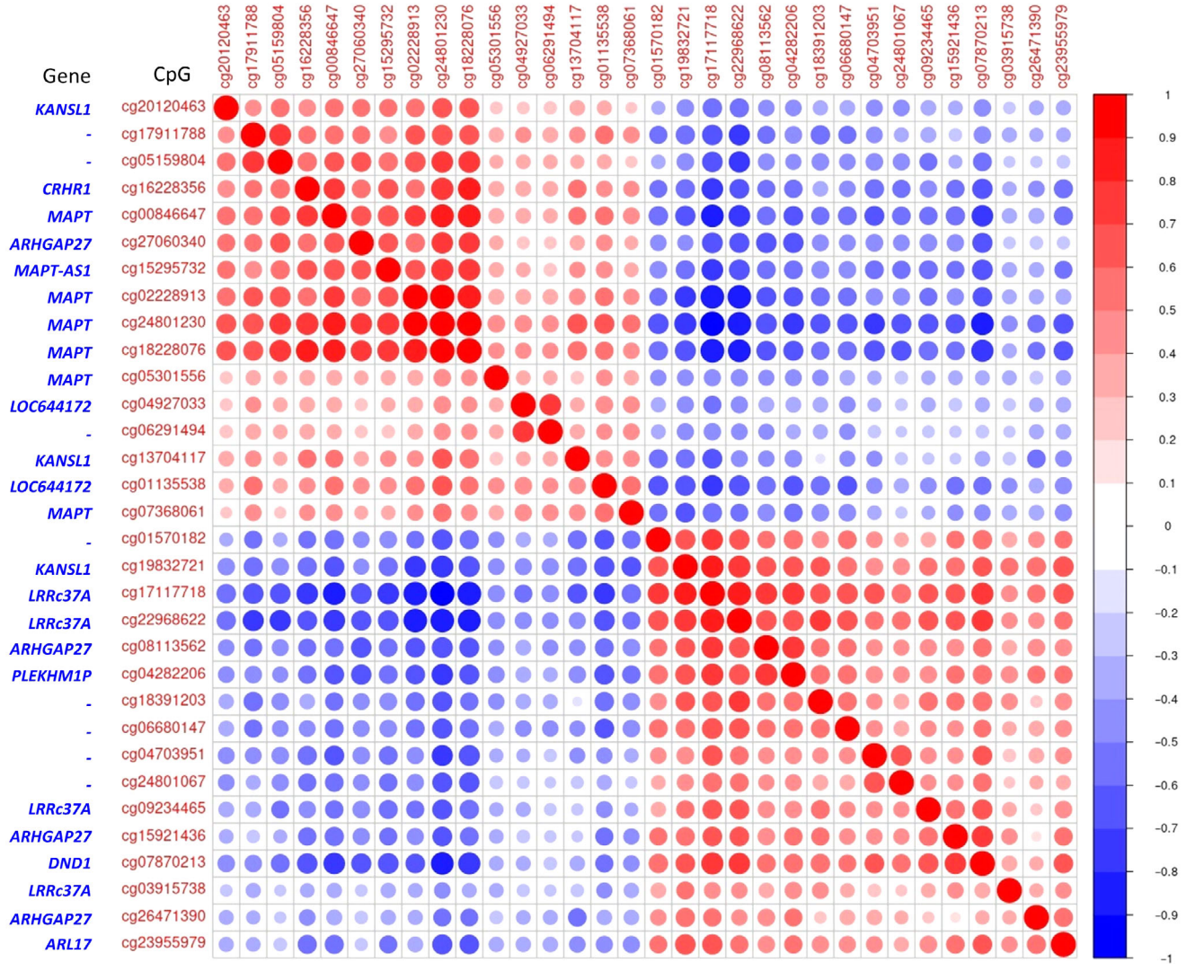

Figure 1 Intercorrelations between CpG sites in module1

driven analyses did not reveal associations between DNAm in clock genes and sleep parameters. Second, we found that DNAm patterns were only weakly associated with mental health outcomes. These associations did not overlap with those identified for sleep outcomes, suggesting comethylation modules associated with sleep and mental health are largely independent.

Self-reported and actigraphic sleep assess distinct sleep domains (Gregory & Sadeh, 2016; Meltzer et al., 2012), as reflected in the weak correlations between these metrics found in the present study. Of note, self-reported measures capture sleep perception and reports may be biased by subject characteristics. Interestingly, we found here that DNAm associated with actigraphic sleep duration but not with self-reported dyssomnia. This could be due to the fact that actigraphic sleep shows greater variability in the general population and has less measurement error (Sadeh, 2011). Furthermore, we did not find associations between DNAm and actigraphic determined midpoint sleep. Nights assessed in our sample have been constrained by school schedules, limiting variability in midpoint. Since circadian preference changes during adolescence (Crowley, Wolfson, Tarokh, & Carskadon, 2018), future research should study the longitudinal association between DNAm, and sleep and circadian rhythm across this age period.

Most epigenetic research on sleep in humans has focused on sleep deprivation (Gaine et al., 2018). In this study, we show that DNAm patterns associate with *typical* variation in sleep in 10-year-old children. Specifically, one DNAm module was found to associate with actigraphic sleep duration. This association was generalizable to a smaller, independent sample of Generation R participants at age 14 years. The lack of significance could be due to low power in this smaller sample. The fact that we found a generally comparable effect size supports the robustness of our findings.

The sleep-associated module contained 32 CpG sites spanning a large region on chromosome 17. Based on accessible databases, we found that

several of the sites in the module were located directly on SNPs, and over a third were linked to known mQTLs. Intriguingly, the one CpG site in this module on chromosome 5 was associated with multiple mQTLs located within the chromosome 17 region, supporting a genetically driven link in DNAm patterns between these two chromosomal regions. Genetic influences were further corroborated by twin data showing moderate-to-high heritability estimates for DNAm sites in this module. Together, these findings suggest that underlying genetic variation might largely account for observed associations between DNAm in this region and sleep duration. This is in line with existing literature indicating that variation in DNAm is best explained by genetic influences and gene–environment interactions, as opposed to environmental main effects (Czamara et al., 2019; Teh et al., 2014). Finally, DNAm variability in the identified module showed high blood–brain concordance, highlighting that the signals currently found in blood might be useful proxies for DNAm status in the brain. Future studies will need to test concordance with other brain areas implicated in sleep duration, for example, the hypothalamus, and establish whether the degree of correspondence differs across specific cell types in the brain.

Of the nine genes annotated to our module, several stood out for their role in brain-related processes and previous links to sleep outcomes based on GWAS data. Specifically, a single nucleotide polymorphism (SNP) in *MAPT* was recently identified as a top GWAS hit for self-reported sleep duration (Dashti et al., 2019) and SNPs in *MAPK81P1P2* and *KANSL1-AS1* were identified as top hits in a GWAS on accelerometer-based sleep duration (Doherty et al., 2018). Additionally, a study based on UK Biobank and 23andMe data indicated that variants in *ARHGAP27*, *LRRC37A*, *CRHR1*, *MAPT*, and *KANSL1* associated with various self-reported sleep traits, including sleep duration (Jansen et al., 2019). These findings further support genetic influences on DNAm and sleep duration in this region.

The most strongly associated probe in module1 was annotated to the *MAPT* antisense *RNA 1*, a nonprotein coding RNA gene identified as epigenetic regulator of *MAPT* expression (Coupland et al., 2016), while six sites were annotated to the *MAPT* gene itself. *MAPT* encodes the Tau protein, which is important for neuronal stabilization. Its aberrant aggregation has been frequently linked to Alzheimer's disease and other neurodegenerative diseases (Wang & Mandelkow, 2015) as well as neurodevelopmental disorders (Rankovic & Zweckstetter, 2019). A recent study suggested the involvement of Tau proteins and sleep in the pathogenesis of neurodegenerative diseases, though this process is not yet fully understood (Cantero et al., 2010; Musiek & Holtzman, 2016). Another gene annotated to module1 was *CRHR1* (corticotropin-releasing hormone receptor 1), a pivotal player in hypothalamic–pituitary–adrenal axis functioning (Wasserman, Wasserman, & Sokolowski, 2010) as well as sleep (Romanowski et al., 2010). Our study adds to this growing body of evidence by showing for the first time that, in childhood, epigenetic variation in *MAPT* and surrounding regions are associated with sleep duration.

The epigenetic patterns associated with sleep in this study did not overlap with those associated with mental health. This may be due to several reasons. First, although the link between sleep and mental health is well-established (Gregory & Sadeh, 2016), it is possible that such associations may not be epigenetically mediated. Second, associations between sleep and mental health tend to be stronger for self-report than objective measures (Gregory & Sadeh, 2012). As such, there might be different underlying biological correlates driving the associations between mental health and reported sleep and actigraphic derived sleep. For example, cortisol levels, associated with anxiety and depression, have been linked to self-reported sleep quality but not to actigraphy-derived sleep quantity (Bassett, Lupis, Gianferante, Rohleder, & Wolf, 2015). Third, our population-based cohort may have lacked psychiatric severity to detect shared associations. Future studies are needed to clarify the mechanisms underlying associations between sleep and mental health.

Limitations and future directions

This study has several limitations. First, from our cross-sectional data, we are unable to determine the direction of effect for the association between DNAm and sleep regulation, and we cannot exclude that the observed association may result from a common influence (e.g. environmental or genetic modulation). In the future, the use of longitudinal data on DNAm and sleep, the application of advanced causal inference methods (e.g. two-step Mendelian randomization), and integration with genetic data will mark important steps for furthering our understanding of DNAm-sleep associations. Second, the sample was based on participants of European ancestry. Studies including other ethnicities are necessary to investigate the generalizability of our findings. Third, our independent sample was smaller, limiting statistical power. Fourth, our measure of midpoint sleep, derived from actigraphy, is constrained by school schedules. Studying free nights may better describe underlying circadian processes. Fifth, while we assume that focusing on modules as opposed to single sites may help us to identify broader, functionally meaningful DNAm networks associated with sleep, (a) this does not preclude that there may be important sleep-associated single CpG sites, which might have been missed by using this approach; and (b) integration with gene expression data will be necessary to establish the extent to which the

identified module may play a regulatory role, which we could not do in our study. In addition to the clock genes tested in the current study, it would be interesting to examine associations with CpG sites annotated to genes that have been previously implicated in other sleep parameters, such as sleep duration or chronotype (e.g. by GWAS studies). Sixth, the blood–brain concordance tool we used is based on an elderly population. As such, it is unclear to what extent the identified pattern of concordance extends to the pediatric population, for which there are currently no available tools. Finally, it is unclear whether identified DNAm patterns are functionally relevant. The use of experimental models could inform the biological consequences of these associations. Additionally, it is important to see in future studies whether DNAm levels at these sites change across development. If there is no change in DNAm levels over time, this could indicate that a regulatory process is acting from birth, whereas an epigenetic mark that changes throughout life might indicate that it may be responsive to environmental stimuli.

Conclusion

In summary, the preliminary results of the current study show promising sleep-associated DNAm patterns in the pediatric population. Specifically, we identified an association between sleep duration and DNAm in the 17q21.31 region, spanning multiple genes previously linked to sleep by GWAS studies, including *MAPT*. These epigenetic patterns did not overlap with those associated with self-reported sleep problems, midpoint sleep, or mental health. Future studies are needed to replicate our findings and establish causality. Overall, our findings offer novel insights into epigenetic patterns associated with typical variation in sleep duration in children.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Figure S1. Blood-brain associations of *MAPT* CpG sites.
Figure S2. Correlation matrix of DNAm modules, sleep and mental health.
Table S1. Selection of clock and clock-related genes based on van den Berg et al., 2017 (ngenes = 39; nCpGs = 939).
Table S2. Correlations between dyssomnia symptoms, actigraphic sleep, and mental health.
Table S3. WGCNA-derived modules and number of CpG sites.
Table S4. Functional characterization of module1.

Acknowledgements

The authors thank all participants and parents, students, practitioners, hospitals, midwives, and pharmacies. The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. The generation and management of the Illumina 450K methylation array data (EWAS data) for the Generation R Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and the Netherlands. The authors thank all colleagues involved in generation and management of methylation data and genotyping.

The general design of the Generation R Study is made possible by financial support from Erasmus MC, Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development and the Ministry of Health, Welfare and Sport. The EWAS data were funded by a grant from the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA; project nr. 050-060-810), by funds from the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by a grant from the National Institute of Child and Human Development (R01HD068437). This study received support from the Erasmus Medical Center Efficiency Grant (MRC-2013-169) to H.T., and H.T. was supported by a grant from NWO (016.VICI.170.200). This project has received funding from the European Union's Horizon 2020 Research And Innovation Programme under the grant agreement number 733206 (LifeCycle) and under the Marie Skłodowska-Curie grant agreement number No 707404 awarded to C.C.. Additionally, from the European Joint Programming Initiative 'A Healthy Diet for a Healthy Life' (JPI HDHL, NutriPROGRAM project, ZonMw the Netherlands no. 529051022). The authors have declared that they have no competing or potential conflicts of interest.

Correspondence

Charlotte A.M. Cecil, Department of Child and Adolescent Psychiatry/Psychology, Erasmus Medical Center, Room NA-2815, P.O. Box 2060, 3000 CB Rotterdam, The Netherlands; Email: c.cecil@erasmusmc.nl

Key points

- Our study is the first to examine the association of genome-wide DNA methylation with sleep and mental health outcomes in the general pediatric population.
- This study utilized multimethod assessments of child sleep, including reported and actigraphy-assessed sleep outcomes.
- We found an association between sleep duration and DNAm in the 17q21.31 region, including genes previously linked to sleep within GWAS studies, such as the *MAPT* gene.
- Epigenetic patterns associated with sleep duration were not associated with mental health outcomes.
- Consistent findings were observed in a smaller, independent sample of children from Generation R.
- Overall, our study offers important new insights into the relationship between DNAm, sleep, and mental health in the pediatric population.

References

Achenbach, T.A., & Rescorla, L.A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Barker, E.D., Walton, E., & Cecil, C.A.M. (2018). Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *59*, 303–322.

Bassett, S.M., Lupis, S.B., Gianferante, D., Rohleder, N., & Wolf, J.M. (2015). Sleep quality but not sleep quantity effects on cortisol responses to acute psychosocial stress. *Stress*, *18*, 638–644.

Botía, J.A., Vandrovcova, J., Forabosco, P., Guelfi, S., D'Sa, K., Hardy, J., … & The United Kingdom Brain Expression Consortium (2017). An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Systems Biology*, *11*, 47.

Breton, C.V., Marsit, C.J., Faustman, E., Nadeau, K., Goodrich, J.M., Dolinoy, D.C., … & Murphy, S.K. (2017). Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: The children's environmental health and disease prevention research center's epigenetics working group. *Environmental Health Perspectives*, *125*, 511–526.

Bruni, O., Ottaviano, S., Guidetti, V., Romoli, M., Innocenzi, M., Cortesi, F., & Giannotti, F. (1996). The Sleep Disturbance Scale for Children (SDSC) Construction and validation of an instrument to evaluate sleep disturbances in childhood and adolescence. *Journal of Sleep Research*, *5*, 251–261.

Cantero, J.L., Hita-Yanez, E., Moreno-Lopez, B., Portillo, F., Rubio, A., & Avila, J. (2010). Tau protein role in sleep-wake cycle. *Journal of Alzheimer's Disease*, *21*, 411–421.

Cedernaes, J., Osler, M.E., Voisin, S., Broman, J.E., Vogel, H., Dickson, S.L., … & Benedict, C. (2015). Acute sleep loss induces tissue-specific epigenetic and transcriptional alterations to circadian clock genes in men. *Journal of Clinical Endocrinology and Metabolism*, *100*, E1255–E1261.

Chen, Y.-A., Lemire, M., Choufani, S., Butcher, D.T., Grafodatskaya, D., Zanke, B.W., … & Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*, *8*, 203–209.

Chuang, Y.H., Paul, K.C., Bronstein, J.M., Bordelon, Y., Horvath, S., & Ritz, B. (2017). Parkinson's disease is associated with DNA methylation levels in human blood and saliva. *Genome Medicine*, *9*, 76–88.

R Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. Available from: http://www.R-project.org/

Coupland, K.G., Kim, W.S., Halliday, G.M., Hallupp, M., Dobson-Stone, C., & Kwok, J.B.J. (2016). Role of the long non-coding RNA MAPT-AS1 in regulation of microtubule associated protein tau (MAPT) expression in Parkinson's disease. *PLoS ONE*, *11*, e0157924.

Crowley, S.J., Wolfson, A.R., Tarokh, L., & Carskadon, M.A. (2018). An update on adolescent sleep: New evidence informing the perfect storm model. *Journal of Adolescence*, *67*, 55–65.

Czamara, D., Eraslan, G., Page, C.M., Lahti, J., Lahti-Pulkkinen, M., Hämäläinen, E., … & Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (2019). Integrated analysis of environmental and genetic influences on cord blood DNA methylation in new-borns. *Nature Communications*, *10*, 2548–2566.

Dashti, H.S., Jones, S.E., Wood, A.R., Lane, J.M., van Hees, V.T., Wang, H., … & Saxena, R. (2019). Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nature Communications*, *10*, 1100–1112.

Doherty, A., Smith-Byrne, K., Ferreira, T., Holmes, M.V., Holmes, C., Pulit, S.L., & Lindgren, C.M. (2018). GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nature Communications*, *9*, 5257–5265.

Gaine, M.E., Chatterjee, S., & Abel, T. (2018). Sleep deprivation and the epigenome. *Frontiers in Neural Circuits*, *12*, 14–24.

Gaunt, T.R., Shihab, H.A., Hemani, G., Min, J.L., Woodward, G., Lyttleton, O., … & Relton, C.L. (2016). Systematic identification of genetic influences on methylation across the human life course. *Genome Biology*, *17*, 61–75.

Gregory, A.M., & Sadeh, A. (2012). Sleep, emotional and behavioral difficulties in children and adolescents. *Sleep Medicine Reviews*, *16*, 129–136.

Gregory, A.M., & Sadeh, A. (2016). Annual Research Review: Sleep problems in childhood psychiatric disorders–a review of the latest science. *Journal of Child Psychology and Psychiatry*, *57*, 296–317.

Hannon, E., Knox, O., Sugden, K., Burrage, J., Wong, C.C.Y., Belsky, D.W., … & Mill, J. (2018). Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *Plos Genetics*, *14*, e1007544–e1007571.

Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation across blood, cortex, and cerebellum: Implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, *10*, 1024–1032.

Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., … & Kelsey, K.T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, *13*, 1–16.

Huang, H., Zhu, Y., Eliot, M.N., Knopik, V.S., McGeary, J.E., Carskadon, M.A., & Hart, A.C. (2017). Combining human epigenetics and sleep studies in *Caenorhabditis elegans*: A cross-species approach for finding conserved genes regulating sleep. *Sleep*, *40*, 1–13.

Ivanova, M.Y., Achenbach, T.M., Rescorla, L.A., Harder, V.S., Ang, R.P., Bilenberg, N., … & Verhulst, F.C. (2010). Preschool psychopathology reported by parents in 23 societies: Testing the seven-syndrome model of the child behavior checklist for ages 1.5-5. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*, 1215–1224.

Jansen, E.C., Dolinoy, D.C., O'Brien, L.M., Peterson, K.E., Chervin, R.D., Banker, M., … & Goodrich, J.M. (2019). Sleep duration and fragmentation in relation to leukocyte DNA methylation in adolescents. *Sleep*, *42*, 1–9.

Jansen, P.R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A.R., … & The 23andMe Research Team (2019). Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nature Genetics*, *51*, 394–403.

Kooijman, M.N., Kruithof, C.J., van Duijn, C.M., Duijts, L., Franco, O.H., van IJzendoorn, M.H., … & Jaddoe, V.W. (2016). The Generation R Study: Design and cohort update 2017. *European Journal of Epidemiology*, *31*, 1243–1264.

Koopman-Verhoeff, M.E., Bolhuis, K., Cecil, C.A.M., Kocevska, D., Hudziak, J.J., Hillegers, M.H.J., … & Tiemeier, H. (2018). During day and night: Childhood psychotic experiences and objective and subjective sleep problems. *Schizophrenia Research*, *206*, 127–134.

Koopman-Verhoeff, M.E., Serdarevic, F., Kocevska, D., Bodrij, F.F., Mileva-Seitz, V.R., Reiss, I., … & Luijk, M. (2019). Preschool family irregularity and the development of sleep problems in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *60*, 857–865.

Lahtinen, A., Puttonen, S., Vanttola, P., Viitasalo, K., Sulkava, S., Pervjakova, N., … & Paunio, T. (2019). A distinctive DNA methylation pattern in insufficient sleep. *Scientific Reports*, *9*, 1193–1202.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*, 559–572.

Lehne, B., Drong, A.W., Loh, M., Zhang, W., Scott, W.R., Tan, S.-T., … & Elliott, P. (2015). A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology*, *16*, 37–49.

Lovato, N., & Gradisar, M. (2014). A meta-analysis and model of the relationship between sleep and depression in adolescents: Recommendations for future research and clinical practice. *Sleep Medicine Reviews*, *18*, 521–529.

Massart, R., Freyburger, M., Suderman, M., Paquet, J., El Helou, J., Belanger-Nelson, E., … & Mongrain, V. (2014). The genome-wide landscape of DNA methylation and hydroxymethylation in response to sleep deprivation impacts on synaptic plasticity genes. *Transl Psychiatry*, *4*, e347–e355.

Meltzer, L.J., Biggs, S., Reynolds, A., Avis, K.T., Crabtree, V.M., & Bevans, K.B. (2012). The children's report of sleep patterns – sleepiness scale: A self-report measure for school-aged children. *Sleep Medicine*, *13*, 385–389.

Morales-Lara, D., De-la-Pena, C., & Murillo-Rodriguez, E. (2018). Dad's snoring may have left molecular scars in your DNA: The emerging role of epigenetics in sleep disorders. *Molecular Neurobiology*, *55*, 2713–2724.

Musiek, E.S., & Holtzman, D.M. (2016). Mechanisms linking circadian clocks, sleep, and neurodegeneration. *Science (New York, NY)*, *354*, 1004–1008.

Owens, J.A. (2005). The ADHD and sleep conundrum: A review. *Journal of Developmental and Behavioral Pediatrics*, *26*, 312–322.

Rankovic, M., & Zweckstetter, M. (2019). Upregulated levels and pathological aggregation of abnormally phosphorylated Tau-protein in children with neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews*, *98*, 1–9.

Reinius, L.E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.-E., Greco, D., … & Kere, J. (2012). Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLoS ONE*, *7*, e41361.

Romanowski, C.P.N., Fenzl, T., Flachskamm, C., Wurst, W., Holsboer, F., Deussing, J.M., & Kimura, M. (2010). Central deficiency of corticotropin-releasing hormone receptor type 1 (CRH-R1) abolishes effects of CRH on NREM but not on REM sleep in mice. *Sleep*, *33*, 427–436.

Romens, S.E., McDonald, J., Svaren, J., & Pollak, S.D. (2015). Associations between early life stress and gene methylation in children. *Child Development*, *86*, 303–309.

Sadeh, A. (2011). The role and validity of actigraphy in sleep medicine: An update. *Sleep Medicine Reviews*, *15*, 259–267.

Teh, A.L., Pan, H., Chen, L., Ong, M.-L., Dogra, S., Wong, J., … & Holbrook, J.D. (2014). The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Research*, *24*, 1064–1074.

van den Berg, C.B., Chaves, I., Herzog, E.M., Willemsen, S.P., van der Horst, G.T.J., & Steegers-Theunissen, R.P.M. (2017). Early- and late-onset preeclampsia and the DNA methylation of circadian clock and clock-controlled genes in placental and newborn tissues. *Chronobiology International*, *34*, 921–932.

van Hees, V.T., Fang, Z., Langford, J., Assah, F., Mohammad, A., da Silva, I.C., … & Brage, S. (2014). Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: An evaluation on four continents. *Journal of Applied Physiology*, *1985*(117), 738–744.

Verhoeff, M.E., Blanken, L.M.E., Kocevska, D., Mileva-Seitz, V.R., Jaddoe, V.W.V., White, T., … & Tiemeier, H. (2018). The bidirectional association between sleep problems and autism spectrum disorder: A population-based cohort study. *Molecular Autism*, *9*, 1–9.

Wang, Y., & Mandelkow, E. (2015). Tau in physiology and pathology. *Nature Reviews Neuroscience*, *17*, 5–21.

Wasserman, D., Wasserman, J., & Sokolowski, M. (2010). Genetics of HPA-axis, depression and suicidality. *European Psychiatry*, *25*, 278–280.

Wolfson, A.R., & Carskadon, M.A. (1998). Sleep schedules and daytime functioning in adolescents. *Child Development*, *69*, 875–887.

Wong, C.C., Parsons, M.J., Lester, K.J., Burrage, J., Eley, T.C., Mill, J., … & Gregory, A.M. (2015). Epigenome-wide DNA methylation analysis of monozygotic twins discordant for diurnal preference. *Twin Research and Human Genetics*, *18*, 662–669.

Accepted for publication: 1 April 2020