ORIGINAL ARTICLE

Evaluation of polygenic risks for narcolepsy and essential hypersomnia

Maria Yamasaki[1], Taku Miyagawa[1,2], Hiromi Toyoda[1], Seik-Soon Khor[1], Xiaoxi Liu[3], Hitoshi Kuwabara[4], Yukiko Kano[4], Takafumi Shimada[5], Toshiro Sugiyama[6], Hisami Nishida[7], Nagisa Sugaya[8], Mamoru Tochigi[9], Takeshi Otowa[10], Yuji Okazaki[11], Hisanobu Kaiya[12], Yoshiya Kawamura[13], Akinori Miyashita[14], Ryozo Kuwano[14], Kiyoto Kasai[10], Hisashi Tanii[15], Tsukasa Sasaki[16], Yutaka Honda[17], Makoto Honda[2,17] and Katsushi Tokunaga[1]

In humans, narcolepsy is a sleep disorder that is characterized by sleepiness, cataplexy and rapid eye movement (REM) sleep abnormalities. Essential hypersomnia (EHS) is another type of sleep disorder that is characterized by excessive daytime sleepiness without cataplexy. A human leukocyte antigen (HLA) class II allele, *HLA-DQB1*06:02*, is a major genetic factor for narcolepsy. Almost all narcoleptic patients are carriers of this *HLA* allele, while 30–50% of EHS patients and 12% of all healthy individuals in Japan carry this allele. The pathogenesis of narcolepsy and EHS is thought to be partially shared. To evaluate the contribution of common single-nucleotide polymorphisms (SNPs) to narcolepsy onset and to assess the common genetic background of narcolepsy and EHS, we conducted a polygenic analysis that included 393 narcoleptic patients, 38 EHS patients with *HLA-DQB1*06:02*, 119 EHS patients without *HLA-DQB1*06:02* and 1582 healthy individuals. We also included 376 individuals with panic disorder and 213 individuals with autism to confirm whether the results were biased. Polygenic risks in narcolepsy were estimated to explain 58.1% ($P_{HLA\text{-}DQB1*06:02} = 2.30 \times 10^{-48}$, $P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 6.73 \times 10^{-2}$) including *HLA-DQB1*06:02* effects and 1.3% ($P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 2.43 \times 10^{-2}$) excluding *HLA-DQB1*06:02* effects. The results also indicated that small-effect SNPs contributed to the development of narcolepsy. Reported susceptibility SNPs for narcolepsy in the Japanese population, *CPT1B* (carnitine palmitoyltransferase 1B), *TRA@* (T-cell receptor alpha) and *P2RY11* (purinergic receptor P2Y, G-protein coupled, 11), were found to explain 0.8% of narcolepsy onset ($P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 9.74 \times 10^{-2}$). EHS patients with *HLA-DQB1*06:02* were estimated to have higher shared genetic background to narcoleptic patients than EHS patients without *HLA-DQB1*06:02* even when the effects of *HLA-DQB1*06:02* were excluded (EHS with *HLA-DQB1*06:02*: 40.4%, $P_{HLA\text{-}DQB1*06:02} = 7.02 \times 10^{-14}$, $P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 1.34 \times 10^{-1}$, EHS without *HLA-DQB1*06:02*: 0.4%, $P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 3.06 \times 10^{-1}$). Meanwhile, the polygenic risks for narcolepsy could not explain the onset of panic disorder and autism, suggesting that our results were reasonable.
Journal of Human Genetics (2016) 61, 873–878; doi:10.1038/jhg.2016.65; published online 16 June 2016

INTRODUCTION

Narcolepsy is a sleep disorder characterized by excessive daytime sleepiness, cataplexy (sudden loss of muscle tone in response to strong emotions), and pathological manifestations of rapid eye movement (REM) sleep, such as hypnagogic hallucinations, sleep paralysis and sleep-onset REM sleep. Patients usually develop narcolepsy around adolescence. It affects both men and women equally, and its prevalence is 0.16–0.18% in Japan and 0.02–0.06% in the United States and Europe.[1,2] The relative risk of narcolepsy in first-degree family members of patients with narcolepsy is 10- to 40-fold higher than that in the general population.[1]

[1]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [2]Sleep Disorders Project, Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan; [3]Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Kanagawa, Japan; [4]Department of Child Psychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [5]Division for Counseling and Support, The University of Tokyo, Tokyo, Japan; [6]Department of Child and Adolescent Psychiatry, Hamamatsu University School of Medicine, Shizuoka, Japan; [7]Asunaro Hospital for Child and Adolescent Psychiatry, Mie, Japan; [8]Department of Epidemiology and Public Health, Yokohama City University Graduate school of Medicine, Kanagawa, Japan; [9]Teikyo University Hospital, Tokyo, Japan; [10]Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [11]Tokyo Metropolitan Matsuzawa Hospital, Tokyo, Japan; [12]Panic Disorder Research Center, Warakukai Med Corp, Tokyo, Japan; [13]Department of Psychiatry, Shonan Kamakura General Hospital, Kanagawa, Japan; [14]Department of Molecular Genetics, Bioresource Science Branch, Center for Bioresources, Brain Research Institute, Niigata University, Niigata, Japan; [15]Department of Psychiatry, Graduate School of Medicine, Mie University, Mie, Japan; [16]Department of Physical and Health Education, Graduate School of Education, The University of Tokyo, Tokyo, Japan and [17]Japan Somnology Center, Neuropsychiatric Research Institute, Tokyo, Japan
Correspondence: Dr T Miyagawa, Sleep Control Project, Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya-ku, Tokyo 156-8506, Japan.
E-mail: miyagawa-taku@umin.ac.jp.
Received 23 November 2015; revised 21 April 2016; accepted 28 April 2016; published online 16 June 2016

Multiple factors are associated with the development of narcolepsy, including genetic variations at several loci. Initially, it was discovered that narcolepsy was closely associated with a human leukocyte antigen (*HLA*) class II allele, *HLA-DQB1*06:02*.[3–7] Almost all narcoleptic patients carry *HLA-DQB1*06:02*. This *HLA* allele is thought to be a requisite for the development of narcolepsy, but it cannot fully explain the onset of narcolepsy because 10–40% of individuals in the general population carry it as well. Genome-wide association studies have identified several narcolepsy susceptibility loci: a locus near *CPT1B* (carnitine palmitoyltransferase 1B), which may be involved in a new pathogenic mechanism related to fatty acid oxidation,[8] as well as *TRA@* (T-cell receptor alpha) and *P2RY11* (purinergic receptor P2Y, G-protein coupled, 11).[9–12] These discoveries provided evidence that narcolepsy is caused by the immune system attacking brain cells.

The destruction of hypothalamic neurons that regulate wake-promoting neuropeptide hypocretin (orexin) causes narcolepsy.[13,14] The amount of hypocretin-1 in the cerebrospinal fluid of narcoleptic patients is low or undetectable.[15,16] However, mutations and polymorphisms in prepro-hypocretin and hypocretin-receptor genes do not contribute considerably to the onset of narcolepsy, except for rare cases.[17–20]

Essential hypersomnia (EHS) is another type of sleep disorder that is characterized by excessive daytime sleepiness. However, EHS patients do not exhibit cataplexy. Both genetic and environmental factors are considered to contribute to the development of EHS.[15,21] Diagnosis of EHS is made based on the following criteria in central nervous system (CNS) hypersomnias: (i) recurrent daytime sleep episodes that occur basically everyday over a period of at least 6 months; (ii) absence of cataplexy; and (iii) the hypersomnia is not better explained by another sleep disorder, medical or neurological disorder, mental disorder, medication use or substance use disorder.[22–25] If we applied criteria of International Classification of Sleep Disorders 2nd edition (AASM2005) to EHS patients, they would be classified either two CNS hypersomnia such as narcolepsy without cataplexy or a part of idiopathic hypersomnia without long sleep time. Previous studies have reported that EHS and narcolepsy share several susceptibility genes of note,[22–26] approximately 30–50% of EHS patients carry *HLA-DQB1*06:02*.[22,23] In addition, EHS is associated with narcolepsy susceptibility single-nucleotide polymorphisms (SNPs) in *CPT1B* and *TRA@*,[22,26] and it has been reported that EHS and narcolepsy share partially similar pathogenic mechanisms.[25]

In this study, we evaluated the contribution of common variants associated with narcolepsy onset. The estimations resulting from this study will be useful for designing future strategies to identify the indeterminate genetic factors involved in narcolepsy onset.[27,28] Also, we assessed the extent of genetic share between narcolepsy and EHS. This assessment of the genetic share between narcolepsy and EHS can enable phenotypic classification based on the genetic architecture. In addition, the reliability of our results was assessed by including individuals with panic disorder and autism.

MATERIALS AND METHODS
Subjects
The participants in this study were all Japanese and included 426 narcoleptic patients,[29] 46 EHS patients with *HLA-DQB1*06:02*, 125 EHS patients without *HLA-DQB1*06:02*,[30] 432 individuals with panic disorder,[31] 246 individuals with autism[32] and 2072 healthy individuals. All narcoleptic patients carried *HLA-DQB1*06:02*. Among the 171 EHS patients, 46 individuals possessed *HLA-DQB1*06:02* (EHS with *HLA-DQB1*06:02*), while the remaining 125 EHS patients did not (EHS without *HLA-DQB1*06:02*).[22] Each patient with narcolepsy or EHS had received a diagnosis at the Neuropsychiatric Research

Institute by specialists for sleep disorders. Data from individuals with neuropsychiatric disorders, including panic disorder and autism, were obtained in previous studies[31–34] and were included for analysis in this study.

Data from healthy individuals genotyped in previous studies[31–35] were also included for analysis in this study as healthy controls. The control subjects did not have a history of narcolepsy or CNS hypersomnia. Age and gender were not matched between the cases and the controls. Ethical approval was obtained from the local institutional review boards of all participating organizations. All individuals provided written informed consent for participation in the study.

Genotyping and quality control
All samples were genotyped for 906 622 SNPs by the Affymetrix Genome-Wide SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) (http://www.affymetrix.com). Genotype calling was performed with the Birdseed algorithm in Birdsuite (http://www.broadinstitute.org/).[36] Quality control procedures were performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) (Supplementary Figure 1).[37] Samples with overall call rates lower than 97% were excluded. Samples from individuals who had reported a family relationship with other participants or with a mean probability of identity-by-descent (PIHAT) value greater than 0.125, as calculated by PLINK, were excluded, until only one sample remained out of the reported or estimated family members. Outliers in principal component analysis were also excluded to eliminate population stratification using EIGENSOFT (http://www.hsph.harvard.edu/alkes-price/software/).[38] In the principal component analysis, data from 91 JPT (Japanese in Tokyo, Japan), 90 CHB (Han Chinese in Beijing, China), 180 CEU (Utah residents with Northern and Western European ancestry) and 180 YRI (Yoruba in Ibadan, Nigeria) obtained from the HapMap Project (http://hapmap.ncbi.nlm.nih.gov/)[39] were also included. Data from the HapMap populations and the present sample sets were combined after the quality control steps described above were performed. SNPs were excluded if (1) the minor allele frequencies (MAFs) were less than 0.05; (2) P-values from the Hardy–Weinberg Equilibrium (HWE) test either for patient groups or healthy control groups were less than 0.001; (3) SNP call rates were less than 99%; or (4) SNPs were located within the *HLA* region (chr6: 27 539 703–35 377 701), hg18 or sex chromosomes. After the quality control steps, 476 446 SNPs in 393 narcoleptic patients, 38 EHS patients with *HLA-DQB1*06:02*, 119 EHS patients without *HLA-DQB1*06:02*, 376 individuals with panic disorder, 213 individuals with autism and 1582 healthy individuals were analyzed in this study (Supplementary Figure 1).

Statistical analysis
Polygenic risks were calculated using SNPs with $r^2 < 0.25$ within a 200-kb window in order to exclude secondary associations due to linkage disequilibrium. The model of logistic regression analysis was

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1$$

where x_1 is the collective polygenic risk score (0–1) for each individual. Simply, collective polygenic score for an individual is sum of the number of risk alleles of SNPs which the individual is possessing. Actually, when we summed the number of risk alleles, risk alleles were weighted based on the chi-square values of them. To calculate this collective polygenic score, '—score' function in PLINK was utilized. Nagelkerke's pseudo R^2 was utilized to estimate the ability of polygenic risks to explain the disease onset. As narcolepsy is closely associated with *HLA-DQB1*06:02*, the region (chr6: 27 539 703–35 377 701), hg18 linked to *HLA* risk[3–7] was excluded to eliminate the *HLA* region linkage complexity. Instead, SNP rs7744293, which is in high linkage disequilibrium with *HLA-DQB1*06:02* ($r^2 = \sim 0.8$), was included in the logistic regression model as a surrogate marker so that estimation of polygenic risks of narcolepsy reflects *HLA-DQB1*06:02* effects,

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

where x_1 is the collective polygenic risk score (0–1) for each individual and x_2 is 0 or 1 for positivity of the risk allele of SNP rs7744293, which surrogates positivity of the *HLA-DQB1*06:02* allele. In addition, to assess risks explained by reported susceptibility SNPs, that is, *CPT1B*, *TRA@* and *P2RY11*, in the

Table 1 Phenotypic variance explained by polygenic risks for narcolepsy, including *HLA-DQB1*06:02* effects

		P < 0.001	P < 0.01	P < 0.1	P < 0.2	P < 0.3	P < 0.4	P < 0.5
Narcolepsy	Nagelkerke_R^2	57.8%	57.9%	58.1%	58.1%	58.1%	58.1%	58.1%
	$P_{HLA-DQB1*06:02}$	2.31E−48	2.31E−48	2.55E−48	2.36E−48	2.40E−48	2.23E−48	2.30E−48
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	2.32E−01	2.05E−01	9.79E−02	8.00E−02	7.11E−02	6.78E−02	6.73E−02
EHS with *HLA-DQB1*06:02*	Nagelkerke_R^2	39.9%	39.9%	40.1%	40.2%	40.3%	40.3%	40.4%
	$P_{HLA-DQB1*06:02}$	7.91E−14	7.65E−14	7.08E−14	7.02E−14	7.01E−14	7.02E−14	7.02E−14
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	2.84E−01	2.92E−01	1.99E−01	1.64E−01	1.48E−01	1.38E−01	1.34E−01

Abbreviations: EHS, essential hypersomnia; HLA, human leukocyte antigen; SNP, single-nucleotide polymorphism.
HLA-DQB1*06:02 effects were reflected by including SNP rs7744293, which was in the highest correlation with *HLA-DQB1*06:02* ($r^2 = \sim 0.8$), into the regression model. $P_{HLA-DQB1*06:02}$ and $P_{whole\ genome\ without\ HLA-DQB1*06:02}$ were P-values for regression coefficients of *HLA-DQB1*06:02* and polygenic risk scores other than the *HLA* region, respectively. Nagelkerke's pseudo R^2 reflected coefficient of determination. Nagelkerke's pseudo R^2 and P-values were the averages of one thousand permutation analysis, randomizing sample division of narcoleptic patients and healthy controls into discovery stage and test stage.
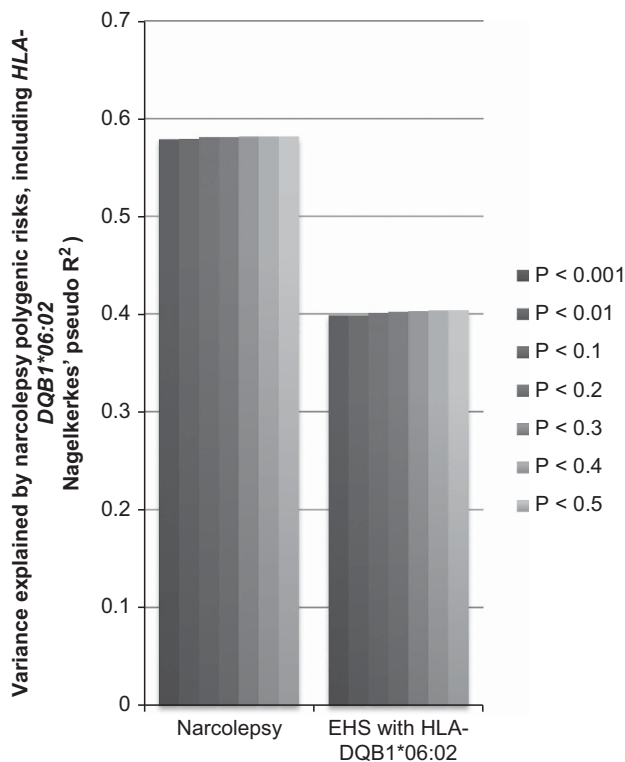

Figure 1 Phenotypic variance explained by polygenic risks for narcolepsy, including *HLA-DQB1*06:02* effects. Nagelkerke's pseudo R^2 was utilized to estimate the ability of polygenic risks to explain the disease onset. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

Japanese population, SNPs that were 500 kb upstream and downstream of the reported SNPs were included.[8–10] The logistic regression analysis was conducted using PredictABEL packages (http://www.genabel.org/) provided in R (http://www.r-project.org/).[40,41] In this assessment, we did not control the difference of case–control ratios between the sample set in this study and the Japanese population. In addition, GCTA (http://cnsgenomics.com/software/gcta/)[42] was utilized to enforce reliability of our analysis. For GCTA software, a total of 1975 samples, consisting of 393 narcoleptic patients and 1582 healthy individuals, were analyzed. We also estimated the required sample size using the power.prop.test3 function provided in R (http://aoki2.si.gunma-u.ac.jp/R/power_prop_test2.html). All statistical procedures were conducted using PLINK (http://pngu.mgh.harvard.edu/ ~ purcell/plink/) and R (http://www.r-project.org/).[37,41]

The study design is shown in Supplementary Figure 2. A permutation analysis was conducted in order to randomize samples used in each stage, the discovery stage and the test stage. The case–control status of samples was kept in this permutation analysis. The permutation was repeated 1000 times. In addition, in the test stage, the same healthy controls across different diseases within one permutation were included.

RESULTS

A total of 3347 samples were genotyped using Affymetrix Genome-Wide SNP Array 6.0 and the Birdseed algorithm. Samples were excluded based on overall call rate and family relationships. SNPs were excluded based on MAF, HWE, SNP call rates, their locations and linkage disequilibria, the *HLA* region and sex chromosomes. In the present study, 476 446 SNPs in narcoleptic patients ($n = 393$), EHS patients with *HLA-DQB1*06:02* ($n = 38$), EHS patients without *HLA-DQB1*06:02* ($n = 119$), individuals with panic disorder ($n = 376$), individuals with autism ($n = 213$) and healthy controls ($n = 1582$) were analyzed after the quality control procedures (Supplementary Figure 1).

One thousand permutations were performed in the discovery stage for genome-wide association studies of narcolepsy. The average lambda value of the 1000 permutations was 1.004, suggesting no population stratification. For each permutation, the narcolepsy risk effects of the SNPs were defined.

Polygenic risks were computed for each set of narcoleptic patients and healthy controls in the test stage at seven different P-value thresholds ($P < 0.001, 0.01, 0.1, 0.2, 0.3, 0.4$ and 0.5). As a result of the 1000-permutation analysis, narcolepsy polygenic risks were estimated to be 58.1% ($P_{HLA-DQB1*06:02} = 2.30 \times 10^{-48}$, $P_{whole\ genome\ without\ HLA-DQB1*06:02} = 6.73 \times 10^{-2}$) (Table 1 and Figure 1). In addition, estimation by GCTA was 59.7% (standard error = 7.6%, $P < 0.05$). Polygenic risks of narcolepsy other than the *HLA* region were estimated to be 1.3% ($P_{whole\ genome\ without\ HLA-DQB1*06:02} = 2.43 \times 10^{-2}$) (Table 2 and Figure 2). As SNPs with smaller P-values were included into the analysis, total risk continued to increase until SNPs with a P-value of < 0.5 were included. This suggested that small-effect SNPs contributed to the onset of narcolepsy. In addition, the proportion of risks, which has been explained by reported susceptibility SNPs in the Japanese population, was calculated. Using only SNPs located 500 kb upstream and downstream of the reported narcolepsy susceptibility genes *CPT1B*, *TRA@* and *P2RY11*, the result was 0.8% ($P_{whole\ genome\ without\ HLA-DQB1*06:02} = 9.74 \times 10^{-2}$) (Table 2 and Figure 2).

Next, we aimed to assess the ability of polygenic risks for narcolepsy to explain the onset of EHS with and without *HLA-DQB1*06:02*. Overall, no statistical significant estimates of narcolepsy polygenic effects on EHS either with or without *HLA-DQB1*06:02* were observed when the *HLA* effects were excluded (Table 2 and Figure 2). However, EHS patients with *HLA-DQB1*06:02* tended to have higher polygenic risks for narcolepsy than EHS patients without *HLA-DQB1*06:02* (EHS with *HLA-DQB1*06:02*: 1.4%, $P_{whole\ genome}$

Table 2 Phenotypic variance explained by polygenic risks for narcolepsy in each disease

		P < 0.001	P < 0.01	P < 0.1	P < 0.2	P < 0.3	P < 0.4	P < 0.5
Narcolepsy	Nagelkerke_R^2	0.7%	0.7%	1.1%	1.2%	1.3%	1.3%	1.3%
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	1.66E−01	1.41E−01	4.78E−02	3.37E−02	2.72E−02	2.51E−02	2.43E−02
Narcolepsy by CPT1B,	Nagelkerke_R^2			0.9%	0.9%	0.8%	0.8%	0.8%
TRA@ and P2RY11	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$			8.28E−02	8.72E−02	9.33E−02	9.52E−02	9.74E−02
EHS with HLA-DQB1*06:02	Nagelkerke_R^2	1.0%	1.0%	1.1%	1.2%	1.3%	1.4%	1.4%
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	2.93E−01	3.03E−01	2.34E−01	1.93E−01	1.74E−01	1.61E−01	1.56E−01
EHS without HLA-DQB1*06:02	Nagelkerke_R^2	0.3%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	4.50E−01	3.70E−01	3.15E−01	3.12E−01	3.07E−01	3.08E−01	3.06E−01
Panic disorder	Nagelkerke_R^2	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	5.04E−01	4.98E−01	4.95E−01	5.02E−01	5.12E−01	5.20E−01	5.20E−01
Autism	Nagelkerke_R^2	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.2%
	$P_{whole\ genome\ without\ HLA-DQB1*06:02}$	5.09E−01	5.31E−01	5.45E−01	5.17E−01	4.95E−01	4.81E−01	4.75E−01

Abbreviations: CPT1B, carnitine palmitoyltransferase 1B; EHS, essential hypersomnia; HLA, human leukocyte antigen; P2RY11, purinergic receptor P2Y, G-protein coupled, 11; TRA@, T-cell receptor alpha.
HLA-DQB1*06:02 effects were not taken into account. $P_{whole\ genome\ without\ HLA-DQB1*06:02}$ was P-values for regression coefficients of polygenic risk scores other than the HLA region. Nagelkerke's pseudo R^2 and P-values were the averages of one thousand permutations, randomizing sample division of narcolepsy and healthy controls into discovery stage and test stage.
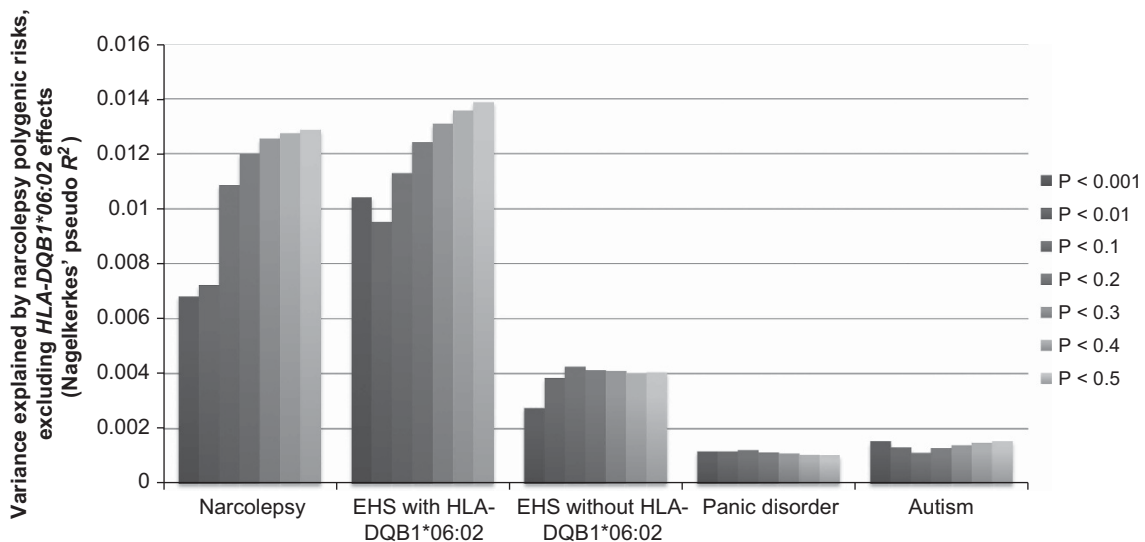

Figure 2 Phenotypic variance explained by polygenic risks for narcolepsy, excluding HLA-DQB1*06:02 effects. Nagelkerke's pseudo R^2 was utilized to estimate the ability of polygenic risks to explain the disease onset. A full color version of this figure is available at the Journal of Human Genetics journal online.

without $HLA-DQB1*06:02 = 1.56 \times 10^{-1}$, EHS without HLA-DQB1*06:02: 0.4%, $P_{whole\ genome\ without\ HLA-DQB1*06:02} = 3.06 \times 10^{-1}$) (Table 2 and Figure 2). We found a significant similarity between narcolepsy and EHS with HLA-DQB1*06:02 including HLA-DQB1*06:02 effects (narcolepsy: 58.1%, $P_{HLA-DQB1*06:02} = 2.30 \times 10^{-48}$, $P_{whole\ genome\ without\ HLA-DQB1*06:02} = 6.73 \times 10^{-2}$, EHS with HLA-DQB1*06:02: 40.4%, $P_{HLA-DQB1*06:02} = 7.02 \times 10^{-14}$, $P_{whole\ genome\ without\ HLA-DQB1*06:02} = 1.34 \times 10^{-1}$) (Table 1 and Figure 1).

To test disease specificity, polygenic analyses were extended to other neuropsychiatric diseases, including panic disorder and autism. Polygenic risks for narcolepsy did not explain the phenotypic variance of these neuropsychiatric diseases (panic disorder: 0.1%, $P_{whole\ genome\ without\ HLA-DQB1*06:02} = 5.20 \times 10^{-1}$, autism: 0.2%, $P_{whole\ genome\ without\ HLA-DQB1*06:02} = 4.75 \times 10^{-1}$) (Table 2 and Figure 2).

We also assessed the possibility that the healthy controls in the present study led to spurious estimations. All healthy controls were divided into pseudo-case and pseudo-control groups; and the polygenic risks of these groups were estimated using the same methodology in addition to shuffling 'pseudo' case–control status. Results showed that there was no polygenic contribution among healthy controls (0.2%, $P = 5.10 \times 10^{-1}$).

We found that small-effect SNPs contributed to the onset of narcolepsy; thus, the sample sizes were estimated to detect remaining small-effect susceptibility SNPs (Supplementary Figure 3). A total of 42 000 narcoleptic patients would be required to obtain an odds ratios of 1.1 for the targeted SNPs, a risk allele frequency of 0.2 in the controls, a case/control ratio of 0.25 and an α-level of 5.00×10^{-8}.

DISCUSSION

In this study, a polygenic analysis of narcolepsy and EHS was performed. We evaluated the extent of genetic factors for narcolepsy, both including and excluding HLA-DQB1*06:02 effects. We utilized 426 narcoleptic patients,[29] 46 EHS patients with HLA-DQB1*06:02 and 125 EHS patients without HLA-DQB1*06:02,[30] which derived from previous reports of our team. Polygenic risks of narcolepsy were calculated to be approximately 58.1% ($P_{HLA-DQB1*06:02} = 2.30 \times 10^{-48}$,

$P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 6.73 \times 10^{-2}$) (Table 1 and Figure 1) and 1.3% excluding *HLA* effects ($P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 2.43 \times 10^{-2}$) (Table 2 and Figure 2). In addition, estimation by GCTA was 59.7% (standard error = 7.6%, $P < 0.05$), being concordant between the different methodologies. *HLA-DQB1*06:02* is thought to be essential in narcolepsy because almost all narcoleptic patients possess this allele. However, the allele itself cannot fully explain the onset of narcolepsy, because 10–40% of individuals in the general populations carry this allele whereas the prevalence of the disease is around 0.1%.[1] More than 50% of the first-degree relatives of individuals with narcolepsy possess this allele, but the risk of narcolepsy in these first-degree relatives is only 1–2%.[43] In addition, two-thirds of monozygotic twins are discordant for the development of narcolepsy.[1] These reports show that the onset of narcolepsy is associated with other genetic and environmental factors. In this study, the effects of genetic factors were evaluated, including the effects of genes located outside the *HLA* region.

Several genetic factors other than the *HLA* region have been reported to be related to narcolepsy in the Japanese population; these include the reported SNPs *CPT1B*, *TRA@* and *P2RY11*. In this study, the extent of these factors to explain polygenic risks of narcolepsy was calculated (Table 2 and Figure 2). SNPs located within 500 kb upstream and downstream of the reported SNPs were included in the analysis. Polygenic risks were estimated to be 0.8% ($P_{whole\ genome\ without\ HLA\text{-}DQB1*06:02} = 9.74 \times 10^{-2}$). The difference between polygenic risks explained by regions other than *HLA* (1.3%) and risks explained by *CPT1B*, *TRA@* and *P2RY11* (0.8%) was approximately 0.5%, suggesting the presence of some other genetic factors for narcolepsy susceptibility that have not yet been discovered.

EHS patients with and without *HLA-DQB1*06:02* were evaluated to determine whether polygenic risks for narcolepsy contribute to the onset of EHS. No statistically-significant polygenic risk factors for narcolepsy were observed in EHS patients either with or without *HLA-DQB1*06:02*. However, EHS patients with *HLA-DQB1*06:02* tended to have higher polygenic risks for narcolepsy than EHS patients without *HLA-DQB1*06:02* (Table 2 and Figure 2): in total, approximately 40.4% of polygenic risks in EHS patients with *HLA-DQB1*06:02* could be explained by a genetic background of narcolepsy (Table 1 and Figure 1). These results suggest that EHS with *HLA-DQB1*06:02* is more similar to narcolepsy than EHS without *HLA-DQB1*06:02*, which is concordant with the previous studies.[22,26] In the reported association between EHS and *TRA@*, only EHS with *HLA-DQB1*06:02* showed a significant association, while EHS without *HLA-DQB1*06:02* did not.[26] The definition of EHS with and without *HLA-DQB1*06:02* did not follow that of ICSD-2. However, the classification of EHS with and without *HLA-DQB1*06:02* seems to truly reflect the genetic features of the disease. Thus, genetic variants discovered to be associated with narcolepsy should be tested for an association with EHS with *HLA-DQB1*06:02*. Combined analyses for narcolepsy and EHS with *HLA-DQB1*06:02* may be performed as appropriate.

Other neuropsychiatric disorders, that is, panic disorder and autism, were also evaluated in this study. Polygenic risks for narcolepsy could not explain the onset of these neuropsychiatric disorders, suggesting that narcolepsy does not share many genetic factors with these neuropsychiatric disorders. In future studies, comparisons between narcolepsy and several autoimmune diseases are needed as narcolepsy is strongly associated with *HLA*.

In studies of polygenic risk, it is essential to assess whether the data of healthy control samples affected the estimations of polygenic risk. Thus, the genetic features among the control samples were validated.

Control samples were divided into pseudo-case and pseudo-control groups; and the same polygenic analysis was performed. There was no significant risk contribution among the healthy controls, indicating that the healthy controls were unbiased and usable for this analysis.

Next, the sample sizes needed to identify the remaining narcoleptic genetic factors were estimated (Supplementary Figure 3). In the future, an effort should be made to collect data from approximately 42 000 narcoleptic patients. At the same time, the sample size for the healthy controls must be four times larger than that of the patients. Therefore, it will be necessary to build a consortium of both patients and healthy controls.

Limitations of this study include assumptions inherent to the methods used. This polygenic analysis assumed an additive effect of each SNP across the genome. Dominant and epistasis effects were not taken into consideration for the estimations even though susceptibility SNPs with dominant effects have been reported in mammals[44–46] and epistasis effects have been reported in human.[47] To account for this limitation, the genetic backgrounds that contributed to the development of these diseases were not fully described in this study.

In this study, we demonstrated the existence of uncovered risks in narcolepsy. Also, nominal polygenic similarity between narcolepsy and EHS with *HLA-DQB1*06:02* was found. However, much effort, such as drastically increasing the sample sizes, is needed to identify the remaining genetic factors of narcolepsy. Therefore, collaboration is necessary to collect sufficient samples from both patients and healthy controls for further studies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Mignot, E. Genetic and familial aspects of narcolepsy. *Neurology* **50**, S16–S22 (1998).
2 Overeem, S., Mignot, E., van Dijk, J. G. & Lammers, G. J. Narcolepsy: clinical features, new pathophysiologic insights, and future perspectives. *J. Clin. Neurophysiol.* **18**, 78–105 (2001).
3 Juji, T., Satake, M., Honda, Y. & Doi, Y. HLA antigens in Japanese patients with narcolepsy. All the patients were DR2 positive. *Tissue Antigens* **24**, 316–319 (1984).
4 Langdon, N., Welsh, K. I., van Dam, M., Vaughan, R. W. & Parkes, D. Genetic markers in narcolepsy. *Lancet* **2**, 1178–1180 (1984).
5 Matsuki, K., Juji, T., Tokunaga, K., Naohara, T., Satake, M. & Honda, Y. Human histocompatibility leukocyte antigen (HLA) haplotype frequencies estimated from the data on HLA class I, II, and III antigens in 111 Japanese narcoleptics. *J. Clin. Invest.* **76**, 2078–2083 (1985).
6 Miyagawa, T., Hohjoh, H., Honda, Y., Juji, T. & Tokunaga, K. Identification of a telomeric boundary of the HLA region with potential for predisposition to human narcolepsy. *Immunogenetics* **52**, 12–18 (2000).
7 Mignot, E., Lin, L., Rogers, W., Honda, Y., Qiu, X., Lin, X. et al. Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. *Am. J. Hum. Genet.* **68**, 686–699 (2001).
8 Miyagawa, T., Kawashima, M., Nishida, N., Ohashi, J., Kimura, R., Fujimoto, A. et al. Variant between CPT1B and CHKB associated with susceptibility to narcolepsy. *Nat. Genet.* **40**, 1324–1328 (2008).
9 Hallmayer, J., Faraco, J., Lin, L., Hesselson, S., Winkelmann, J., Kawashima, M. et al. Narcolepsy is strongly associated with the T-cell receptor alpha locus. *Nat. Genet.* **41**, 708–711 (2009).
10 Kornum, B. R., Kawashima, M., Faraco, J., Lin, L., Rico, T. J., Hesselson, S. et al. Common variants in P2RY11 are associated with narcolepsy. *Nat. Genet.* **43**, 66–71 (2011).

11 Faraco, J., Lin, L., Kornum, B. R., Kenny, E. E., Trynka, G., Einen, M. et al. ImmunoChip study implicates antigen presentation to T cells in narcolepsy. PLoS Genet. 9, e1003270 (2013).

12 Han, F., Faraco, J., Dong, X. S., Ollila, H. M., Lin, L., Li, J. et al. Genome wide analysis of narcolepsy in China implicates novel immune loci and reveals changes in association prior to versus after the 2009 H1N1 influenza pandemic. PLoS Genet. 9, e1003880 (2013).

13 Chemelli, R. M., Willie, J. T., Sinton, C. M., Elmquist, J. K., Scammell, T., Lee, C. et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. Cell 98, 437–451 (1999).

14 Lin, L., Faraco, J., Li, R., Kadotani, H., Rogers, W., Lin, X. et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. Cell 98, 365–376 (1999).

15 Mignot, E., Lammers, G. J., Ripley, B., Okun, M., Nevsimalova, S., Overeem, S. et al. The role of cerebrospinal fluid hypocretin measurement in the diagnosis of narcolepsy and other hypersomnias. Arch. Neurol. 59, 1553–1562 (2002).

16 Nishino, S., Ripley, B., Overeem, S., Lammers, G. J. & Mignot, E. Hypocretin (orexin) deficiency in human narcolepsy. Lancet 355, 39–40 (2000).

17 Gencik, M., Dahmen, N., Wieczorek, S., Kasten, M., Bierbrauer, J., Anghelescu, I. et al. A prepro-orexin gene polymorphism is associated with narcolepsy. Neurology 56, 115–117 (2001).

18 Olafsdottir, B. R., Rye, D. B., Scammell, T. E., Matheson, J. K., Stefansson, K. & Gulcher, J. R. Polymorphisms in hypocretin/orexin pathway genes and narcolepsy. Neurology 57, 1896–1899 (2001).

19 Peyron, C., Faraco, J., Rogers, W., Ripley, B., Overeem, S., Charnay, Y. et al. A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. Nat. Med. 6, 991–997 (2000).

20 Hungs, M., Lin, L., Okun, M. & Mignot, E. Polymorphisms in the vicinity of the hypocretin/orexin are not associated with human narcolepsy. Neurology 57, 1893–1895 (2001).

21 Anderson, K. N., Pilsworth, S., Sharples, L. D., Smith, I. E. & Shneerson, J. M. Idiopathic hypersomnia: a study of 77 cases. Sleep 30, 1274–1281 (2007).

22 Miyagawa, T., Honda, M., Kawashima, M., Shimada, M., Tanaka, S., Honda, Y. et al. Polymorphism located between CPT1B and CHKB, and HLA-DRB1*1501-DQB1*0602 haplotype confer susceptibility to CNS hypersomnias (essential hypersomnia). PLoS ONE 4, e5394 (2009).

23 Honda, Y., Juji, T., Matsuki, K., Naohara, T., Satake, M., Inoko, H. et al. HLA-DR2 and Dw2 in narcolepsy and in other disorders of excessive somnolence without cataplexy. Sleep 9, 1–2 (1986).

24 Honda, Y., Takahashi, Y., Honda, M., Watanabe, Y., Sato, T., Miki, T. et al. in Sleep and Sleep Disorders: From Molecule to Behavior. (eds Inoue, S. & Hayashi, O.) 341–358 (Academic Press, New York, NY, USA, 1997).

25 Komada, Y., Inoue, Y., Mukai, J., Shirakawa, S., Takahashi, K. & Honda, Y. Difference in the characteristics of subjective and objective sleepiness between narcolepsy and essential hypersomnia. Psychiatry Clin Neurosci. 59, 194–199 (2005).

26 Miyagawa, T., Honda, M., Kawashima, M., Shimada, M., Tanaka, S., Honda, Y. et al. Polymorphism located in TCRA locus confers susceptibility to essential hypersomnia with HLA-DRB1*1501-DQB1*0602 haplotype. J. Hum. Genet. 55, 63–65 (2010).

27 Maher, B. Personal genomes: the case of the missing heritability. Nature 456, 18–21 (2008).

28 Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J. et al. Finding the missing heritability of complex diseases. Nature 461, 747–753 (2009).

29 Toyoda, H., Miyagawa, T., Koike, A., Kanbayashi, T., Imanishi, A., Sagawa, Y. et al. A polymorphism in CCR1/CCR3 is associated with narcolepsy. Brain Behav. Immun. 49, 148–155 (2015).

30 Khor, S. S., Miyagawa, T., Toyoda, H., Yamasaki, M., Kawamura, Y., Tanii, H. et al. Genome-wide association study of HLA-DQB1*06:02 negative essential hypersomnia. PeerJ 1, e66 (2013).

31 Otowa, T., Kawamura, Y., Nishida, N., Sugaya, N., Koike, A., Yoshida, E. et al. Meta-analysis of genome-wide association studies for panic disorder in the Japanese population. Transl. Psychiatry 2, e186 (2012).

32 Liu, X. X., Kawamura, Y., Shimada, T., Otowa, T., Koishi, S., Sugiyama, T. et al. Association of the oxytocin receptor (OXTR) gene polymorphisms with autism spectrum disorder (ASD) in the Japanese population. J. Hum. Genet. 55, 137–141 (2010).

33 Shiota, S., Tochigi, M., Shimada, H., Ohashi, J., Kasai, K., Kato, N. et al. Association and interaction analyses of NRG1 and ERBB4 genes with schizophrenia in a Japanese population. J. Hum. Genet. 53, 929–935 (2008).

34 Tochigi, M., Zhang, X., Ohashi, J., Hibino, H., Otowa, T., Rogers, M. et al. Association study between the TNXB locus and schizophrenia in a Japanese population. Am. J. Med. Genet. B Neuropsychiatr. Genet. 144B, 305–309 (2007).

35 Adachi, S., Tajima, A., Quan, J., Haino, K., Yoshihara, K., Masuzaki, H. et al. Meta-analysis of genome-wide association scans for genetic susceptibility to endometriosis in Japanese population. J. Hum. Genet. 55, 816–821 (2010).

36 McCarroll, S. A., Kuruvilla, F. G., Korn, J. M., Cawley, S., Nemesh, J., Wysoker, A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat. Genet. 40, 1166–1174 (2008).

37 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575 (2007).

38 Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. PLoS Genet. 2, 2074–2093 (2006).

39 International HapMap Consortium. The International HapMap Project. Nature 426, 789–796 (2003).

40 Kundu, S., Aulchenko, Y. S., van Duijn, C. M. & Janssens, A. C. PredictABEL: an R package for the assessment of risk prediction models. Eur. J. Epidemiol. 26, 261–264 (2011).

41 Austin, G. The New S Language - a programming environment for data-analysis and graphics. Econ. J. 100, 650–652 (1990).

42 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88, 76–82 (2011).

43 Pollmacher, T., Schulz, H., Geisler, P., Kiss, E., Albert, E. D., Schwarzfischer, F. et al. Discordant for narcolepsy. Sleep 13, 336–343 (1990).

44 Gengler, N., VanVleck, L. D., MacNeil, M. D., Misztal, I. & Pariacote, F. A. Influence of dominance relationships on the estimation of dominance variance with sire-dam subclass effects. J. Anim. Sci. 75, 2885–2891 (1997).

45 Palucci, V., Schaeffer, L. R., Miglior, F. & Osborne, V. Non-additive genetic effects for fertility traits in Canadian Holstein cattle (Open Access publication). Genet. Sel. Evol. 39, 181–193 (2007).

46 Norris, D., Varona, L., Ngambi, J. W., Visser, D. P., Mbajiorgu, C. A. & Voordewind, S. F. Estimation of the additive and dominance variances in SA Duroc pigs. Livest. Sci. 131, 144–147 (2010).

47 Kirino, Y., Bertsias, G., Ishigatsubo, Y., Mizuki, N., Tugal-Tutkun, I., Seyahi, E. et al. Genome-wide association analysis identifies new susceptibility loci for Behcet's disease and epistasis between HLA-B*51 and ERAP1. Nat. Genet. 45, 202–207 (2013).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)