Multiomics analysis of narcolepsy T cells: global hypomethylation in solo-WCGW motif linked to T cell proliferation

*Mihoko Shimada[1, 2, 3], Makoto Honda [3, 4], Yoshiko Honda [3], Tohru Kodama [3], Yuki Hitomi [5], Katsushi Tokunaga[1], *Taku Miyagawa [3]

Running title: **Hypomethylation in narcolepsy T cells**

[1] Genome Medical Science Project (Toyama), National Center for Global Health and Medicine (NCGM), Tokyo, Japan

[2] Center for Clinical Sciences, National Center for Global Health and Medicine (NCGM), Tokyo, Japan

[3] Sleep Disorders Project, Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

[4] Seiwa Hospital, Institute of Neuropsychiatry, Tokyo, Japan

[5] Department of Human Genetics, Research Institute, National Center for Global Health and Medicine (NCGM), Tokyo, Japan

*Address correspondence to:

Taku Miyagawa, Ph.D.

Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya, Tokyo 156-8506, Japan

Tel: +81-3-6834-2389 E-mail: miyagawa-tk@igakuken.or.jp

Mihoko Shimada, Ph.D.

Genome Medical Science Project (Toyama), National Center for Global Health and Medicine (NCGM), 1-21-1 Toyama, Shinjuku-ku, Tokyo 162-8655, Japan

Tel: +81- 3-6205-6572 E-mail: mihshimada@ri.ncgm.go.jp

Abstract

Narcolepsy type 1 (NT1) is a chronic sleep disorder caused by the loss of orexin-producing cells in the lateral hypothalamus. The involvement of immune system abnormalities has been suggested, as indicated by the presence of autoreactive $CD4^+$ and $CD8^+$ T cells. However, it is difficult to sample precisely when this immune abnormality occurs, and previous studies examining cytokines and gene expression have not elucidated the mechanism of the immune abnormality. By focusing our analysis on epigenetic memory encoded in DNA methylation, we performed a genome-wide DNA methylation analysis using $CD4^+$ and $CD8^+$ T cells of NT1 patients. Analysis of differentially methylated regions as well as integrative analyses with genomic and transcriptomic data obtained from the same samples suggested that cell chemotaxis pathways are implicated in the etiology of NT1. Although causality in disease from methylation changes is generally unclear, changes linked to SNP genotypes in chemotaxis pathways suggested a likely disease cause. Additionally, we found global hypomethylation in both the $CD4^+$ and $CD8^+$ T cells of NT1 cases ($CD4^+$: P = 1.69E-67; $CD8^+$: P = 4.83E-12). The hypomethylation level in NT1 was correlate well with hypoSC, an index related to cell division (R^2 = 0.64, P = 1.61E-12). Further, these NT1-associated hypomethylated sites were significantly more abundant in solo-WCGW (sequences without neighboring CpGs, where W is a purine base, P = 9.87E-194). Solo-WCGW tends to lose DNA methylation over the course of cell divisions, suggesting enhanced T cell proliferation in NT1.

Introduction

Narcolepsy type 1 (NT1) is a chronic sleep disorder characterized by frequent lapses into sleep and cataplexy. Both genetic and environmental factors contribute to the development of NT1 (Mignot 1998; Khatami et al. 2004; Mignot 2017). Selective loss of orexin (hypocretin)-producing cells has been observed in the lateral hypothalamus of NT1 postmortem brains, indicating that the loss of these cells is involved in the onset of NT1 (Peyron et al. 2000; Thannickal et al. 2000; Honda et al. 2009); however, the mechanism has not yet been clarified. NT1 is strongly associated with human leukocyte antigen (*HLA*), and more than 90% of patients with NT1 carry the *HLA- DQB1*06:02* allele (Mignot et al. 2001; Roh et al. 2006; Hong et al. 2007; Han et al. 2012). Genome-wide association studies (GWASs) of NT1 have also identified other genetic factors; most of these factors are immune-related (Miyagawa et al. 2008; Hallmayer et al. 2009; Kornum et al. 2011; Faraco et al. 2013; Han et al. 2013; Toyoda et al. 2015). Exposure to Pandemrix, an H1N1 ASO3-adjuvanted vaccine against a pandemic influenza, was found to be an environmental factor for the development of NT1 (Nohynek et al. 2012; Dauvilliers et al. 2013), and some immunological abnormalities have been reported to trigger the development of NT1 (Miyagawa and Tokunaga 2019). Recent studies have also reported the involvement of $CD4^+$ T cells and $CD8^+$ T cells in NT1. In a study using sensitive cellular screens, hypocretin-specific autoreactive $CD4^+$ T cells were found in all of the NT1 patients, and hypocretin-specific autoreactive $CD8^+$ T cells were found in some of the NT1 patients; in contrast, these cells were found in only a small number of the healthy controls (Latorre et al. 2018). Another study has reported that the frequency of autoreactive $CD8^+$ T cells specific for NT1-related peptides, which are presented primarily by *HLA-DQB1*06:02*, was lower in healthy individuals with the disease susceptibility allele *HLA-DQB1*06:02* than in patients with NT1 or *HLA-DQB1*06:02*-negative healthy individuals (Pedersen et al. 2019). In addition, it was reported that a higher percentage of $CD4^+$ and $CD8^+$ T cells release cytokines in response to hypocretin in children with NT1 (Cogswell et al. 2019), and that variations in the memory subsets of $CD4^+$ and $CD8^+$ T cells are

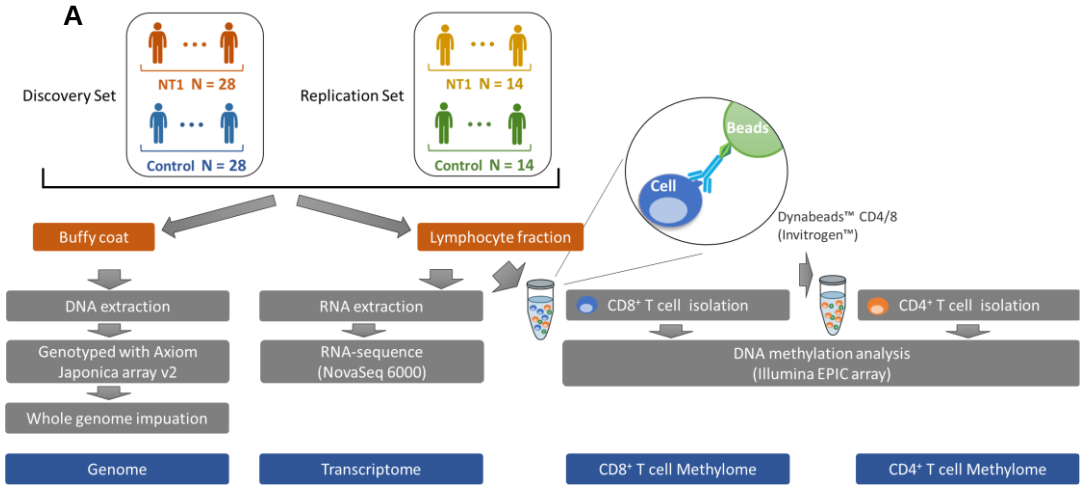
73 higher in patients with NT1 (Moresco et al. 2019). These results suggest that an autoimmune basis
74 involving $CD4^+$ and $CD8^+$ T cells is associated with the pathogenesis of NT1.
75 DNA methylation is one of the major epigenetic mechanisms. The DNA methylation status
76 affects gene expression, and it has been investigated in many diseases, because it is a relatively
77 stable modification and is more readily measurable than gene expression, which is susceptible to
78 various factors and prone to change (Flanagan 2015; Berdasco and Esteller 2019). Causal
79 mutations in DNA methyltransferase 1 (*DNMT1*) have been found in families with atypical
80 narcolepsy (Winkelmann et al. 2012), and patients with *DNMT1* mutations showed a genome-
81 wide methylation abnormality (Kernohan et al. 2016). In addition, a GWAS of NT1 has reported
82 a significant association between single-nucleotide polymorphisms (SNPs) near *DNMT1* and NT1
83 (Kornum et al. 2011). We previously performed an epigenome-wide association study (EWAS) of
84 DNA methylation using peripheral blood samples, and we found that methylation around *CCR3*
85 (C-C motif chemokine receptor 3) were associated with NT1; in addition, we also found that NT1-
86 related methylation sites were significantly located outside of CpG islands, 95% of which were
87 hypomethylated in NT1 (Shimada et al. 2018). The array we used in that study was the Infinium
88 HumanMethylation27 BeadChip (27K array; Illumina, San Diego, CA, USA), which could only
89 measure methylation in about 27,000 locations and was considered to be insufficient for
90 understanding the entire genome. Furthermore, the DNA methylation patterns are known to vary
91 widely among cells and tissues (Schultz et al. 2015), and our previous study using only peripheral
92 blood cells might have been insufficient for detecting subtle changes in methylation that occur in
93 individual immune cells.
94 In the present study, we performed EWASs of DNA methylation using $CD4^+$ and $CD8^+$ T cells
95 isolated from whole blood samples to detect NT1-related differentially methylated sites, as these
96 two types of immune cells have been suggested to play an important role in the pathogenesis of
97 NT1. We also performed RNA sequencing (RNA-seq) using whole blood from the same subjects
98 for methylation analysis to examine NT1-associated gene expression, at least that at the time of
99 blood sampling. Genome-wide genotyping was also carried out, and we performed multi-omics

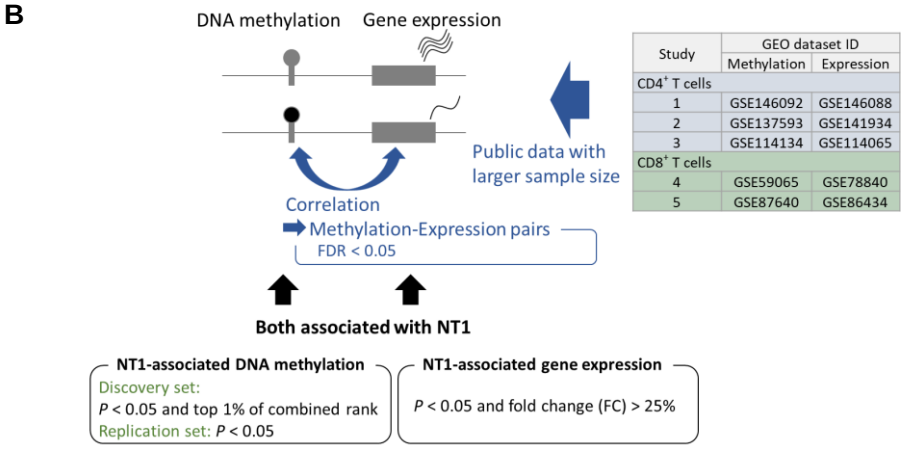
analysis using genome, methylome, and transcriptome data collected from the same subjects. Although gene expression is the most phenotypic of these omics data, we focused our analysis on methylation with consideration of the potential of DNA methylation as a form of epigenetic memory (Thiagalingam 2020). Although it is difficult to capture the gene expression changes that trigger the development of NT1 in real-time since there is an inevitable time lag between disease onset and diagnosis (including sleep studies and blood sampling), DNA methylation may be a particularly useful tool for studying disease onset as the effects of such abnormal immune events may be retained as epigenetic memory.

Results

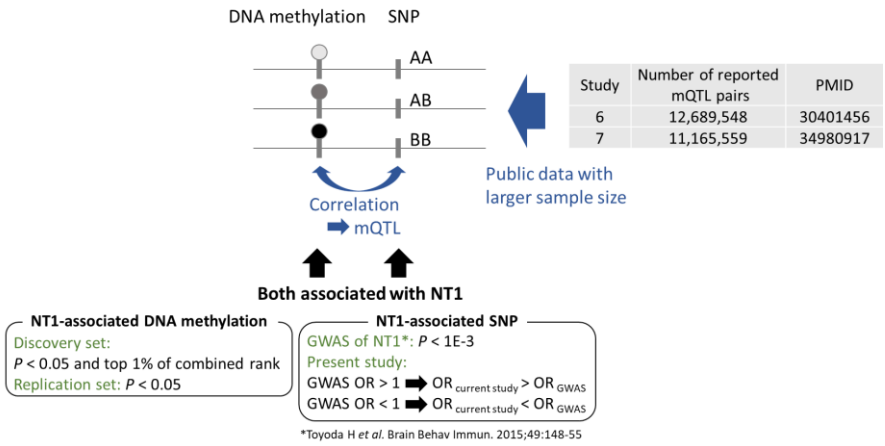
DNA methylation analysis and characteristics of the samples based on DNA methylation

We examined the DNA methylation profiles of CD4$^+$ and CD8$^+$ T cells using the EPIC methylation array. Our research design is shown in Fig. 1.





Study	GEO dataset ID	
	Methylation	Expression
CD4$^+$ T cells		
1	GSE146092	GSE146088
2	GSE137593	GSE141934
3	GSE114134	GSE114065
CD8$^+$ T cells		
4	GSE59065	GSE78840
5	GSE87640	GSE86434



Study	Number of reported mQTL pairs	PMID
6	12,689,548	30401456
7	11,165,559	34980917

Figure 1. Overview of the study. A. Study design and graphical abstract for multiomics analysis. The discovery set consists of 28 cases of NT1 and 28 controls, while the replication set consists of 14 cases of NT1 and 14 controls. From each sample, the buffy coat and leukocyte fractions were collected. DNA was extracted from the buffy coat and genotyped using an array. From one of the two divided leukocyte fractions, RNA was first extracted, then RNA sequencing was performed to obtain transcriptome data. Using the remaining leukocyte fraction, $CD8^+$ T cells were isolated using beads, followed by the isolation of $CD4^+$ T cells. DNA was extracted from each of these isolated T cell types, and the DNA methylation rate was measured using an array. **B.** Methodology of the integrative analysis of methylome and transcriptome data. Initially, we identified pairs of methylation rate and gene expression that correlated with each other (Methylation-Expression pairs) using data from previous studies with larger sample sizes. Among the identified pairs, we then explored pairs where both the methylation rate and gene expression were associated with NT1 in our EWAS and RNA-seq. **C.** Methodology of the integrative analysis of methylome and genome data. From among the mQTLs detected in previous studies with larger sample sizes, we searched for mQTLs that had both a methylation rate and genotype that were associated with NT1 in our EWAS and GWAS.

After probe filtering and data normalization, 746,913 probes were selected for the following analysis (Fig. S1). We performed principal component analysis (PCA) analysis using normalized β values for all methylation sites, and confirmed that the $CD4^+$ T cell-derived and $CD8^+$ T cell-derived samples fell into widely different clusters, and that there was no significant bias in the distribution of the cases and controls within each cluster (Fig. S2A, B). We next examined whether the methylation rate of the master transcription factor binding site of T cells explained most of the variations in the PCA, as has been reported in previous studies (Hachiya et al. 2017). Variation in PC1 was found to be well-explained by the methylation rate of the T-bet binding sites, and variation in PC2 was found to be well-explained by the methylation rate of the GATA3 binding sites (Fig. S2C, D). These results indicated that the variability in the present sample mainly reflected differences in the T cell subtypes. Although the PCA results showed no relationship between T cell subtype variation and case or control bias, we further examined in more detail the presence of T cell subtypes within each of the $CD4^+$ and $CD8^+$ T cell samples by estimating the composition of these subtypes in the discovery set. Estimation using the MeDeCom method revealed the lowest cross-validation error when assuming six subtypes for $CD4^+$ T cells and four subtypes for $CD8^+$ T cells (Fig. S3). We then compared the proportion of each estimated subtype between the NT1 cases and controls, and found that only the proportion of latent methylation component (LMC) 4 in the $CD4^+$ T cell analysis was significantly smaller in the NT1 cases than

130 in the controls ($P = 0.0032$; Fig. S4A, B). Furthermore, we verified whether the samples analyzed
131 as $CD4^+$ and $CD8^+$ T cells were correctly estimated to be the respective T cells using a reference-
132 based estimation method. The results showed that in most of the samples, the estimated T cell
133 types ($CD4^+$ T cells or $CD8^+$ T cells) were largely consistent with the sorted T cell types, but in
134 three control samples of $CD8^+$ T cells, we found that $CD8^+$ T cells were not estimated to be the
135 major component since there was a higher percentage of other leukocytes (Fig. S4C). These three
136 control samples were excluded from the subsequent association analyses, because they could not
137 be properly evaluated for $CD8^+$ T cell methylation. Similarly, estimation of the T cell proportions
138 was also performed for the replication samples. In five NT1 samples and four control samples,
139 we found that $CD8^+$ T cells were not correctly estimated to be the major component, and these
140 samples were thus excluded from the subsequent association analyses (Fig. S4C).
141 We also estimated the methylation age in the discovery set, and confirmed a strong correlation
142 between the DNA methylation age and chronological age ($r = 0.913$; Fig. S5A, B, C). The
143 biological age was not increased in the NT1 cases when compared to the controls (Fig. S5C). The
144 DNA methylation age and chronological age were also strongly correlated in the replication
145 samples ($r = 0.949$; Fig. S5D).
146
147 *Differentially methylated positions and regions of NT1*
148 We devised two strategies to identify methylation associated with NT1 to minimize false
149 positives. One was to identify differentially methylated regions (DMRs), which are not just single-
150 point associations like differentially methylated positions (DMPs), but regions that have a chain
151 of related methylation sites. The other was to investigate the DMPs that have been indicated to be
152 associated with NT1 by other omics data.

First, we examined DMPs to explore the overall trends of the associations. We found DMPs that were significantly associated with NT1 after Bonferroni correction and were replicated with $P < 0.05$ in three loci in the $CD4^+$ T cell analysis; they were all localized to methylation sites within the *HLA* region (Fig. 2A), and no corresponding significant DMPs were detected in $CD8^+$ T cell analysis (Fig. 2B).

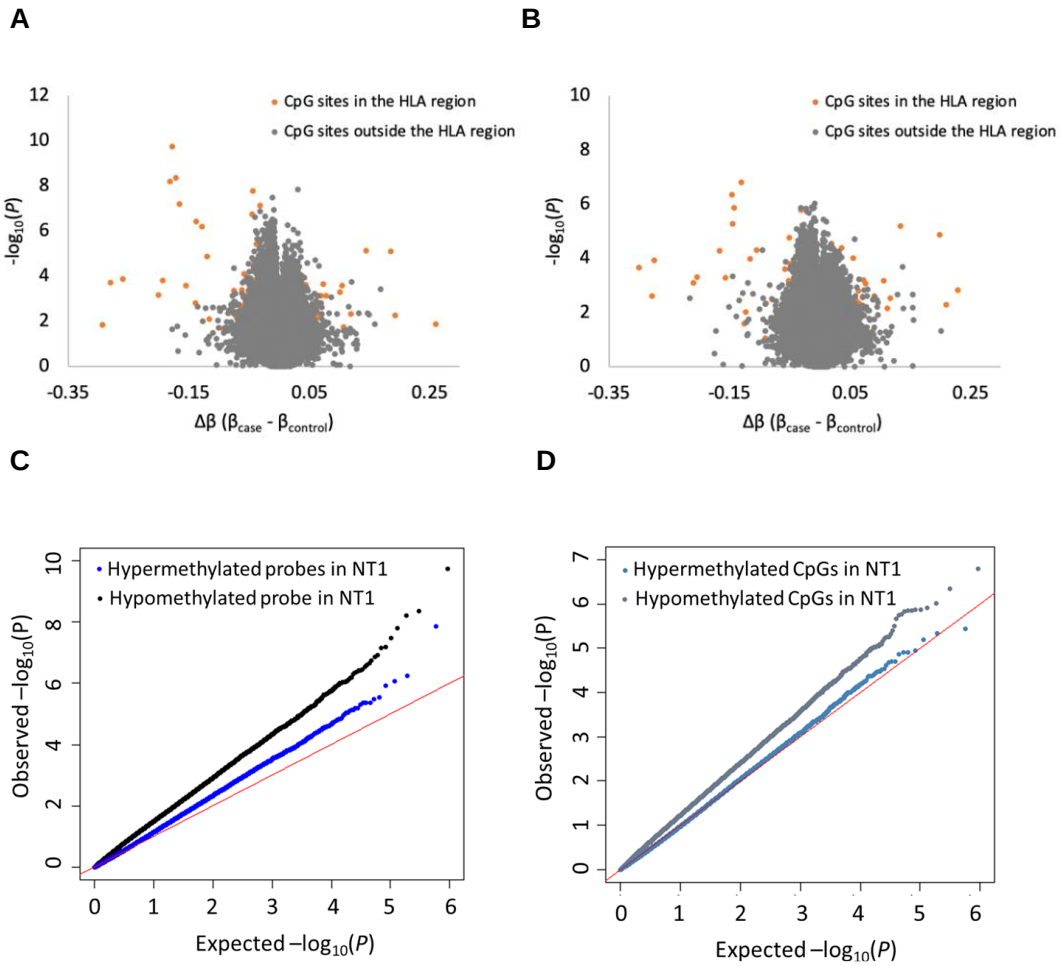


Figure 2. Results of the association analysis in $CD4^+$ and $CD8^+$ T cells. A. Volcano plot of the DMP analysis in the $CD4^+$ T cells. **B.** Volcano plot of the DMP analysis in the $CD8^+$ T cells. **C.** Q-Q plot of the DMP analysis in the $CD4^+$ T cells. We plotted the methylation sites separately, categorizing them into those that are hypomethylated and those that are hypermethylated in NT1. **D.** Q-Q plot of the DMP analysis in the $CD8^+$ T cells. Similar to the $CD4^+$ T cell analysis, we plotted the results of the methylation sites separately as hypomethylated or hypermethylated in NT1.

In our analysis, probes with poor performance were excluded; however, the *HLA* region, being highly polymorphic and challenging to analyze due to variations in probe affinity caused by the

160 polymorphisms, was analyzed separately, and the findings have been reported in a dedicated
161 manner elsewhere (Yoshida-Tanaka et al. 2023). Despite the lack of substantial differences in the
162 proportions of $CD4^+$ T cell and $CD8^+$ T cell subtypes between the case and control groups,
163 inflation of the observed *P*-values was evident in the quantile-quantile (Q-Q) plots for the analyses
164 involving both $CD4^+$ and $CD8^+$ T cells. When we categorized differentially methylated sites into
165 those that were hypermethylated or hypomethylated in NT1, it was discernible that the inflation
166 was more pronounced at sites of hypomethylation (Fig. 2C, D).

167 Next, we focused our investigation on DMRs to identify plausible methylation sites associated
168 with NT1. We identified 15 replicated DMRs in the $CD4^+$ T cell analysis, and five replicated
169 DMRs in the $CD8^+$ T cell analysis (Table 1). Of the DMRs detected in the $CD8^+$ T cell analysis,
170 three were also identified in $CD4^+$ T cells. One of these three DMRs was located in a region
171 extending from upstream to the first exon of the C-C motif chemokine ligand 5 (*CCL5*) gene on
172 chromosome 17 (Table 1). Another was located in an area near the cytoplasmic linker associated
173 protein 2 (*CLASP2*) gene, which included a methylation site that was found to be associated with
174 C-C motif chemokine receptor 4 (*CCR4*) gene expression in $CD4^+$ T cells (Table 1).

175 **Table 1. DMRs in NT1 in the CD4⁺ and CD8⁺ T cell analyses**

CHR	Position		Length (bp)	Hyper- and hypo-methylated in NT1	Bumphunter_P		Neighboring genes	Genes whose expression is associated with methylation*	Detected in both CD4⁺ and CD8⁺ T cells
	Start	End			Discovery	Replication			
CD4⁺ T cells									
8	81787438	81788187	749	Hypo	4.7.E-04	5.7.E-03	ZNF704	PAG1_N	
7	27137427	27138974	1547	Hyper	3.5.E-03	2.6.E-04	HOTAIRM1	NA	
10	81044831	81045166	335	Hypo	6.5.E-03	4.5.E-04	ZMIZ1	NA	
3	33700962	33701223	261	Hyper	8.3.E-03	1.0.E-03	CLASP2	CCR4_P;CRTAP_P	●
1	32264223	32264703	480	Hypo	9.2.E-03	9.3.E-03	SPOCD1	NA	
17	6899085	6899380	295	Hyper	7.1.E-03	5.1.E-04	ALOX12	NA	
1	159046937	159047177	240	Hypo	1.2.E-02	3.0.E-04	AIM2	PYHIN1_N;AIM2_N;TAGLN2_N;SPTA1_P;KIRREL1_N;DUSP23_N	
8	38831681	38832728	1047	Hyper	1.4.E-02	1.1.E-03	HTRA4	NA	
18	77723195	77723602	407	Hypo	1.4.E-02	4.6.E-04	HSBP1L1	NA	
3	1.84E+08	1.84E+08	677	Hypo	1.5.E-02	6.5.E-03	EIF4G1	NA	
17	27048437	27049557	1120	Hypo	1.5.E-02	1.9.E-03	RPL23A;SNORD42A;SNORD4B;SNORD4A	SDF2_N;ALDOC_N;SPAG5_N;CORO6_N;TLCD1_P;SEZ6_P;TMEM199_P;PIPOX_P;SEZ6_P;ABHD15_N;NLK_P	
20	47896888	47897451	563	Hypo	1.7.E-02	1.1.E-04	MIR1259;SNORD12B;C20orf199;SNORD12	B4GALT5_N;PREX1_N;ZNFX1_P	●
2	74727000	74727321	321	Hyper	1.6.E-02	1.2.E-04	LBX2	MRPL53_N	

17	34207332	34207663	331	Hypo	2.1.E-02	1.3.E-03	CCL5	SLFN11_N;ENSG00000267648_P;ENSG00000141150_N;ENSG00000267102_P;PIGW_P	●
11	1892038	1892307	269	Hypo	1.7.E-02	8.2.E-04	LSP1	NA	
CD8$^+$ T cells									
20	47896448	47897451	1003	Hypo	6.7.E-03	2.7.E-03	C20orf199;SNORD12B;SNORD12;MIR1259	B4GALT5_N;PREX1_N;PEDS1_N;STAU1_N;UBE2V1_N;ZNFX1_P	●
6	39758343	39758998	655	Hypo	8.6.E-03	1.3.E-03	DAAM2	NA	
22	31318179	31318546	367	Hypo	9.8.E-03	2.1.E-03	C22orf27	NA	
17	34207332	34208012	680	Hypo	1.1.E-02	1.9.E-03	CCL5	SLFN11_N;ENSG00000267648_P;ENSG00000267102_P;ENSG00000141150_N;PIGW_P	●
3	33701101	33701223	122	Hyper	1.3.E-02	1.5.E-02	CLASP2	CCR4_P;CRTAP_P	●

* Genes for which methylation sites correlated with expression were identified based on Methylation-Expression information from the integrated analysis of methylation and expression data utilizing data from previous studies. '_N' and '_P' indicate a negative and positive correlation, respectively, between the methylation rate and gene expression.

Integrative analysis of methylation and gene expression

To reduce type 1 errors in the detection of DMPs associated with NT1 from genome-wide data, we examined DMPs that have been indicated to be associated with NT1 by gene expression (Fig. 1B). As it is well-known that the expression levels of numerous genes are related to the DNA methylation levels of nearby methylation sites, we first searched for methylation status and gene expression pairs that correlated with each other (Methylation-Expression pairs) from previous studies with larger samples sizes that examined both DNA methylation and gene expression in $CD4^+$ and $CD8^+$ T cells throughout the entire genome (Table S1) (Tserel et al. 2015; Ventham et al. 2016; Martino et al. 2018; Clark et al. 2020; Li et al. 2020). The results of the EWAS for DMPs indicated that the majority of NT1-related methylation sites are common to both $CD4^+$ and $CD8^+$ T cells, although the strength of the association with NT1 was stronger in $CD4^+$ T cells (Fig. 2); consequently, our analysis mainly focused on $CD4^+$ T cells. We analyzed the results for $CD4^+$ T cells from the three previous studies (Study 1, 2, and 3) listed in Table S1. In total, 3,191 and 78,881 Methylation-Expression pairs with a false discovery rate (FDR) < 0.05 were detected in Study 1 and 2, respectively. In Study 3, results for both naive $CD4^+$ T cells and $CD4^+$ T cells after stimulation were available, so we analyzed both, and detected 4,272 and 26,725 Methylation-Expression pairs, respectively (Table S1). These Methylation-Expression pairs were combined for a total of 110,713 Methylation-Expression pairs. Utilizing these pairs, we searched for combinations where both the methylation status and gene expression were associated with NT1, and the direction of this relation aligned with the direction of the relation in the Methylation-Expression pairs detected from the previous studies. We set the following criteria for each analysis: NT1-related methylation with $P < 0.05$ and a combined rank within the top 1% in the NT1 EWAS for DMP; and NT1-related gene expression with $P < 0.05$ and fold change (FC) > 25% in the RNA-seq analysis. As a result, four Methylation-Expression pairs were detected (Table 2). Among these, two pairs were related to natural killer (NK) cell granule protein 7 (*NKG7*), and the two other pairs were related to programmed cell death 1 (*PDCD1*) and junctional adhesion molecule 2 (*JAM2*). All of these genes were genes that showed a negative correlation between

206 their expression levels and the methylation levels of NT1-associated methylation sites, and in the
207 NT1 cases, the expression levels of all of these genes increased and the methylation rate at the
208 methylation sites decreased.

209 Although we also conducted an analysis for $CD8^+$ T cells, there were no previous studies with
210 data were available on the detection of methylation using whole-genome bisulfite sequencing or
211 the EPIC array and the measurement of gene expression. Therefore, we conducted the analysis of
212 $CD8^+$ T cells using the results of two previous studies (Study 4 and 5) that utilized the Infinium
213 Human Methylation 450K BeadChip (450K array; Illumina; Table S1), which preceded the EPIC
214 array (850K array), making it difficult to compare the $CD8^+$ T cell analysis results with the $CD4^+$
215 T cell analysis results. When the analysis was conducted under conditions similar to those for the
216 $CD4^+$ T cell analysis, no Methylation-Expression pair that met the criteria was detected in the
217 $CD8^+$ T cell analysis. We then relaxed the conditions for detecting correlated Methylation-
218 Expression pairs from FDR < 0.05 to a nominal significance level of $P < 0.05$, and conducted the
219 analysis. As a result, we identified three Methylation-Expression pairs, one of which was related
220 to *NKG7,* which was also detected in the analysis of $CD4^+$ T cells. Similar to the $CD4^+$ T cell
221 analysis, its expression level was increased in the $CD8^+$ T cells in NT1 (Table S2).

Table 2. Results of the integrative analysis of methylation and gene expression data in CD4⁺ T cells

Methylation-Expression detection			Expression analysis of NT1				Methylation analysis of NT1					
							Discovery set			Replication set		
							Average beta		*P*-value	Average beta		*P*-value
Methylation-Expression	beta	FDR	Gene name	logFC	logCPM	*P*-value	Control	NT1		Control	NT1	
cg10044336: ENSG00000105374	-14.5	4.11E-05	*NKG7*	0.416	4.191	0.0131	0.699	0.665	8.81.E-04	0.751	0.708	3.47.E-03
cg21289857: ENSG00000105374	-15.4	3.31.E-03	*NKG7*	0.416	4.191	0.0131	0.820	0.802	6.87.E-04	0.855	0.830	2.81.E-02
cg01203136: ENSG00000188389	-3.8	0.0119	*PDCD1*	0.414	1.370	0.0251	0.642	0.613	3.70.E-03	0.699	0.627	2.41.E-04
cg00565412: ENSG00000154721	-572.6	0.0497	*JAM2*	0.435	0.623	0.0127	0.783	0.761	7.87.E-03	0.782	0.730	6.93.E-03

Abbreviations: FDR, false discovery rate; logFC, log fold change; logCPM, log counts per million

Integrative analysis of methylation and genotype data

To integrate the genotype and methylation data, we used information of the methylation quantitative trait locus (mQTL; Fig. 1C). Although tissue- and cell-specific mQTLs were present, in the context of QTL analysis, it has been demonstrated that larger sample sizes enhance the reliability of the results and enable the detection of a greater number of QTLs (Wang et al. 2022). Furthermore, prior research has reported that 84% and 80% of the effects of mQTLs detected in $CD4^+$ and $CD8^+$ T cells, respectively, exhibit directions concordant with those detected in whole blood. Consequently, we employed data pertaining to mQTLs identified in studies focusing on whole blood with sample sizes exceeding 1,000, as detailed in Table S3, for our integrative analysis. As shown in Table S3, Study 6 and 7 reported 12,689,548 and 11,165,559 mQTLs, respectively. Among these mQTLs, we searched for mQTLs related to NT1 that met the following conditions for both the involved SNP and methylation: for the SNP, a $P < 1E\text{-}3$ in the GWAS of NT1 (Toyoda et al. 2015) and an odds ratio (OR) that was in the same direction and greater than the OR in the GWAS in the samples of the present study (discovery and replication samples, NT1: N = 42 and control: N = 42); and for the methylation, a combined rank within the top 1% in the discovery samples that was replicated. As a result, in $CD4^+$ T cells, four mQTLs from Study 6 and two mQTLs from Study 7 were detected, with no mQTLs commonly detected in both studies (Table 3 and Table S4). The four mQTLs detected in Study 6 were all pairs consisting of two nearby SNPs and two methylation sites, which were found to negatively correlate with the expression of S100 calcium binding protein A4 (*S100A4*; Table 3 and Table S4).

When a similar analysis was conducted for $CD8^+$ T cells, no mQTL showing relationships between NT1 and both methylation and SNP was found. When the criteria for SNPs related to NT1 in the GWAS were relaxed from $P < 1E\text{-}3$ to $P < 0.01$, two mQTLs were detected (Table S5). However, in $CD8^+$ T cells, the correlation of SNPs and methylation pairs was weaker than that in the $CD4^+$ T cells.

Table 3. Results of the integrative analysis of methylation and genotype data in CD4⁺ T cells

mQTL analysis				Methylation analysis of NT1					Genotype analysis		
SNP	Methylation	beta	*P*-value	Discovery set			Nearby gene	Related gene expression in CD4⁺ T cell*	Current study	NT1 GWAS	
				Average beta		*P*-value					
				Control	NT1				OR	*P*-value	OR
Study 6 mQTL analysis											
chr1.154139381.SNP_C	cg25064552	-0.012	9.19E-20	0.658	0.626	1.08.E-02	SLC39A1	S100A4_N; CHRNB2_P	1.78	9.28.E-04	1.31
chr1.154134557.SNP_C	cg25064552	-0.012	1.42E-19	0.658	0.626	1.08.E-02	SLC39A1	S100A4_N; CHRNB2_P	1.78	9.86.E-04	1.31
chr1.154134557.SNP_C	cg23017879	-0.010	1.76E-14	0.677	0.650	2.69.E-04	NUP210L	S100A4_N; PYGO2_N	1.78	9.86.E-04	1.31
chr1.154139381.SNP_C	cg23017879	-0.010	2.04E-14	0.677	0.650	2.69.E-04	NUP210L	S100A4_N; PYGO2_N	1.78	9.28.E-04	1.31
Study 7 mQTL analysis											
chr8.11417150.SNP_G	cg21701351	0.0190	1.14E-99	0.811	0.788	7.73.E-03	BLK		NA	3.73.E-05	0.19
chr10.133934200.SNP_G	cg18650367	0.0547	5.26E-112	0.218	0.165	6.10.E-03			0.64	6.18.E-04	0.69

* Genes for which methylation sites correlated with expression were identified based on Methylation-Expression information from the integrated analysis of methylation and expression data utilizing data from previous studies. '_N' and '_P' indicate a negative and positive correlation, respectively, between the methylation rate and gene expression.

Abbreviation: OR, odds ratio

255 *Pathway analysis of genes affected by methylation sites associated with NT1*

256 Through the detection of DMRs and the integrative analysis of gene expression and genotype
257 data, various associated methylation regions or sites were identified, especially in the analysis of
258 $CD4^+$ T cells. Consequently, we conducted a pathway analysis of genes associated with the
259 methylation rates at these methylation sites by using the Methylation-Expression information
260 obtained from the $CD4^+$ T cell integrative analysis. The results revealed the involvement of
261 pathways related to immune responses, chemotaxis, and cell migration (Fig. 3 and Table S6).

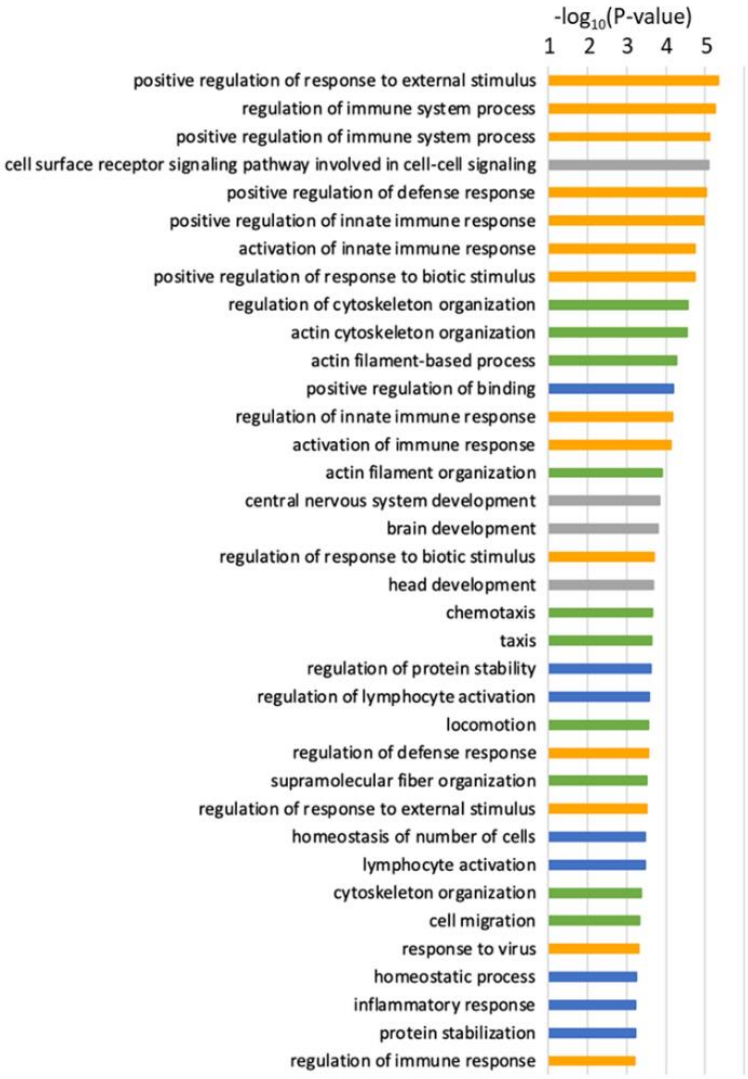


Figure 3. Results of the pathway analysis using genes suggested to be associated with NT1 in $CD4^+$ T cells. For easier interpretation, pathways were clustered into four groups using the k-means method. The groups were color-coded as follows: yellow, pathways related to the immune response; green, pathways related to cell migration; gray, neural pathways; and blue, others.

Characterization of NT1-related hypomethylated sites

In the process of conducting the association analyses for the identification of DMPs, a significant inflation in *P*-values was observed across the analyses of both cell types, with this phenomenon being more pronounced in $CD4^+$ T cells. Additionally, it was found that this inflation predominantly occurred at methylation sites exhibiting hypomethylation in NT1. Therefore, we explored the characteristics of these hypomethylated sites in NT1 to determine why such a pronounced trend of hypomethylation was observed in NT1.

In the examination of the prevalence of hypomethylation, the results indicated that for the discovery set, across all thresholds ranging from a nominally significant $P = 0.05$ to 1E-5, more than 97% of the replicated NT1-associated methylation sites exhibited hypomethylation in the $CD4^+$ T cells (Table 4).

Table 4. The proportion of hypomethylated sites among the NT1-related methylation sites in $CD4^+$ T cells

Discovery set	Replicated with $P < 0.05$*		
Association test	Number of methylation sites		*P*-value**
P-value	All	Hypomethylated in NT1 (%)	
1.00E-05	5	5 (100%)	8.34E-01
0.0001	45	44 (97.8%)	6.36E-02
0.001	364	359 (98.6%)	8.24E-09
0.01	3332	3259 (97.8%)	1.69E-67
0.05	14151	13728 (97.0%)	1.89E-273
Analyzed CpG sites***			
	746913	455481 (61.0%)	

* Replicated probes only include probes for which the direction of methylation was consistent in both the discovery set and the replication set. In other words, we targeted probes that are hypomethylated at NT1 in the discovery set and are also hypomethylated

at NT1 in the replication set, as well as those that are hypermethylated at NT1 in the discovery set and are also hypermethylated at NT1 in the replication set.
** Hypothesis testing for the difference in the population proportions.
*** Probes located in the *HLA* region were excluded.

In contrast, 61% of all analyzed methylation sites were hypomethylated in NT1. When comparing this ratio to the proportion of NT1-associated methylation sites that were hypomethylated in NT1, the results showed that the latter proportion was significantly larger (discovery set: probes with $P < 0.01$; $P = 1.69E-67$; Table 4). A similar trend was observed in $CD8^+$ T cells, in which over 80% of the NT1-associated methylation sites were hypomethylated in NT1 (Table 5).

Table 5. The proportion of hypomethylated sites among the NT1-related methylation sites in $CD8^+$ T cells

Discovery set	Replicated with $P < 0.05$*		
Association test	Number of methylation sites		*P*-value**
P-value	All	Hypomethylated in NT1 (%)	
1.00E-05	4	4 (100%)	9.38E-01
0.0001	23	21 (91.3%)	2.92E-01
0.001	147	123 (83.7%)	7.92E-03
0.01	1116	897 (80.4%)	4.83E-12
0.05	4650	3749 (80.6%)	4.13E-46
Analyzed CpG sites***			
	735,841	448,555 (61.8%)	

* Replicated probes only include probes for which the direction of methylation was consistent in both the discovery set and the replication set.
** Hypothesis testing for the difference in the population proportions.
*** Probes located in the *HLA* region were excluded.

298 We then examined the regions in which NT1-associated hypomethylated sites were common with
299 a $P < 0.01$ in the discovery set and were replicated with a $P < 0.05$. In both T cell analyses, we
300 found that NT1-related hypomethylated sites were significantly less frequent within CpG islands
301 ($CD4^+$ T cells: P = 3.03E-115 and $CD8^+$ T cells: P = 9.30E-34), and that they became more
302 frequent with increasing distance from the CpG islands in the shores and shelves (Fig. 4A and
303 Table S7). Furthermore, NT1-associated hypomethylated sites were particularly scarce in exon
304 and promoter regions (exons, $CD4^+$ T cells: P = 6.70E-26 and $CD8^+$ T cells: P = 7.98E-13;
305 Promoter_inactive, $CD4^+$ T cells: P = 2.99E-22 and $CD8^+$ T cells: P = 1.61E-09; Fig. 4B and
306 Table S8). When compared to the related regions from ChIP-seq in the T cells, NT1-related
307 hypomethylated sites were significantly less frequent in regions marked by histone modifications
308 associated with transcriptional activation, such as H3K4me3 and H3K27ac (Fig. 4C and Table
309 S9). Interestingly, despite the low overlap between NT-associated hypomethylated sites in the
310 $CD4^+$ T cell analysis (N = 3,259) and those in the $CD8^+$ T cell analysis (N = 897), with only 38
311 sites in common, both shared the characteristic that they were less common in regions likely to
312 influence gene expression, such as CpG islands, promoter regions, and open chromatin.
313 To investigate the relationships between NT1-associated hypomethylated sites, we generated
314 heatmaps of the methylation rates for the hypomethylated sites in each sample. The analysis
315 revealed a consistent pattern: individuals with a higher degree of hypomethylation showed this
316 trait across almost all of these methylation sites, while those with a lower degree of
317 hypomethylation displayed higher methylation across the board (Fig. S6). Consequently, the
318 average methylation rate of these hypomethylated sites was used as an indicator to represent the
319 degree of hypomethylation in each sample. We investigated the association between the degree
320 of hypomethylation for each sample and clinical information, including age, sex, and body mass
321 index (BMI). Sex and BMI did not show an association with the average value of hypomethylation
322 in the discovery set (Fig. S7A, B).

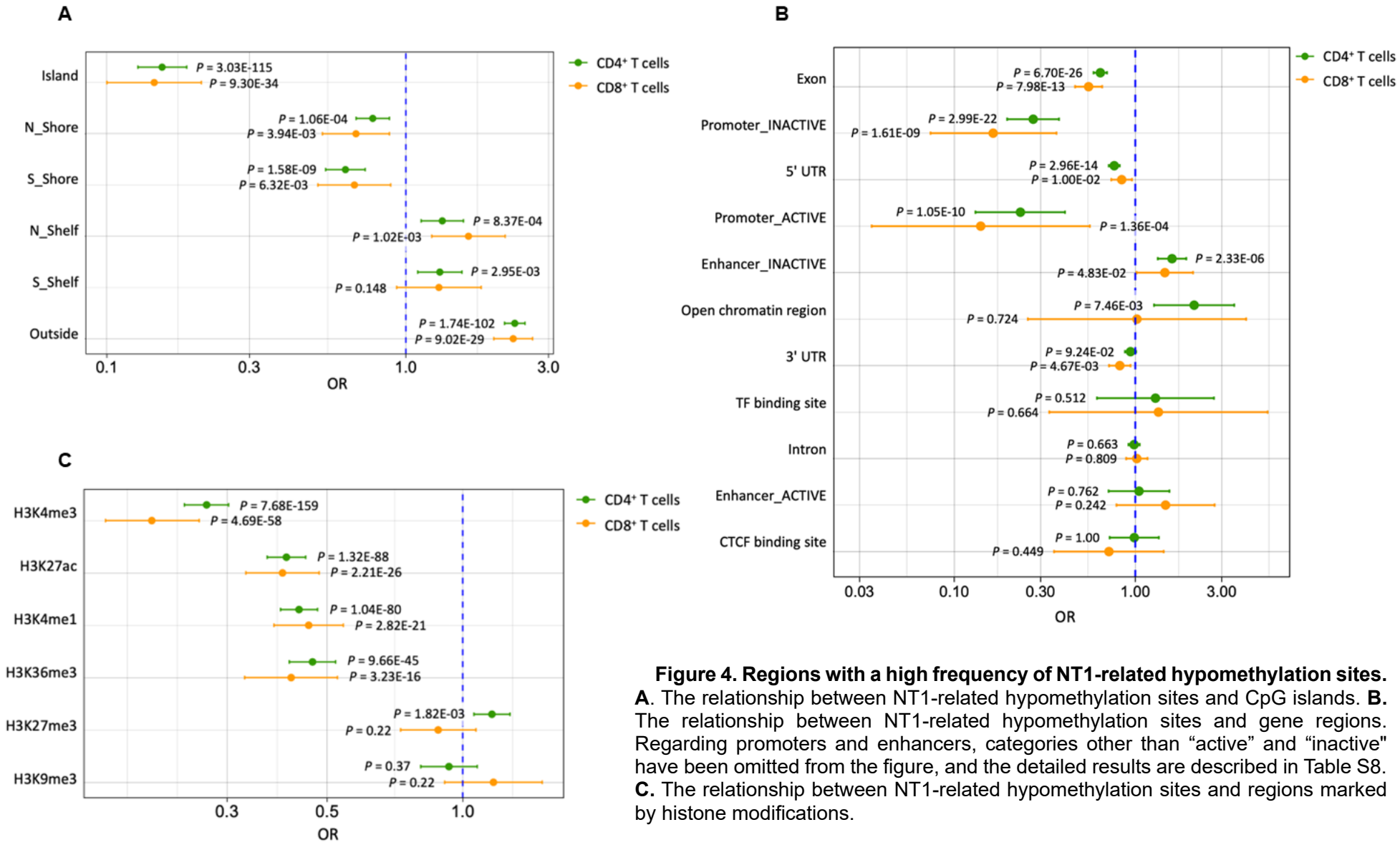
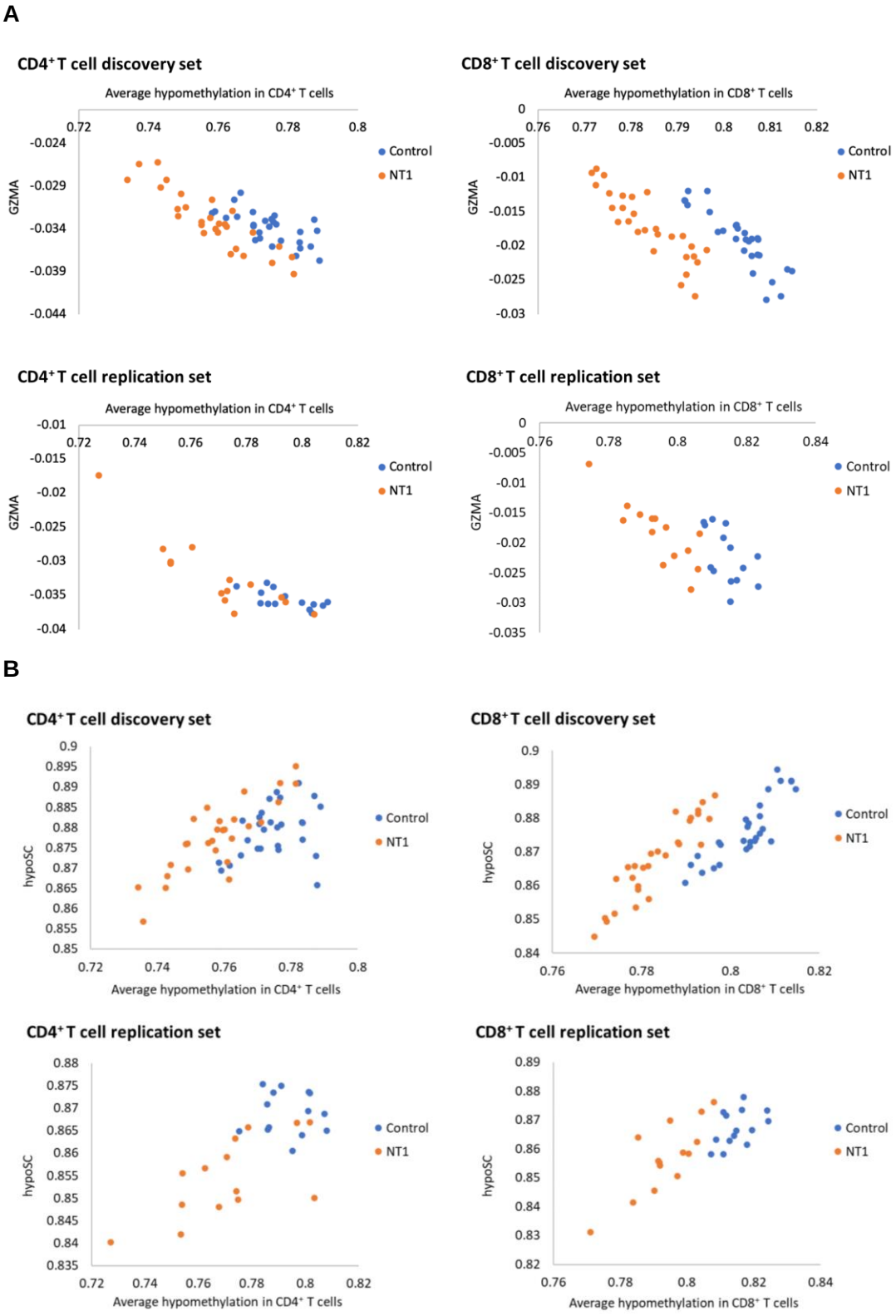


Figure 4. Regions with a high frequency of NT1-related hypomethylation sites. **A**. The relationship between NT1-related hypomethylation sites and CpG islands. **B.** The relationship between NT1-related hypomethylation sites and gene regions. Regarding promoters and enhancers, categories other than “active” and “inactive" have been omitted from the figure, and the detailed results are described in Table S8. **C.** The relationship between NT1-related hypomethylation sites and regions marked by histone modifications.

323

In the replication set, the BMI appeared to be weakly correlated with the average value of hypomethylation (R^2 = 0.396); this may be attributable to the fact that the BMI was significantly higher in the NT1 cases than in the controls in the replication set. Age was weakly correlated with the average value of hypomethylation, and a tendency was observed for increased hypomethylation with higher age (R^2 = 0.123-0.467; Fig. S7C). However, since age alone cannot explain the hypomethylation seen in NT1, we next examined the relationship between hypomethylation and various indicators that could be calculated with genome-wide methylation data using the DNA Methylation Age Calculator (Horvath 2013) and MethylDetectR (Hillary and Marioni 2020; Gadd et al. 2022) (Table S10). The analysis was adjusted for age. It should be noted that some of these indicators are based on values calculated assuming the input data is from whole blood, but our methylation data was obtained from fractionated T cells; thus, the reliability of these estimates was verified in the following section using whole blood methylation analysis data that we previously obtained with the 27K array. The analysis results showed that in the $CD4^+$ T cells, granzyme A (GZMA) demonstrated the strongest association with hypomethylation (P = 2.26E-10), followed by correlations with hypoSC (a methylation-based mitotic index; P = 6.08E-06; Fig. 5A, B and Table S11). In $CD8^+$ T cells, hypoSC showed the strongest association with hypomethylation (P = 1.61E-12), followed by a correlation with GZMA (P = 3.44E-08; Fig. 5A, B and Table S12).



341 **Figure 5. Indices that showed a correlation with the degree of hypomethylation.** **A**. Correlation between the estimated values of GZMA and the degree of hypomethylation (adjusted for age). **B.** Correlation between hypoSC and the degree of hypomethylation (adjusted for age).

342 *Investigation of NT1-related hypomethylation in whole blood*

343 As described in the previous section, when exploring the indicators showing an association with
344 hypomethylation in NT1 in the $CD4^+$ and $CD8^+$ T cells, some of the indicators were based on the
345 assumption that the input data for estimation are methylation data derived from whole blood; thus,
346 we also conducted studies using methylation data obtained from whole blood. We had previously
347 reported that many NT1-related methylation sites were hypomethylated in NT1 when using whole
348 blood for the analysis (Shimada et al. 2018). In the present study, to examine other details more
349 thoroughly, we reanalyzed the data. As shown in Fig. S8A, the *P*-values of the methylation sites
350 hypomethylated in NT1 were significantly inflated when compared to the expected values. In a
351 manner similar to our analysis of T cells, we then examined the relationship between the average
352 methylation rates of hypomethylated sites associated with NT1 and the indicators calculated using
353 the DNA Methylation Age Calculator (Horvath 2013) and MethylDetectR (Hillary and Marioni
354 2020; Gadd et al. 2022) to confirm the reproducibility of the results in the T cells. In our analysis,
355 we focused particularly on hypoSC and GZMA. As a result, no significant association with the
356 hypomethylation rate was observed when only age was adjusted for (Tables S11, S12 and Fig. S8B,
357 C). However, upon reviewing the plots and adjusting for the effect of NT1, nominally significant
358 associations between hypomethylation in NT1 and both hypoSC and GZMA were observed
359 (hypoSC: P = 1.14E-4 and GZMA: P = 1.56E-2; Table S13 and Fig. S8D, E).

360

361 *Prevalence of solo-WCGW CpGs in NT1-associated hypomethylated sites*

362 The hypoSC, an indicator related to cell mitosis, was among the indicators that showed an
363 association with hypomethylation at NT1-related methylation sites. Previous studies have reported
364 that DNA methylation at CpG sites is more likely to be lost in regions where there are fewer
365 neighboring CpG sites, particularly in CpGs within the "WCGW" context (where W represents an
366 A or T base; Solo-WCGW) (Zhou et al. 2018). Following the previous study, we investigated the
367 relationship between NT1-related hypomethylated sites and the number of CpG within windows of

368 ±35 bp as well as the sequence context, including the bases on either side of the CpG (Table S14).
369 We found that in both types of T cells, NT1-associated hypomethylated CpG sites were significantly
370 more prevalent in solo-WCGW ($CD4^+$ T cells: P = 9.87E-194, OR = 3.07 and $CD8^+$ T cells: P =
371 1.46E-22, OR = 2.20; Table 6). Furthermore, upon detailed examination, it was found that NT1-
372 associated hypomethylated sites were particularly prevalent within CpG sites located in the
373 sequence of ACGT (Fig. S9).

Table 6. NT1-associated hypomethylation sites are abundant in solo-WCGW

	Rate in all analyzed CpGs	CD4$^+$ T cells					CD8$^+$ T cells				
		Rate in hypomethylated NT1 CpGs	*P*-value	OR	95%CI		Rate in hypomethylated NT1 CpGs	*P*-value	OR	95%CI	
					Lower	Upper				Lower	Upper
Solo-WCGW	0.11	0.27	9.87E-194	3.069	2.839	3.317	0.21	1.46E-22	2.199	1.872	2.584
Solo	0.34	0.55	1.12E-150	2.446	2.283	2.622	0.52	9.64E-32	2.154	1.89	2.456
WCGW	0.23	0.45	2.85E-198	2.77	2.585	2.969	0.38	1.89E-29	2.136	1.867	2.444

Abbreviation: OR, odds ratio

Discussion

In this study, we conducted an EWAS in $CD4^+$ and $CD8^+$ T cells from patients with NT1, and found that in both types of T cells, methylation sites associated with chemokine-related genes, such as *CCR4* and *CCL5*, were related to NT1. Particularly in $CD4^+$ T cells, genes with expression levels that correlate with the methylation sites were predominantly involved in pathways of immune responses and cell chemotaxis. Additionally, we discovered the global hypomethylation of NT1-associated methylation sites in both cell types, and that the degree of hypomethylation was well-correlated with the estimated values of GZMA and hypoSC, indicators of cell mitosis. NT1-associated hypomethylated sites were significantly more prevalent in solo-WCGW context, which is known to lose methylation modifications easily during cell division, suggesting that cell division might be enhanced in the T cells of NT1 patients.

When we searched for methylation sites associated with NT1, to minimize false positives, we adopted two strategies. Firstly, instead of focusing only on single-point associations, *i.e.,* DMPs, we focused on DMRs. Secondly, even for single-point DMP associations, we sought to identify the DMPs that have been indicated to be associated with NT1 by other omics data. In the analysis for DMR identification, more DMRs were detected in the analysis of $CD4^+$ T cells than in the analysis of $CD8^+$ T cells (Table 1). However, most DMRs were found to be commonly associated in both cell types when including those that showed an association in either the discovery or replication analysis. Among these, especially in the replicated DMRs of both cell types, there were DMRs near *CCL5* and DMRs affecting the gene expression of *CCR4*. Both of these genes are chemokine-related, and considering that the previous GWAS of NT1 reported associations with a SNP, *i.e.,* rs3181077, located between other chemokine-related genes, *i.e., CCR1* and *CCR3*, and considering that the observed decrease in their expression levels in NT1, *CCL5* and *CCR4* represent highly intriguing candidate genes. *CCL5*, which is also known as regulated on activation, normal T cell expressed and secreted (RANTES), is a ligand for *CCR1*, *CCR3,* and *CCR5*, and it acts as a chemoattractant for monocytes, memory T-helper cells, and eosinophils (Korbecki et al. 2020). In this study, we were unable to find an association between the upstream DMR of *CCL5* and the gene expression level of *CCL5* itself, which was likely due to the very low expression levels of *CCL5* in the analyzed data. An investigation into the expression in immune cells using DICE (the database of immune cell expression, expression quantitative trait loci (eQTLs),

and epigenomics) (Schmiedel et al. 2018) revealed that the average transcripts per million (TPM) was less than 1 in all cell types. However, the single-cell analysis results showed expression primarily in $CD4^+$ T cells (Schmiedel et al. 2018). If we consider that the expression levels of genes with low expression levels, or in cases where *CCL5* expression levels are increased by some kind of immune system activation, there is a reasonable possibility of an association between gene expression and the NT1-associated DMRs in the *CCL5* region identified in this study. *CCR4* is a chemokine receptor for TARC/CCL17 and MDC/STCP-1/CCL22, which are mainly expressed by T cells (Yoshie 2021). *CCR4* plays a crucial role in regulating immune balance, and is highly expressed in activated regulatory T cells; thus, it is increasingly being recognized as a potential target for cancer therapy (Maeda et al. 2021; Yoshie 2021).

The involvement of chemotaxis in NT1 was also suggested from the observed association with *S100A4* in the results of the integrated analysis of mQTLs using genotype data of $CD4^+$ T cells (Table 3). *S100A4* is a member of the Ca(2+)-dependent S100 protein family and a transcription factor that regulates cellular motility. It is known that *S100A4* promotes filament depolymerization and chemotaxis through its direct interaction with soluble myosin-IIA, leading to the formation of stable protrusions (Li and Bresnick 2006). The gene expression of *S100A4* showed a negative correlation with methylation sites that were hypomethylated in the NT1 patient group, suggesting the possibility that, at least at the onset of NT1, the expression of *S100A4* was increased in the $CD4^+$ T cells of the patients with NT1. A common issue in normal DNA methylation analysis is the difficulty in determining whether the associated factors are causes or consequences of the disease. In our analysis, it was also challenging to distinguish whether the associations found with many methylation sites were causes or consequences of NT1; however, the results of the integrated analysis with genotype data indicated causal effects, because genetic factors can be considered as causal for the disease. Therefore, it is highly likely that changes in the methylation rates affected by genetic variants (mQTLs) contributed to the disease onset. Considering this in conjunction with previous reports suggesting associations with chemokines derived from GWAS SNP associations (Toyoda et al. 2015), chemokine-related pathways could be a significant cause of NT1 development. Further detailed examinations of the chemokine pathways are necessary.

In this study, we observed global hypomethylation at NT1-related methylation sites in both $CD4^+$ and $CD8^+$ T cells (Fig. 2 and Tables 4, 5). This trend was not only seen in the replication set, but was also confirmed in the results of the methylation analysis for NT1 using whole blood that we had conducted previously (Shimada et al. 2010). In our investigation of hypomethylated sites, we found that they were particularly scarce in CpG islands, and that they became more prevalent further from the islands. It was also evident that they were significantly less common in gene promoter regions and intragenic areas (Fig. 4). Although the methylation of CpG islands in the upstream promoter regions of genes greatly influences gene expression, much remains unknown about the functions of the methylation sites in such outside regions. While exploring why such hypomethylation occurs and what factors are involved, we discovered that the degree of hypomethylation was highly correlated with the levels of GZMA estimated from methylation data, as well as with the hypoSC cell division index (Fig. 5 and Tables S11, S12). We confirmed that there was very little overlap between the methylation sites used to calculate the degree of hypomethylation ($CD4^+$ T cells: N = 3,259 and $CD8^+$ T cells: N = 897) and those used to estimate indices, such as GZMA and hypoSC (N = 3,629), with only five sites in $CD4^+$ T cells and one site in $CD8^+$ T cells that overlapped. GZMA is one of the most abundant proteases present in the cytoplasmic granules of cytotoxic T cells and NK cells, and it is known for its crucial role in inducing the death of infected cells (van Daalen et al. 2020). However, in recent years, there has been debate regarding its role in cytotoxicity, and it has been shown that extracellular GZMA is involved in the regulation of inflammation and inflammatory diseases (van Daalen et al. 2020). The potential of GZMA to exacerbate inflammation and cause excessive immune responses, sepsis, tumor promotion, and rheumatoid arthritis has been suggested (van Daalen et al. 2020). In the present study, the association of NT1 with cytotoxic granule exocytosis and inflammation was also suggested by the association of *NKG7* identified through the integrated analysis of methylation and gene expression data (Table 2). NK cell granule protein 7 is a factor that controls lymphocyte granule exocytosis and subsequent inflammation in a wide range of diseases, and it has been reported that *NKG7*, which is particularly expressed in $CD4^+$ and $CD8^+$ T cells, plays a crucial role in promoting inflammation in infections (Ng et al. 2020). Although determining whether the greater impact on hypomethylation stems from events like inflammation, associated with GZMA, or from the promotion of cell division,

associated with hypoSC, is challenging due to the high correlation between their estimated values ($CD4^+$ T cells: $R^2 = 0.528$ and $CD8^+$ T cells: $R^2 = 0.706$), it is entirely plausible that T cell proliferation progressed due to the accelerated inflammation (Moro-García et al. 2018). The possibility of increased hypomethylation as a result of increased cell division was supported by the observation that NT1-related hypomethylated sites were especially prevalent in solo-WCGW motifs (Table 6). Previous studies have shown that hypomethylation occurs in only about half of the genome, especially in CG sequences not accompanied by surrounding CpGs and flanked by purine bases, such as in WCGW motifs (Zhou et al. 2018). It is also known that hypomethylation is more likely to occur in areas with low gene density, in areas with low GC density, and in so-called compartment B regions (Quante and Bird 2016; Zhou et al. 2018). These genomic characteristics overlap well with the features of regions where NT1-related hypomethylated sites were predominantly found. In other autoimmune diseases, particularly in systemic lupus erythematosus (SLE) and multiple sclerosis, global hypomethylation in immune cells has been reported (Hurtado et al. 2020; Ma et al. 2021). In the case of SLE, there have been reports suggesting that hypomethylated sites are often found upstream of immune-related genes, indicating that the mechanism may differ slightly from that of the hypomethylation we observed in NT1 (Hurtado et al. 2020). In multiple sclerosis, the global hypomethylation sites were particularly numerous within promoter regions in B cells, and they were reported to influence the expression of genes involved in the activation of B cells, which in turn further activates B cells (Ma et al. 2021). In contrast, the hypomethylation observed in NT1 was thought to possibly result from the progression of T cell division. Therefore, even though global hypomethylation is observed across different immune-related disorders, the mechanisms involved appear to be different.

There are several limitations to this study. First, we obtained gene expression data from the same samples from which we acquired methylation data and conducted an integrated analysis, but the expression data was derived from the mRNA in whole blood. If we had more samples available, it would have been preferable to obtain gene expression data from separated $CD4^+$ and $CD8^+$ T cells. Thus, we might have missed associations for some genes, especially those with varying expression patterns in different immune cells. Second, among the indicators showing a relationship with global hypomethylation in NT1, the estimated values for proteins, including GZMA, were based on predictions

488 intended for methylation data derived from whole blood, but our methylation data was obtained from
489 isolated T cells, leaving questions about the accuracy of these estimates. Although we have validated
490 the factors of interest using methylation data from whole blood, this data was obtained using the 27K
491 array, which has fewer probes. Therefore, it is desirable to confirm the reproducibility in the future
492 using EPIC array data derived from whole blood. Thirdly, although we conducted a comprehensive
493 methylation analysis using arrays, we cannot deny the possibility that methylation sites not covered by
494 the arrays may be associated with NT1. In the future, when whole-genome bisulfite sequencing with a
495 deeper depth than the sequencing depths possible now becomes available, more associated factors may
496 be identified. Lastly, in this study, we first isolated $CD8^+$ T cells from the buffy coat using magnetic
497 beads, and then separated $CD4^+$ T cells from the remaining blood components, resulting in the
498 possibility that the $CD8^+$ T cells included $CD4^+CD8^+$ double-positive T cells. Furthermore, we did not
499 examine the different subtypes of $CD4^+$ and $CD8^+$ T cells. The global hypomethylation associated with
500 NT1 showed a high correlation with indicators, such as GZMA and hypoSC, suggesting that
501 inflammation and increased cell division could explain much of the hypomethylation. However, there
502 were still differences in the degree of hypomethylation between the case and control groups that could
503 not be fully explained. Namely, when comparing the same hypoSC and GZMA values, the NT1 cases
504 tended to show more hypomethylation than the controls, particularly in the results of the $CD8^+$ T cells
505 and in the replication analysis with the 27K array using whole blood data. These data suggest the
506 possibility of hypomethylation in NT1 that cannot be explained by inflammation or increased cell
507 division alone. Particularly, the significant fact that hypomethylation cannot be explained by GZMA
508 and hypoSC alone in the whole blood data indicates that the presence of different subtypes of immune
509 cells in NT1 compared to control individuals may be a contributing factor. Future single-cell analyses
510 in NT1 could potentially reveal the details of the immune abnormalities that contribute to the onset of
511 the disease.

512 In conclusion, the results of our DNA methylation analysis of $CD4^+$ and $CD8^+$ T cells in NT1 suggest
513 a high likelihood that abnormalities in chemokine-related pathways are a major cause of disease onset.
514 Furthermore, our results indicate that increased T cell division in NT1, possibly driven by heightened
515 inflammatory responses, may lead to hypomethylation in regions of DNA where methylation is more

easily disrupted. Future investigations, including single-cell analyses and single-cell DNA methylation analyses, are expected to reveal more details of the mechanism of NT1 onset.

Methods

The details of the methods are described in the supplementary materials.

Subjects in the discovery set

Samples from 28 patients with NT1 and 28 control subjects without significant sleep-disturbing events who underwent diagnostic sleep studies (polysomnogram and the multiple sleep latency test) at Seiwa Hospital (Shinjuku, Tokyo, Japan) or Koishikawa Tokyo Hospital (Bunkyo, Tokyo, Japan) were included. We extracted DNA and RNA from the samples. This study received ethical approval from all relevant agencies, and written informed consent was obtained from all participants.

Isolation of $CD4^+$ and $CD8^+$ T cells and DNA extraction

$CD4^+$ and $CD8^+$ T cells were extracted from the lymphocyte fraction using the Dynabeads CD4 Positive Isolation Kit and Dynabeads CD8 Positive Isolation Kit (Invitrogen, Waltham, MA, USA) according to the manufacturer's protocols. First, $CD8^+$ T cells were isolated from the buffy coat, and $CD4^+$ T cells were isolated from the remaining blood cell components. DNA was extracted from both types of T cells using the QIAamp DNA Mini Kit (QIAGEN).

DNA methylation analysis

DNA was first bisulfite-converted, then the genome-wide methylation profile was examined with a DNA methylation array (EPIC array; Infinium® Methylation EPIC BeadChip, Illumina).

Association analysis of the methylation data

To detect NT1-associated methylation sites, we searched for DMPs and DMRs using the *t* test and bump hunting method (Jaffe et al. 2012), respectively. For DMPs, since not only the statistical significance from P values, but also a large difference in the methylation rate is important for biological significance, we calculated a combined rank considering both indices (the *P* value and $\Delta\beta$, which is the average β for the cases – average β for the controls), and we considered the top-ranked NT1-associated

methylation sites to be particularly important. Association analyses and visualization of the data, such as Q-Q plots, were conducted using R software (ver. 3.6.2).

Replication samples and analysis

We obtained replication samples from 14 patients with NT1 and 14 control subjects, (Table S16).

RNA sequencing

We extracted RNA from the whole blood samples of the 42 patients with NT1 and 42 control subjects included in the discovery and replication sets for DNA methylation analysis and performed RNA-seq. Sequencing libraries were prepared using TruSeq Stranded Total RNA Ribo Zero (Illumina), and paired-end RNA-seq was performed on a NovaSeq 6000 (Illumina). The FASTQ read data were trimmed using a quality score of 0.05 and a minimum length of 15 nt in CLC Genomics Workbench 21 (QIAGEN). Read mapping was performed with the following parameters: mismatch cost, 2; insertion cost, 3; deletion cost, 3; length fraction, 0.8; similarity fraction, 0.8; and maximum number of hits for a read, 10. The read count data were corrected by the trimmed mean of M values (TMM) method implemented in the edgeR package (Robinson et al. 2010; McCarthy et al. 2012; Chen et al. 2016). Association analysis of the expression data was performed using the likelihood ratio test with age, sex, and BMI as covariates in the design matrix.

Genotyping

DNA samples that were included in the DNA methylation analysis and RNA-seq (NT1: N = 42 and control: N = 42) were genotyped with Axiom Japonica array v2 (Toshiba, Tokyo, Japan). Imputation analysis was performed using BEAGLE5.1 (Browning et al. 2018) with reference panels from the 1000 Genomes Project (phase 3) and from the Tohoku Medical Megabank 14KJPN (Fuse et al. 2019). SNP filtering was performed with the following criteria: information score ≥ 0.5; minor allele frequency ≥ 0.01; SNP call rate ≥ 95%; and Hardy-Weinberg equilibrium P value > 0.01. Genotype data from a previous GWAS of NT1 were also used in the analysis. The details of the GWAS data are described elsewhere (Toyoda et al. 2015; Shimada et al. 2020).

RNA sequencing and Genotyping

We extracted RNA from the whole blood samples of the 42 patients with NT1 and 42 control subjects included in the discovery and replication sets for DNA methylation analysis and performed RNA-seq. DNA samples that were included in the DNA methylation analysis and RNA-seq were also genotyped with Axiom Japonica array v2 (Toshiba, Tokyo, Japan).

Integrative analysis of methylation and gene expression

To perform an integrative analysis of the methylation and gene expression data, we first detected pairs of methylation rate and gene expression that correlated with each other (Methylation-Expression pairs) using data from previous studies with larger sample sizes (Table S1) (Tserel et al. 2015; Ventham et al. 2016; Martino et al. 2018; Clark et al. 2020; Li et al. 2020). Next, we searched for pairs in which the correlated methylation rate and gene expression were both associated with NT1 in the samples of this study, and in which the direction of the association was consistent with the Methylation-Expression pair (Fig. 1B).

Integrative analysis of methylation and genotype data

To integrate the genotype and methylation data, we used information from the mQTL. An mQTL is where the methylation rate correlates with the genotype of the SNP, and it has been reported that the mQTL profiles are nearly identical between whole blood and each immune cell (Hawe et al. 2022).

Investigation of NT1-related hypomethylated sites

Since a global hypomethylation trend was observed in NT1, we examined the NT1-associated hypomethylation sites with a $P < 0.01$ in the discovery set that were replicated with a $P < 0.05$. To explore the kinds of genomic regions where the hypomethylated sites are commonly present, Ensembl Human Regulatory Features (GRCh38.p14) was used to examine the positional relationship between hypomethylated sites and genomic regulatory features, such as enhancers and promoters (Cunningham et al. 2022). We also used the functional genomics data obtained from the ENCODE project (Moore et

al. 2020) to investigate genomic features by Histone ChIP-seq. The positional information for CpG islands was based on Illumina's annotation file for the EPIC array.

We generated heat maps, and confirmed that the methylation rates were generally low in samples with low levels of methylation in the NT1-associated hypomethylation sites. Since some subjects showed a generally high methylation status and others showed a generally low hypomethylation status across the selected DMPs, the average methylation rates of NT1-associated hypomethylation sites were calculated for each sample, serving as an indicator of the degree of hypomethylation. To identify factors that could explain this degree of hypomethylation, correlations with age, sex, and BMI were examined.

To examine factors associated with the degree of hypomethylation, we calculated various indicators that could be determined from genome-wide methylation data using the DNA Methylation Age Calculator (Horvath 2013) and MethylDetectR (Hillary and Marioni 2020; Gadd et al. 2022) (Table S10). We then conducted a linear regression analysis using age as a covariate to explore indices indicating an association with the degree of hypomethylation. Some indices were designed assuming that the methylation data were derived from whole blood as the input. Therefore, for the validation of the results, similar estimations were also conducted using data from an EWAS of whole blood from patients with NT1 that we had previously performed (Shimada et al. 2018). The prior study included 26 cases of NT1 and 20 controls, and the methylation levels were measured using the 27K array (Illumina); the details of the analysis are described elsewhere (Shimada et al. 2018).

Acknowledgements

This study was supported by Grants-in-Aid for Scientific Research (Grant numbers 15H04709, 17J01616, 18K15053, 19H03588, 21H02856, and 22H03006) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. This study was also supported by the Practical Research Project for Rare/Intractable Diseases and the Strategic Research Program for Brain Sciences from the Japan Agency for Medical Research and Development (AMED). We thank all participants who cooperated in this study. We also thank the staff of Seiwa Hospital and Koishikawa Tokyo Hospital for their support in blood sampling.

Authors' contributions

TM and MS designed the study. MH collected the blood samples. MS and YH performed the experiments. MS and TM carried out the statistical analyses. TM, YH, KT, TK and MH interpreted the results. MS, TM, and MH drafted the manuscript, and all authors contributed to the final version of the paper.

Disclosures and competing interests

The authors have no financial arrangements or connections related to the content of this article to disclose. The authors have no conflicts of interest to declare.

Data availability

The genome-wide methylation dataset supporting the conclusions of this article has been deposited in the DNA Data Bank of Japan (DDBJ; https://www.ddbj.nig.ac.jp/index-e.html), and is available from the European Nucleotide Archive operated by the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI; https://www.ebi.ac.uk/ena/browser/home) under the accession number PRJDB15980.

References

Berdasco M, Esteller M. 2019. Clinical epigenetics: seizing opportunities for translation. *Nat Rev Genet* **20**: 109-127.

Browning BL, Zhou Y, Browning SR. 2018. A One-Penny Imputed Genome from Next-Generation Reference Panels. *Am J Hum Genet* **103**: 338-348.

Chen Y, Lun AT, Smyth GK. 2016. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* **5**: 1438.

Clark AD, Nair N, Anderson AE, Thalayasingam N, Naamane N, Skelton AJ, Diboll J, Barton A, Eyre S, Isaacs JD et al. 2020. Lymphocyte DNA methylation mediates genetic risk at shared immune-mediated disease loci. *J Allergy Clin Immunol* **145**: 1438-1451.

Cogswell AC, Maski K, Scammell TE, Tucker D, Orban ZS, Koralnik IJ. 2019. Children with Narcolepsy type 1 have increased T-cell responses to orexins. *Ann Clin Transl Neurol* **6**: 2566-2572.

Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov AG, Barnes I, Bennett R et al. 2022. Ensembl 2022. *Nucleic Acids Res* **50**: D988-D995.

Dauvilliers Y, Arnulf I, Lecendreux M, Monaca Charley C, Franco P, Drouot X, d'Ortho MP, Launois S, Lignot S, Bourgin P et al. 2013. Increased risk of narcolepsy in children and adults after pandemic H1N1 vaccination in France. *Brain* **136**: 2486-2496.

Faraco J, Lin L, Kornum BR, Kenny EE, Trynka G, Einen M, Rico TJ, Lichtner P, Dauvilliers Y, Arnulf I et al. 2013. ImmunoChip study implicates antigen presentation to T cells in narcolepsy. *PLoS Genet* **9**: e1003270.

Flanagan JM. 2015. Epigenome-wide association studies (EWAS): past, present, and future. *Methods Mol Biol* **1238**: 51-63.

Fuse N, Sakurai-Yageta M, Katsuoka F, Danjoh I, Shimizu R, Tamiya G, Nagami F, Kawame H, Higuchi S, Kinoshita K et al. 2019. Establishment of Integrated Biobank for Precision Medicine and Personalized Healthcare: The Tohoku Medical Megabank Project. *JMA J* **2**: 113-122.

Gadd DA, Hillary RF, McCartney DL, Zaghlool SB, Stevenson AJ, Cheng Y, Fawns-Ritchie C, Nangle C, Campbell A, Flaig R et al. 2022. Epigenetic scores for the circulating proteome as tools for disease prediction. *Elife* **11**.

Hachiya T, Furukawa R, Shiwa Y, Ohmomo H, Ono K, Katsuoka F, Nagasaki M, Yasuda J, Fuse N, Kinoshita K et al. 2017. Genome-wide identification of inter-individually variable DNA methylation sites improves the efficacy of epigenetic association studies. *NPJ Genom Med* **2**: 11.

680 Hallmayer J, Faraco J, Lin L, Hesselson S, Winkelmann J, Kawashima M, Mayer G, Plazzi G,
681 Nevsimalova S, Bourgin P et al. 2009. Narcolepsy is strongly associated with the T-cell receptor
682 alpha locus. *Nat Genet* **41**: 708-711.
683 Han F, Faraco J, Dong XS, Ollila HM, Lin L, Li J, An P, Wang S, Jiang KW, Gao ZC et al. 2013.
684 Genome wide analysis of narcolepsy in China implicates novel immune loci and reveals
685 changes in association prior to versus after the 2009 H1N1 influenza pandemic. *PLoS Genet*
686 **9**: e1003880.
687 Han F, Lin L, Li J, Dong SX, An P, Zhao L, Liu NY, Li QY, Yan H, Gao ZC et al. 2012. HLA-DQ
688 association and allele competition in Chinese narcolepsy. *Tissue Antigens* **80**: 328-335.
689 Hawe JS, Wilson R, Schmid KT, Zhou L, Lakshmanan LN, Lehne BC, Kühnel B, Scott WR,
690 Wielscher M, Yew YW et al. 2022. Genetic variation influencing DNA methylation provides
691 insights into molecular mechanisms regulating genomic function. *Nat Genet* **54**: 18-29.
692 Hillary RF, Marioni RE. 2020. MethylDetectR: a software for methylation-based health profiling.
693 *Wellcome Open Res* **5**: 283.
694 Honda M, Arai T, Fukazawa M, Honda Y, Tsuchiya K, Salehi A, Akiyama H, Mignot E. 2009.
695 Absence of ubiquitinated inclusions in hypocretin neurons of patients with narcolepsy.
696 *Neurology* **73**: 511-517.
697 Hong SC, Lin L, Lo B, Jeong JH, Shin YK, Kim SY, Kweon Y, Zhang J, Einen M, Smith A et al.
698 2007. DQB1*0301 and DQB1*0601 modulate narcolepsy susceptibility in Koreans. *Hum*
699 *Immunol* **68**: 59-68.
700 Horvath S. 2013. DNA methylation age of human tissues and cell types. *Genome Biol* **14**: R115.
701 Hurtado C, Acevedo Sáenz LY, Vásquez Trespalacios EM, Urrego R, Jenks S, Sanz I, Vásquez G.
702 2020. DNA methylation changes on immune cells in Systemic Lupus Erythematosus.
703 *Autoimmunity* **53**: 114-121.
704 Jaffe AE, Murakami P, Lee H, Leek JT, Fallin MD, Feinberg AP, Irizarry RA. 2012. Bump hunting
705 to identify differentially methylated regions in epigenetic epidemiology studies. *Int J*
706 *Epidemiol* **41**: 200-209.
707 Kernohan KD, Cigana Schenkel L, Huang L, Smith A, Pare G, Ainsworth P, Boycott KM, Warman-
708 Chardon J, Sadikovic B, Consortium CRC. 2016. Identification of a methylation profile for
709 DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin*
710 *Epigenetics* **8**: 91.
711 Khatami R, Maret S, Werth E, Retey J, Schmid D, Maly F, Tafti M, Bassetti CL. 2004. Monozygotic
712 twins concordant for narcolepsy-cataplexy without any detectable abnormality in the
713 hypocretin (orexin) pathway. *Lancet* **363**: 1199-1200.
714 Korbecki J, Kojder K, Barczak K, Simińska D, Gutowska I, Chlubek D, Baranowska-Bosiacka I.
715 2020. Hypoxia Alters the Expression of CC Chemokines and CC Chemokine Receptors in a
716 Tumor-A Literature Review. *Int J Mol Sci* **21**.

717 Kornum BR, Kawashima M, Faraco J, Lin L, Rico TJ, Hesselson S, Axtell RC, Kuipers H, Weiner
718 K, Hamacher A et al. 2011. Common variants in P2RY11 are associated with narcolepsy. *Nat*
719 *Genet* **43**: 66-71.
720 Latorre D, Kallweit U, Armentani E, Foglierini M, Mele F, Cassotta A, Jovic S, Jarrossay D, Mathis
721 J, Zellini F et al. 2018. T cells in patients with narcolepsy target self-antigens of hypocretin
722 neurons. *Nature* **562**: 63-68.
723 Li T, Ortiz-Fernández L, Andrés-León E, Ciudad L, Javierre BM, López-Isac E, Guillén-Del-
724 Castillo A, Simeón-Aznar CP, Ballestar E, Martin J. 2020. Epigenomics and transcriptomics
725 of systemic sclerosis CD4+ T cells reveal long-range dysregulation of key inflammatory
726 pathways mediated by disease-associated susceptibility loci. *Genome Med* **12**: 81.
727 Li ZH, Bresnick AR. 2006. The S100A4 metastasis factor regulates cellular motility via a direct
728 interaction with myosin-IIA. *Cancer Res* **66**: 5173-5180.
729 Ma Q, Caillier SJ, Muzic S, Wilson MR, Henry RG, Cree BAC, Hauser SL, Didonna A, Oksenberg
730 JR, Team UoCSFM-E. 2021. Specific hypomethylation programs underpin B cell activation in
731 early multiple sclerosis. *Proc Natl Acad Sci U S A* **118**.
732 Maeda Y, Wada H, Sugiyama D, Saito T, Irie T, Itahashi K, Minoura K, Suzuki S, Kojima T, Kakimi
733 K et al. 2021. Depletion of central memory CD8+ T cells might impede the antitumor therapeutic effect
734 of Mogamulizumab. *Nat Commun* **12**: 7280.
735 Martino D, Neeland M, Dang T, Cobb J, Ellis J, Barnett A, Tang M, Vuillermin P, Allen K, Saffery
736 R. 2018. Epigenetic dysregulation of naive CD4+ T-cell activation genes in childhood food
737 allergy. *Nat Commun* **9**: 3308.
738 McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq
739 experiments with respect to biological variation. *Nucleic Acids Res* **40**: 4288-4297.
740 Mignot E. 1998. Genetic and familial aspects of narcolepsy. *Neurology* **50**: S16-22.
741 Mignot E. 2017. Narcolepsy: Genetics, Immunology, and Pathophysiology. In: Kryger M, Roth T,
742 Dement W, editors. Principles and Practice of Sleep Medicine. 6th ed., pp. p. 855-872. Elsevier.
743 Mignot E, Lin L, Rogers W, Honda Y, Qiu X, Lin X, Okun M, Hohjoh H, Miki T, Hsu S et al. 2001.
744 Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic
745 groups. *Am J Hum Genet* **68**: 686-699.
746 Miyagawa T, Kawashima M, Nishida N, Ohashi J, Kimura R, Fujimoto A, Shimada M, Morishita
747 S, Shigeta T, Lin L et al. 2008. Variant between CPT1B and CHKB associated with
748 susceptibility to narcolepsy. *Nat Genet* **40**: 1324-1328.
749 Miyagawa T, Tokunaga K. 2019. Genetics of narcolepsy. *Hum Genome Var* **6**: 4.
750 Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, Kawli T, Davis CA, Dobin A,
751 Kaul R et al. 2020. Expanded encyclopaedias of DNA elements in the human and mouse
752 genomes. *Nature* **583**: 699-710.

Moresco M, Lecciso M, Ocadlikova D, Pizza F, Curti A, Plazzi G. 2019. Flow cytometry T cell profiling in a recent-onset narcoleptic type 1 child: a case report. *Sleep Med* **68**: 21-23.

Moro-García MA, Mayo JC, Sainz RM, Alonso-Arias R. 2018. Influence of Inflammation in the Process of T Lymphocyte Differentiation: Proliferative, Metabolic, and Oxidative Changes. *Front Immunol* **9**: 339.

Ng SS, De Labastida Rivera F, Yan J, Corvino D, Das I, Zhang P, Kuns R, Chauhan SB, Hou J, Li XY et al. 2020. The NK cell granule protein NKG7 regulates cytotoxic granule exocytosis and inflammation. *Nat Immunol* **21**: 1205-1218.

Nohynek H, Jokinen J, Partinen M, Vaarala O, Kirjavainen T, Sundman J, Himanen SL, Hublin C, Julkunen I, Olsen P et al. 2012. AS03 adjuvanted AH1N1 vaccine associated with an abrupt increase in the incidence of childhood narcolepsy in Finland. *PLoS One* **7**: e33536.

Pedersen NW, Holm A, Kristensen NP, Bjerregaard AM, Bentzen AK, Marquard AM, Tamhane T, Burgdorf KS, Ullum H, Jennum P et al. 2019. CD8+ T cells from patients with narcolepsy and healthy controls recognize hypocretin neuron-specific antigens. *Nat Commun* **10**: 837.

Peyron C, Faraco J, Rogers W, Ripley B, Overeem S, Charnay Y, Nevsimalova S, Aldrich M, Reynolds D, Albin R et al. 2000. 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.

Quante T, Bird A. 2016. Do short, frequent DNA sequence motifs mould the epigenome? *Nat Rev Mol Cell Biol* **17**: 257-262.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139-140.

Roh EY, Park MH, Park H, Park DH, Choi JB, Kim SJ, Jeong DU. 2006. Association of HLA-DR and -DQ genes with narcolepsy in Koreans: comparison with two control groups, randomly selected subjects and DRB1*1501-DQB1*0602--positive subjects. *Hum Immunol* **67**: 749-755.

Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G et al. 2018. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell* **175**: 1701-1715.e1716.

Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H et al. 2015. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**: 212-216.

Shimada M, Miyagawa T, Kawashima M, Tanaka S, Honda Y, Honda M, Tokunaga K. 2010. An approach based on a genome-wide association study reveals candidate loci for narcolepsy. *Hum Genet* **128**: 433-441.

Shimada M, Miyagawa T, Takeshima A, Kakita A, Toyoda H, Niizato K, Oshima K, Tokunaga K, Honda M. 2020. Epigenome-wide association study of narcolepsy-affected lateral hypothalamic brains, and overlapping DNA methylation profiles between narcolepsy and multiple sclerosis. *Sleep* **43**.

Shimada M, Miyagawa T, Toyoda H, Tokunaga K, Honda M. 2018. Epigenome-wide association study of DNA methylation in narcolepsy: an integrated genetic and epigenetic approach. *Sleep* **41**.

Thannickal TC, Moore RY, Nienhuis R, Ramanathan L, Gulyani S, Aldrich M, Cornford M, Siegel JM. 2000. Reduced number of hypocretin neurons in human narcolepsy. *Neuron* **27**: 469-474.

Thiagalingam S. 2020. Epigenetic memory in development and disease: Unraveling the mechanism. *Biochim Biophys Acta Rev Cancer* **1873**: 188349.

Toyoda H, Miyagawa T, Koike A, Kanbayashi T, Imanishi A, Sagawa Y, Kotorii N, Kotorii T, Hashizume Y, Ogi K et al. 2015. A polymorphism in CCR1/CCR3 is associated with narcolepsy. *Brain Behav Immun* **49**: 148-155.

Tserel L, Kolde R, Limbach M, Tretyakov K, Kasela S, Kisand K, Saare M, Vilo J, Metspalu A, Milani L et al. 2015. Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. *Sci Rep* **5**: 13107.

van Daalen KR, Reijneveld JF, Bovenschen N. 2020. Modulation of Inflammation by Extracellular Granzyme A. *Front Immunol* **11**: 931.

Ventham NT, Kennedy NA, Adams AT, Kalla R, Heath S, O'Leary KR, Drummond H, Wilson DC, Gut IG, Nimmo ER et al. 2016. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat Commun* **7**: 13507.

Wang T, Liu Y, Yin Q, Geng J, Chen J, Yin X, Wang Y, Shang X, Tian C, Peng J. 2022. Enhancing discoveries of molecular QTL studies with small sample size using summary statistic imputation. *Brief Bioinform* **23**.

Winkelmann J, Lin L, Schormair B, Kornum BR, Faraco J, Plazzi G, Melberg A, Cornelio F, Urban AE, Pizza F et al. 2012. Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum Mol Genet* **21**: 2205-2210.

Yoshida-Tanaka K, Shimada M, Honda Y, Fujimoto A, Tokunaga K, Honda M, Miyagawa T. 2023. Narcolepsy type I-associated DNA methylation and gene expression changes in the human leukocyte antigen region. *Sci Rep* **13**: 10464.

Yoshie O. 2021. CCR4 as a Therapeutic Target for Cancer Immunotherapy. *Cancers (Basel)* **13**.

Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, Laird PW, Berman BP. 2018. DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat Genet* **50**: 591-602.