CHEST Postgraduate Education Corner

CONTEMPORARY REVIEWS IN SLEEP MEDICINE

Genome-Wide Association Studies of Sleep Disorders

David M. Raizen, MD, PhD; and Mark N. Wu, MD, PhD

Sleep disorders tend to be complex diseases, with multiple genes and environmental factors interacting to contribute to phenotypes. Our understanding of the genetic underpinnings of sleep disorders has benefited from recent genome-wide association studies (GWAS). We review principles underlying GWAS and discuss recent GWAS for restless legs syndrome and narcolepsy. These studies have identified four gene variants associated with restless legs syndrome (*BTBD9*, *MEIS1*, *MAP2K5/LBXCOR1*, and *PTPRD*) and two variants associated with narcolepsy (one in the T-cell receptor α locus and another between *CPT1B* and *CHKB*). These discoveries have opened new lines of research to understand the pathophysiology of these disorders. In addition to GWAS, we expect that new technologies, such as next-generation sequencing, and continued use of animal models will provide important contributions to our understanding of the genetic basis of sleep disorders. ***CHEST 2011; 139(2):446–452***

Abbreviations: GWAS = genome-wide association studies; HLA = human leukocyte antigen; MHC = major histocompatibility; REM = rapid eye movement; RLS = restless legs syndrome; SNP = single-nucleotide polymorphism; TRA@ = T-cell receptor α locus

Over the past decade, scientific and technological advances have made possible new approaches for understanding the genetic basis of human diseases, including disorders affecting sleep. The sequencing of the human genome,[1,2] the HapMap project,[3] high-density single-nucleotide polymorphism (SNP) chips, and next-generation sequencing[4,5] have made the identification of genes underlying sleep disorders a realizable goal. Understanding the genetic basis of sleep disorders should lead to insights into their pathogenesis, could lead to new tests for diagnosis and prognosis, allow for treatments tailored to specific patients (ie, personalized medicine), and may lead to the development of novel treatments. An example of these concepts follows from the identification of mutations affecting the hypocretin/orexin pathway as a cause of narcolepsy.[6-8] The focus on this pathway has resulted in a deeper understanding of the circuitry regulating sleep,[9] diagnostic tests examining hypocretin/orexin levels in cerebrospinal fluid,[10] and potential novel treatments for insomnia using hypocretin/orexin receptor antagonists.[11,12]

This review introduces the clinician to how genome-wide association studies (GWAS) are furthering our understanding of the genetic underpinnings of sleep disorders. Our goal is not to be comprehensive regarding the pathophysiology of particular diseases or to cover all approaches to understanding the genetics of sleep disorders. Several recent reviews have dealt with the former[13-15] and the latter.[16,17] Instead, we review some of the principles related to GWAS and focus on recent GWAS for restless legs syndrome (RLS) and narcolepsy. In addition, we look ahead

Manuscript received May 24, 2010; revision accepted August 9, 2010.

Affiliations: From the Department of Neurology (Dr Raizen), University of Pennsylvania School of Medicine, Philadelphia, PA; and Department of Neurology (Dr Wu), Johns Hopkins University, Baltimore, MD.

Funding/Support: This work was supported by the National Institutes of Health [R01NS064030 to Dr Raizen and K08NS059671 to Dr Wu], the National Alliance for Research on Schizophrenia and Depression [Young Investigator Award to Dr Raizen], and the Burroughs-Wellcome Fund [Career Award for Medical Scientists to Dr Wu].

Correspondence to: Mark Wu, MD, PhD, Department of Neurology, Johns Hopkins University, Meyer 5-166, 600 N Wolfe St, Baltimore, MD 21287; e-mail: mwu38@jhmi.edu

DOI: 10.1378/chest.10-1313

to how new sequencing technologies and the use of model organisms will affect future studies of the genetics of sleep disorders.

Sleep Disorders Are Complex Genetic Diseases

Genetic disorders can be classified as having simple or complex inheritance. Simple genetic disorders refer to diseases where a single gene exerts a powerful effect on determining whether the patient exhibits the phenotype. These disorders have little environmental influence and, in general, are rare. In contrast, complex genetic disorders include those diseases in which multiple genes and the environment interact to cause the phenotype, and they are far more common. With few exceptions, all sleep disorders fall into the category of complex genetic diseases.

A powerful way to assess the relative contribution of genetics and environment in sleep disorders is to study twins. For instance, the twin concordance rate in narcolepsy is ~30%,[18] ranges from 54% to 83% in RLS,[19,20] and has been reported as variable in insomnia but may be as high as 47%,[21-23] indicating both genetic and environmental contributions to the pathogenesis or expression of these sleep disorders. In general, complex genetic diseases do not lend themselves to genetic analysis by traditional methods such as linkage studies, which typically require large families in which the disease has nearly 100% penetrance and segregates with a clear Mendelian inheritance pattern. Although this approach has been attempted in RLS, the disease segregated with a different part of the genome in each of the described families, indicating high genetic heterogeneity. Furthermore, in no case was the underlying gene identified using this approach alone.

Genome-wide Association Studies

GWAS represent an unbiased way to study complex genetic disorders, including most sleep disorders. Rather than studying families, one examines in GWAS a large population of unrelated individuals and samples 500,000 to 1 million DNA polymorphisms (typically SNPs) scattered throughout the genome. One asks whether certain SNPs are associated with the disease at a frequency greater than chance. Because 500,000 or more hypotheses are being tested simultaneously, care must be taken to statistically control for multiple testing. In practice, significant association usually is seen with nominal P values of $< 10^{-8}$.[24,25] Because GWAS are population based, they are powered to detect relatively common variants.[24,26]

Particular attention also must be paid to matching the ethnicities of the cases and the controls because one of the greatest sources of error in these studies is population stratification.[27] That is, if a disease is more common in a genetically defined subpopulation, there may be an artifactual association with other polymorphisms more common in the subpopulation. The presence of the variant identified in the initial association then should be replicated in an independent sample set of at least the same size and, ideally, in a different ethnic population.

The magnitude of the effect of the polymorphism on the disease phenotype can be quantified by the OR, which is the probability of having the disease given the designated genotype divided by the probability of having the disease given the other genotype. Despite the initial promise of using GWAS to find gene variants with large effects,[28-30] the ORs for variants identified by GWAS in general have been modest, typically < 2.0 and frequently < 1.5.[31] Thus, GWAS are thought to identify common gene variants, each of which has only a small effect on the likelihood of disease. However, even when multiple susceptibility loci are identified, each with a small contribution to disease, consideration of all of these loci together rarely explains more than a fraction of the heritability of the disease. Therefore, it has been proposed that a significant amount of the heritability of diseases depends on rare variants with large effects, which may be missed by GWAS (the common disease, rare variant hypothesis).[26,31,32]

Evidence suggesting that the gene variant or its associated loci are causally related to the disease should be sought. For example, are there changes in transcript or protein levels related to the implicated gene? The association is strengthened if one can propose a plausible relationship between variation in the identified gene and the disease. The strongest evidence of causality would be the demonstration that manipulating the gene affects disease phenotype in vivo in a model organism.

GWAS of RLS

The first successful use of GWAS in the field of sleep disorders was for RLS. Traditional linkage studies using families had previously implicated six loci, 12q, 14q, 9p, 2q, 20p, and 19p (RLS1-6, respectively),[33-38] but were unable to identify specific genes segregating with the disease.

The results of two GWAS for RLS were reported in 2007. In one study, the researchers used periodic limb movements during sleep as an endophenotype for RLS.[39] In an Icelandic population, the investigators identified a strong association (OR, ~1.8) of periodic limb movements during sleep with SNP rs3923809, which is located on chromosome 6p in the fifth intron of *BTBD9*, a gene of unknown function. They then replicated this finding, first in another Icelandic

population and then in a cohort of patients and controls from the United States. Interestingly, they found that the A allele of the SNP rs3923809 (the allele associated with increased RLS risk) also was associated with lower body iron stores as reflected by decreased serum ferritin levels. Given the well-described association between RLS and low body iron stores,[40,41] one can question whether the identified association is with periodic limb movements (and RLS) or whether it is with reduced serum iron, which in turn predisposes to RLS. Future GWAS that would control a priori for serum ferritin levels between patients with RLS and controls thus may identify additional loci involved in RLS pathogenesis.

A second GWAS of RLS independently identified the *BTBD9* intronic SNP rs3923809 as associated with RLS.[42] This study was performed in Europe and identified two additional genomic regions associated with RLS. One was on chromosome 2p, located within the *MEIS1* gene (OR, ~1.8), and the other was on chromosome 15q, located in a region that includes both the genes *MAP2K5* and *LBXCOR1*. MEIS1 is a transcription factor that functions in limb development[43]; MAP2K5 is a protein kinase; and LBXCOR1 is a repressor of LBX1,[44] which is a homeobox gene involved in sensory pathways in the dorsal horn of the spinal cord. However, the relationship of these genes to disease pathogenesis remains unclear. In an independent study, an *MEIS1* intronic variant was found to be associated with reduced levels of *MEIS1* mRNA and protein in both peripheral blood and postmortem thalamus samples,[45] suggesting that variants in *MEIS1* can contribute to RLS pathogenesis through a reduction of function. A subsequent case-control study in the United States independently confirmed an association between both *MEIS1* and *BTBD9* and RLS.[46] A European study confirmed an association between *BTBD9* with both sporadic and familial RLS but confirmed an association between *MEIS1* and *MAPK25/LBXCOR1* only with familial RLS.[47] There are two noteworthy points to the RLS GWAS results. First, the identified ORs are higher than those identified in most GWAS (see "Genome-wide Association Studies"), suggesting a larger contribution of the gene variant to disease. Second, there is no overlap among the three loci identified by GWAS and the six loci identified by traditional linkage analysis. This finding might be explained by the notion that affected gene variants detectable by linkage analysis are too infrequent to be detected by GWAS, despite having a large effect on odds of disease.

Winkelmann and colleagues[48] more recently took an approach that combined knowledge obtained from the linkage studies and genome-wide scans. They chose to sample SNPs in a large population but only in the critical regions defined by prior linkage studies. In this sense, these studies are not truly genome wide because they focus on a specific part of the genome. In one study, they focused on a 31-Mb region on chromosome 9p that had been identified as associated with RLS (RLS3) in prior linkage studies.[35,49-51] By testing for an association with only the 3,270 SNPs found in this region, they increased their power for detection of significance by several orders of magnitude and were able to detect significance with nominal P values of only 10^{-3}. They detected significant associations with two SNPs in introns 8 and 10 of the gene *PTPRD*, which encodes a receptor-like protein tyrosine phosphatase. Both SNPs are located in the 5′ untranslated region of the gene and, therefore, might have transcriptional or posttranscriptional effects. How the function of PTPRD might contribute to the pathogenesis of RLS remains unknown. Surprisingly, despite focusing the analysis on a genomic region defined by prior linkage studies, extensive sequence and gene dosage analysis of *PTPRD* coding and noncoding exons in members of RLS3-linked families revealed no differences from control subjects.[48] Therefore, it remains an open question about whether the function of PTPRD is affected in RLS3-linked family members.

The same group applied a similar approach for identifying RLS-associated SNPs in a 21-Mb region of chromosome 12q,[52] another region previously implicated in linkage studies.[33,53,54] They further reduced the number of tested SNPs and the size of the explored genome to 12.5 Mb by excluding SNPs located in regions between genes. They applied a three-stage approach, which included an exploratory stage, a replication stage, and a high SNP density mapping stage focusing on the gene implicated in the replication stage. The significant SNP identified by this approach was located in the gene neuronal nitric oxide synthetase, an exciting discovery given the known role of nitric oxide as a neurotransmitter, including in sensory pathways.[55] But they found that particular SNP alleles were associated with increased RLS risk in one population but with reduced RLS risk in another. These results are somewhat difficult to interpret, and further studies are needed in independent populations in order to understand whether and which allelic variant in neuronal nitric oxide synthetase is associated with RLS.

GWAS of Narcolepsy

Two recent studies have implicated variants in *CPT1B*, *CHKB*, and the T-cell receptor as associated with narcolepsy. In 2008, Miyagawa et al[56] reported a GWAS on 222 Japanese patients with narcolepsy with clear cataplexy and 389 healthy Japanese control subjects. Narcolepsy is known to be tightly associated with human leukocyte antigen HLA-DQB1*0602,[18,57] and

indeed, the investigators found the expected strong association with the HLA region on chromosome 6 ($P < 10^{-46}$). By removing the HLA-associated SNP from subsequent analysis, they were able to identify an association of the C allele of the SNP rs5770917 with narcolepsy with cataplexy ($P = 4 \times 10^{-7}$; OR, ~1.8). SNP rs 5770917 is located between the *CPT1B* and *CHKB* genes, and this association was replicated in a sample that was independent, but drawn from the same Japanese ethnic group consisting of 159 cases and 190 controls. When they looked for the association in other ethnic groups, they were able to find a significant association with this SNP in Koreans but not Europeans or blacks which was attributed to reduced allele frequency in the latter groups. Using quantitative polymerase chain reaction, the authors also demonstrated a reduction in expression of both *CPT1B* and *CHKB* in healthy individuals carrying the C allele of rs5770917 as compared with healthy individuals homozygous for the T allele. Both *CPT1B* and *CHKB* are biologically plausible candidates. CPT1B is the rate-limiting enzyme for β-oxidation of long-chain fatty acids, and previous studies have implicated carnitine transport, an important step in β-oxidation of fatty acids, in narcoleptic-like phenotypes.[58,59] CHKB is a kinase involved in the production of cytidine 5′diphospho choline, which has been shown to enhance acetylcholine release, a neurotransmitter known to promote wakefulness and rapid eye movement (REM) sleep.[60]

A second study by Hallmayer et al[61] implicated the T-cell receptor in narcolepsy. This study was larger than the previous study, starting with 1,830 cases and 2,164 controls of mixed European descent. Like the previous study, the study by Hallmayer et al[61] included patients with narcolepsy with definite cataplexy who were HLA-DQB1 * 0602 positive. However, in contrast to the previous study, Hallmayer et al[61] only included controls who also had the HLA-DQB1 * 0602 haplotype, thereby removing the effect of that haplotype on disease association prior to the start of their analysis. They found a strong association with the T-cell receptor α locus (*TRA@*) ($P < 10^{-21}$; OR, ~1.7). This finding was replicated in whites and Asians but not in blacks. Interestingly, in their significantly larger sample, the researchers did not observe the previously described association with *CPT1B* and *CHKB*, suggesting a possible effect of ancestry in the first Japanese study. The study by Hallmayer et al[61] was particularly powerful in its a priori control for the HLA haplotype, its large size, and its replication in multiple different ethnic groups.

The product of the *TRA@* locus is the α chain of T-cell antigen receptor αβ, which binds to major histocompatibility (MHC) molecules. This finding is striking, given the known association of narcolepsy with the HLA locus, which encodes MHC molecules. The authors speculate that the form of TRA@ associated with narcolepsy may show preferential binding to MHC molecules encoded by the HLA-DQB1 * 0602 isoform, providing a specific hypothesis regarding disease pathogenesis. Further, the identification of the association of the *TRA@* locus with narcolepsy provides additional support for the notion that narcolepsy is an autoimmune disorder.[62]

For a summary of GWAS for sleep disorders, refer to Table 1. The table lists each study discussed herein, its discovery group, and the implicated genes.

Now What? Hypothesis Testing in Model Organisms

With the notable exception of the *TRA@* association with narcolepsy, where a clear mechanistic explanation can be postulated, the relationship of other genes implicated by the narcolepsy and RLS GWAS to their respective diseases is less clear. The finding of association of SNPs with disease remains just that: an association. Only in a few instances has the SNP been shown to be associated with a functional change, such as reduced mRNA or protein levels.[45,56,63] But even in these cases, causality of gene dysfunction to disease has not been proven.

We propose that the use of model organisms that are easy to manipulate genetically would be one approach to this problem. One would determine

Table 1—*GWAS for Sleep Disorders*

Study/Sleep Disorder	Discovery Group	Implicated Gene (Chromosomal Location)
Stefansson et al[39]/RLS	306 Icelandic subjects with RLS and PLMS	*BTBD9* (6p)
Winkelmann et al[42]/RLS	401 German subjects with familial RLS	*BTBD9* (6p), *MEIS1* (2p), *MAP2K5/LBXCOR1* (15q)
Schormair et al[48]/RLS	628 European subjects with RLS	*PTPRD* (9p)
Winkelmann et al[52]/RLS	367 European subjects with RLS	*nNOS* (12q)
Miyagawa et al[56]/narcolepsy	222 Japanese subjects with narcolepsy with cataplexy	HLA locus (6p), *CPT1B/CHKB* (22q)
Hallmayer et al[61]/narcolepsy	807 North American and European subjects with narcolepsy and cataplexy who were also HLA-DQB1 * 0602 positive	*TRA@* (14q)

GWAS = genome-wide association studies; HLA = chuman leukocyte antigen; PLMS = periodic limb movements during sleep; RLS = restless legs syndrome.

whether extreme change in function of the implicated gene has an effect on a defined phenotype. Inherent in this approach is the need to study a clearly defined phenotype. There is emerging work demonstrating that not only is sleep conserved in many model organisms, such as worms, flies, fish, and mice, but many of the pathways that regulate sleep appear to be conserved in these organisms as well.[64,65] Therefore, genes may have similar functions in these organisms as they do in humans. For example, narcolepsy is a disease characterized by instability of behavioral state and with transitions between states (wake to REM) that are not normally observed. Such behavioral state instability can be readily modeled in mice and, indeed, was the basis for ascribing a narcoleptic phenotype to mice lacking orexin.[6] Even in nonmammalian genetic systems such as flies, fish, or worms, such a narcoleptic phenotype may be modeled with sleep-state instability, even if these organisms do not experience REM sleep. For instance, zebrafish lacking the hypocretin receptor have fragmented sleep at night,[66] and overexpression of hypocretin in fish causes increased wakefulness[67]; both phenotypes are predicted by the role of hypocretin in normal sleep and in narcolepsy.

Another example of this model organism approach has been demonstrated for flies and mice in work on familial sleep timing and duration phenotypes.[68-70] Verifying the effects of mutations on these circadian disorders was arguably simpler because the mutations studied were in protein coding regions and were predicted to have large effects on protein function. However, although gene variants implicated from GWAS have modest effects, this may reflect the nature of the mutation rather than the importance of the gene. In other words, these variants may have mild effects on mRNA or protein expression, but a complete knockout of such genes in model organisms may lead to a stronger phenotype. Furthermore, as discussed next, much of the heritability of common diseases may be due to rare variants with large effects, which may be easier to model in animals. Finally, in addition to their use as a testing ground for genetic variants found in humans, model organisms have been and will continue to be used to discover and understand new genetic pathways regulating sleep. The power of this approach is clearly demonstrated in circadian biology, where virtually all we know comes from model organism genetic research.[71]

The Sequencer Cometh: Next-Generation Sequencing

The technical details of the massively parallel sequencing technologies that have arrived and continue to evolve have been reviewed.[4,5] From the perspective of identifying disease genes, two specific techniques likely will be most useful. In exomic sequencing, the sequence of all the coding regions of the genome (ie, the exome) is determined and scanned for mutations. The power of this approach is that mutations that change the sequence of a protein are more likely to be causative; however, the main disadvantage of this approach is that mutations that affect gene expression and splicing would not be detected. In general, exomic sequencing will be most useful for identifying rare genetic variants with large effects and has been used successfully to identify mutations causing human diseases.[72-74] The other major approach is sequencing the entire human genome, which consists of $\sim$3 billion base pairs, in affected and unaffected individuals. The advantage of this approach is that every single base pair will be interrogated for changes compared with a reference sequence, but the disadvantage of this approach is the effort it will take to determine which of the many changes is relevant. Recently, this whole-genome sequencing approach has been used to identify mutations for Miller syndrome and Charcot-Marie-Tooth disease.[75,76] Like exomic sequencing, whole-genome sequencing would be well suited for identifying rare variants with large effects, which would be missed by GWAS.[26] In the future, as the costs of whole-genome sequencing drop, it likely will become the dominant method for identifying disease-related mutations.

Conclusions

Recent scientific and technological achievements in human genetics have paved the way toward personalized medicine and likely will influence the way medicine is practiced in the future. In this review, we have discussed how the genome-wide association approach has been applied to sleep disorders, which in most cases are complex genetic disorders. Properly designed GWAS allow one to potentially identify several genes in the pathway to disease with a single experiment. Our focus was on RLS and narcolepsy, the two sleep disorders in which GWAS have been successfully applied, but we eagerly await the results of using this approach to study obstructive sleep apnea and insomnia, common disorders with tremendous health burden. GWAS are more appropriate for finding common variants with small effects on disease pathogenesis. In the near future, whole-genome and exome sequencing will be used to complement GWAS and to identify rare variants that have larger effects. We anticipate that such approaches, combined with continued use of model organisms for gene discovery and for understanding variants found in humans, will comprise the principal approach to understanding the genetics of sleep disorders for the foreseeable future.

Acknowledgments

Financial/nonfinancial disclosures: The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

References

1. Lander ES, Linton LM, Birren B, et al; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
2. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304-1351.
3. International HapMap Consortium. The International HapMap Project. *Nature*. 2003;426(6968):789-796.
4. Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods*. 2008;5(1):16-18.
5. Pettersson E, Lundeberg J, Ahmadian A. Generations of sequencing technologies. *Genomics*. 2009;93(2):105-111.
6. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell*. 1999;98(4):437-451.
7. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell*. 1999;98(3):365-376.
8. Peyron C, Faraco J, Rogers W, et al. A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nat Med*. 2000;6(9):991-997.
9. Saper CB, Scammell TE, Lu J. Hypothalamic regulation of sleep and circadian rhythms. *Nature*. 2005;437(7063):1257-1263.
10. Nishino S, Ripley B, Overeem S, et al. Low cerebrospinal fluid hypocretin (orexin) and altered energy homeostasis in human narcolepsy. *Ann Neurol*. 2001;50(3):381-388.
11. Brisbare-Roch C, Dingemanse J, Koberstein R, et al. Promotion of sleep by targeting the orexin system in rats, dogs and humans. *Nat Med*. 2007;13(2):150-155.
12. Neubauer DN. Almorexant, a dual orexin receptor antagonist for the treatment of insomnia. *Curr Opin Investig Drugs*. 2010;11(1):101-110.
13. Mignot E. Sleep, sleep disorders and hypocretin (orexin). *Sleep Med*. 2004;5(suppl 1):S2-S8.
14. Trenkwalder C, Högl B, Winkelmann J. Recent advances in the diagnosis, genetics and treatment of restless legs syndrome. *J Neurol*. 2009;256(4):539-553.
15. Raizen DM, Mason TB, Pack AI. Genetic basis for sleep regulation and sleep disorders. *Semin Neurol*. 2006;26(5):467-483.
16. Bamne MN, Mansour H, Monk TH, Buysse DJ, Nimgaonkar VL. Approaches to unravel the genetics of sleep. *Sleep Med Rev*. 2010;14(6):397-404.
17. Tafti M. Genetic aspects of normal and disturbed sleep. *Sleep Med*. 2009;10(suppl 1):S17-S21.
18. Mignot E. Genetic and familial aspects of narcolepsy. *Neurology*. 1998;50(2)(suppl 1):S16-S22.
19. Ondo WG, Vuong KD, Wang Q. Restless legs syndrome in monozygotic twins: clinical correlates. *Neurology*. 2000;55(9):1404-1406.
20. Xiong L, Jang K, Montplaisir J, et al. Canadian restless legs syndrome twin study. *Neurology*. 2007;68(19):1631-1633.
21. Heath AC, Kendler KS, Eaves LJ, Martin NG. Evidence for genetic influences on sleep disturbance and sleep pattern in twins. *Sleep*. 1990;13(4):318-335.
22. McCarren M, Goldberg J, Ramakrishnan V, Fabsitz R. Insomnia in Vietnam era veteran twins: influence of genes and combat experience. *Sleep*. 1994;17(5):456-461.
23. Watson NF, Goldberg J, Arguelles L, Buchwald D. Genetic and environmental influences on insomnia, daytime sleepiness, and obesity in twins. *Sleep*. 2006;29(5):645-649.
24. Hardy J, Singleton A. Genomewide association studies and human disease. *N Engl J Med*. 2009;360(17):1759-1768.
25. Mullen SA, Crompton DE, Carney PW, Helbig I, Berkovic SF. A neurologist's guide to genome-wide association studies. *Neurology*. 2009;72(6):558-565.
26. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010;11(6):415-425.
27. Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet*. 2010;11(7):459-463.
28. Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308(5720):385-389.
29. Haines JL, Hauser MA, Schmidt S, et al. Complement factor H variant increases the risk of age-related macular degeneration. *Science*. 2005;308(5720):419-421.
30. Edwards AO, Ritter R III, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. *Science*. 2005;308(5720):421-424.
31. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747-753.
32. Goldstein DB. Common genetic variation and human traits. *N Engl J Med*. 2009;360(17):1696-1698.
33. Desautels A, Turecki G, Montplaisir J, Sequeira A, Verner A, Rouleau GA. Identification of a major susceptibility locus for restless legs syndrome on chromosome 12q. *Am J Hum Genet*. 2001;69(6):1266-1270.
34. Bonati MT, Ferini-Strambi L, Aridon P, Oldani A, Zucconi M, Casari G. Autosomal dominant restless legs syndrome maps on chromosome 14q. *Brain*. 2003;126(pt 6):1485-1492.
35. Chen S, Ondo WG, Rao S, Li L, Chen Q, Wang Q. Genomewide linkage scan identifies a novel susceptibility locus for restless legs syndrome on chromosome 9p. *Am J Hum Genet*. 2004;74(5):876-885.
36. Pichler I, Marroni F, Volpato CB, et al. Linkage analysis identifies a novel locus for restless legs syndrome on chromosome 2q in a South Tyrolean population isolate. *Am J Hum Genet*. 2006;79(4):716-723.
37. Levchenko A, Provost S, Montplaisir JY, et al. A novel autosomal dominant restless legs syndrome locus maps to chromosome 20p13. *Neurology*. 2006;67(5):900-901.
38. Kemlink D, Plazzi G, Vetrugno R, et al. Suggestive evidence for linkage for restless legs syndrome on chromosome 19p13. *Neurogenetics*. 2008;9(2):75-82.
39. Stefansson H, Rye DB, Hicks A, et al. A genetic risk factor for periodic limb movements in sleep. *N Engl J Med*. 2007;357(7):639-647.
40. Nordlander NB. Therapy in restless legs. *Acta Med Scand*. 1953;145(6):453-457.
41. Ekbom KA. Restless legs syndrome. *Neurology*. 1960;10:868-873.
42. Winkelmann J, Schormair B, Lichtner P, et al. Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nat Genet*. 2007;39(8):1000-1006.
43. Mercader N, Leonardo E, Azpiazu N, et al. Conserved regulation of proximodistal limb axis development by Meis1/Hth. *Nature*. 1999;402(6760):425-429.
44. Mizuhara E, Nakatani T, Minaki Y, Sakamoto Y, Ono Y. Corl1, a novel neuronal lineage-specific transcriptional corepressor for the homeodomain transcription factor Lbx1. *J Biol Chem*. 2005;280(5):3645-3655.

45. Xiong L, Catoire H, Dion P, et al. MEIS1 intronic risk haplotype associated with restless legs syndrome affects its mRNA and protein expression levels. *Hum Mol Genet*. 2009; 18(6):1065-1074.
46. Vilariño-Güell C, Farrer MJ, Lin SC. A genetic risk factor for periodic limb movements in sleep. *N Engl J Med*. 2008; 358(4):425-427.
47. Kemlink D, Polo O, Frauscher B, et al. Replication of restless legs syndrome loci in three European populations. *J Med Genet*. 2009;46(5):315-318.
48. Schormair B, Kemlink D, Roeske D, et al. PTPRD (protein tyrosine phosphatase receptor type delta) is associated with restless legs syndrome. *Nat Genet*. 2008;40(8):946-948.
49. Liebetanz KM, Winkelmann J, Trenkwalder C, et al. RLS3: fine-mapping of an autosomal dominant locus in a family with intrafamilial heterogeneity. *Neurology*. 2006;67(2): 320-321.
50. Kemlink D, Polo O, Montagna P, et al. Family-based association study of the restless legs syndrome loci 2 and 3 in a European population. *Mov Disord*. 2007;22(2):207-212.
51. Lohmann-Hedrich K, Neumann A, Kleensang A, et al. Evidence for linkage of restless legs syndrome to chromosome 9p: are there two distinct loci? *Neurology*. 2008;70(9):686-694.
52. Winkelmann J, Lichtner P, Schormair B, et al. Variants in the neuronal nitric oxide synthase (nNOS, NOS1) gene are associated with restless legs syndrome. *Mov Disord*. 2008;23(3): 350-358.
53. Desautels A, Turecki G, Montplaisir J, et al. Restless legs syndrome: confirmation of linkage to chromosome 12q, genetic heterogeneity, and evidence of complexity. *Arch Neurol*. 2005;62(4):591-596.
54. Winkelmann J, Lichtner P, Pütz B, et al. Evidence for further genetic locus heterogeneity and confirmation of RLS-1 in restless legs syndrome. *Mov Disord*. 2006;21(1):28-33.
55. Schmidtko A, Tegeder I, Geisslinger G. No NO, no pain? The role of nitric oxide and cGMP in spinal pain processing. *Trends Neurosci*. 2009;32(6):339-346.
56. Miyagawa T, Kawashima M, Nishida N, et al. Variant between CPT1B and CHKB associated with susceptibility to narcolepsy. *Nat Genet*. 2008;40(11):1324-1328.
57. Langdon N, Welsh KI, van Dam M, Vaughan RW, Parkes D. Genetic markers in narcolepsy. *Lancet*. 1984;324(8413): 1178-1180.
58. Yoshida G, Li MX, Horiuchi M, et al. Fasting-induced reduction in locomotor activity and reduced response of orexin neurons in carnitine-deficient mice. *Neurosci Res*. 2006;55(1): 78-86.
59. Kuwajima M, Fujihara H, Sei H, et al. Reduced carnitine level causes death from hypoglycemia: possible involvement of suppression of hypothalamic orexin expression during weaning period. *Endocr J*. 2007;54(6):911-925.
60. Dixon CE, Ma X, Marion DW. Effects of CDP-choline treatment on neurobehavioral deficits after TBI and on hippocampal and neocortical acetylcholine release. *J Neurotrauma*. 1997;14(3):161-169.
61. Hallmayer J, Faraco J, Lin L, et al. Narcolepsy is strongly associated with the T-cell receptor alpha locus. *Nat Genet*. 2009;41(6):708-711, corrected 2009;41(7):859.
62. Hor H, Kutalik Z, Dauvilliers Y, et al. Genome-wide association study identifies new HLA class II haplotypes strongly protective against narcolepsy. *Nature Genetics*. 2010;42: 786-789.
63. Altshuler D, Daly M. Guilt beyond a reasonable doubt. *Nat Genet*. 2007;39(7):813-815.
64. Zimmerman JE, Naidoo N, Raizen DM, Pack AI. Conservation of sleep: insights from non-mammalian model systems. *Trends Neurosci*. 2008;31(7):371-376.
65. Allada R, Siegel JM. Unearthing the phylogenetic roots of sleep. *Curr Biol*. 2008;18(15):R670-R679.
66. Yokogawa T, Marin W, Faraco J, et al. Characterization of sleep in zebrafish and insomnia in hypocretin receptor mutants. *PLoS Biol*. 2007;5(10):e277.
67. Prober DA, Rihel J, Onah AA, Sung RJ, Schier AF. Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish. *J Neurosci*. 2006;26(51):13400-13410.
68. Xu Y, Padiath QS, Shapiro RE, et al. Functional consequences of a CKIdelta mutation causing familial advanced sleep phase syndrome. *Nature*. 2005;434(7033):640-644.
69. Xu Y, Toh KL, Jones CR, Shin JY, Fu YH, Ptácek LJ. Modeling of a human circadian mutation yields insights into clock regulation by PER2. *Cell*. 2007;128(1):59-70.
70. He Y, Jones CR, Fujiki N, et al. The transcriptional repressor DEC2 regulates sleep length in mammals. *Science*. 2009; 325(5942):866-870.
71. Allada R, Emery P, Takahashi JS, Rosbash M. Stopping time: the genetics of fly and mouse circadian clocks. *Annu Rev Neurosci*. 2001;24:1091-1119.
72. Jones S, Hruban RH, Kamiyama M, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science*. 2009;324(5924):217.
73. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. 2009;461(7261):272-276.
74. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA*. 2009;106(45):19096-19101.
75. Roach JC, Glusman G, Smit AF, et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*. 2010;328(5978):636-639.
76. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med*. 2010;362(13):1181-1191.