Table 2. Thirty Microsatellite Markers Showing Significant Association in the First and Second Screenings and in the Subsequent Individual Genotyping

		P					
		Pooled DNA[a]		95 Cases and 95 Controls		220 Cases and 420 Controls	
Marker	Chromosome	First	Second	2×2[b]	$2 \times m$[c]	2×2[b]	$2 \times m$[c]
D1S0500i	1q23.1	.046	.00024	.017	.28	...	...
D1S0583i	1q32.2	.038	.049	.018	.11	...	...
D1S1208i	1q32.2	.045	.045	.029	.034	.0010	.0027
D1S1148i	1q31.2	.025	.047	.018	.073	...	...
D2S303	2p14	.045	.017	.021	.19	...	...
D2S0878i	2p16.2	.0011	.025	.0023	.0020	.0010	.00073
D3S0502i	3p22.2	.019	.036	.037	.15	...	...
D3S0971i	3q22.2	.037	.016	.031	.23	...	...
D3S0978i	3q24	.0031	.0047	.0044	.041	.0081	.13
D3S1174i	3p14.1	.044	.027	.022	.0084	.0021	.0013
G08391	4p15.1	.041	.047	.044	.16	...	...
D4S0140i	4p15.1	.0027	.037	.0023	.0067	.0020	.0070
D4S0424i	4p16.3	.038	.036	.049	.63	...	...
D5S0022i	5q32	.012	.0084	.0040	.016	.011	.0017
D5S0565i	5p13.3	.0051	.013	.0031	.052	...	...
D7S0486i	7p15.2	.036	.0020	.049	.68	...	...
D7S0760i	7p11.2	.0044	.0040	.016	.086	...	...
D7S1066i	7q11.22	.0089	.028	.011	.095	...	...
D8S0068i	8	.018	.034	.011	.30	...	...
D8S0584i	8q24.21	.012	.012	.018	.11	...	...
D14S0284i	14q31.3	.0030	.049	.0058	.018	.027	.38
D15S150	15q21.3	.020	.049	.024	.21	...	...
D15S0157i	15q21.3	.019	.0088	.010	.16	...	...
D17S1300	17q24.3	.022	.0071	.036	.13	...	...
D17S0179i	17	.0044	.050	.018	.0090	.0016	.0060
D17S0306i	17p11.2	.0060	.021	.011	.014	.0069	.076
D20S0027i	20p11.21	.0055	.014	.024	.18	...	...
D21S0098i	21q21.1	.023	.0044	.00029	.0051	.00031	.011
D21S0241i	21q22.3	.018	.00043	.0098	.018	.0012	.048
DXS0660i	Xp22.13	.00061	.0080	.0025	.065	...	...

[a] *P* values calculated by Fisher's exact test, based on 2×2 contingency tables with estimated allele frequencies. The smallest *P* value was selected. The alleles that showed the smallest *P* values in the pooled DNA genotypings were reflected in the individual genotyping.

[b] *P* values calculated by Fisher's exact test, based on 2×2 contingency tables. The smallest *P* value was selected.

[c] *P* values calculated by Fisher's exact test, based on $2 \times m$ contingency tables.

microsatellite marker *D21S0012m* were in the same LD block, whereas *D21S0241i* was not. In addition, the *rs13046884* g allele, the *rs13048981* t allele, and the *D21S0012m* $(AC)_{10}$ allele were found to be in strong LD ($r^2 > 0.94$), and the estimated haplotype frequency was 4.5% in the controls.

The region around the three polymorphisms contains three predicted genes registered in the UCSC Genome Browser, each of which is supported by between two and five mRNAs or ESTs (fig. 2*C*). These predicted genes are on the reverse strand on chromosome 21q22.3, with positions as follows: *NLC1-A* 45234209–4523842, *NLC1-B* 45238709–45239923, and *NLC1-C* 45243550–45249070 (Genome Browser accession numbers BC036902, BC009635, and BC027456, respectively). According to the UCSC Genome Browser, *NLC1-A* produces two alternatively spliced transcripts encoding different protein isoforms; the position of the short isoform is 45235619–45238383. The functions of these predicted genes are currently unknown.

We performed further variation screening on the three genes by direct sequencing with 16 samples, and 26 polymorphisms were observed. Fourteen of the polymorphisms were new: eight SNPs in *NLC1-A*, two in *NLC1-B*, and four in *NLC1-C*. Next, these new polymorphisms were subjected to association analysis with 190 cases and 190 controls (fig. 2*C*). Four SNPs reached significance in the analysis, but none was stronger than *rs13048981* or *rs13046884*, indicating that these two SNPs, as well as *D21S0012m*, are associated primarily in this region.

Table 3. Association Analyses of *D21S0241i, D21S0012m, rs13048981,* and *rs13046884* with 370 Patients with Narcolepsy and 610 Unaffected Controls

Marker or SNP and Allele	No. (%) of Individuals		
	Patients with Narcolepsy (n = 370)	Control Individuals (n = 610)	DRB1*1501-Positive Control Individuals (n = 125)
D21S0241i:			
$(AAGG)_7$	0 (.0)	1 (.1)	...
$(AAGG)_8$	0 (.0)	1 (.1)	...
$(AAGG)_9$	1 (.1)	0 (.0)	...
$(AAGG)_{10}$[a]	51 (6.9)	42 (3.4)	...
$(AAGG)_{11}$	65 (8.8)	90 (7.4)	...
$(AAGG)_{12}$	155 (20.9)	258 (21.1)	...
$(AAGG)_{13}$	154 (20.8)	249 (20.4)	...
$(AAGG)_{14}$	195 (26.4)	359 (29.4)	...
$(AAGG)_{15}$	94 (12.7)	180 (14.8)	...
$(AAGG)_{16}$	24 (3.2)	35 (2.9)	...
$(AAGG)_{17}$	1 (.1)	5 (.4)	...
D21S0012m:			
$(AC)_8$	127 (17.2)	168 (13.8)	...
$(AC)_9$	447 (60.4)	762 (62.5)	...
$(AC)_{10}$[b]	12 (1.6)	54 (4.4)	...
$(AC)_{11}$	36 (4.9)	48 (3.9)	...
$(AC)_{12}$	117 (15.8)	188 (15.4)	...
$(AC)_{13}$	1 (.1)	0 (.0)	...
rs13048981:			
Genotype[c,d]:			
CC	360 (97.3)	561 (92.0)	115 (92.0)
CT	10 (2.7)	47 (7.7)	10 (8.0)
TT	0 (.0)	2 (.3)	0 (.0)
Allele[e,f]:			
C	730 (98.6)	1,169 (95.8)	240 (96.0)
T	10 (1.4)	51 (4.2)	10 (4.0)
rs13046884:			
Genotype[g,h]:			
AA	359 (97.0)	559 (91.6)	115 (92.0)
AG	11 (3.0)	49 (8.0)	10 (8.0)
GG	0 (.0)	2 (.3)	0 (.0)
Allele[i,j]:			
A	729 (98.5)	1,167 (95.7)	240 (96.0)
G	11 (1.5)	53 (4.3)	10 (4.0)

[a] OR 2.08; 95% CI 1.4–3.1; P = .00064.

[b] OR 0.36; 95% CI 0.2–0.7; P = .00068.

[c] P = .00095 (patients with narcolepsy compared with controls, by Fisher's exact test based on a 2 × 3 contingency table).

[d] P = .016 (patients with narcolepsy compared with DRB1*1501-positive controls, by Fisher's exact test based on a 2 × 3 contingency table).

[e] OR 0.31; 95% CI 0.16–0.60; P = .00039 (patients with narcolepsy compared with controls, by Fisher's exact test based on a 2 × 2 contingency table).

[f] OR 0.33; 95% CI 0.14–0.77; P = .017 (patients with narcolepsy compared with DRB1*1501-positive controls, by Fisher's exact test based on a 2 × 2 contingency table).

[g] P = .0011 (patients with narcolepsy compared with controls, by Fisher's exact test based on a 2 × 3 contingency table).

[h] P = .022 (patients with narcolepsy compared with DRB1*1501-positive controls, by Fisher's exact test based on a 2 × 3 contingency table).

[i] OR 0.33; 95% CI 0.18–0.62; P = .00036 (patients with narcolepsy compared with controls, by Fisher's exact test based on a 2 × 2 contingency table).

[j] OR 0.36; 95% CI 0.16–0.83; P = .023 (patients with narcolepsy compared with DRB1*1501-positive controls, by Fisher's exact test based on a 2 × 2 contingency table).

Expression Analyses

We assessed the expression of these three predicted genes in the human brain, hypothalamus, and other organs by RT-PCR, using specific primers (fig. 4*A*). Products with the expected size were amplified for *NLC1-A* and *NLC1-C* in whole brain and hypothalamus (fig. 4*B* and 4*C*). Moreover, direct sequencing of the products confirmed that the correct sequence was amplified. However, for *NLC1-B*, the amplified band was from genomic DNA, not from cDNA. These observations indicate that *NLC1-A* and *NLC1-C* were expressed in human whole brain and hypothalamus, whereas *NLC1-B* was not. Notably, *NLC1-A* was also expressed in human spleen, lung, kidney, and skeletal muscle, and *NLC1-C* was also expressed in human spleen, pancreas, lung, and sperm (fig. 4*C*), but neither was expressed in peripheral blood (data not shown). SNP *rs13046884* is in *NLC1-A* intron 1, and *D21S0012m* is 424 bp upstream of the transcriptional start site of *NLC1-A*, which suggests that *NLC1-A* may be a susceptibility/resistance gene for human narcolepsy. SNP *rs13048981* is located 2,602 bp upstream of *NLC1-B*, which was not expressed in human brain, and its position, 4,164 bp upstream of *NLC1-A*, suggests that its association with narcolepsy resulted merely from the LD with *rs13046884* and *D21S0012m*. Thus, it is unlikely that *NLC1-B* is a susceptibility/resistance gene for human narcolepsy.

Table 4. Association Analyses with High-Density Microsatellite Markers around Marker *D21S0241i*

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

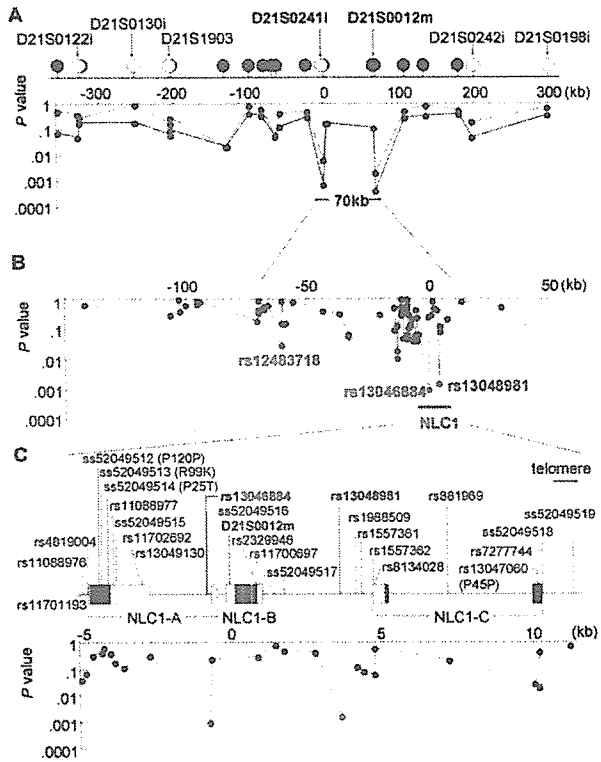

Figure 2. High-density mapping with additional microsatellite markers and SNPs. *A*, Association analyses using high-density microsatellite markers with 220 cases and 420 controls. Unblackened circles indicate microsatellite markers used in the first and second screenings. Blackened circles indicate microsatellite markers newly developed for the high-density mapping. The dark line shows the *P* values calculated by Fisher's exact test based on 2 × 2 contingency tables, whereas the lighter line shows those from 2 × *m* contingency tables. *B*, Association analyses using SNPs with 190 cases and 190 controls. The *X*-axis indicates the distance from *D21S0012m*. SNP *rs12483718* is located near *D21S0241i*. The *Y*-axis shows the *P* values calculated by Fisher's exact test based on 2 × 2 contingency tables. Two SNPs, *rs13048981* and *rs13046884*, showed the strongest associations in the NLC1 region. *C*, Variation screening and high-density association analyses in the NLC1 region. *Top*, Exon-intron structures of *NLC1-A*, *NLC1-B*, and *NLC1-C*. Boxes indicate exons, with unblackened boxes indicating untranslated regions and blackened boxes indicating coding regions. Predicted gene regions and 1 kb of upstream region were screened for sequence variations. Fourteen additional polymorphisms, including two nonsynonymous substitutions, were detected and were examined for possible associations, but no polymorphisms showed stronger association than *D21S0012m*, *rs13048981*, and *rs13046884*. *Bottom*, *P* values for individual SNPs.

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure 3. LD block structure. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics.*

To test whether polymorphisms *rs13046884* and *D21S0012m* directly influence the transcription level of *NLC1-A,* we performed reporter-gene assays. Six constructs carrying different alleles of *D21S0012m* or *rs13046884* were prepared from individuals with *D21S0012m* $(CA)_8$, $(CA)_9$, $(CA)_{10}$, and $(CA)_{12}$ repeats or the *rs13046884* a/g genotype. These constructs were introduced into NB-1 or HeLa cells, and the expression of luciferase was examined in three independent experiments. The differences of transcriptional activity were assessed by *t* test. The luciferase activities of each construct were divided by the ones of empty vector. These values were used for the *t* test. Figure 5 shows that the luciferase activities of reporters carrying the resistance alleles (g allele of *rs13046884* and $[AC]_{10}$ allele of *D21S0012m*) were 1.5- to 2-fold lower than those of other reporters in both NB-1 and HeLa cells, and the differences assessed by *t* test reached statistical significance (for NB-1 cell, $t = 2.4–6.7$ and $P = .039–.0010$; for HeLa cell, $t = 6.9–74.7$ and $P = .0034–.000000096$). Thus, the promoter activity of *NLC1-A* is likely to be reduced in individuals who possess the haplotype *D21S0012m* $(AC)_{10}$–*rs13046884* g.

Discussion

We have systematically performed the first genomewide association analyses, to our knowledge, for detecting susceptibility or resistance genes to human narcolepsy, using 23,244 microsatellite markers. After two separate screenings with pooled DNA samples, followed by individual genotyping with 95 case and 95 control samples of 80 initial candidate markers located outside chromosome 6, 30 microsatellite markers remained as candidates for association with narcolepsy. Among them, one marker (*D21S0241i*) was further analyzed with a third set of cases and controls, to confirm the association. Although the difference between cases and controls in the third set did not reach statistical significance, the allele frequencies were similar to those in the first and second sets. Moreover, a significant association was detected in an analysis of all the available samples (370 cases and 610 controls). In an analysis of the region surrounding *D21S0241i,* one microsatellite marker (*D21S0012m*) and eight nearby SNPs, all located ~70 kb from *D21S0241i,* were significantly associated with narcolepsy. *D21S0012m* and two of the SNPs were the markers most strongly associated with narcolepsy (all $P < .0005$); these three polymorphisms are in strong LD. The genomic region including these three polymorphisms is, therefore, a candidate region for human narcolepsy, which we tentatively designated "NLC1." For each of the three strongly associated polymorphisms, a minor allele displayed significantly reduced frequency in patients with narcolepsy compared with controls (OR 0.19–0.33), which suggests that these alleles confer resistance to narcolepsy.

NLC1 is located on 21q22.3, 2.6 Mb away from a locus recently reported as a candidate region for French familial narcolepsy.[12] According to the SNP genotype data of 45 unrelated Japanese living in the Tokyo area registered in the HapMap project database, there is no LD between NLC1 and the region reported in the French family study. Therefore, the association of NLC1 with human narcolepsy is considered a novel observation.

The NLC1 region contains no known genes, but data-

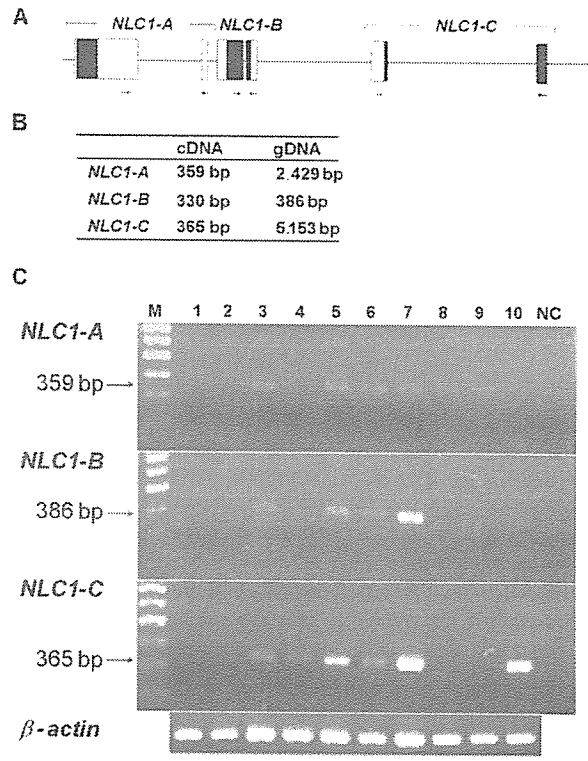

Figure 4. Expression analysis for *NLC1-A, NCL1-B,* and *NCL1-C,* with the use of RT-PCR. *A,* Schematic drawing of the specific primers for RT-PCR. *B,* Expected size of RT-PCR products from cDNA or genomic DNA. On the basis of the UCSC Genome Browser, products with the expected size were amplified from cDNA for *NLC1-A* and *NLC1-C* in samples of whole brain, hypothalamus, and several other organs, but, for *NLC1-B,* only the products from genomic DNA were observed (*C*). Amplified products were confirmed by direct sequencing. *Lane 1,* Heart; *lane 2,* liver; *lane 3,* spleen; *lane 4,* pancreas; *lane 5,* lung; *lane 6,* whole brain; *lane 7,* hypothalamus; *lane 8,* kidney; *lane 9,* skeletal muscle; *lane 10,* sperm. NC = negative control. M = 100-bp ladder size marker.

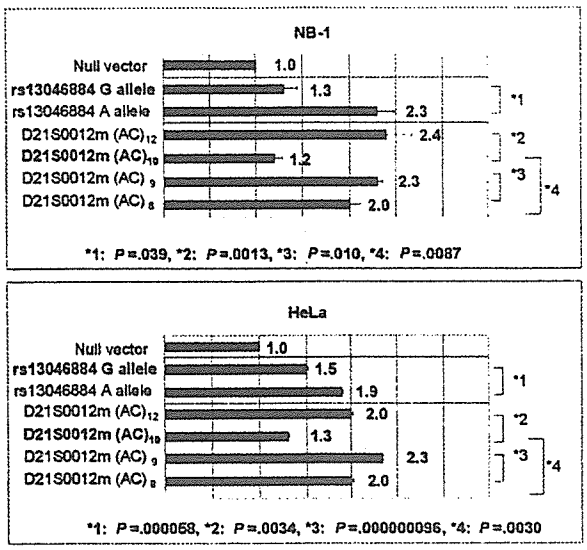

Figure 5. Effects of the microsatellite marker *D21S0012m* in the promoter region and of SNP *rs13046884* in the intron 1 of *NLC1-A* on transcriptional activity. Reporter-gene constructs contained the sequences from IVS1+31 to IVS1+327 for *rs13046884* or 80–987 nt upstream of the transcription initiation site for *D21S0012m*. The chart shows luciferase expression from each reporter in transfected HeLa cells or NB-1 cells, relative to empty vector. Data are means of at least three independent experiments. Error bars represent SDs.

bases show three predicted genes, which we tentatively named "*NLC1-A*," "*NLC1-B*," and "*NLC1-C*." Because of the locations of the three most strongly associated polymorphisms (*D21S0012m* in intron 1 of *NLC1-A*, *rs13046884* 424 bp upstream of *NLC1-A* and in the 3′ UTR of *NLC1-B*, and *rs13048981* 2,602 bp upstream of *NLC1-B*), we focused on *NLC1-A* and *NLC1-B*. In RT-PCR analysis, *NLC1-A*, but not *NLC1-B*, was expressed in human hypothalamus, which also expresses preprohypocretin,[42] a protein important in orchestrating the sleep-wake cycle.[43] Therefore, we finally focused on *NLC1-A*, and we tested whether the *D21S0012m* and *rs13046884* polymorphisms affect gene expression. In a reporter-gene assay, *NLC1-A* fragments containing the alleles for narcolepsy resistance (*D21S0012m* $[CA]_{10}$ allele and *rs13046884* g allele) were less transcriptionally active than were those of other alleles. This finding supports the hypothesis that the polymorphisms of *NLC1-A* may be directly involved in resistance to human narcolepsy.

A motif search of the putative NLC1-A protein, with the use of MOTIF (GenomeNet) and Motif-Finder (RIKEN), revealed a domain known as "binding-protein–dependent transport systems inner membrane component." Binding-protein–dependent transport systems have been characterized as members of a superfamily of transporters found not only in bacteria but also in humans, and they include both import and export systems.[44] Therefore, NLC1-A might function as a transporter of certain substances (amino acids, sugars, large polysaccharides, or proteins). A motif search of the cDNA sequence of *NLC1-A* was also performed using MOTIF and Motif-Finder, and NLC1-A includes domains known as integrin β-chain cysteine-rich domain, anaphylatoxin domain, and epidermal growth factor–1 domain signatures. Furthermore, the amino acid sequence of *NLC1-A* was subjected to secondary structure prediction (SOSUI program). NLC1-A has a long loop (residues 78–125) with high hydrophilicity, flexibility, and surface probability, which suggests that NLC1-A may be a membrane protein. No carbohydrate-modification region was predicted. The UCSC Genome Browser showed a chimpanzee gene with 98% sequence identity to *NLC1-A*. In contrast, there was no homologous gene in rodent or canine genomes. Thus, *NLC1-A* is likely to exist only in primates.

Recently, genomewide association analysis with hundreds of thousands of SNPs has become realistic, but such a systematic product was not available when we started the present study. Therefore, we took a unique approach—genomewide association analyses with highly polymorphic microsatellite markers that were selected every ~100 kb throughout the human genome.[33] Because pooled DNAs were used in the first and second screenings, the typing cost was reasonable, even when 23,244 markers were used.

Because human narcolepsy is a multifactorial disorder for which the relative risks of individual associated genes may not be particularly high, we hypothesize that several more susceptibility/resistance genes remain to be elucidated. Thirty microsatellite markers displayed association with human narcolepsy in both first and second screenings. The observed associations of the microsatellite markers were not strong, and the markers were similar to each other in the strength of association. Therefore, the remaining 29 uncharacterized regions may include other susceptibility/resistance loci for narcolepsy. Some false-positive results may still survive after both screenings with the use of pooled DNA samples, but most of them can be excluded in subsequent high-density mapping and association analysis with additional cases and controls. An association study with an entirely separate set of cases and controls or replication studies in other populations and transmission disequilibrium test may be preferred to completely eliminate false-positive associations, although the detection power is decreased because additional association studies lead to an increase in false-negative associations.

In conclusion, a genomewide association study with the use of a dense set of microsatellite markers and pooled DNA can be useful for the systematic search for candidate regions of multifactorial disorders—such as human narcolepsy, rheumatoid arthritis (RA [MIM 180300]), type II diabetes (NIDDM [MIM 125853]), hypertension (MIM 145500), psoriasis (MIM 177900), and schizophrenia

(SCZD [MIM 181500])—for which pathophysiological mechanisms remain unclear. We were able to detect 30 candidate microsatellite markers, among which one narcolepsy resistance gene, *NLC1-A*, was identified successfully. Functional analyses of *NLC1-A* are in progress, and the remaining 29 candidate markers will be further analyzed.

Acknowledgments

We sincerely thank the patients with narcolepsy who participated in this study. We are indebted to Dr. Jun Ohashi (Department of Human Genetics, Graduate School of Medicine, University of Tokyo), for helpful discussion on statistical analysis, and to Drs. Tomoki Ikuta, Minori Shinya, and Satoshi Makino (Department of Molecular Life Science, Tokai University School of Medicine), for technical help in the genomewide screening. This work was supported by a Grant-in-Aid for Science Research on Priority Areas (Comprehensive Genomics) from the Japanese Ministry of Education, Culture, Sports, Science, and Technology and by a grant from the Kato Memorial Trust for Nambyo Research.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Celera database, http://www.celera.com/
dbSNP, http://www.ncbi.nlm.nih.gov/SNP/
GOLD program, http://www.sph.umich.edu/csg/abecasis/GOLD/
HapMap, http://www.hapmap.org/
MOTIF, http://motif.genome.ad.jp/
Motif-Finder, http://gibk26.bse.kyutech.ac.jp/jouhou/HOMOLOGY/dbsearch/pdb/pdb_seq.html
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for narcolepsy, *HLA-DRB1*, *HLA-DQB1*, *TNFA*, *TNFR2*, *HCRTR2*, prepohypocretin, RA, NIDDM, hypertension, psoriasis, and SCZD)
RepeatMasker program, http://www.repeatmasker.org/
SOSUI program, http://sosui.proteome.bio.tuat.ac.jp/sosuiframe0.html
UCSC Genome Browser (November 2002 version, based on NCBI Build 31), http://genome.ucsc.edu/ (for *NLC1-A* [accession number BC036902], *NLC1-B* [accession number BC009635], and *NLC1-C* [accession number BC027456])

References

1. Honda Y, Matsuki K (1979) Census of narcolepsy, cataplexy and sleep life among teen-agers in Fujisawa city. Sleep Res 8: 191
2. Mignot E (1998) Genetic and familial aspects of narcolepsy. Neurology 50:S16–S22
3. Juji T, Satake M, Honda Y, Doi Y (1984) HLA antigens in Japanese patients with narcolepsy: all the patients were DR2 positive. Tissue Antigens 24:316–319
4. Matsuki K, Juji T, Tokunaga K, Naohara T, Satake M, Honda Y (1985) 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
5. Mignot E, Lin L, Rogers W, Honda Y, Qiu X, Lin X, Okun M, Hohjoh H, Miki T, Hsu S, Leffell M, Grumet F, Fernandez-Vina M, Honda M, Risch N (2001) Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. Am J Hum Genet 68:686–699
6. Vyse TJ, Todd JA (1996) Genetic analysis of autoimmune disease. Cell 85:311–318
7. Hohjoh H, Terada N, Honda Y, Juji T, Tokunaga K (2001) Negative association of the *HLA-DRB1*1502-DQB1*0601* haplotype with human narcolepsy. Immunogenetics 52:299–301
8. Ohashi J, Yamamoto S, Tsuchiya N, Hatta Y, Komata T, Matsushita M, Tokunaga K (2001) Comparison of statistical power between 2 × 2 allele frequency and allele positivity tables in case-control studies of complex disease genes. Ann Hum Genet 65:197–206
9. James JW (1971) Frequency in relatives for an all-or-none trait. Ann Hum Genet 35:47–49
10. Risch N (1987) Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet 40:1–14
11. Nakayama J, Miura M, Honda M, Miki T, Honda Y, Arinami T (2000) Linkage of human narcolepsy with HLA association to chromosome 4p13-q21. Genomics 65:84–86
12. Dauvilliers Y, Blouin JL, Neidhart E, Carlander B, Eliaou JF, Antonarakis SE, Billiard M, Tafti M (2004) A narcolepsy susceptibility locus maps to a 5 Mb region of chromosome 21q. Ann Neurol 56:382–388
13. Wieczorek S, Jagiello P, Arning L, Dahmen N, Epplen JT (2004) Screening for candidate gene regions in narcolepsy using a microsatellite based approach and pooled DNA. J Mol Med 82:696–705
14. Aldrich MS (1990) Narcolepsy. N Engl J Med 323:389–394
15. Chabas D, Taheri S, Renier C, Mignot E (2003) The genetics of narcolepsy. Annu Rev Genomics Hum Genet 4:459–483
16. Hohjoh H, Nakayama T, Ohashi J, Miyagawa T, Tanaka H, Akaza T, Honda Y, Juji T, Tokunaga K (1999) Significant association of a single nucleotide polymorphism in the tumor necrosis factor-alpha (*TNF-α*) gene promoter with human narcolepsy. Tissue Antigens 54:138–145
17. Hohjoh H, Terada N, Kawashima M, Honda Y, Tokunaga K (2000) Significant association of the tumor necrosis factor receptor 2 (*TNFR2*) gene with human narcolepsy. Tissue Antigens 56:446–448
18. Kawashima M, Hohjoh H, Terada N, Komata T, Honda Y, Tokunaga K (2001) Association studies of the tumor necrosis factor-alpha (TNFA) and its receptor 1 (TNFR1) and 2 (TNFR2) genes with human narcolepsy. Korean J Genet 23:365–370
19. Wieczorek S, Dahmen N, Jagiello P, Epplen JT, Gencik M (2003) Polymorphisms of the tumor necrosis factor receptors: no association with narcolepsy in German patients. J Mol Med 81:87–90
20. Wieczorek S, Gencik M, Rujescu D, Tonn P, Giegling I, Epplen JT, Dahmen N (2003) TNFA promoter polymorphisms and narcolepsy. Tissue Antigens 61:437–442
21. Lin L, Faraco J, Li R, Kadotani H, Rogers W, Lin X, Qiu X, de Jong PJ, Nishino S, Mignot E (1999) The sleep disorder canine narcolepsy is caused by a mutation in the *hypocretin (orexin) receptor 2* gene. Cell 98:365–376
22. Chemelli RM, Willie JT, Sinton CM, Elmquist JK, Scammell T, Lee C, Richardson JA, Williams SC, Xiong Y, Kisanuki Y, Fitch TE, Nakazato M, Hammer RE, Saper CB, Yanagisawa M (1999) Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. Cell 98:437–451
23. Nishino S, Ripley B, Overeem S, Lammers GJ, Mignot E (2000)

Hypocretin (orexin) deficiency in human narcolepsy. Lancet 355:39–40
24. Peyron C, Faraco J, Rogers W, Ripley B, Overeem S, Charnay Y, Nevsimalova S, Aldrich M, Reynolds D, Albin R, Li R, Hungs M, Pedrazzoli M, Padigaru M, Kucherlapati M, Fan J, Maki R, Lammers GJ, Bouras C, Kucherlapati R, Nishino S, Mignot E (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
25. 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
26. Sakurai T, Moriguchi T, Furuya K, Kajiwara N, Nakamura T, Yanagisawa M, Goto K (1999) Structure and function of human prepro-orexin gene. J Biol Chem 274:17771–17776
27. Hungs M, Lin L, Okun M, Mignot, E (2001) Polymorphisms in the vicinity of the hypocretin/orexin are not associated with human narcolepsy. Neurology 57:1893–1895
28. Olafsdottir BR, Rye DB, Scammell TE, Matheson JK, Stefansson K, Gulcher JR (2001) Polymorphisms in hypocretin/orexin pathway genes and narcolepsy. Neurology 57:1896–1899
29. Smith AJ, Jackson MW, Neufing P, McEvoy RD, Gordon TP (2004) A functional autoantibody in narcolepsy. Lancet 364: 2122–2124
30. Black JL 3rd, Silber MH, Krahn LE, Avula RK, Walker DL, Pankratz VS, Fredrickson PA, Slocumb NL (2005) Studies of humoral immunity to preprohypocretin in human leukocyte antigen DQB1*0602-positive narcoleptic subjects with cataplexy. Sleep 28:1191–1192
31. Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517
32. Ohashi J, Tokunaga K (2001) The power of genome-wide association studies of complex disease genes: statistical limitations of indirect approaches using SNP markers. J Hum Genet 46:478–482
33. Tamiya G, Shinya M, Imanishi T, Ikuta T, Makino S, Okamoto K, Furugaki K, et al (2005) Whole genome association study of rheumatoid arthritis using 27,039 microsatellites. Hum Mol Gen 14:2305–2321
34. Ohashi J, Tokunaga K (2003) Power of genome-wide linkage disequilibrium testing by using microsatellite markers. J Hum Genet 48:487–491
35. Barcellos LF, Klitz W, Field LL, Tobias R, Bowcock AM, Wilson R, Nelson MP, Nagatomi J, Thomson G (1997) Association mapping of disease loci, by use of a pooled DNA genomic screen. Am J Hum Genet 61:734–747
36. Kawashima M, Ikuta T, Tamiya G, Hohjoh H, Honda Y, Juji T, Tokunaga K, Inoko H (2004) Genome-wide association study of narcolepsy: initial screening on chromosome 6. In: Dupont J, Hansen JA (eds) HLA 2004: immunobiology of the human MHC. Proceedings of the 13th International Histocompatibility Workshop and Conference. IHWG Press, Seattle
37. Collins HE, Li H, Inda SE, Anderson J, Laiho K, Tuomilehto J, Seldin MF (2000) A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. Hum Genet 106:218–226
38. Kirov G, Williams N, Sham P, Craddock N Owen MJ (2000) Pooled genotyping of microsatellite markers in parent-offspring trios. Genome Res 10:105–115
39. Lau J, Ioannidis JP, Schmid CH (1997) Quantitative synthesis in systematic reviews. Ann Intern Med 127:820–826
40. Lewontin RC (1964) The interaction of selection and linkage. II. Optimum models. Genetics 50:757–782
41. Abecasis GR, Cookson WO (2000) GOLD: graphical overview of linkage disequilibrium. Bioinformatics 16:182–183
42. Sakurai T, Amemiya A, Ishii M, Matsuzaki I, Chemelli RM, Tanaka H, Williams SC, Richardson JA, Kozlowski GP, Wilson S, Arch JRS, Buckingham RE, Haynes AC, Carr SA, Annan RS, McNulty DE, Liu W-S, Terrett JA, Elshourbagy NA, Bergsma DJ, Yanagisawa M (1998) Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. Cell 92:573–585
43. Hagan JJ, Leslie RA, Patel S, Evans ML, Wattam TA, Holmes S, Benham CD, Taylor SG, Routledge C, Hemmati P, Munton RP, Ashmeade TE, Shah AS, Hatcher JP, Hatcher PD, Jones DN, Smith MI, Piper DC, Hunter AJ, Porter RA, Upton N (1999) Orexin A activates locus coeruleus cell firing and increases arousal in the rat. Proc Nat Acad Sci USA 96:10911–10916
44. Higgins CF, Hyde SC, Mimmack MM, Gileadi U, Gill DR, Gallagher MP (1990) Binding protein-dependent transport systems. J Bioenerg Biomembr 22:571

Tissue Antigens ISSN 0001-2815

Estimation of the species-specific mutation rates at the *DRB1* locus in humans and chimpanzee

J. Ohashi[1,*], I. Naka[1], A. Toyoda[2], M. Takasu[1], K. Tokunaga[1], T. Ishida[3], Y. Sakaki[2] & H. Hohjoh[4,*]

1 Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
2 Sequence Technology Team, RIKEN Genomic Science Center, RIKEN Yokohama Institute, Yokohama, Japan
3 Department of Biological Science, Division of Evolutionary Life Systems, Graduate School of Science, The University of Tokyo, Tokyo, Japan
4 National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan

Key words
chimpanzee; *DRB1*; humans; nucleotide substitution rate; mutation rate

Correspondence
Jun Ohashi, PhD
Department of Human Genetics
Graduate School of Medicine,
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113-0033
Japan
Tel: +81-3-5841-3693
Fax: +81-3-5802-8619
e-mail: juno-tky@umin.ac.jp
or
Hirohiko Hohjoh, PhD
National Institute of Neuroscience
NCNP
4-1-1 Ogawahigashi, Kodaira
Tokyo 187-8502
Japan
Tel: +81-42-341-2711
Fax: +81-42-346-1755
e-mail: hohjohh@ncmp.go.jp

Received 19 August 2006; accepted 24 August 2006

doi: 10.1111/j.1399-0039.2006.00688.x

Abstract

To estimate the species-specific mutation rates at the *DRB1* locus in humans and chimpanzee, we analyzed the nucleotide sequence of a 37.6-kb chimpanzee chromosomal segment containing the entire *Patr-DRB1*0701* allele and the flanking nongenic region and we compared it with two corresponding human sequences containing the *HLA-DRB1*070101* allele using the sequence of *HLA-DRB1*04011* as an outgroup. Because the allelic pair of *HLA-DRB1*070101* and *Patr-DRB1*0701* shows the lowest number of substitutions between the two species, it appears that these sequences diverged close to the time of the humans–chimpanzee divergence (6 million years ago). Alignment of the nucleotide sequences for *HLA-DRB1*070101* and *Patr-DRB1*0701* alleles showed that they share a high degree of similarity, suggesting that the studied chromosomal segments with these sequences have not been subjected to recombination since the humans–chimpanzee divergence. Comparison of the flanking 10.6 kb of nongenic sequences revealed an average of 41.5 and 83 single nucleotide substitutions in humans and chimpanzee, respectively. Thus, the species-specific nucleotide substitution rates in the flanking nongenic region were estimated to be 6.53×10^{-10} and 1.31×10^{-9} per site per year in humans and chimpanzee, respectively. Unexpectedly, the estimated rate in humans was twofold lower than in chimpanzee ($P < 10^{-3}$, Tajima's relative rate test) and lower than the average substitution rate in the human genome. Because the nucleotide substitution rate in nongenic regions free from selection is expected to be equal to the mutation rate, the estimated substitution rate should correspond to the species-specific mutation rate at the *DRB1* locus. Our results strongly suggest that the mutation rate at *DRB1* locus differs among species.

Introduction

A large number of alleles (>400) have been found at the major histocompatibility complex (MHC) class II *DRB1* locus in humans. This high degree of polymorphism is considered to be due to strong balancing selection such as overdominant selection (1, 2) and frequency-dependent selection (2–4), while the high allelic diversity may have been achieved partly by a high mutation rate. Although it is difficult to estimate the mutation rate directly, it can be inferred from the substitution rate (k), which is calculated from the nucleotide difference (n) between two sequences of different species whose divergence time (t) is known (i.e. $k = n/2t$). Specifically, the divergence time of the two sequences is assumed to be equal to t in this case; however, it is not easy to use this formula to estimate the substitution rate at *DRB1* locus. Because the divergence of most allelic lineages

*These authors contributed equally to this work.

predates the humans–chimpanzee divergence, the species divergence time cannot be used as the divergence time for two randomly selected *DRB1* sequences from humans and chimpanzee.

To overcome this problem, we applied the minimum-minimum method proposed by Satta and co-workers (5, 6), which compares the most closely related sequences from two different species. The human-specific substitution rate in the *HLA-DRB1* region can be assessed only using the minimum-minimum method to compare the two sequences with the smallest difference between humans and chimpanzee along with an outgroup sequence. Among the humans and chimpanzee *DRB1* alleles, the allelic pair of *HLA-DRB1*0701* and *Patr-DRB1*0701* is one of the most similar pairs (7–9). Thus, these alleles appear to have diverged close to the time of the humans–chimpanzee divergence, so that the above formula can be used to estimate the substitution rate.

In this study, the nucleotide sequences of the genomic region containing the entire *HLA-DRB1*070101* and *Patr-DRB1*0701* alleles were compared using the sequence of *HLA-DRB1*04011* as an outgroup. *HLA-DRB1*04011* and *HLA-DRB1*070101* alleles belong to the *DR53* haplotype group. Although the substitution rate at the *DRB1* locus has been analyzed based on the number of the synonymous substitutions (5, 6, 10), the synonymous substitution rate may be different from the actual mutation rate because synonymous sites are known to be subjected to weak purifying selection. Thus, we considered that the flanking nongenic region was more suitable for estimating the mutation rate at *DRB1* unless recombination has occurred in the studied chromosomal segments to be compared since the divergence of humans and chimpanzee.

Materials and methods

To estimate the species-specific mutation rates at the *DRB1* locus in humans and chimpanzee, we analyzed the nucleotide sequence of a 37.6-kb DNA of the chimpanzee chromosomal region containing the entire *Patr-DRB1*0701* allele and the flanking nongenic region. The 37.6-kb DNA fragment detected in our previous study (11) was cloned using the pWE15 cosmid vector (Stratagene, Cedar Creek, TX, USA), and the clone was sequenced according to the methods previously described (12, 13). The GenBank accession number for the analyzed sequence is AP006503. The corresponding human genomic sequences containing the entire *HLA-DRB1* locus and the flanking regions were obtained from GenBank under accession numbers CR753835 (*HLA-DRB1*070101*), CR753309 (*HLA-DRB1*070101*), and AL137064 (*HLA-DRB1*04011*).

The four sequences were first aligned by VISTA (14) after the repetitive sequences were masked using RepeatMasker (AFA Smit, R Hubley and P Green; RepeatMasker at http://repeatmasker.org). Next, after excluding the masked repetitive sequences and the flanking sites to avoid inclusion of the misaligned nucleotides as point mutations, multiple alignments were performed manually.

To evaluate the possibility of recombination between the studied chromosomal segments, we further calculated the proportions of nucleotide difference between *HLA-DRB1*070101* and *HLA-DRB1*04011* (denoted by π_h) and between *Patr-DRB1*0701* and *HLA-DRB1*04011* (denoted by π_c) using SNPs-GRAPHIC (available at http://bioinformatica.uab.es/dpdb/diversity.asp), where window size was set to 200 bp and step size was 50 bp.

To examine whether the substitution rate is different between humans and chimpanzee, Tajima's relative rate test (15, 16) was performed using MEGA version 3.1 (17) based on both transitions and transversions.

Results and discussion

Aligned sequences with the same lengths were visualized by VISTA to compare the chimpanzee sequence containing the *Patr-DRB1*0701* allele with the human sequences containing *HLA-DRB1*070101* (CR753835 and CR753309) and *HLA-DRB1*04011* (AL137064) alleles (Figure 1). VISTA plots (14) showed that *Patr-DRB1*0701* is more similar to *HLA-DRB1*070101* (CR753835 and CR753309) than to *HLA-DRB1*04011* (AL137064). Because the similarity

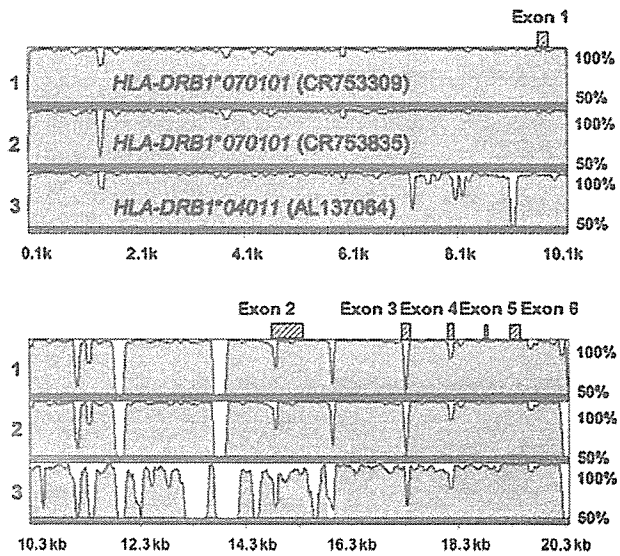

Figure 1 VISTA plots showing the alignments between the chimpanzee sequence containing *Patr-DRB1*0701* and the two human sequences containing *HLA-DRB1*070101* (CR753835 and CR753309), with the sequence *HLA-DRB1*04011* (AL137064) as an outgroup. The sequence conservation (per cent nucleotide identity) relative to *Patr-DRB1*0701* allele was evaluated in 100-bp stretches. For each plot, the lowest mapped score is 50% and the maximum is 100%. The exons of *DRB1* are indicated by shaded boxes above the plots. It should be noted that *DRB1* is the only locus in the genomic region studied here.

between the sequences containing *HLA-DRB1*070101* and *Patr-DRB1*0701* alleles was the same as that between the *DRB1* locus and the flanking nongenic region, it appears that the nongenic region also diverged close to the time of humans–chimpanzee divergence.

Figure 2 shows the difference in the proportion of nucleotide difference, $\pi_c - \pi_h$. If recombination has occurred between the studied chromosomal segments containing *HLA-DRB1*070101* and *HLA-DRB1*04011* alleles since the divergence of *HLA-DRB1*070101* and *Patr-DRB1*0701* allele, a long sequential region with positive $\pi_c - \pi_h$ values would be observed. However, no such region is shown in Figure 2. Taken together with the similarity of sequences between *HLA-DRB1*070101* and *Patr-DRB1*0701* alleles observed for the entire region in Figure 1, we can say that recombination has not occurred between the studied chromosomal segments containing *HLA-DRB1*070101* and *HLA-DRB1*04011* since the divergence of *HLA-DRB1*070101* and *Patr-DRB1*0701* alleles.

To estimate the species-specific substitution rates of the *DRB1* locus, using *HLA-DRB1*04011* as an outgroup, we identified the sequence-specific nucleotide differences between the *HLA-DRB1*070101* (CR753835 and CR753309) and the *Patr-DRB1*0701* alleles (Figure 3). Because parallel mutation is unlikely to occur, the number of unique nucleotide differences can be regarded as the number of substitutions that occurred in the sequence. Of 18,806 bp, we observed 68 ([66 + 70]/2) and 128 single nucleotide substitutions specific to the *HLA-DRB1*070101* and *Patr-DRB1*0701* sequences, respectively. Here, only the regions showing a high similarity for the three sequences were used to estimate the nucleotide substitution rates, which allowed us to consider only point mutations occurred after the sequence divergence. Assuming that this allelic pair diverged at the time of the humans–chimpanzee divergence

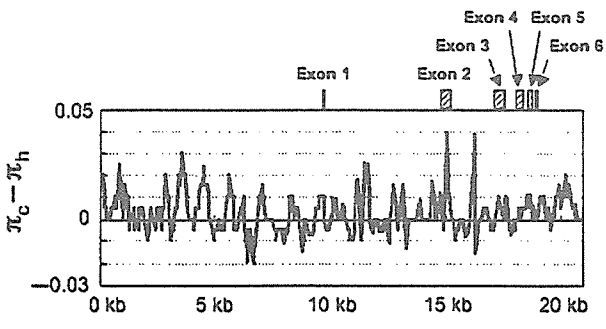

Figure 2 Plots of the difference in the proportion of nucleotide difference, $\pi_c - \pi_h$. The proportions of nucleotide difference between *Patr-DRB1*070143* and *HLA-DRB1*04011* and between *HLA-DRB1*070101* and *HLA-DRB1*04011* alleles are denoted by π_c and π_h, respectively. Window size is 200 bp, and step size is 50 bp. Alignment gaps are excluded from the analyses. The exons of *DRB1* are indicated by shaded boxes above the plots.

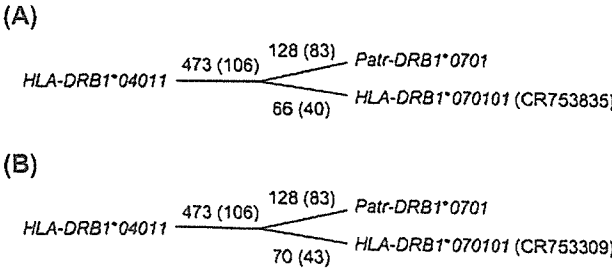

Figure 3 The number of unique nucleotide differences in each sequence (A) from the comparison of *Patr-DRB1*0701*, *HLA-DRB1*070101* (CR753835), and *HLA-DRB1*04011* (AL137064) (B) and from the comparison of *Patr-DRB1*0701*, *HLA-DRB1*070101* (CR753309), and *HLA-DRB1*04011* (AL137064) alleles. The number of unique nucleotide differences in the flanking region of the *DRB1* locus is given in parentheses.

(6 million years ago), the species-specific nucleotide substitution rates in this region are estimated to be 6.03×10^{-10} and 1.13×10^{-9} per site per year for humans and chimpanzee, respectively. We performed the same analyses for the flanking nongenic region. Comparison of *HLA-DRB1*070101* (CR753835) and *Patr-DRB1*0701* showed 40 unique single nucleotide substitutions of 10,595 bp, and comparison of *HLA-DRB1*070101* (CR753309) and *Patr-DRB1*0701* showed 43 single nucleotide substitutions of 10,590 bp (Figure 3). Thus, the average specific substitution rates in the flanking nongenic region were 6.53×10^{-10} and 1.31×10^{-9} per site per year for humans and chimpanzee, respectively. Satta et al. (5) estimated the synonymous substitution rate at the *DRB1* locus (1.18×10^{-9} per site per year) using the minimum-minimum method for the synonymous substitutions. The estimated synonymous substitution rate is close to the species-specific nucleotide substitution rate for the flanking nongenic region in chimpanzee (1.31×10^{-9} per site per year). Although the estimated rates are largely dependent on the assumed divergence time between two species to be compared, we may say that the synonymous sites of the *DRB1* locus are not subjected to strong purifying selection. The nucleotide difference between humans and chimpanzee of 1.23% (18, 19) corresponds to the average substitution rate of 1.03×10^{-9} per site per year. Thus, the estimated species-specific substitution rates in the entire (6.03×10^{-10} per site per year) and flanking nongenic regions (6.53×10^{-10} per site per year) in humans are much lower than the average for the entire genome.

Mutation is the ultimate source of allelic diversity at the *HLA-DRB1* locus, whereas the rate of mutation is not fully understood. According to the neutral theory of molecular evolution (20), the nucleotide substitution rate in a nongenic region free from selection (e.g., positive diversifying selection, balancing selection, and purifying selection) is expected to be equal to the mutation rate. Therefore, the

mutation rate at the nongenic region flanking the *DRB1* locus can be regarded as a mutation rate at the *DRB1* locus. Because the mutation rate at the *DRB1* locus is unlikely to be markedly different from that in the flanking regions, we conclude that the mutation rates in the *HLA-DRB1* region in humans is approximately 6.53×10^{-10} per site per year. This low mutation rate implies that a large number of alleles observed at the *HLA-DRB1* locus have been maintained not by frequent mutation but rather by strong balancing selection such as overdominant selection (1, 2) and frequency-dependent selection (2–4). In fact, the selection coefficient of *HLA-DRB1* has been estimated to be 0.019 under the assumption of symmetric overdominant selection, which is the second highest of seven *HLA* loci examined (21).

We observed a remarkable difference in nucleotide substitution rate or mutation rate at *DRB1* region between humans and chimpanzee. This observation does not come from recombination. The differences in the substitution rate between humans and chimpanzee for the entire and the flanking nongenic regions were highly significant according to Tajima's relative rate test ($P < 10^{-3}$ for both regions). Of particular interest, the estimated nucleotide substitution rate at the *DRB1* locus and in the flanking region was approximately twofold lower in humans than in chimpanzee.

The difference in the substitution rate between humans and chimpanzee may be due to the difference in intensity of natural selection at the *DRB1* locus in the two species because a higher substitution rate is the result of a stronger balancing selection (22). Such balancing selection operating at the antigen recognition sites of the *MHC* locus appears not to influence the substitution rate at the linked neutral locus (22), but the estimated substitution rate in the flanking region of the *DRB1* locus was also shown to be twofold lower in humans than in chimpanzee. Therefore, it appears that the difference in substitution rate cannot be explained by the difference in selection intensity between chimpanzee and humans.

The present data suggested that the mutation rate at the *DRB1* region differed between humans and chimpanzee. The mutation rate may also differ among *DRB1* alleles because the genomic structure is very different for the various *DR* haplotypes (*DR52*, *DR1*, *DR51*, *DR53*, and *DR8*) in humans. To address these questions, it will be necessary to analyze several sequences containing *DRB1* and its flanking region from different species.

Acknowledgments

We are grateful to all technical staff of the Sequence Technology Team in RIKEN – Genomic Science Center for their contribution of technical assistance to this work. We would like to thank three anonymous reviewers for their valuable comments and suggestions. This study was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

1. Hughes AL, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci U S A* 1989: **86**: 958–62.
2. Takahata N, Nei M. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* 1990: **124**: 967–78.
3. Borghans JA, Beltman JB, De Boer RJ. MHC polymorphism under host-pathogen coevolution. *Immunogenetics* 2004: **55**: 732–9.
4. De Boer RJ, Borghans JA, van Boven M, Kesmir C, Weissing FJ. Heterozygote advantage fails to explain the high degree of polymorphism of the MHC. *Immunogenetics* 2004: **55**: 725–31.
5. Satta Y, O'HUigin C, Takahata N, Klein J. The synonymous substitution rate of the major histocompatibility complex loci in primates. *Proc Natl Acad Sci U S A* 1993: **90**: 7480–84.
6. Satta Y, Takahata N, Schonbach C, Gutknecht J, Klein J. Calibrating evolutionary rates at major histocompatibility complex loci. In: Klein D, Klein J, eds. *Molecular Evolution of the Major Histocompatibility Complex. Heidelberg: Springer, 1991, 51–62.*
7. Bergstrom TF, Engkvist H, Erlandsson R et al. Tracing the origin of HLA-DRB1 alleles by microsatellite polymorphism. *Am J Hum Genet* 1999: **64**: 1709–18.
8. Kenter M, Otting N, Anholts J, Jonker M, Schipper R, Bontrop RE. Mhc-DRB diversity of the chimpanzee (Pan troglodytes). *Immunogenetics* 1992: **37**: 1–11.
9. Mayer WE, O'HUigin C, Zaleska-Rutczynska Z, Klein J. Trans-species origin of Mhc-DRB polymorphism in the chimpanzee. *Immunogenetics* 1992: **37**: 12–23.
10. Hughes AL, Nei M. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. *Mol Biol Evol* 1990: **7**: 491–514.
11. Hohjoh H, Ohashi J, Takasu M, Nishioka T, Ishida T, Tokunaga K. Recent divergence of the HLA-DRB1*04 allelic lineage from the DRB1*0701 lineage after the separation of the human and chimpanzee species. *Immunogenetics* 2003: **54**: 856–61.
12. Hattori M, Tsukahara F, Furuhata Y et al. A novel method for making nested deletions and its application for sequencing of a 300 kb region of human APP locus. *Nucleic Acids Res* 1997: **25**: 1802–8.
13. Toyoda A, Noguchi H, Taylor TD et al. Comparative genomic sequence analysis of the human chromosome 21 Down syndrome critical region. *Genome Res* 2002: **12**: 1323–32.
14. Mayor C, Brudno M, Schwartz JR et al. VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* 2000: **16**: 1046–7.

15. Tajima F. Unbiased estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* 1993: **10**: 677–88.
16. Tajima F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 1993: **135**: 599–607.
17. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 2004: **5**: 150–63.
18. Fujiyama A, Watanabe H, Toyoda A et al. Construction and analysis of a human-chimpanzee comparative clone map. *Science* 2002: **295**: 131–4.
19. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 2005: **437**: 69–87.
20. Kimura M. Evolutionary rate at the molecular level. *Nature* 1968: **217**: 624–6.
21. Satta Y, O'HUigin C, Takahata N, Klein J. Intensity of natural selection at the major histocompatibility complex loci. *Proc Natl Acad Sci U S A* 1994: **91**: 7184–8.
22. Ohashi J, Tokunaga K. Sojourn times and substitution rate at overdominant and linked neutral loci. *Genetics* 2000: **155**: 921–7.

Huntington病のsiRNAによる治療研究

Study on the siRNA treatment for Huntington's disease

和田圭司
Keiji WADA
国立精神・神経センター神経研究所疾病研究第4部

◎ポリグルタミン病に関し，RNAiを活用した先端的治療法開発が世界的にも検討されている．本稿ではポリグルタミン病のひとつであるHuntington病について，著者らの研究を含めRNAi治療開発の現状を紹介した．モデルマウスでの検討では，Huntingtin遺伝子に選択的なsiRNAはHuntington病原因遺伝子発現を特異的に抑制し，運動機能障害の進行が対照に比べ軽減するだけでなく，病理学的にも原因遺伝子産物の凝集形成を抑制することが見出されている．RNAiはHuntington病の有望な治療手段のひとつであると考えられる．

Key word：Huntington病，モデル動物，RNAi，治療，ノックダウン

Huntington病は，1872年にアメリカのDr. George Huntingtonによりはじめて報告された神経変性疾患である[1]．常染色体優性遺伝形式をとり，40歳前後に発病することが多い[2,3]．慢性進行性で，舞踏病様不随意運動を主体とする神経症状と知的障害を主体とする精神症状が病態を形成する[2,3]．有病率は，欧米では人口10万人当り4〜7人と多いが，わが国では0.4人くらいと少なく人種差がある[2,3]．病理学的には線条体の萎縮が特徴的で，小細胞が脱落し，グリアの増生，側脳室の拡大を伴う[4]．1993年にDr. James Gusellaをはじめとする研究チームが長年の努力の後に第4染色体短腕にある原因遺伝子を同定し，ハンチンチン(huntingtin)と命名した[5]．Huntingtinは分子量が約350kDaの巨大な蛋白質でその機能は不明であるが，患者症例ではhuntingtin遺伝子の第一エクソンの蛋白質翻訳領域にあるCAG部分が伸長している．そのため，Huntington病はトリプレットリピート病のひとつに位置づけられている．1997年にDr. Gillian Batesらのグループにより伸長したエクソン一部分を発現するトランスジェニックマウスが作成され，Huntington病に類似した病理像を呈するモデル動物であることが報告された[6]．以来，このモデルマウスあるいは他グループにより開発されたモデルマウス[7]などを用いた治療研究が世界的に展開されるようになり，トレハロースの有効性が見出されるなど，治療法開発に関して輝かしい成果が生みだされている[8]．

他方，これまでの研究から，神経変性疾患の発症には病因関連遺伝子・蛋白質の細胞内動態変化や転写・DNA修復など核内現象の変化が大きくかかわり，これまでの神経細胞死という概念だけでなく神経細胞機能不全という状態が発症を左右する重要な因子であることが示されてきている．この神経細胞機能不全は可逆的な状態にあると考えられており，原因遺伝子産物の動態制御あるいはその除去による機能不全の修復こそが，神経変性疾患の根本的治療実現に向けたひとつの扉であると考えられるようになっている．したがって，治療法の開発の方向性も従前とは異なるあらたな視点が要求されるようになってきており，これらの潮流のなかからRNAiの活用など，神経変性疾

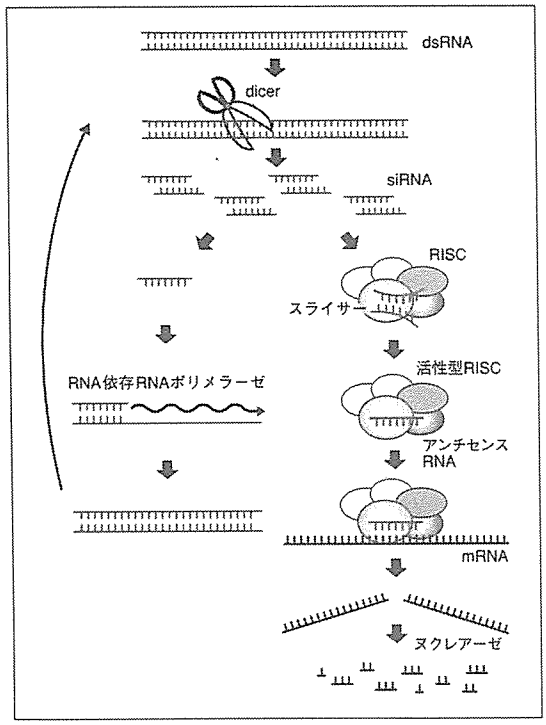

図 1　RNAiの概略

患の根本的治療に迫る展開があらたに生みだされてきている．

本稿では著者らの成果も含めて，RNAi を用いた Huntington 病治療研究の現状と将来を紹介する．

RNAiとは

RNA interference（RNAi）は近年発見された細胞内遺伝子発現調節現象で，二本鎖 RNA（dsRNA）によって配列特異的に mRNA が分解され，遺伝子産物である蛋白質の翻訳が阻害される機序のことである（図 1）．もともと植物で共抑制として報告された現象[9]まで研究の歴史はたどられると思われるが，線虫でアンチセンス RNA による interference の研究を行っていた Dr. Andrew Fire らがはじめて dsRNA の有効性を示し RNA interference という名称を使用して以来，RNAi という言葉が今日では広く使われている[10,11]．Dr. Fire らの論文により，細胞内に導入された dsRNA が，一本鎖 RNA を導入した場合よりもはるかに効果的に mRNA 発現を抑制することがはじめて明確に示された．その後，dsRNA の導入が哺乳類細胞においても有効であることが確認されたことから[12,13]，RNAi は医学生物学をはじめとするさまざまな分野において遺伝子機能を解析するためのツールだけでなく，あらたな治療法としていまやその研究が世界中で展開されている．

現在までの知見に基づき RNAi の機序を要約すると，図 1 に示したように細胞内に取り込まれた dsRNA は，RNase Ⅲに似た dicer とよばれる二本鎖 RNA 特異的エンドヌクレアーゼにより 20〜25 塩基の小さな二本鎖 RNA（small interfering RNA：siRNA）に分解される[14]．siRNA は，argonaute というエンドリボヌクレアーゼ活性をもつ蛋白質（スライサー）を含む蛋白質核酸複合体である RISC（RNA induced silencing complex）に取り込まれ，解きほぐされて一本鎖 RNA になる[15]．この一本鎖 RNA を含む RISC（活性型 RISC）は，一本鎖 RNA に対して相補的配列をもつ標的 mRNA を認識・結合し，スライサー蛋白質が標的 mRNA を結合中央部分で切断する[16]．切断された標的 mRNA はさらにヌクレアーゼによって分解され，結果として遺伝子発現が抑制される．また，RISC と結合していない一本鎖 RNA は相補的な配列をもつ mRNA に結合し，RNA 依存的 RNA ポリメラーゼのプライマーとして作用することも見出されている．RNA 依存的 RNA ポリメラーゼにより合成された二本鎖 RNA は，dicer の基質となり siRNA が生成される．生成された siRNA は，RISC に取り込まれて標的 mRNA の切断を促すか，一本鎖に解きほぐされてふたたび RNA 依存的 RNA ポリメラーゼのプライマーとして作用する．

Huntington病治療用ツールとしてのsiRNAの作成（自験例）

RNAi が哺乳類細胞で有効であることが報告されて以来，RNAi 技術に関する研究が急速に進展し，siRNA の構造的特性に関する知見が多数集積した[17,18]．その結果，効果的な配列の設計，siRNA 発現の構築が進み，モデル動物を用いた治療実験にも応用されるようになった．哺乳類では siRNA が細胞内に存在すると RISC 以降の反応が誘導されることから，siRNA を細胞内に導入する手段として 21〜23 塩基の合成二本鎖 RNA のトランス

```
5'-GCCUUCGAGUCCCUCAAGUCC-3'
   |||||||||||||||||||
3'-UCCGGAAGCUCAGGGAGUUCA-5'
```

図 2　Huntington病原因遺伝子に対するsiRNA-HDエクソン1の塩基配列

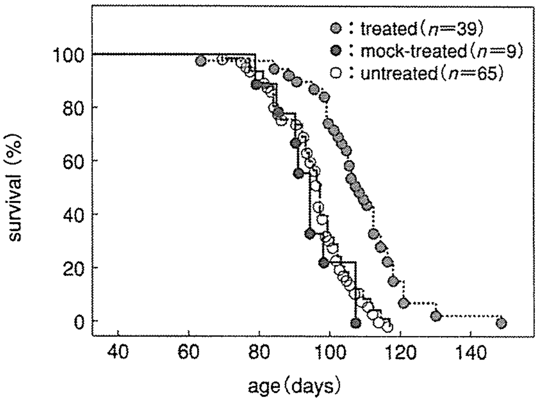

図 3　Huntington病モデルマウスを用いたRNAi治療——生存曲線[20]

siRNA-HD エクソン 1 の投与を受けたモデルマウスは，無処置マウス，コントロール siRNA 投与を受けたマウスに比べ生存が延長した．( )内は使用匹数を示す．

フェクション，siRNA の配列を shRNA(short hairpin RNA)として発現するプラスミドベクターのトランスフェクション，あるいはウイルスベクターを用いた shRNA の導入などが使用されている．著者らは huntingtin 遺伝子エクソン一部分に対する siRNA をいくつか合成し，さらにその配列をもとに U6 プロモーターの制御下で当該配列を含む shRNA として発現するプラスミドベクターを数種作成し，治療用ツールとしての有効性を細胞ならびに Dr. Bates らが報告したモデルマウス(R6/2 マウス[6])において検証した[19,20]．

Huntington病モデルのRNAiの効果（自験例）

裸二重鎖 siRNA と 72 回の CAG リピートを含む Huntingtin 遺伝子エクソン 1 翻訳部分と GFP の融合蛋白質を HEK293 細胞あるいは Neuro2a 細胞に一過性に発現させ，huntingtin 遺伝子発現のノックダウン効果を検討したところ，siRNA-HD エクソン 1 と命名された siRNA(図 2)が濃度依存性に huntingtin 融合蛋白質の発現を特異的に抑制した[19]．その効果は 40nM という低濃度から認められ，huntingtin 遺伝子を含まない GFP 蛋白質発現には影響を与えなかった．

そこで生後 2 日の Huntington 病モデルマウスに siRNA-HD エクソン 1(200 ng)を含むリポフェクタミン溶液 5 μl を注入し，一定期間の後，行動科学的評価と病理学的検討を行った．siRNA-HD エクソン 1 を脳内投与されたモデルマウスは，無処置対照マウス，コントロール siRNA 投与マウスに比べ発症時期が遅れ，体重減少が少なく生存期間は有意に延長した(図 3)[20]．尾吊り下げ試験，rotor rod 試験およびオープンフィールド試験においても行動障害の程度は対照に比べ進行が遅くなった．病理学的には神経細胞死が抑制され，線条体での huntingtin 陽性あるいはユビキチン陽性の細胞内封入体の出現が減少した(図 4)．線条体における huntingtin mRNA レベルの抑制効果は，注入後約 2 週間まで確認でき，蛋白レベルでは huntingtin 凝集に対する抑制効果は 8 週齢まで確認された．発現抑制が持続的でなく短期的であっても症状の進行の抑制が可能である機序については不明であるが，一過性の抑制が継続的な治療効果をもたらしたことは Huntington 病の治療法の実用化を考えるうえで重要な点と考えられる[20]．

ついで，siRNA-HD エクソン 1 の配列を含む shRNA 発現プラスミドを作成し，同様に細胞レベルで Huntingtin 遺伝子に対する抑制効果を検討した．その結果，U6 プロモーター下流に標的配列に対するアンチセンス配列が最初に位置するように挿入された発現プラスミド U6-shHD-3 がもっとも効果を示した．すなわち，U6-shHD-3 は huntingtin＋GFP 融合蛋白質の発現を濃度依存性に抑制し，プラスミド量において 2 ng から有効性を示した．ついで生後 2 日の Huntington 病モデルマウス R6/2 脳内に U6-shHD-3 を注入し，Huntington 遺伝子の発現抑制と病態の改善程度を検討したところ，U6-shHD-3 を投与した R6/2 個体はコントロールプラスミドを投与した対照群と比べ体重減少が少なく，生存期間は有意に延長した．その効果は siRNA-HD エクソン 1 の直接投与よりも優

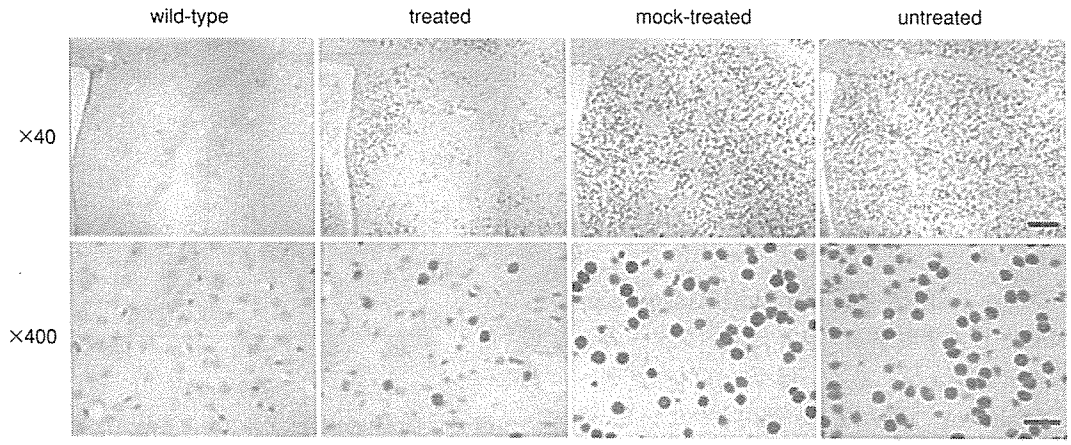

図 4　RNAi治療による細胞内huntingtin凝集体の減少
siRNA-HD エクソン 1 の投与を受けたモデルマウスは，無処置マウス，コントロール siRNA 投与を受けたマウスに比べ，細胞内 huntingtin 凝集体が減少した．野生型マウスでは凝集体形成は認められなかった．

れていた(投稿中)．

RNAiによるHuntington病治療の現状と展望

以上の自験例の結果は RNAi がヒトの Huntington 病治療の有効な手段である可能性が高いことを示している．今回の研究では生後 2 日という早期の治療がモデルマウスで有効であることが示されたが，成体に投与した場合の治療効果についても今後検討する予定である．なお，R6/2 マウスはヒト huntingtin 遺伝子エクソン一部分の過剰発現マウスであり，また著者らが作成した siRNA はヒト huntingtin 遺伝子に選択的であるため，著者らの今回の方法では normal allele 由来の huntingtin の発現抑制が生体に及ぼす影響については検討がされていない．Huntingtin 遺伝子欠損マウスは胎性致死であるため[21]，Huntingtin 遺伝子は成体にとって重要な役割を担っていると推察される．東京医大の金子博士らは，内在性 Huntingtin 遺伝子の抑制が小胞体機能に影響を及ぼす可能性を細胞で見出していることから[22]，今後 RNAi による治療の実用化を考えた場合，内在性 huntingtin 遺伝子の発現抑制の程度と生体機能に与える影響の関連について個体で解析することが必須であろう．また，siRNA の導入法についても継続的な検討が要求される．

著者らは AAV ベクターを用いた導入についてもその効果を解析中であるが，Dr. Beverly Davidson[23]，ならびに理研脳センター・貫名信行博士らのグループ[24]がそれぞれ，shRNA を発現するウイルスベクターを用いたモデル動物の治療実験について報告している．Dr. Davidson らは huntingtin 遺伝子エクソン 2 部分に対する siRNA を作成し，著者らが使用した R6/2 マウスと異なり，HD-N171-82Q マウス[7]でその効果を検討した．4 週齢のマウスに U6 プロモーター制御下で shRNA を発現するように作成された AAV ベクターを投与したところ，トランスジーンであるヒト huntingtin 遺伝子の発現が抑制され，運動機能障害の進行が軽減したと報告している．成体に投与し効果があった点で貴重な報告である．貫名博士らは，独自に開発したモデルマウス[25]に対して同様に U6 プロモーター制御下で shRNA を発現する AAV ベクターを投与したところ，対照に比べて神経病理学的な進行が軽減することを見出している．病理学的に凝集体の形成が確認できる生後 12 週で投与され，効果があった点で貴重な報告である．また，貫名博士らの場合はトランスジーンである huntingtin＋GFP 融合遺伝子のうち GFP 配列部分に対して作成された RNAi であり，Huntington 病のみならずポリグルタミン病全般の RNAi 治療開拓を考えるうえで重要な報告と考えられる．

おわりに

自験例を含め Huntington 病モデルマウスの病態進行抑制に関して，RNAi が有効であることが見出された．安全性の検討，投与ルートの開拓など実用化に向けて克服すべき課題はいまだ多いが，Huntington 病の根本的治療をめざすうえで，RNAi を活用した治療法開発はきわめて有望な選択肢と考えられる．

謝辞：本稿をまとめるにあたってご指導いただいた金澤一郎先生に深謝申し上げます．

文献

1) Huntington, G.：On chorea. *Med. Surg. Rep.*, **26**：317-321, 1872.
2) Bates, G. P.：The molecular genetics of Huntington disease—a history. *Nat. Rev. Genet.*, **6**：766-773, 2005.
3) 金澤一郎：ハンチントン病—臨床と病態．日本内科学会雑誌，**87**：1647-1657，1998.
4) Vonsattel, J. P.：Neuropathology of Huntington's disease. *Neurosci. News*, **3**：45-53, 2000.
5) The Huntington's Disease Collaborative Research Group：A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, **26**：971-983, 1993.
6) Davies, S. W. et al.：Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell*, **90**：537-548, 1997.
7) Schilling, G. et al.：Intranuclear inclusions and neuritic aggregates in transgenic mice expressing a mutant *N*-terminal fragment of huntingtin. *Hum. Mol. Genet.*, **8**：397-407, 1999.
8) Tanaka, M. et al.：Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease. *Nat. Med.*, **10**：148-154, 2004.
9) Napoli, C. et al.：Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell*, **2**：279-289, 1990.
10) Fire, A. et al.：Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle. *Development*, **113**：503-514, 1991.
11) Fire, A. et al.：Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**：806-811, 1998.
12) Elbashir, S. M. et al.：Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**：494-498, 2001.
13) Caplen, N. J. et al.：Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl. Acad. Sci. USA*, **98**：9742-9747, 2001.
14) Elbashir, S. M. et al.：RNA interference is mediated by 21-and 22-nucleotide RNAs. *Genes Dev.*, **15**：188-200, 2001.
15) Hammond, S. M. et al.：An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, **404**：293-296, 2000.
16) Zamore, P. D. et al.：RNAi：double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101**：25-33, 2000.
17) Elbashir, S. M. et al.：Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.*, **20**：6877-6888, 2001.
18) Holen, T. et al.：Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Res.*, **30**：1757-1766, 2002.
19) Liu, W. et al.：Specific inhibition of Huntington's disease gene expression by siRNAs in cultured cells. *Proc. Jpn. Acad.*, **79**(SerB)：293-298, 2003.
20) Wang, Y. L. et al.：Clinico-pathological rescue of a model mouse of Huntington's disease by siRNA. *Neurosci. Res.*, **53**：241-249, 2005.
21) Nasir, J. et al.：Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell*, **81**：811-823, 1995.
22) Omi, K. et al.：siRNA-mediated inhibition of endogenous Huntington disease gene expression induces an aberrant configuration of the ER network *in vitro*. *Biochem. Biophys. Res. Commun.*, **338**：1229-1235, 2005.
23) Harper, S. Q. et al.：RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. *Proc. Natl. Acad. Sci. USA*, **102**：5820-5825, 2005.
24) Machida, Y. et al.：rAAV-mediated shRNA ameliorated neuropathology in Huntington disease model mouse. *Biochem. Biophys. Res. Commun.*, **343**：190-197, 2006.
25) Kotliarova, S. et al.：Decreased expression of hypothalamic neuropeptides in Huntington disease transgenic mice with expanded polyglutamine-EGFP fluorescent aggregates. *J. Neurochem.*, **93**：641-653, 2005.

*　　*　　*

神経・筋疾患治療へのRNAi応用*

北條　浩彦**

Key Words: mammalian RNAi, interferon response, siRNA duplex, duration of RNAi activity

I．哺乳動物RNAiとインターフェロン応答

RNA interference（RNAi；RNA干渉）は，二本鎖RNA（dsRNA）によって誘導される配列特異的な遺伝子発現の転写後抑制現象である．この不思議な現象は，1998年，Fireらによって線虫を用いた研究から発見された[1]．その後，RNAiは，線虫をはじめ，ショウジョウバエ，原生動物，脊椎動物，さらに植物とさまざまな生物種で観察される保存された現象であることが明らかになった[2~5]．そのメカニズムは，1）細胞内に生じた長いdsRNAが，Dicerと呼ばれるRNase III酵素によって消化され，約21～25bpの短い二本鎖RNA，small interfering RNA（siRNA）二量体となり，次に，2）RNA-induced silencing complex（RISC）と呼ばれる複合体にそのsiRNAが取り込まれ，3）取り込まれたsiRNAと相補的なメッセンジャーRNA（mRNA）がRISCによって特異的に切断されるという反応である．このような反応過程を経て，RNAiの配列特異的な遺伝子発現の転写後抑制が起こる．

ヒトを含めた哺乳類細胞のRNAiは，当初，初期胚や未分化細胞といった限られた細胞でのみ観察される特殊な現象であると思われていた．これは，哺乳類細胞が持つもう一つの二本鎖RNAに対する応答経路が密接に関わっている．初期胚や未分化細胞など，一部の細胞を除いた殆どの哺乳動物細胞は，30bp以上の長いdsRNAが細胞内に生じると，RNAiよりも速くインターフェロン応答（抗ウイルス反応）を引き起こし，細胞死が誘導される（**Fig. 1a**）．この反応には，主に二つの経路が関わっている．ひとつは，dsRNA依存的タンパク質キナーゼ（interferon-inducible, ds-RNA-activated protein kinase：PKR）を活性化し，翻訳因子であるeIF2aをリン酸化して翻訳阻害を誘導する経路，もう一つは，2′-5′オリゴアデニル酸合成酵素（2′-5′ oligoadenylate synthetase：2-5AS）を活性化し，それによって合成されたポリアデニル酸を介して非特異的なRNase Lを活性化する経路である．これらの応答経路は，dsRNA導入後直ちに活性化されるため，Dicerを介したRNAi誘導を観察することができなかった．しかしながら，2001年，この応答経路を回避する画期的なブレークスルーが，Tuschlらのグループによって報告された[6]．彼らは，化学合成したsiRNA二量体を直接哺乳動物細胞内に導入し，細胞死を引き起こさず，RNAiだけを誘導することに成功したのである（**Fig. 1b**）．この方法によって，ほぼ全ての哺乳動物細胞にRNAiを誘導することが可能となり，哺乳類のRNAi研究

* Medical Application of RNA Interference to Neuronal and Muscular Diseases.

** 国立精神・神経センター神経研究所　Hirohiko HOHJOH：National Institute of Neuroscience, NCNP

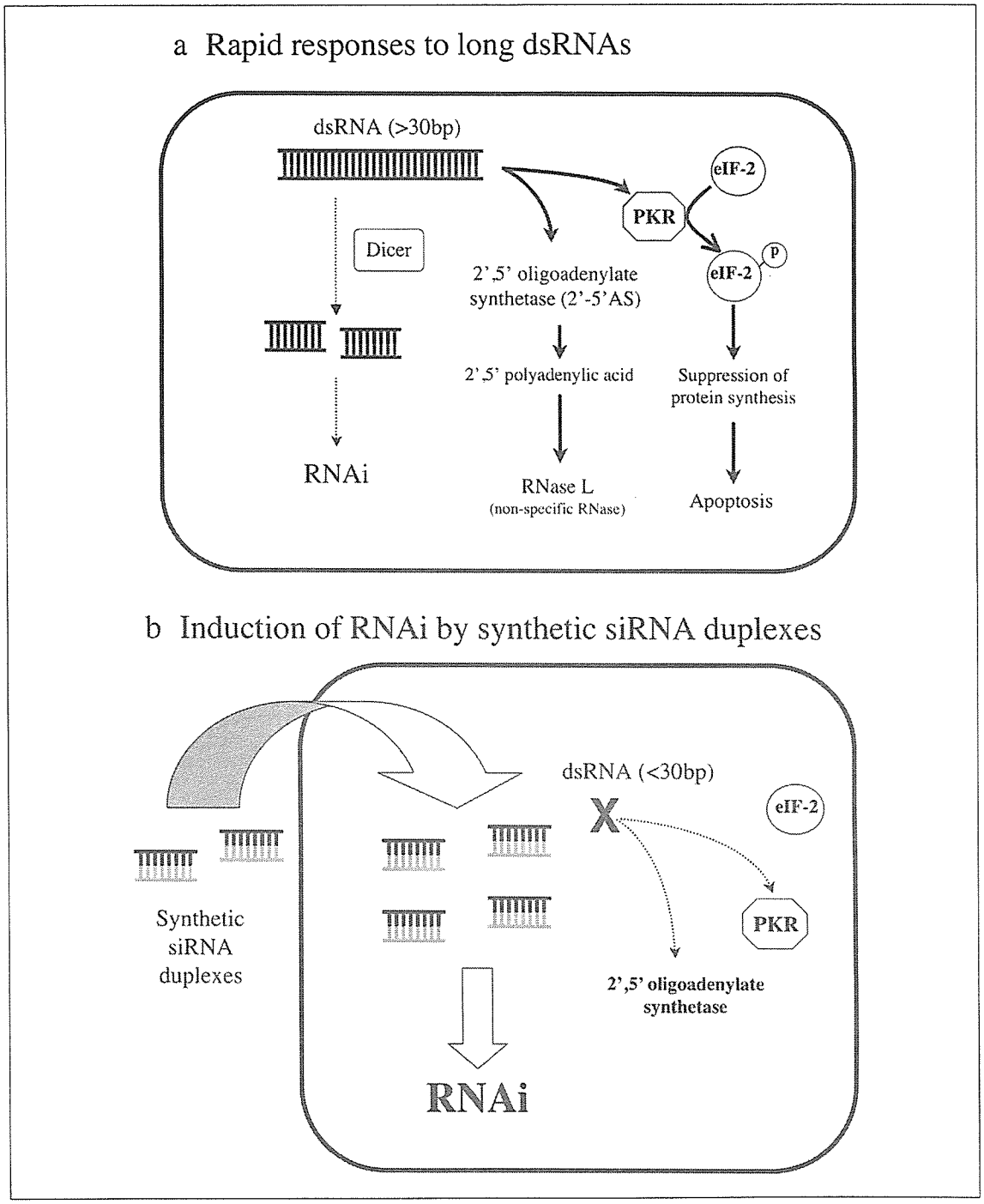

Fig. 1 Schematic drawing of interferon responses (a) and RNAi induction by synthetic siRNA duplexes (b). (a) Long double-stranded RNAs (>30bp) trigger a rapid translation inhibition involving the interferon-inducible, dsRNA-activated protein kinase, PKR, and a rapid and non-specific RNA degradation involving the sequence-non-specific RNase, RNase L, in most of mammalian cells except for some undifferentiated cells. Consequently, these rapid responses to long dsRNAs may mask the sequence-specific RNA interference (RNAi) activity. (b) Chemically synthesized siRNA duplexes (19～27bp) introduced into cells can induce sequence-specific RNAi without triggering the rapid and non-specific RNA degradation and translation inhibition.

Table 1

Disease	Target (genes) of RNAi	References
Triplet repeat diseases : Huntington's disease Spinocerebellar ataxia type 1 (SCA1)	 Huntingtin SCA1	 13～15 16
Parkinsonism	Tau (V337M variant)	17
Alzheimer's disease	APP (Swedish variant) Tau (V337M variant)	17, 18
Slow channel congenital myasthenic syndrome (SCCMS)	Muscle acetylcholine receptor (AChR) subunits (αS226F variant)	19
Amyotrophic lateral sclerosis (ALS)	SOD1	20, 21

が飛躍的に発展した．現在では，合成siRNAを用いて直接RNAiを誘導する方法の他に，発現プラスミドを用いて細胞内で短いヘアピン型RNA（shRNA）を発現させ，その後Dicerによるプロセスを経てsiRNA二量体を生じさせる方法も用いられている．

II. 神経筋疾患治療に向けた神経細胞・筋細胞でのRNAi誘導

RNAiは，その不思議なメカニズムに対する学問的な興味だけでなく，目的の遺伝子発現を簡単に，しかも強力に抑制できることから，簡便な遺伝子機能阻害方法（遺伝子ノックダウン方法）としても注目されている．特に，医療方面・創薬方面へのRNAi技術の応用に大きな期待が持たれている．

RNAiの治療方面への応用については，その作用機序から，機能獲得（gain of function）が関わる疾患など（変異遺伝子が関連する優性の遺伝病など）に対して有効であると考えられる．そして，すでに様々な疾患に対して，RNAiを用いた疾患原因遺伝子をターゲットとするノックダウンが試みられている．特に神経筋疾患では，神経変性疾患をはじめ多くの疾患がgain of functionに関連するものであり，RNAi技術を用いた様々な治療への取り組みがなされている．**Table 1**に，その一部ではあるが，神経筋疾患に関連する遺伝子をターゲットとしたRNAiノックダウンの報告をまとめた．

上記のような異常な遺伝子（機能獲得）によって疾患が惹起されるものに対して，筋ジストロフィー症のような遺伝子機能の喪失（loss of function）が原因で引き起こされる疾患もある．そのような疾患に対しては，RNAi技術を用いた遺伝子ノックダウンによって様々なタイプの疾患モデル生物を作出することも可能であると考えられる．そして，それらのモデル生物を用いた新しい治療法や治療薬の開発に貢献できると考えられる．

さて，疾患モデル生物の作出を目的としたRNAiでも，疾患原因遺伝子をターゲットとするRNAiでも，目的とする細胞や組織でRNAiが効果的に誘導されなければならない．また，その特徴も知らなければならない．哺乳動物のRNAiは，その誘導物質であるsiRNA二量体に依存したRNAi活性があることが既に知られている[7]．したがって，ターゲット遺伝子に対して効果的なノックダウンを行うためには，まず，強いRNAi活性を誘導するポテンシャルを持ったsiRNA二量体を設計しなければならない．今日では，優れたアルゴリズムを備えた予測プログラムによってそのようなsiRNA二量体を設計することが可能となっている[8]．

RNAiの機能面だけでなく，RNAiを誘導する細胞の特性もRNAi誘導の重要な情報となる．筋肉組織では，RNAiに関わるDicerの発現が他の組織に比べて低いことが既に知られている[9,10]．われわれはさらに，Dicer以外にも，RISCの重要な構成タンパク質であるAgo2（eIF2C2）遺伝子やその遺伝子ファミリーに属するeIF2C1,