Table 1 Clinical characteristics of study populations. Data are mean ± SD. *NA* data not available, *Ctrl* control, *T2D* type 2 diabetes

Trait	Study population											
	Study 1		Study 2		Study 3		Horikawa et al. (2003)		Shima et al. (2003)		Daimon et al. (2003)	
Subjects	Ctrl[a]	T2D	Ctrl	T2D	Ctrl[a]	T2D	Ctrl[a]	T2D	Ctrl	T2D	Ctrl	T2D
n	208	205	281	...	192	454	172	177	276	10	81	81
Gender (F/M)	68/140	84/121	135/146	...	70/122	206/248	90/82	63/114	NA	NA	46/35	45/36
Age-at-study (years)	67.9 ± 5.5	59.1 ± 13.0	44.4 ± 15.9	...	67.8 ± 5.6	59.9 ± 11.7	68.0 ± 5.7	62.0 ± 11.0	NA	NA	62.3 ± 8.2	65.1 ± 10.7
Age-at-diagnosis (years)	...	45.9 ± 12.7	...	...	...	51.0 ± 12.0[b]	...	49.8 ± 11.4	...	NA	...	NA
BMI	23.1 ± 2.4	23.5 ± 3.6	22.3 ± 3.0	...	23.1 ± 2.8	24.2 ± 4.2	22.8 ± 3.3	23.9 ± 3.3	NA	NA	23.8 ± 3.6	25.6 ± 3.9
HbA_{1c} (%)	5.0 ± 0.3	8.2 ± 2.0	4.9 ± 0.4	...	4.9 ± 0.3	7.9 ± 1.9	5.0 ± 0.4	6.7 ± 1.0	NA	NA	5.3 ± 0.3	6.7 ± 1.1
Fasting glucose (mg/dl)	95.0 ± 8.7	160.5 ± 48.5	92.8 ± 9.4	...	NA	NA	NA	NA	NA	NA	NA	NA
Treatment (diet/oral agents/insulin)	...	42/81/82	...	...	...	NA	...	46/70/61	...	NA	...	NA

[a] All subjects were > 60 years old
[b] Data are available for only 246 subjects

nosis ≥50 years included 87 patients from study 1, 143 from study 3, and 85 from Horikawa et al. (2003). There was no significant difference in SNP-43, Indel-19, and SNP-63 allele or genotype frequencies between the type 2 diabetic group with age-at-diagnosis < 50 years and the controls (Table 4). The SNP-43 and SNP-63 frequencies were also not different between the type 2 diabetic group with age-at-diagnosis ≥50 years and the controls. However, there was a small but significant difference in Indel-19 allele frequency (Table 4). The 3R allele at Indel-19 (allele 2 in the haplotype) was associated with lower risk of type 2 diabetes (OR = 0.82, $P = 0.04$).

The 121 haplotype was associated with significantly decreased risk (OR = 0.80, $P = 0.02$) of type 2 diabetes in the group of patients with age-at-diagnosis ≥50 years (Table 5). The 111 haplotype had the highest risk (OR = 1.33, $P = 0.046$) in the older group of patients. This effect of the 111 haplotype likely reflects the contribution of SNP-44 to type 2 diabetes risk (Weedon et al. 2003) since 88% of the 111 haplotypes in Japanese carry the at-risk C-allele at SNP-44. The rare 111/221 haplogenotype was associated with increased risk of type 2 diabetes irrespective of age-at-diagnosis (Table 6). The 121/121 haplogenotype had a protective effect against type 2 diabetes that approached significance in patients with age-at-diagnosis ≥50 years (OR = 0.76, $P = 0.06$). Individuals with haplotypes 111/111, 111/112, and 111/221 had the highest risk of type 2 diabetes (OR = 1.83, 1.25, and 4.12, respectively) although the increase in risk was not significant because of the small numbers of individuals studied. If the nondiabetic subjects in study 2 (Table 1) are included in the pooled control group, the results are similar, including the effects of the 121/121 haplogenotype on risk in patients with age-at-diagnosis ≥50 years (OR = 0.75, $P = 0.04$).

Discussion

The results suggest that genetic variation in CAPN10 may affect risk of type 2 diabetes in the Japanese population, especially in older individuals. Interestingly, the common 121 haplotype appears to be protective in Japanese, suggesting the overall effect of CAPN10 in this population is to reduce the risk of diabetes rather than increase it. It is important to note, though, that the statistical significance of the comparison is marginal ($P = 0.01$–0.04), and none of the comparisons would be significant if corrected for multiple testing. Thus, the results presented here need to be confirmed through studies of a much larger dataset. However, if our results are correct, they suggest an interaction between genetic (CAPN10) and nongenetic (age) factors to modify risk of type 2 diabetes. In this regard, recent studies have shown that calpain-10 is part of a novel apoptotic pathway in insulin-secreting pancreatic beta cells and thus may

Table 2 Genotype and allele frequencies of CAPN10 polymorphisms in Japanese. The number of subjects of each genotype are indicated. All genotypic distributions are in Hardy-Weinberg equilibrium. *NA* not available. *Ctrl* control, *T2D* type 2 diabetes

Marker	Subjects	Genotype	This study 1	This study 2	This study 3	Horikawa et al. (2000)	Shima et al. (2002)	Daimon et al. (2002)	*P* for heterogeneity
SNP-43	Ctrl	G/G	188	251	165	154	252	76	
		G/A	20	29	24	18	24	5	
		A/A	0	1	0	0	0	0	0.75
		Allele frequency	G: 0.95	G: 0.94	G: 0.94	G: 0.95	G: 0.96	G: 0.97	
			A: 0.05	A: 0.06	A: 0.06	A: 0.05	A: 0.04	A: 0.03	0.62
	T2D	G/G	184	NA	389	158	8	76	
		G/A	21	NA	57	19	2	5	
		A/A	0	NA	1	0	0	0	0.75
		Allele frequency	G: 0.95	NA	G: 0.93	G: 0.95	G: 0.90	G: 0.97	
			A: 0.05		A: 0.07	A: 0.05	A: 0.10	A: 0.03	0.37
Indel-19[a]	Ctrl	2R/2R	27	42	35	23	42	NA	
		2R/3R	99	126	78	78	126	NA	
		3R/3R	82	113	73	71	108	NA	0.90
		Allele frequency	2R: 0.37	2R: 0.37	2R: 0.40	2R: 0.36	2R: 0.38	NA	
			3R: 0.63	3R: 0.63	3R: 0.60	3R: 0.64	3R: 0.62		0.86
	T2D	2R/2R	32	NA	63	28	1	NA	
		2R/3R	104	NA	209	82	3	NA	
		3R/3R	69	NA	176	67	6	NA	0.62
		Allele frequency	2R: 0.41	NA	2R: 0.37	2R: 0.39	2R: 0.25	NA	
			3R: 0.59		3R: 0.63	3R: 0.61	3R: 0.75		0.34
SNP-63	Ctrl	C/C	111	151	99	90	151	NA	
		C/T	81	106	65	70	103	NA	
		T/T	16	24	18	12	22	NA	0.99
		Allele frequency	C: 0.73	C: 0.73	C: 0.72	C: 0.73	C: 0.73	NA	
			T: 0.27	T: 0.27	T: 0.28	T: 0.27	T: 0.27		0.99
	T2D	C/C	90	NA	255	93	6	NA	
		C/T	92	NA	165	74	3	NA	
		T/T	23	NA	30	10	1	NA	0.09
		Allele frequency	C: 0.66	NA	C: 0.75	C: 0.73	C: 0.75	NA	
			T: 0.34		T: 0.25	T: 0.27	T: 0.25		0.03

[a] Indel-19 is a diallelic insertion/deletion polymorphism with alleles of two repeats (2R) or three repeats (3R) of a 32-bp sequence

Table 3 Clinical and metabolic characteristics of normal glucose tolerant subjects (study 2) by SNP-63 genotype. Data are mean ± SD. Subjects underwent a standard 75-g oral glucose tolerance test with glucose and insulin determined at 0′, 30′, 60′, and 120′. The number of individuals in each group for determination of insulinogenic index and HOMA are noted in parentheses. *BMI* body mass index, *AUC* area under the curve

Trait	Genotype			*P*[a]
	C/C	C/T	T/T	
n	150	105	25	
Gender (M/F)	72/78	59/46	14/11	0.40
Age (years)	45.0 ± 14.9	43.8 ± 16.7	44.4 ± 18.7	0.72
BMI (kg/m2)	22.3 ± 2.5	22.2 ± 3.6	22.6 ± 3.0	0.70
HbA1c (%)	4.9 ± 0.4	4.9 ± 0.4	5.0 ± 0.4	0.79
Plasma glucose (mg/dl)				
0 min	93.1 ± 9.6	92.2 ± 9.1	94.1 ± 8.8	0.51
120 min	104.1 ± 18.5	106.1 ± 18.8	108.2 ± 14.5	0.30
AUC 0-120′	14,510.6 ± 2706.2	15,233.4 ± 2937.1	15,043.2 ± 3056.1	0.03
Plasma insulin (μU/ml)				
0 min	6.4 ± 2.2	6.4 ± 2.6	6.5 ± 4.4	0.81
120 min	30.6 ± 16.9	32.0 ± 16.8	30.1 ± 20.2	0.74
AUC 0-120′	3,972.4 ± 2030.8	4,267.4 ± 1816.9	4,049.2 ± 1969.1	0.44
Insulinogenic index	0.89 ± 0.89 (147)	0.96 ± 3.62 (101)	1.14 ± 1.85 (24)	0.08
HOMA	1.43 ± 0.53 (146)	1.41 ± 0.54 (101)	1.30 ± 0.63 (23)	0.35

[a] *P* value by ANCOVA with institution, gender, and genotype as independent factors and age and BMI as covariates

affect the response of the beta cell to aging or its ability to compensate in response to an increasing demand for insulin (Johnson et al. 2004).

The observation that individuals homozygous for the 121 haplotype may be at increased risk of type 2 diabetes in some European populations (Orho-

Table 4 SNP-43, Indel-19, and SNP-63 and type 2 diabetes in Japanese—a pooled analysis. The number of subjects of each genotype are indicated. All genotypic distributions are in Hardy-Weinberg equilibrium. The cases were divided into two groups based on the median age-at-diagnosis in the pooled sample—50 years. Note that age-at-diagnosis was not available for all subjects

Marker	Genotype	Overall			Age-at-diagnosis <50 years		Age-at-diagnosis ≥50 years	
		Ctrl	T2D	*P*	T2D	*P*	T2D	*P*
SNP-43	G/G	833	811		255		276	
	G/A	91	104		28	0.98	32	0.78
	A/A	0	1	0.35	0		0	
	Allele frequency	G: 0.95	G: 0.94		G: 0.95		G: 0.95	
		A: 0.05	A: 0.06	0.25	A: 0.05	0.98	A: 0.05	0.79
Indel-19[a]	2R/2R	127	124		36		58	
	2R/3R	381	396		143		151	0.11
	3R/3R	334	316	0.67	107	0.33	105	
	Allele frequency	2R: 0.38	2R: 0.39		2R: 0.38		2R: 0.43	
		3R: 0.62	3R: 0.61	0.69	3R: 0.62	0.91	3R: 0.57	0.04
SNP-63	C/C	451	443		149		151	
	C/T	319	333		125		129	
	T/T	68	62	0.72	14	0.09	28	0.35
	Allele frequency	C: 0.73	C: 0.73		C: 0.73		C: 0.71	
		T: 0.27	T: 0.27	0.94	T: 0.27	0.79	T: 0.29	0.17

[a] Indel-19: 2R, 2 repeats of 32-bp sequence; 3R, 3 repeats

Table 5 CAPN10 haplotype frequency and risk of type 2 diabetes in Japanese—a pooled analysis. The haplotypes are those defined by SNP-43, Indel-19, and SNP-63, and the specific alleles are: SNP-43, allele 1, G and allele 2, A; Indel-19, allele 1, 2 repeats of 32-bp sequence, and allele 2, 3 repeats; and SNP-63, allele 1, C, and allele 2, T. The cases were divided into two groups based on the median age-at-diagnosis in the pooled sample—50 years. Note that age-at-diagnosis was not available for all subjects. *Ctrl* control. *T2D* type 2 diabetes

Haplotype	Overall				Age-at-diagnosis <50 years			Age-at-diagnosis ≥50 years		
	Ctrl (n=825)	T2D (n=827)	OR (95% CI)[a]	*P*	T2D (n=277)	OR (95% CI)[a]	*P*	T2D (n=305)	OR (95% CI)[a]	*P*
111	0.106	0.113	1.07 (0.86–1.34)	0.52	0.117	1.12 (0.83–1.52)	0.46	0.136	1.33 (1.00–1.75)	0.05
121	0.572	0.553	0.93 (0.81–1.07)	0.29	0.567	0.98 (0.81–1.19)	0.85	0.515	0.80 (0.66–0.96)	0.02
112	0.270	0.274	1.02 (0.88–1.19)	0.79	0.265	0.98 (0.78–1.21)	0.82	0.297	1.14 (0.93–1.40)	0.21
221	0.052	0.059	1.15 (0.85–1.54)	0.37	0.051	0.97 (0.62–1.50)	0.88	0.052	1.01 (0.66–1.53)	0.97

[a] The OR and 95% CI of each haplotype relative to other haplotypes as a group are shown

Table 6 CAPN10 haplogenotype and risk of type 2 diabetes in Japanese—a pooled analysis. The haplotypes are those defined by SNP-43, Indel-19, and SNP-63, and the specific alleles are indicated in the legend to Table 5. The number of individuals with each haplogenotype is indicated

Haplogenotype	Overall				Age-at-diagnosis <50 years			Age-at-diagnosis ≥50 years		
	Ctrl	T2D	OR (95% CI)[a]	*P*	T2D	OR (95% CI)[a]	*P*	T2D	OR (95% CI)[a]	*P*
111/111	15	18	1.20 (0.60–2.40)	0.60	5	0.99 (0.36–2.76)	0.99	10	1.83 (0.82–4.07)	0.14
111/121	97	93	0.95 (0.70–1.29)	0.74	31	0.95 (0.62–1.45)	0.80	37	1.04 (0.69–1.55)	0.86
111/112	44	44	1.00 (0.65–1.53)	0.99	18	1.23 (0.70–2.17)	0.47	20	1.25 (0.72–2.15)	0.43
111/221	4	14	3.53 (1.24–10.1)	0.02	6	4.54 (1.42–14.5)	0.01	6	4.12 (1.27–13.3)	0.02
112/112	66	62	0.93 (0.65–1.34)	0.70	13	0.57 (0.31–1.04)	0.06	27	1.12 (0.70–1.78)	0.64
112/121	247	254	1.04 (0.84–1.28)	0.73	90	1.13 (0.84–1.51)	0.43	97	1.09 (0.82–1.45)	0.55
112/221	23	32	1.40 (0.82–2.41)	0.22	13	1.72 (0.86–3.41)	0.12	10	1.18 (0.56–2.51)	0.66
121/121	270	259	0.94 (0.76–1.15)	0.54	92	1.02 (0.77–1.37)	0.88	82	0.76 (0.56–1.01)	0.06
121/221	59	50	0.84 (0.57–1.23)	0.37	9	0.44 (0.22–0.87)	0.02	16	0.72 (0.41–1.27)	0.25
221/221	0	1	-	-	0			0		

[a] The OR and 95% CI of each haplogenotype relative to the other haplotype combinations as a group are shown

Melander et al. 2002; Malecki et al. 2002) but at decreased risk in older Japanese raises the possibility that additional genetic variation may distinguish high- and low-risk subtypes of the 121 haplotype. Transpopulation mapping may be a useful strategy for identifying this variation.

Acknowledgements The authors thank Ms. A. Nogami, Ms. M. Y. Sagisaka, and Mr. M. Ikeda for their skillful technical assistance. This study was supported by Grants-in-Aid for Scientific Research C (10671084, 10470234) and for Scientific Research on Priority Areas Medical Genome Science from the Japan Ministry of Science, Education, Sports, Culture and Technology (12204102, 13204082, 14013059, 15012250), Novo Nordisk Foundation, the Naito Foundation (to N.I.) and Grants-in-Aid for Scientific Research B (13470223, 13557091) (to Y.H.), and U.S. Public Health Service (Grants DK-20595, −47486 and −55889). G.I.B. is an Investigator of the Howard Hughes Medical Institute.

References

Abecasis GR, Cookson WO (2000) GOLD–graphical overview of linkage disequilibrium. Bioinformatics 16:182–183

Baier LJ, Permana PA, Yang X, Pratley RE, Hanson RL, Shen GQ, Mott D, Knowler WC, Cox NJ, Horikawa Y, Oda N, Bell GI, Bogardus C (2000) A 1-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. J Clin Invest 106:R69-R73

Daimon M, Oizumi T, Saitoh T, Kameda W, Yamaguchi H, Ohnuma H, Igarashi M, Manaka H, Kato T (2002) Calpain-10 gene polymorphisms are related, not to type 2 diabetes, but to increased serum cholesterol in Japanese. Diabetes Res Clin Pract 56:147–152

Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, Lindner TH, Mashima H, Schwarz PE, del Bosque-Plata L, Oda Y, Yoshiuchi I, Colilla S, Polonsky KS, Wei S, Concannon P, Iwasaki N, Schulze J, Baier LJ, Bogardus C, Groop L, Boerwinkle E, Hanis CL, Bell GI (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 26:163–175

Horikawa Y, Oda N, Yu L, Imamura S, Fujiwara K, Makino M, Seino Y, Itoh M, Takeda J (2003) Genetic variations in calpain-10 gene are not a major factor in the occurrence of type 2 diabetes in Japanese. J Clin Endocrinol Metab 88:244–247

Iwasaki N, Cox NJ, Wang YQ, Schwarz PE, Bell GI, Honda M, Imura M, Ogata M, Saito M, Kamatani N, Iwamoto Y (2003) Mapping genes influencing type 2 diabetes risk and BMI in Japanese subjects. Diabetes 52:209–213

Johnson JD, Han Z, Otani K, Ye H, Zhang H, Wu H, Horikawa Y, Misler S, Bell GI, Polonsky KS (2004) RyR2 and calpain-10 delineate a novel apoptosis pathway in pancreatic islets. J Biol Chem 279:24794–24802

Malecki MT, Moczulski DK, Klupa T, Wanic K, Cyganek K, Frey J, Sieradzki J (2002) Homozygous combination of calpain 10 gene haplotypes is associated with type 2 diabetes mellitus in a Polish population. Eur J Endocrinol 146:695–699

Mori H, Ikegami H, Kawaguchi Y, Seino S, Yokoi N, Takeda J, Inoue I, Seino Y, Yasuda K, Hanafusa T, Yamagata K, Awata T, Kadowaki T, Hara K, Yamada N, Gotoda T, Iwasaki N, Iwamoto Y, Sanke T, Nanjo K, Oka Y, Matsutani A, Maeda E, Kasuga M (2001) The Pro12 → Ala substitution in PPAR-γ is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. Diabetes 50:891–894

Mori Y, Otabe S, Dina C, Yasuda K, Populaire C, Lecoeur C, Vatin V, Durand E, Hara K, Okada T, Tobe K, Boutin P, Kadowaki T, Froguel P (2002) Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate Loci on 7p and 11p. Diabetes 51:1247–1255

Orho-Melander M, Klannemark M, Svensson MK, Ridderstrale M, Lindgren CM, Groop L (2002) Variants in the calpain-10 gene predispose to insulin resistance and elevated free fatty acid levels. Diabetes 51:2658–2664

Seino S on behalf of the Study Group of Comprehensive Analysis of Genetic Factors in Diabetes Mellitus (2001) S20G mutation of the amylin gene is associated with Type II diabetes in Japanese. Diabetologia 44:906–909

Shima Y, Nakanishi K, Odawara M, Kobayashi T, Ohta H (2003) Association of the SNP-19 genotype 22 in the calpain-10 gene with elevated body mass index and hemoglobin A1c levels in Japanese. Clin Chim Acta 336:89–96

Song Y, Niu T, Manson JE, Kwiatkowski DJ, Liu S (2004) Are variants in the CAPN10 gene related to risk of type 2 diabetes? A quantitative assessment of population and family-based association studies. Am J Hum Genet 74:208–222

Weedon MN, Schwarz PE, Horikawa Y, Iwasaki N, Illig T, Holle R, Rathmann W, Selisko T, Schulze J, Owen KR, Evans J, Del Bosque-Plata L, Hitman G, Walker M, Levy JC, Sampson M, Bell GI, McCarthy MI, Hattersley AT, Frayling TM (2003) Meta-analysis and a large association study confirm a role for calpain-10 variation in type 2 diabetes susceptibility. Am J Hum Genet 73:1208–1212

DNA RESEARCH 11, 315–323 (2004)

Construction of a Multi-Functional cDNA Library Specific for Mouse Pancreatic Islets and Its Application to Microarray

Motoi NISHIMURA,[1,†] Norihide YOKOI,[2] Takashi MIKI[2,3] Yukio HORIKAWA[3,4] Hirokazu YOSHIOKA,[5] Jun TAKEDA,[3,6] Osamu OHARA,[7,8] and Susumu SEINO[1,2,*]

Department of Cellular and Molecular Medicine, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan.[1] *Division of Cellular and Molecular Medicine, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan.*[2] *Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), 4-1-8 Honcho, Kawaguchi 332-0012, Japan.*[3] *Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15 Showa-machi, Maebashi 371-8512, Japan.*[4] *Kakengeneqs Corporation, 439-1 Matsuhidai, Matsudo 270-2214, Japan.*[5] *Department of Endocrinology, Diabetes and Rheumatology, Division of Bioregulatory Medicine, Gifu University School of Medicine, 1-1 Yanagido, Gifu 501-1194, Japan.*[6] *Department of Human Gene Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan.*[7] *and Laboratory for Immunogenomics, RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan*[8]

(Received 19 January 2004; revised 24 August 2004)

Abstract

We have constructed a high-quality and multi-applicable cDNA library specific for mouse pancreatic islets. This is the first pancreatic islet cDNA library created using a recombination-based method, which can readily be converted into other applications including yeast two-hybrid and mammalian expression libraries. Based on sequence data of the library, we constructed a sequence database specific for mouse pancreatic islets. Among the 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium, 10% of which were "molecular function unknown" genes. We also developed cDNA microarray membranes with 8108 non-redundant clones. Analyses of expression profiles of three different cell lines and of MIN6 cells with or without overexpression of transcription factor NeuroD1 established the usefulness and applicability of our microarrays. The mouse pancreatic islet cDNA library, sequence database, set of clones, and microarrays developed in this study should be useful resources for studies of pancreatic islets and related diseases including diabetes mellitus.

Key words: cDNA library; microarray; pancreatic islet; recombination-based method; sequence database

1. Introduction

Pancreatic islets are small inner parts of pancreata, and play the central role in glucose homeostasis.[1] The majority of islet cells are β-cells, which secrete the hypoglycemic hormone insulin; other islet cells (α-cells, δ-cells and PP cells) also secrete hormones affecting glucose homeostasis. Impairment of pancreatic β-cell function readily causes disorders of glucose homeostasis such as diabetes.

Studies of knockout mice[2–6] and a subtype of human diabetes[7–12] have revealed that impairment of the transcription network in pancreatic β-cells is involved in the pathogenesis of diabetes. In these models, diabetes occurs due to disruptions or mutations in transcription factors that regulate gene expression in pancreatic β-cells. Thus, construction of a gene expression database for pancreatic islets and comprehensive analysis of expression profiles in the pancreatic islets might provide valuable information on the molecular mechanisms of the disease.

To investigate the molecules involved in normal function of pancreatic islets, several groups have constructed cDNA libraries from human pancreatic islets and sequenced clones from the libraries (In the GenBank database (www.ncbi.nlm.nih.gov/genbank/index.html), nucleotide sequences of 4559 clones were deposited by

Communicated by Michio Oishi

* To whom correspondence should be addressed. Kobe University Tel. +81-78-382-5360, Fax. +81-78-382-5370. E-mail: seino@med.kobe-u.ac.jp

† Present address: Division of Cellular and Molecular Medicine, Kobe University Graduate School of Medicine, Kobe, Japan

Table 1. Characteristics of the mouse pancreatic islet cDNA library.

Starting material	About 5,000 pancreatic islets from C57BL/6 male mice
Complexity without amplification	3.12×10^7

Fraction by insert size[1)]	1	2	3	4	5	6
Proportion of chimeric clones (%)[2)]	4.23	2.76	4.60	6.45	1.84	1.10
Proportion of clones with full ORFs (%)[3)]	64	56	48	76	72	52

1) Insert sizes of the fractions are: 1, >8 kbp; 2, 6–8 kbp; 3, 4–6 kbp; 4, 3–4 kbp; 5, 1.5–3 kbp; 6, <1.5 kbp.
2) Proportions of chimeric clones are calculated on the analysis of 3264 clones sequenced from both ends.
3) Proportions of clones with full ORFs are estimated from 25 randomly selected clones from each fraction.

Bell and colleagues, Chicago University, and those of 2055 clones by the I.M.A.G.E. consortium). As mice are widely used as experimental animals in studies of pancreatic islets,[1] a cDNA library specific for mouse pancreatic islets should be useful. Recently, a recombination-based method for cDNA library construction has been developed.[13] With size fractionation of the cDNA inserts, this method provides several improvements.[13,14] The characteristics of the library are a high degree of complexity, an abundance of full-length sequences, and multiple applications. Libraries constructed by this method can readily be converted into other libraries including yeast two-hybrid and mammalian expression libraries.[15]

In the present study, we constructed a recombination-based cDNA library specific for mouse pancreatic islets. Based on the sequence data of the library, we built a mouse pancreatic islet sequence database, and also developed cDNA microarray membranes.

2. Materials and Methods

2.1. Construction of a cDNA library specific for mouse pancreatic islets

All animal procedures were approved by the Chiba University Animal Care Committee. Pancreatic islets were isolated from 8- to 10-week-old male C57BL/6 mice by hand-picking under the microscope and collagenase digestion method as described previously.[16] Poly(A)$^+$ RNA was obtained from the islets using RNeasy Mini kit (Qiagen GmbH, Hilden, Germany) and μMACS mRNA isolation kit (Miltenyi Biotech, Bergisch Gladbach, Germany). Two micrograms of poly(A)$^+$ RNA was isolated from approximately 5000 islets, from which cDNA synthesis was carried out as described previously.[14] The synthesized cDNA was size-fractionated by agarose gel electrophoresis into three segments (<1.5, 1.5–3, and >3 kbps, respectively), and subjected to recombination reaction with BP Clonase (Invitrogen, Carlsbad, CA) and plasmid attP pSP73.[14] Recombined plasmids were introduced into *Escherichia coli* cells (ElectroMAX DH10B cells, Invitrogen) by electroporation. The numbers of transformants from each fraction were counted to estimate the complexity of the library without any amplification (Table 1).

After transformants were grown, the retrieved plasmids from the largest size fraction (>3 kbps) were further size-fractionated into four segments on agarose gel. Thus, the extracted plasmids were finally size-separated into six segments (<1.5, 1.5–3, 3–4, 4–6, 6–8, and >8 kbps) and subjected to recombination reaction with LR clonase (Invitrogen) and *BamHI*-linearized plasmid attR pBC.[14] Recombined plasmids were again introduced into *E. coli* cells (ElectroMAX DH10B cells, Invitrogen) by electroporation. The direction of the cDNA inserts was T7-attB1-cDNA insert-attB2-T3.

2.2. Plasmid preparation and DNA sequencing

Plasmids for sequencing the cDNA inserts and spotting on microarray membranes were prepared with MAGNIA robot (TOYOBO, Osaka, Japan). With ABI3700 sequencers (Applied Biosciences, Mountain View, CA), 21,018 plasmid clones were successfully sequenced with T3 sequencing primer, and 3,264 clones were sequenced with T7 sequencing primer after elimination of redundant clones.

2.3. Construction of a sequence database

The sequence data obtained using T3 or T7 primers were subjected to BLAST search against the GenBank/DDBJ/EMBL nucleotide sequence database (www.ncbi.nlm.nih.gov/genbank/index.html) with a threshold E-value of $1.0e^{-50}$. The sequences were also subjected to BLAST search against the Ensembl mouse genome database (www.ensembl.org) to determine their locations on mouse chromosomes. All the sequence data were submitted to the DDBJ/EMBL/GenBank nucleotide sequence database (Accession nos. BP753069-BP777127), and are available from the ftp site at ftp://ftp.kazusa.or.jp/pub/pancreas/. The summary of the homology searches and the genome mapping data of the analyzed cDNAs are also available (Supplemental Information 1, http://www.dna-res.kazusa.or.jp/

11/5/01/supplemental/information1.html).

2.4. *cDNA microarray analysis*

The method of cDNA microarray analysis was described previously.[17] A total of 8108 cDNA clones were spotted on the microarray membranes. For single microarray analysis, 10 μg of total RNA from each sample was reverse-transcribed in the presence of SuperScriptII Reverse Transcriptase (Invitrogen), [^{35}S]dCTP, oligo(dT)$_{12-18}$, and oligo(dT)$_{25}$. The signal intensities in each analysis were normalized against the housekeeping gene Gapd, and the average signal intensities were calculated from more than two independent experiments.

2.5. *Cell culture and recombinant adenoviruses*

Mouse pancreatic islet β-cell line MIN6 cells were cultured as described previously.[18] Mouse ES cell line R1 cells[19] were cultured on mitomycin C-treated embryonic fibroblasts in a gelatinized dish with complete ES medium: high-glucose Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 20% fetal calf serum (FCS) (Invitrogen), 2 mM L-glutamine (Invitrogen), 1×nonessential amino acids (Invitrogen), 1 μM 2-mercaptoethanol (Invitrogen), 100 mg/ml streptomycin sulfate, 60.5 mg/l penicillin G (Invitrogen), and 1000 U/ml leukemia inhibitory factor (Chemicon, Temecula, CA). Mouse pituitary cell line AtT-20 cells[20] were grown in Ham's F10 medium (Invitrogen) with 5% FBS (Sigma, St. Louis, MO), and 15% heat-inactivated horse serum (Invitrogen).

Neurogenic differentiation 1 (NeuroD1/BETA2) recombinant adenovirus (Ad.CMVNeuroD1) and LacZ recombinant adenovirus (Ad.CMVLacZ) were constructed to express human NeuroD1 mRNA or *E. coli* LacZ (β-galactosidase) mRNA under cytomegalovirus (CMV) promoter, using an adenovirus expression vector kit (Takara Bio Inc., Otsu, Japan). MIN6 cells were infected with Ad.CMVNeuroD1 or Ad.CMVLacZ at a range of 5, 10, 30 and 50 multiplicity of infections (MOI), and cultured for another 2 days. Expression of NeuroD1 was confirmed by immunoblot analysis using antibodies specific for NeuroD1/BETA2 (Santa Cruz Biotechnology, Santa Cruz, CA) (Ishizuka et al. unpublished data), and expression of LacZ was confirmed using a galactosidase reporter gene staining detection kit (Sigma).

3. Results and Discussion

3.1. *Construction of a mouse pancreatic islet cDNA library based on a recombination-based method*

To construct a high quality cDNA library specific for normal mouse pancreatic islets, we applied a recombination-based method for library construction. Using 2 μg of poly(A)$^+$ RNA extracted from approximately 5000 mouse pancreatic islets, we successfully constructed a cDNA library having a complexity of more than 3.12×10^7 without amplification (Table 1). The library consists of six size fractions, which minimizes the size-bias effect on the population of cDNAs and results in improved variety.[14] The library is also characterized by an abundance of full-length cDNA clones and fewer chimeric clones. Since each cDNA insert has specific sequences for recombination (compatible with Invitrogen GATEWAY^{TM} system) at both ends, the library can readily be converted into other libraries such as mammalian expression and yeast two-hybrid libraries. This is the first pancreatic islet cDNA library created based on a recombination-based method, and should be a useful tool for pancreatic islet studies.

3.2. *Construction of a cDNA sequence database specific for mouse pancreatic islets*

We then sequenced a portion of the cDNA library to construct a sequence database specific for mouse pancreatic islets. We collected 4608 plasmid clones from each fraction, resulting in a total of 27,648 clones, and successfully sequenced 21,018 clones from the 3′ region of the cDNAs. This raw sequence data contained about 13.7% of insulin 1 or 2 (Ins1/2) sequences and only 0.03% amylase sequences (representative gene of exocrine pancreas), indicating that the library was constructed primarily from pancreatic islets.

After eliminating redundant clones, we obtained a total of 8882 non-redundant clones (Supplemental Information 2, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information2.html). Among them, we also sequenced 3264 clones from the 5′ region of the cDNAs, and detected only 114 chimeric clones by comparison with the Ensembl mouse genome database (www.ensemble.org). By analyzing 150 randomly selected clones, we found that about 60% of the clones contained full open reading frames (ORFs) (Table 1). We collected these results and built a mouse pancreatic islet sequence database.

We compared 8882 non-redundant sequence data with the "all genes" database (www.allgenes.org), and found that 261 had no match in the database. Of 8621 sequences, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology (GO) Consortium (www.geneontology.org) (Fig. 1, Table 2 and Supplemental Information 3, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information3.html), while 2822 have no GO annotation, suggesting that these were derived from non-coding RNAs or rare transcripts. Of the 5799 sequences having GO annotation, about 42% belong to the largest category "binding," which includes nucleic acid binding and peptide/protein binding. The second category "catalytic activity" comprises about 25%, and the third, "molecular function un-

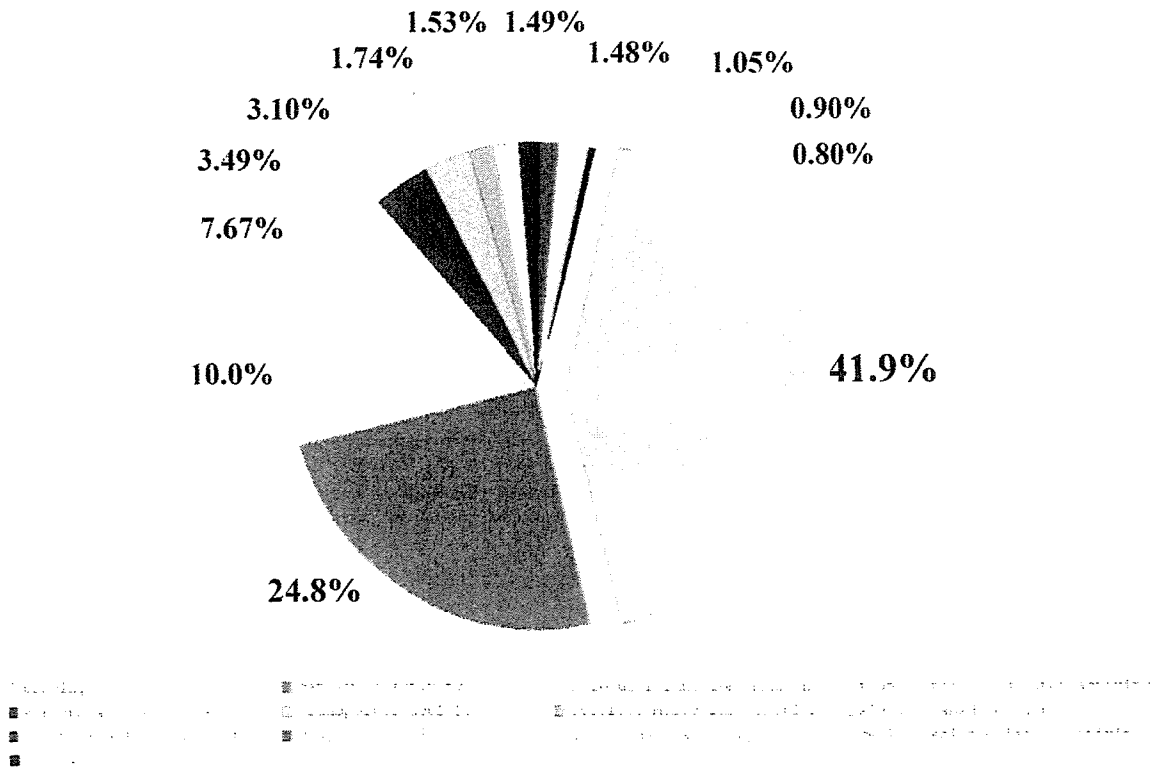

Figure 1. Functional classification of the mouse pancreatic islet cDNA clones. Out of 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium. Representative genes for each functional category are listed in Table 2. A complete list is available in Supplemental Information 3. http://www.dna-res.kazusa.or.jp/11/5/03/supplemental/information3.html.

known." about 10%.

Although the distributions of GO annotations for mRNAs expressed in mouse whole pancreas (including exocrine and endocrine glands) have previously been reported,[21,22] those for pancreatic islets have not been available. Our present data are the first to show the distributions of GO annotations for mRNAs from mouse pancreatic islets. Possibly due to differences in materials or the method of library construction, our data exhibit some differences with previous reports.[21,22] For example, we found that "antioxidant" genes showed 1.1% distribution rather than 0.2%,[21] and "signal transducer" genes showed 1.6% rather than 10%.[22] In contrast, we detected "catalytic activity" genes at 24.8% distribution, which was similar to their "enzyme" classifications of 26.9%[21] and 31%.[22] We found 10% "molecular function unknown" genes, while they report only 3%.[22] Thus, gene expression profiles of pancreatic islets could provide valuable information on the molecules involved in normal function of the endocrine pancreas.

3.3. Application to microarray analyses

We then applied our cDNA library to microarray analyses. We selected a total of 8108 cDNA clones from 8882 non-redundant clones and spotted them on microarray membranes with 12 control clones (Supplemental Information 2, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information2.html). To estimate the reproducibility of our microarray analyses, we compared membranes of different lots hybridized with the same probe. Scattered plot analysis showed very similar signal intensities from the two different membranes, indicating that our microarray analyses are reproducible (Supplemental Information 4A, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information4.html).

To evaluate the usefulness and applicability of our microarray membranes, we performed the following two experiments. First, we compared the expression profiles of three cell lines: MIN6 cells (mouse insulin-secreting pancreatic islet β-cell line), mouse embryonic stem (ES) cells, and AtT-20 cells (mouse pituitary endocrine cell line), representing endocrine pancreatic islet cells, undifferentiated cells,

Table 2. A list of functional classifications of the mouse pancreatic islet cDNA clones.

GO function group	(subgroup)	annotation	GENBANK accesion No (gb)	gene index (gi)	unigene cluster (ug)
Binding	Nucleic acid binding	Mm#S2205607 Mus musculus H2B histone family, member [illegible], clone MGC:19259 IMAGE:3989862, mRNA, complete cds	gb=BC011440	gi=15030325	ug=Mm.21579
		Mm#S937714 Mus musculus H3 histone, family 3A (H3f3a), mRNA	gb=NM_008210	gi=6680158	ug=Mm.89136
		Mm#S947247 Mus musculus poly A binding protein, cytoplasmic 1 (Pabpc1) mRNA	gb=NM_008774	gi=6679198	ug=Mm.2642
		Mm#S2608854 Mus musculus dicer-like protein (Dicer1) mRNA, complete cds	gb=AF430845	gi=20385912	ug=Mm.21523
		Mm#S978931 Mus musculus poly(rC) binding protein 2 (Pcbp2) mRNA	gb=NM_011042	gi=6997238	ug=Mm.111
	Peptide Protein binding	Mm#S1085421 Mus musculus Ran binding protein 5 mRNA, partial cds	gb=AF294327	gi=12034715	ug=Mm.151529
		Mm#S2204674 Mus musculus, TAF binding protein, clone MGC:13789 IMAGE:4223922, mRNA, complete cds	gb=BC015074	gi=15929261	ug=Mm.14697
		Mm#S939589 Mus musculus syntaxin binding protein 2 (Stxbp2) mRNA	gb=NM_011503	gi=6755687	ug=Mm.7247
		Mm#S937219 Mus musculus nuclear receptor-binding SET-domain protein 1 (Nsd1), mRNA	gb=NM_008739	gi=6679137	ug=Mm.12964
		Mm#S1600484 Mus musculus piccolo (presynaptic cytomatrix protein) (Pclo) mRNA	gb=NM_011995	gi=15277339	ug=Mm.40880
	others binding	Mm#S939404 Mus musculus retinol binding protein 1, cellular (Rbp1) mRNA	gb=NM_011254	gi=6755299	ug=Mm.2450
	(lipid, carboh, drate, etc.)	Mm#S938389 Mus musculus apolipoprotein E (Apoe) mRNA	gb=NM_009696	gi=6753101	ug=Mm.138866
		Mm#S937544 Mus musculus FK506 binding protein 2 (13 kDa) (Fkbp2) mRNA	gb=NM_008020	gi=6679804	ug=Mm.4234
		Mm#S979142 Mus musculus calcium binding protein 1 (Cabp1), mRNA	gb=NM_013879	gi=7304938	ug=Mm.36551
		Mm#S1598822 Mus musculus lectin, galactose binding, soluble 12 (Lgals12) mRNA	gb=NM_019516	gi=9506759	ug=Mm.87546
Catalytic activity	hydrolase activity	Mm#S937421 Mus musculus protein tyrosine phosphatase, non-receptor type 2 (Ptpn2) mRNA	gb=NM_008977	gi=6679552	ug=Mm.985
		Mm#S937616 Mus musculus acid beta glucosidase (Gba), mRNA	gb=NM_008094	gi=6679954	ug=Mm.5031
		Mm#S978159 Mus musculus carboxypeptidase E (Cpe) mRNA	gb=NM_013494	gi=7304972	ug=Mm.31395
		Mm#S978344 Mus musculus proprotein convertase subtilisin/kexin type 1 (Pcsk1) mRNA	gb=NM_013628	gi=7305376	ug=Mm.1333
		Mm#S937265 Mus musculus proprotein convertase subtilisin/kexin type 2 (Pcsk2) mRNA	gb=NM_008792	gi=6679228	ug=Mm.1247
	ligase activity	Mm#S1059886 Mus musculus cytidine 5'-triphosphate synthase 2 (Ctps2) mRNA	gb=NM_018737	gi=9055197	ug=Mm.2465
		Mm#S2203048 Mus musculus propionyl Coenzyme A carboxylase, beta polypeptide (Pccb), mRNA	gb=NM_025835	gi=13385309	ug=Mm.21079
		Mm#S2203110 Mus musculus phosphoribosylaminoimidazole carboxylase, phosphoribosylaminoribosylaminoimidazole, succinocarboxamide synthetase (Paics), mRNA	gb=NM_025939	gi=13385433	ug=Mm.182931
		Mm#S2534128 Mus musculus, glutamate-cysteine ligase, catalytic subunit, clone MGC:30487 IMAGE:4195425, mRNA, complete cds	gb=BC019374	gi=18043914	ug=Mm.4368
		Mm#S2581467 Mus musculus, similar to alanyl-tRNA synthetase (H. sapiens), clone MGC:37868 IMAGE:4976684, mRNA, complete cds	gb=BC026611	gi=20072364	ug=Mm.24174
	oxidoreductase activity	Mm#S1060389 Mus musculus NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1 (Ndufa1), mRNA	gb=NM_019443	gi=9506919	ug=Mm.34869
		Mm#S1060493 Mus musculus pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyde synthetase) (Pycs), mRNA	gb=NM_019698	gi=9790066	ug=Mm.29751
		Mm#S2202811 Mus musculus sterol-C4-methyl oxidase-like (Sc4mol), mRNA	gb=NM_025436	gi=13384635	ug=Mm.30119
		Mm#S2455829 Mus musculus cytochrome c oxidase, subunit VIc (Cox6c), mRNA	gb=NM_053071	gi=16716342	ug=Mm.548
		Mm#S936625 Mus musculus stearoyl-Coenzyme A desaturase 2 (Scd2), mRNA	gb=NM_009128	gi=6677862	ug=Mm.193096
	transferase activity	Mm#S1427353 Mus musculus histone acetyltransferase (Ne-tl-pending), mRNA	gb=NM_017479	gi=8567375	ug=Mm.30698
		Mm#S1600490 Mus musculus m6a methyltransferase (Ime4-pending), mRNA	gb=NM_019721	gi=9790038	ug=Mm.76983
		Mm#S1085402 Mus musculus prenylcysteine carboxylmethyltransferase mRNA, partial cds	gb=AF209920	gi=12082482	ug=Mm.29356
		Mm#S2455862 Mus musculus, transglutaminase 2, C polypeptide, clone MGC:6152 IMAGE:3256947, mRNA, complete cds	gb=BC016492	gi=16741319	ug=Mm.18843
		Mm#S2003659 Mus musculus glycogen synthase 1, muscle (Gys1), mRNA	gb=NM_030678	gi=13507596	ug=Mm.185247
molecular function unknown		Mm#S1971846 Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone:2900059A22:DiGeorge syndrome chromosome region 6, full insert sequence	gb=AK019346	gi=12859501	ug=Mm.27155
		Mm#S1972121 Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730433K22:unclassifiable, full insert sequence	gb=AK017611	gi=12856645	ug=Mm.158490

Table 2. Continued.

GO function group	(subgroup)	annotation	GENBANK accession No (gb)	gene index (gi)	unigene cluster (ug)
		Mm#S2611167 Mus musculus RIKEN cDNA 2988016D05 gene (2988016D05Rik), mRNA	gb=NM_028381	gi=21312170	ug=Mm.22796
		Mm#S2533711 Mus musculus, Similar to hypothetical protein FLJ10276, clone IMAGE:5344306, mRNA, partial cds	gb=BC021480	gi=18204810	ug=Mm.17918
		Mm#S2049845 Mus musculus, Similar to hypothetical protein, clone MGC:7703 IMAGE:3497634, mRNA, complete cds	gb=BC008544	gi=14250242	ug=Mm.58660
transcription regulator activity		Mm#S937155 Mus musculus myelin transcription factor 1 (Myt1), mRNA	gb=NM_008665	gi=6678997	ug=Mm.20888
		Mm#S2635265 Mus musculus NK6 transcription factor related, locus 1 (Drosophila) (Nkx6-1), mRNA	gb=NM_144955	gi=21450628	ug=Mm.108072
		Mm#S979348 Mus musculus paired box gene 6 (Pax6), mRNA	gb=NM_013627	gi=7305368	ug=Mm.3608
		Mm#S2050432 Mus musculus sterol regulatory element binding protein 1 (Srebp1) mRNA, partial cds	gb=AF374266	gi=14161490	ug=Mm.214958
		Mm#S937735 Mus musculus hairy and enhancer of split 1, (Drosophila) (Hes1) mRNA	gb=NM_008235	gi=6680204	ug=Mm.4451
structural molecule activity		Mm#S121804 Tubulin, alpha 2, mRNA	gb=M13446	gi=202203	ug=Mm.107515
		Mm#S2577772 Mus musculus dystonin (Dst), mRNA	gb=NM_134448	gi=19882220	ug=Mm.25826
		Mm#S937847 Mus musculus integral membrane glycoprotein (Imp), mRNA	gb=NM_008377	gi=6680444	ug=Mm.944
		Mm#S938318 Mus musculus actin, gamma, cytoplasmic (Actg), mRNA	gb=NM_009609	gi=6752955	ug=Mm.196173
		Mm#S979073 Mus musculus ribosomal protein L13a (Rpl13a), mRNA	gb=NM_009438	gi=7110730	ug=Mm.13020
transporter activity		Mm#S979102 Mus musculus ATP-binding cassette, sub-family A (ABC1), member 1 (Abca1), mRNA	gb=NM_013454	gi=7304848	ug=Mm.369
		Mm#S979116 Mus musculus ATP-binding cassette, sub-family C (CFTR/MRP), member 5a (Abcc5a), mRNA	gb=NM_013790	gi=7304855	ug=Mm.20845
		Mm#S1659975 Mus musculus solute carrier family 1 (glutamate/neutral amino acid transporter), member 4 (Slc1a4), mRNA	gb=NM_018861	gi=9256643	ug=Mm.6379
		Mm#S2007281 Mus musculus solute carrier family 2 (facilitated glucose transporter), member 2 (Slc2a2), mRNA	gb=NM_031197	gi=13654261	ug=Mm.18443
		Mm#S1660140 Mus musculus solute carrier family 2 (facilitated glucose transporter), member 5 (Slc2a5), mRNA	gb=NM_019741	gi=9789986	ug=Mm.34150
obsolete molecular function		Mm#S1660236 Mus musculus SEC23B (S. cerevisiae) (Sec23b), mRNA	gb=NM_019787	gi=9790212	ug=Mm.28704
		Mm#S1437100 Mus musculus erythrocyte protein band 7.2 (Epb7.2), mRNA	gb=NM_013515	gi=7710017	ug=Mm.4441
		Mm#S936925 Mus musculus unc5 homolog (C. elegans) 3 (Unc5h3), mRNA	gb=NM_009472	gi=6678514	ug=Mm.24430
		Mm#S937063 Mus musculus mesoderm specific transcript (Mest), mRNA	gb=NM_008590	gi=6679865	ug=Mm.1089
		Mm#S938571 Mus musculus COP9 (constitutive photomorphogenic) homolog subunit 4 (Arabidopsis thaliana) (Cops4), mRNA	gb=NM_012001	gi=6753480	ug=Mm.957
signal transducer activity		Mm#S939854 Mus musculus growth hormone receptor (Ghr), mRNA	gb=NM_010284	gi=6857794	ug=Mm.3986
		Mm#S1659847 Mus musculus G protein-coupled receptor 56 (Gpr56), mRNA	gb=NM_018882	gi=9256530	ug=Mm.12790
		Mm#S2050287 Mus musculus, vitamin D receptor, clone MGC:12147 IMAGE:3710866, mRNA, complete cds	gb=BC006716	gi=13879474	ug=Mm.44170
		Mm#S2609654 Mus musculus notch 1 protein mRNA, complete cds	gb=AF508809	gi=20805340	ug=Mm.31255
		Mm#S979471 Mus musculus prostaglandin E receptor 1 (subtype EP1), 42kD (Ptger1), mRNA	gb=NM_013641	gi=7363446	ug=Mm.213000
enzyme regulator activity		Mm#S939211 Mus musculus neutral sphingomyelinase (N-SMase) activation associated factor (Nsmaf), mRNA	gb=NM_010945	gi=6754807	ug=Mm.3059
		Mm#S2581574 Mus musculus guanine nucleotide exchange factor (Larg) mRNA, complete cds	gb=AF467766	gi=19387125	ug=Mm.101659
		Mm#S938644 Mus musculus tissue inhibitor of metalloproteinase 3 (Timp3), mRNA	gb=NM_011595	gi=6755792	ug=Mm.4871
		Mm#S937002 Mus musculus Rho guanine nucleotide exchange factor (GEF) 1 (Arhgef1), mRNA	gb=NM_008488	gi=6678667	ug=Mm.3181
		Mm#S939398 Mus musculus RAS protein-specific guanine nucleotide-releasing factor 1 (Rasgrf1), mRNA	gb=NM_011245	gi=6755287	ug=Mm.2921
chaperone activity		Mm#S1997753 Mus musculus heat shock 70kD protein 5 (glucose-regulated protein, 78kD) (Hspa5), mRNA	gb=NM_022310	gi=11612488	ug=Mm.918
		Mm#S978930 Mus musculus osmotic stress protein 94 kDa (Osp94), mRNA	gb=NM_011020	gi=6997236	ug=Mm.4150
		Mm#S2007370 Mus musculus DnaJ (Hsp40) homolog, subfamily A, member 3 (Dnaja3), mRNA	gb=NM_023646	gi=13994154	ug=Mm.28330
		Mm#S939134 Mus musculus DnaJ (Hsp40) homolog, subfamily B, member 6 (Dnajb6), mRNA	gb=NM_011847	gi=6754735	ug=Mm.2701
		Mm#S1660317 Mus musculus prefoldin 5 (Pfdn5), mRNA	gb=NM_020031	gi=9910215	ug=Mm.181847

Table 2. Continued.

GO function (subgroup)	annotation	GENBANK accession No. (gb)	gene index (gi)	unigene cluster (ug)
antioxidant activity	Mm#S937[illegible] Mus musculus peroxiredoxin 3 (Prdx3), mRNA	gb=NM_007452	gi=6678660	ug=Mm.20821
	Mm#S14374[illegible] Mus musculus peroxiredoxin 1 (Prdx1), mRNA	gb=NM_011063	gi=6755[illegible]	ug=Mm.[illegible]
	Mm#S82[illegible] Mus musculus thioredoxin 1 (Txn1), mRNA	gb=NM_011660	gi=6755[illegible]	ug=Mm.1275
	Mm#S[illegible]28 Mus musculus thioredoxin 2 (Txn2), mRNA	gb=NM_[illegible]13	gi=6[illegible]8	ug=Mm.3583
	Mm#S2455825 Mus musculus glutaredoxin 1 (thiotransferase) (Glrx1), mRNA	gb=NM_053108	gi=[illegible]	ug=Mm.[illegible]
cell adhesion molecule activity	Mm#S[illegible]5 Mus musculus integrin alpha 6 (Itga6), mRNA	gb=NM_[illegible]	gi=[illegible]	ug=Mm.25222
	Mm#S4156 Human mRNA for integrin beta subunit, mRNA	gb=X[illegible]	gi=52721	ug=Mm.4712
	Mm#S3085 Procollagen, type I, alpha 1, mRNA	gb=[illegible]	gi=47[illegible]	ug=Mm.22621
	Mm#S2080 Procollagen, type IV, alpha 1, mRNA	gb=[illegible]	gi=55[illegible]	ug=Mm.7[illegible]
	Mm#S244281 Procollagen, type XV, mRNA	gb=AF[illegible]	gi=2558824	ug=Mm.4[illegible]

Out of 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium. Representative genes for each functional category are listed in this table. A complete list is available in Supplemental Information 3. http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information3.html.

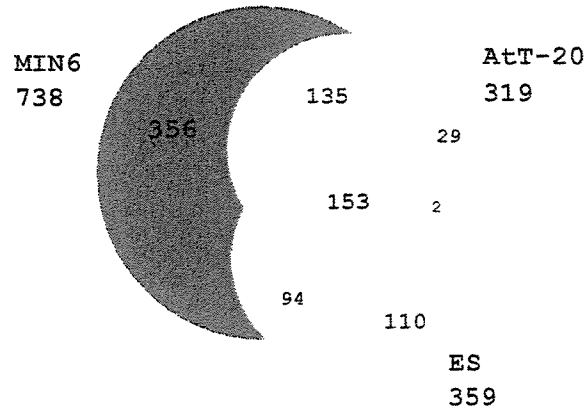

Figure 2. Comparison of expression profiles in MIN6, ES and AtT-20 cells. Genes with signal intensity of one-tenth or more of the housekeeping gene Gapd were compared in the three cell lines. Numbers of genes are shown in this diagram. A complete list of the expressed genes is available in Supplemental Information 5. http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html.

and other endocrine cells, respectively (Supplemental Information 4B and 4C, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information4.html). We used a threshold signal intensity of one-tenth that of the housekeeping gene Gapd, and compared the expressed genes in the three cell lines (Fig. 2). A list of the expressed genes is available (Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html). Among a total of 879 genes, 738 (84.0%) were expressed in MIN6 cells, while only 319 (36.3%) and 359 (40.8%) were expressed in AtT-20 and ES cells, respectively.

Comparison of expression profiles of three cell lines revealed both differentially and similarly expressed genes. A subset of 135 genes was expressed in both MIN6 and AtT-20 cells but not in ES cells. As both MIN6 and AtT-20 cells were derived from endocrine cells and preserve regulated exocytotic ability, the genes involved should be expressed in both endocrine cell types. Indeed, well-known endocrine-specific genes such as chromogranins,[23,24] prohormone convertase 1/3,[25] and carboxypeptidase E[26,27] were in this subset (Fig. 2 and Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html). In addition, 28 genes with unknown functions were contained in this subset, suggesting they may have other roles in regulated exocytosis.

By contrast, a subset of 356 genes was found to be expressed only in MIN6 cells. This subset included known pancreatic β-cell-specific genes, such as Ins1/2, Iapp, and Pdx1 as well as many uncharacterized genes (Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html) that may be involved in phenomena specific to pancreatic β-cells such as insulin synthesis. Thus, comparison of the mRNA expression patterns of various cell types should be useful in investigating the biological function of the clones newly identified in our cDNA library.

As the second experiment, we compared expression profiles between MIN6 cells and those infected with aden-

Table 3. A list of genes upregulated by overexpression of NeuroD1 in MIN6 cells.

No.	Signal intensity +LacZ[1)]	Signal intensity +NeuroD1[2)]	Ratio	Gene name[3)] / GenBank accession number
1	4.70	84.31	17.9324	gi: 26101046 / AK083076
2	14.59	251.26	17.2207	Neurod1 (NeuroD1)[4)] / NM_010894
3	3.60	24.37	6.7660	gi: 38328296 / BC062185
4	1.61	7.60	4.7122	Scnn1b / NM_011325
5	4.26	18.02	4.2319	Ipf1 (Pdx1) / NM_008814
6	16.42	62.79	3.8232	Cdkn1a (p21) / NM_007669
7	2.07	7.51	3.6218	Jund1 / NM_010592
8	7.48	24.16	3.2285	Ddx5 / BC009142
9	5.25	15.33	2.9198	Rangap1 / NM_011241
10	4.98	14.42	2.8952	gi: 12833685 / AK003180
11	4.47	12.80	2.8639	Atp2a2 / NM_009722
12	2.65	7.53	2.8440	Scly / NM_016717
13	7.50	19.71	2.6274	gi: 17391148 / BC018486
14	6.66	17.30	2.5976	Actn4 / NM_021895

Signal intensities are normalized by two independent experiments.
1) Signal intensities from MIN6 cells infected with Ad.CMVLacZ.
2) Signal intensities from MIN6 cells infected with Ad.CMVNeuroD1.
3) gi (gene index number in GenBank) or Gene name registered in NCBI Entrez Gene database.
4) Signal intensity of NeuroD1 from MIN6 cells infected with Ad.CMVNeuroD1 was influenced by the exogenous introduction of NeuroD1.

oviral vector to express the transcription factor NeuroD1 exogenously. Because NeuroD1 is critical in pancreatic islet β-cell development[6] and β-cell functions,[28] and a mutation in the human NeuroD1 gene is responsible for maturity-onset diabetes of the young (MODY6),[11,12] we overexpressed NeuroD1 in MIN6 cells to evaluate the usefulness of our microarray analyses in the search for novel target genes. So far, Pdx1,[29] p21,[30] and Sur1[31] have been reported to be transactivated by NeuroD1.

Infection of the adenoviral vector with a range of 5 or 10 MOI did not alter the expression profiles, but infection with 30 or 50 MOI caused significant alterations (data not shown). The results of infection with 50 MOI are shown in Table 3 and Supplemental Information 4D, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information4.html. We confirmed that the genes known to be induced by NeuroD1, including Pdx1[29] and p21,[30] were upregulated, and also found that expression levels of another 11 genes increased significantly. None of these genes has previously been associated with NeuroD1, and all are candidates for transactivation by NeuroD1. This second experiment shows that our microarray analyses will be useful in the search for novel molecular targets of various β-cell transcription factors and in investigations of pancreatic islets and related diseases including diabetes.

Acknowledgements: We thank Y. Murahashi for assistance with computational analysis of nucleotide sequences. We also thank T. Watanabe and T. Hirano for technical assistance in constructing the mouse pancreatic islet library and in the microarray analysis. This work was supported by Grants-in-Aid for Specially Promoted Research and for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan; by a Grant from the Ministry of Economy, Trade and Industry, Japan, and the Futaba Electronics Memorial Foundation.

References

1. Kahn, C. R., Gordon, C., Weir, G. C. et al. 1994. *Joslin's Diabetes Mellitus, 13th Ed.*, Lippincott Williams & Wilkins, Philadelphia, PA.
2. Jonsson, J., Carlsson, L., Edlund, T., Edlund, H. 1994. Insulin-promoter-factor 1 is required for pancreas development in mice, *Nature*, **13**, 606–609.
3. Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., and Edlund, H. 1998, Beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes, *Genes Dev.*, **12**, 1763–1768.
4. Lee, Y. H., Sauer, B., and Gonzalez, F. J. 1998, Laron dwarfism and non-insulin-dependent diabetes mellitus in the Hnf-1alpha knockout mouse, *Mol Cell Biol.*, **18**, 3059–3068.

5. Shih, D. Q., Screenan, S., Munoz, K. N. et al. 2001, Loss of HNF-1alpha function in mice leads to abnormal expression of genes involved in pancreatic islet development and metabolism. *Diabetes*, **50**, 2472–2480.
6. Naya, F. J., Huang, H. P., Qiu, Y. et al. 1997, Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice, *Genes Dev.*, **11**, 2323–2334.
7. Stoffers, D. A., Ferrer, J., Clarke, W. L., and Habener J. F. 1997, Early-onset type-II diabetes mellitus (MODY4) linked to IPF1, *Nat Genet*, **17**, 138–139.
8. Yamagata, K., Oda, N., Kaisaki, P. J. et al. 1996, Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3), *Nature*, **384**, 455–458.
9. Yamagata, K., Furuta, H., Oda, N. et al. 1996, Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1), *Nature*, **384**, 458–460.
10. Horikawa, Y., Iwasaki, N., Hara, M. et al. 1997, Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY, *Nat Genet*, **17**, 384–385.
11. Malecki, M. T., Jhala, U. S., Antonellis, A. et al. 1999, Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus, *Nat Genet*, **23**, 323–328.
12. Bell G. I. and Polonsky, K. S. 2001, Diabetes mellitus and genetically programmed defects in beta-cell function, *Nature*, **414**, 788–791.
13. Ohara, O. and Temple, G. 2001, Directional cDNA library construction assisted by the in vitro recombination reaction, *Nucleic Acids Res.*, **29**, E22.
14. Ohara, O., Nagase, T., Mitsui, G. et al. 2002, Characterization of size-fractionated cDNA libraries generated by the in vitro recombination-assisted method, *DNA Res.*, **9**, 47–57.
15. Nakayama, M., Kikuno, R., and Ohara, O. 2002, Protein-protein interactions between large proteins: two-hybrid screening using a functionally classified library composed of long cDNAs, *Genome Res.*, **12**, 1773–1784.
16. Miki, T., Nagashima, K., Tashiro, F. et al. 1998, Defective insulin secretion and enhanced insulin action in KATP channel-deficient mice, *Proc. Natl. Acad. Sci. USA*, **95**, 10402–10406.
17. Mitsui, G., Mitsui, K., Hirano, T., Ohara, O., Kato M., and Niwano, Y. 2003, Kinetic profiles of sequential gene expressions for chemokines in mice with contact hypersensitivity, *Immunol Lett.*, **86**, 191–197.
18. Kawaki, J., Nagashima, K., Tanaka, J. et al. 1999, Unresponsiveness to glibenclamide during chronic treatment induced by reduction of ATP-sensitive K^+ channel activity, *Diabetes*, **48**, 2001–2006.
19. Suzuki, M., Mizutani-Koseki, Y., Fujimura, Y. et al. 2002, Involvement of the Polycomb-group gene Ring1B in the specification of the anterior-posterior axis in mice, *Development*, **129**, 4171–4183.
20. Buonassisi, V., Sato, G., and Cohen, AI. 1962, Hormone-producing cultures of adrenal and pituitary tumor origin, *Proc. Natl. Acad. Sci. USA*, **48**, 1184–1190.
21. Scearce, L. M., Brestelli, J. E., McWeeney, S. K. et al. 2002, Functional genomics of the endocrine pancreas: the pancreas clone set and PancChip, new resources for diabetes research, *Diabetes*, **51**, 1997–2004.
22. Kaestner, K. H., Lee, C. S., Scearce, L. M. et al. 2003, Transcriptional program of the endocrine pancreas in mice and humans, *Diabetes*, **52**, 1604–1610.
23. Ehrhart, M., Grube, D., Bader, M. F., Aunis, D., and Gratzl, M. 1986, Chromogranin A in the pancreatic islet: cellular and subcellular distribution, *J. Histochem Cytochem.*, **34**, 1673–1682.
24. Iacangelo, A., Affolter, H. U., Eiden, L. E., Herbert, E., and Grimes, M. 1986, Bovine chromogranin A sequence and distribution of its messenger RNA in endocrine tissues, *Nature*, **323**, 82–86.
25. Zhu, X., Zhou, A., Dey, A. et al. 2002, Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects, *Proc. Natl. Acad. Sci. USA*, **99**, 10293–10298.
26. Fricker, L. D., Adelman, J. P., Douglass, J., Thompson, R. C., von Strandmann, RP., and Hutton, J. 1989, Isolation and sequence analysis of cDNA for rat carboxypeptidase E [EC 3.4.17.10], a neuropeptide processing enzyme, *Mol. Endocrinol.*, **3**, 666–673.
27. Naggert, J. K., Fricker, L. D., Varlamov, O. et al. 1995, Hyperproinsulinaemia in obese fat/fat mice associated with a carboxypeptidase E mutation which reduces enzyme activity, *Nat. Genet*, **10**, 135–142.
28. Huang, H. P. Chu, K. Nemoz-Gaillard, E. Elberg, D. and Tsai, M. J. 2002, Neogenesis of beta-cells in adult BETA2/NeuroD-deficient mice, *Mol Endocrinol.*, **16**, 541–551.
29. Kojima, H., Fujimiya, M., Matsumura, K. et al. 2003, NeuroD-betacellulin gene therapy induces islet neogenesis in the liver and reverses diabetes in mice, *Nat. Med.*, **9**, 596–603.
30. Mutoh, H., Naya, F. J., Tsai, M. J., and Leiter, A. B. 1998, The basic helix-loop-helix protein BETA2 interacts with p300 to coordinate differentiation of secretin-expressing enteroendocrine cells, *Genes Dev.*, **12**, 820–830.
31. Kim J. W., Seghers V., Cho J. H., et al. 2002, Transactivation of the mouse sulfonylurea receptor 1 gene by BETA2/NeuroD, *Mol. Endocrinol.*, **16**, 1097–1107.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 49, Issue of December 3, pp. 51376–51385, 2004
Printed in U.S.A.

Physical and Functional Interaction between Dorfin and Valosin-containing Protein That Are Colocalized in Ubiquitylated Inclusions in Neurodegenerative Disorders*

Received for publication, June 15, 2004, and in revised form, August 31, 2004
Published, JBC Papers in Press, September 29, 2004, DOI 10.1074/jbc.M406683200

Shinsuke Ishigaki‡§¶, Nozomi Hishikawa‡, Jun-ichi Niwa‡, Shun-ichiro Iemura‖, Tohru Natsume‖, Seiji Hori, Akira Kakizuka**‡‡, Keiji Tanaka§, and Gen Sobue‡§§**

*From the ‡Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya 466-8500, Japan, the §Department of Molecular Oncology, Tokyo Metropolitan Institute of Medical Science, Tokyo 113-8613, Japan, the ‖National Institute of Advanced Science and Technology, Biological Information Research Center, Tokyo 135-0064, Japan, the **Laboratory of Functional Biology, Kyoto University Graduate School of Biostudies, Kyoto 606-8502, Japan, and ‡‡CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan*

Dorfin, a RING-IBR type ubiquitin ligase (E3), can ubiquitylate mutant superoxide dismutase 1, the causative gene of familial amyotrophic lateral sclerosis (ALS). Dorfin is located in ubiquitylated inclusions (UBIs) in various neurodegenerative disorders, such as ALS and Parkinson's disease (PD). Here we report that Valosin-containing protein (VCP) directly binds to Dorfin and that VCP ATPase activity profoundly contributes to the E3 activity of Dorfin. High through-put analysis using mass spectrometry identified VCP as a candidate of Dorfin-associated protein. Glycerol gradient centrifugation analysis showed that endogenous Dorfin consisted of a 400–600-kDa complex and was co-immunoprecipitated with endogenous VCP. *In vitro* experiments showed that Dorfin interacted directly with VCP through its C-terminal region. These two proteins were colocalized in aggresomes in HEK293 cells and UBIs in the affected neurons of ALS and PD. VCP^{K524A}, a dominant negative form of VCP, reduced the E3 activity of Dorfin against mutant superoxide dismutase 1, whereas it had no effect on the autoubiquitylation of Parkin. Our results indicate that VCPs functionally regulate Dorfin through direct interaction and that their functional interplay may be related to the process of UBI formation in neurodegenerative disorders, such as ALS or PD.

Amyotrophic lateral sclerosis (ALS)[1] is one of the most common neurodegenerative disorders, characterized by selective motor neuron degeneration in the spinal cord, brain stem, and cortex. Two genes, CuZn-superoxide dismutase (SOD1) and amyotrophic lateral sclerosis 2 have been identified as responsible genes for familial forms of ALS. Using mutant SOD1 transgenic mice, the pathogenesis of ALS has been partially uncovered. The proposed mechanisms of the motor neuron degeneration in ALS include oxidative toxicity, glutamate receptor abnormality, ubiquitin proteasome dysfunction, inflammatory and cytokine activation, dysfunction of neurotrophic factors, damage to mitochondria, cytoskeletal abnormalities, and activation of the apoptosis pathway (1, 2).

In a previous study (3), we identified several ALS-associated genes using molecular indexing. Dorfin was identified as one of the up-regulated genes in ALS, which contains a RING-IBR (in between ring finger) domain at its N terminus and mediated ubiquitin ligase (E3) activity (3, 4). Dorfin colocalized with Vimentin at the centrosome after treatment with a proteasome inhibitor in cultured cells (4). Dorfin physically bound and ubiquitylated various SOD1 mutants derived from familial ALS patients and enhanced their degradation, but it had no effect on the stability of wild-type SOD1 (5). Overexpression of Dorfin protected neural cells against the toxic effects of mutant SOD1 and reduced SOD1 inclusions (5).

Recent findings indicate that the ubiquitin-proteasome system is widely involved in the pathogenesis of Parkinson's disease (PD), Alzheimer's disease, polyglutamine disease, and Prion diseases as well as ALS (6). From this point of view, we previously analyzed the pathological features of Dorfin in various neurodegenerative diseases and found that Dorfin was predominantly localized not only in Lewy body (LB)-like inclusions in ALS but also in LBs in PD, dementia with Lewy bodies, and glial cell inclusions in multiple system atrophy (7). These characteristic intracellular inclusions composed of aggregated, ubiquitylated proteins surrounded by disorganized filaments are the histopathological hallmark of aging-related neurodegenerative diseases (8).

A structure called aggresome by Johnston *et al.* (9) is formed when the cell capacity to degrade misfolded proteins is exceeded. The aggresome has been defined as a pericentriolar, membrane-free, cytoplasmic inclusion containing misfolded ubiquitylated protein ensheathed in a cage of intermediate filaments, such as Vimentin (9). The formation of the aggresome mimics that of ubiquitylated inclusions (UBIs) in the affected neurons of various neurodegenerative diseases (10). Combined with the fact that Dorfin was localized in aggresomes in cultured cells and UBIs in ALS and other neurode-

* This work was supported by a grant for a Center of Excellence from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Research resident of the Japan Foundation for Aging and Health, Psychiatric and Neurological Diseases, and Mental Health.

§§ To whom correspondence should be addressed: Dept. of Neurology, Nagoya University Graduate School of Medicine, Nagoya 466-8500, Japan. Tel.: 81-52-744-2385; Fax: 81-52-744-2384; E-mail: sobueg@med.nagoya-u.ac.jp.

[1] The abbreviations used are: ALS, amyotrophic lateral sclerosis; E3, ubiquitin ligase; ERAD, endoplasmic reticulum-associated degradation; LB, Lewy body; MS, mass spectrometry; LC-MS/MS, liquid chromatography coupled to electrospray tandem mass spectrometry; PD, Parkinson's disease; SOD1, CuZn-superoxide dismutase; UBI, ubiquitylated inclusions; VCP, valosin-containing protein; FLAG-Parkin, pcDNA3.1/FLAG-Parkin; Ub, ubiquitin; MBP, maltose-binding protein; GST, glutathione *S*-transferase; PBS, phosphate-buffered saline; HA, hemagglutinin; WT, wild type.

generative diseases, these observations suggest that Dorfin may have a significant role in the quality control system in the cell. The present study was designed to obtain further clues for the pathophysiological roles of Dorfin. For this purpose, we screened Dorfin-associated proteins using high performance liquid chromatography coupled to electrospray tandem mass spectrometry (LC-MS/MS). The results showed that Valosin-containing protein (VCP), also called p97 or Cdc48 homologue, obtained from the screening, physically and functionally interacted with Dorfin. Furthermore, both Dorfin and VCP proteins colocalized in aggresomes of the cultured cells and in UBIs in various neurodegenerative diseases.

MATERIALS AND METHODS

Plasmids and Antibodies—pCMV2/FLAG-Dorfin vector (FLAG-Dorfin^{WT}) was prepared by PCR using the appropriate design of PCR primers with restriction sites (ClaI and KpnI). The PCR product was digested and inserted into the ClaI-KpnI site in pCMV2 vector (Sigma). pEGFP-Dorfin (GFP-Dorfin), pCMX-VCP^{Wt} (VCP^{Wt}), and pCMX-VCP^{K524A} (VCP^{K524A}) vectors were described previously (5, 11). pcDNA/HA-VCP^{Wt} (HA-VCP^{Wt}) and pcDNA/HA-VCP^{K524A} (HA-VCP^{K524A}) were subcloned from pCMX-VCP^{Wt} and pCMX-VCP^{K524A}, respectively, into pcDNA3.1 vectors (Invitrogen). The HA tag was introduced at the N terminus of VCP. pcDNA3.1/FLAG-Parkin (FLAG-Parkin) was generated by PCR using the appropriate design of PCR primers with restriction sites (EcoRI and NotI) from pcDNA3.1/Myc-Parkin (12). The FLAG tag was introduced at the N terminus of Parkin. To establish the RING mutant plasmid of Dorfin (FLAG-Dorfin$^{C132S/C135S}$), point mutations for Cys at positions 132 and 135 to Ser were generated by PCR-based site-directed mutagenesis using a QuikChange™ site-directed mutagenesis kit (Stratagene, La Jolla, CA). pcDNA3.1/HA-Ub (HA-Ub), pcDNA3.1/Myc-SOD1^{WT} (SOD1^{WT}-Myc), pcDNA3.1/Myc-SOD1^{G93A} (SOD1^{G93A}-Myc), and pcDNA3.1/Myc-SOD1^{G85R} (SOD1^{G85R}-Myc) were described previously (13, 14). Polyclonal anti-Dorfin (Dorfin-30 and Dorfin-41) and monoclonal anti-VCP antibodies were used as in previous reports (5, 15). The following antibodies were used in this study: monoclonal anti-FLAG antibody (M2; Sigma), monoclonal anti-Myc antibody (9E10; Santa Cruz Biotechnology, Santa Cruz, CA), monoclonal anti-HA antibody (12CA5; Roche Applied Science), polyclonal anti-maltose-binding protein (MBP) antibody (New England BioLabs, Beverly, MA), polyclonal anti-Parkin (Cell Signaling, Beverly, MA), and polyclonal anti-SOD1 (SOD-100; Stressgen, San Diego, CA).

Cell Culture and Transfection—All media and reagents for cell culture were purchased from Invitrogen. HEK293 cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum, 5 units/ml penicillin, and 50 μg/ml streptomycin. HEK293 cells at subconfluence were transfected with the indicated plasmids using FuGENE6 reagent (Roche Applied Science). To inhibit cellular proteasome activity, cells were treated with 1 μM MG132 (benzyloxycarbonyl-Leu-Leu-Leu-al; Sigma) for 16 h after overnight post-transfection. Cells were analyzed at 24–48 h after transfection.

Protein Identification by LC-MS/MS Analysis—FLAG-Dorfin^{WT} was expressed in HEK293 cells (semiconfluent in a 10-cm dish) and then immunoprecipitated by anti-FLAG antibody. The immunoprecipitates were eluted with a FLAG peptide and then digested with Lys-C endopeptidase (*Achromobacter* protease I). The resulting peptides were analyzed using a nanoscale LC-MS/MS system as described previously (16). The peptide mixture was applied to a Mightysil-PR-18 (1-μm particle, Kanto Chemical Corp., Tokyo) column (45 × 0.150 mm ID) and separated using a 0–40% gradient of acetonitrile containing 0.1% formic acid over 30 min at a flow rate of 50 nl/min. Eluted peptides were sprayed directly into a quadrupole time-of-flight hybrid mass spectrometer (Q-Tof *Ultima*; Micromass, Manchester, UK). MS and MS/MS spectra were obtained in data-dependent mode. Up to four precursor ions above an intensity threshold of 10 cps were selected for MS/MS analysis from each survey scan. All MS/MS spectra were searched against protein sequences of Swiss Prot and RefSeq (NCBI) using batch processes of the Mascot software package (Matrix Science, London, UK). The criteria for match acceptance were the following: 1) when the match score was 10 over each threshold, identification was accepted without further consideration; 2) when the difference of score and threshold was lower than 10 or when proteins were identified based on a single matched MS/MS spectrum, we manually confirmed the raw data prior to acceptance; 3) peptides assigned by less than three y series ions and peptides with +4 charge state were all eliminated regardless of their scores.

Recombinant Proteins and Pull-down Assay—We used pMALp2 (New England BioLabs) and pMALp2T (Factor Xa cleavage site of pMALp2 was replaced with a thrombin recognition site) to express fusion proteins with MBP. To produce the full-length (residues 1–838) Dorfin (MBP-Dorfin^{full}), N-terminal (residues 1–367) Dorfin (MBP-Dorfin^{N}), and C-terminal (residues 368–838) Dorfin (MBP-Dorfin^{C}), the PCR fragments were amplified from pcDNA4/HisMax-Dorfin (4) by using the appropriate PCR primers with restriction sites (FbaI and HindIII) and then ligated into pMAL-p2 vectors. To produce the MBP-Parkin protein, full-length *PARKIN* cDNA was inserted into the EcoRI-NotI sites of pMALp2T. All of the MBP-tagged recombinant proteins were purified from *Escherichia coli* BL21-codon-plus. The detail of the purification method of MBP-tagged proteins was described previously (17). Recombinant GST fusion VCP^{WT} and VCP^{K524A} proteins were also generated from *E. coli* lysate and purified with glutathione-Sepharose. Recombinant His-VCP^{WT} and His-VCP^{K524A} proteins were purified from insect cells using baculovirus. The detail of purification of these recombinant VCP proteins was described previously (15). Binding experiments were performed with proteins carrying different tags. His- or GST-VCP were mixed with MBP fusion proteins: MBP-Dorfin^{full}, -Dorfin^{N}, -Dorfin^{C}, -Parkin, and -mock. His-VCP and GST-VCP proteins were precipitated by Ni^{2+}-nitrilotriacetic acid-agarose (Qiagen, Valencia, CA), and glutathione-Sepharose (Amersham Biosciences), respectively. Binding was performed with 1–3 μg of each protein in 300 μl of binding buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.5 mg/ml bovine serum albumin, 1 mM dithiothreitol) for 1 h at 4 °C. Then 15 μl of beads were added and incubated for 30 min. The beads were washed by binding buffer three times and eluted with sample buffer and analyzed by SDS-PAGE followed by Western blotting using specific antibodies.

Glycerol Gradient Centrifugation—Cultured cells or mouse tissues were homogenized in 1 ml of PBS with protease inhibitor (Complete Mini; Roche Applied Science). Supernatants (1 mg of protein for cultured cells, 5 mg of protein for mouse tissues, and 0.1 mg of recombinant His-VCP protein) were used as the samples after 10,000 × *g* centrifugation for 20 min. The samples (1.0 ml) were loaded on top of a 34-ml linear gradient of glycerol (10–40%) prepared in 25 mM Tris-HCl buffer, pH 7.5, containing 1 mM dithiothreitol in 40 PA centrifuge tubes (Hitachi, Tokyo), and centrifuged at 4 °C and 80,000 × *g* for 22 h using a Himac CP100α centrifuge system (Hitachi). Thirty fractions were collected from the top of the tubes. Two hundred μl of each fraction was precipitated with acetone, and the remaining pellet was lysed with 50 μl of sample buffer and then used for SDS-PAGE followed by Western blotting.

Immunological Analysis—Cells (4 × 10^6 in a 6-cm dish) were lysed with 500 μl of lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Nonidet P-40, and 1 mM EDTA) with protease inhibitor mixture (Complete Mini) 24–48 h after transfection. The lysate was then centrifuged at 10,000 × *g* for 10 min at 4 °C to remove debris. A 10% volume of the supernatants was used as the "lysate" for SDS-PAGE. When immunoprecipitated, the supernatants were precleared with protein A-Sepharose (Amersham Biosciences), and specific antibodies, anti-FLAG (M2), anti-Myc (9E10), or anti-Dorfin (Dorfin-30) were then added and then incubated at 4 °C with rotation. Immune complexes were then incubated with protein A-Sepharose for 3 h, collected by centrifugation, and washed four times with the lysis buffer. For protein analysis, immune complexes were dissociated by heating in SDS-PAGE sample buffer and loaded onto SDS-PAGE. The samples were separated by SDS-PAGE (12% gel or 4–12% gradient gel) and transferred onto a polyvinylidene difluoride membrane. Finally, Western blotting was performed with specific antibodies.

Immunohistochemistry—HEK293 cells grown on glass coverslips were fixed in 4% paraformaldehyde in PBS for 15 min. Then the cells were blocked for 30 min with 5% (v/v) normal goat serum in PBS, incubated for 1 h at 37 °C with anti-HA antibody (12CA5), washed with PBS, and incubated for 30 min with Alexa 496-nm anti-mouse antibodies (Molecular Probes, Inc., Eugene, OR). The coverslips were washed and mounted on slides. Fluorescence images were obtained using a fluorescence microscope (DMIRE2; Leica, Bannockburn, IL) equipped with a cooled charge-coupled device camera (CTR MIC; Leica). Pictures were taken using Leica Qfluoro software.

Pathological Studies—Pathological studies were carried out on 10% formalin-fixed, paraffin-embedded spinal cords and brain stems filed in the Department of Neurology, Nagoya University Graduate School of Medicine. The specimens were obtained at autopsy from three sporadic cases of ALS and four sporadic PD patients. The spinal cord and brain stem specimens of these ALS and PD cases were immunohistochemically stained with antibodies against Dorfin (Dorfin-41) and VCP. Dou-

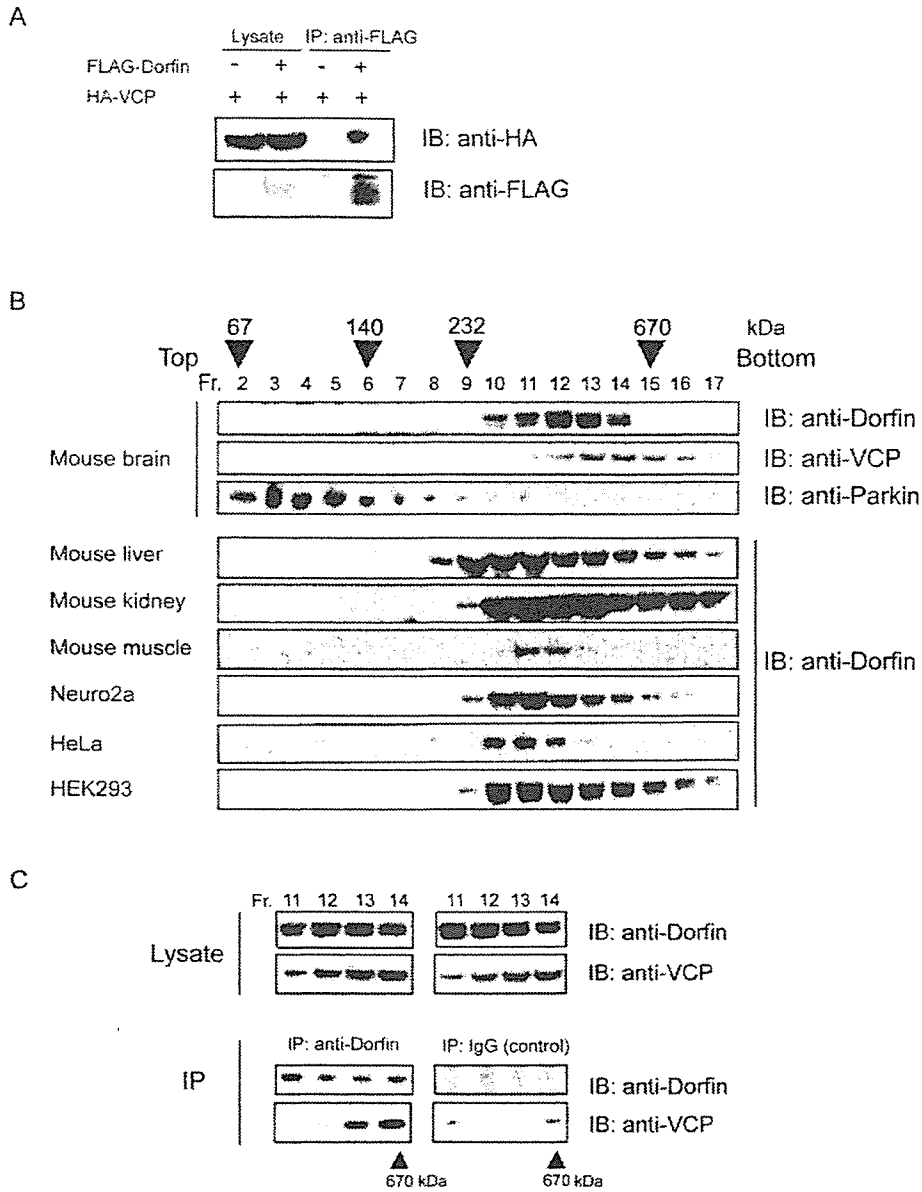

FIG. 1. ***In vivo* interaction between Dorfin and VCP.** *A*, FLAG-Dorfin and HA-VCP are co-expressed in HEK293 cells. FLAG-mock vector was used as a negative control. The amounts of HA-VCP in 10% of the lysate used are shown (*Lysate*); the rest was subjected to immunoprecipitation (*IP*) with anti-FLAG (M2) antibody. Following immunoblotting (*IB*) with anti-HA (12CA5) antibody revealed that HA-VCP was co-immunoprecipitated with FLAG-Dorfin. *B*, 5 mg of protein of various mouse tissues (brain, liver, kidney, and muscle) and 1 mg of protein of cultured cells (HEK293, HeLa, and Neuro2a) were each homogenized in 1 ml of PBS. Supernatants were fractionated by 10–40% glycerol gradient centrifugation followed by separation into 30 fractions using a fraction collector. Immunoblotting using anti-Dorfin, anti-VCP, and anti-Parkin antibodies was performed on the fractions (*Fr.*), including fractions 2–17. Endogenous Dorfin was co-sedimented with VCP in the fractions with a molecular mass of around 400–600 kDa. The positions of co-migrated molecular mass markers are indicated *above* the *panels*. *C*, immunoprecipitation with polyclonal anti-Dorfin antibody (anti-Dorfin-30) was performed on fractions 11–14 collected by glycerol gradient centrifugation analysis, where endogenous Dorfin was seen in *B*. As a negative control, immunoprecipitation with nonimmune rabbit IgG was used on the same fractions.

ble staining of identical sections was performed as described previously (7). In immunofluorescence microscopy, Alexa-488- and Alexa-546-conjugated secondary antibodies (Molecular Probes) were used. All human and animal studies described in this report were approved by the appropriate Ethics Review Committees of the Nagoya University Graduate School of Medicine.

RESULTS

Identification of Dorfin-associated Protein in the Cells—In an effort to identify protein(s) that physically interacts with Dorfin in the cells, FLAG-Dorfin was expressed in HEK293 cells and then immunoprecipitated by anti-FLAG antibody. The immunoprecipitates were eluted with a FLAG peptide and then digested with Lys-C endopeptidase (*Achromobacter* protease I), and the cleaved fragments were directly analyzed using a highly sensitive "direct nanoflow LC-MS/MS" system as described under "Materials and Methods." Following data base search, a total of 13 peptides were assigned to MS/MS spectra obtained from the LC-MS/MS analyses for the FLAG-Dorfin-

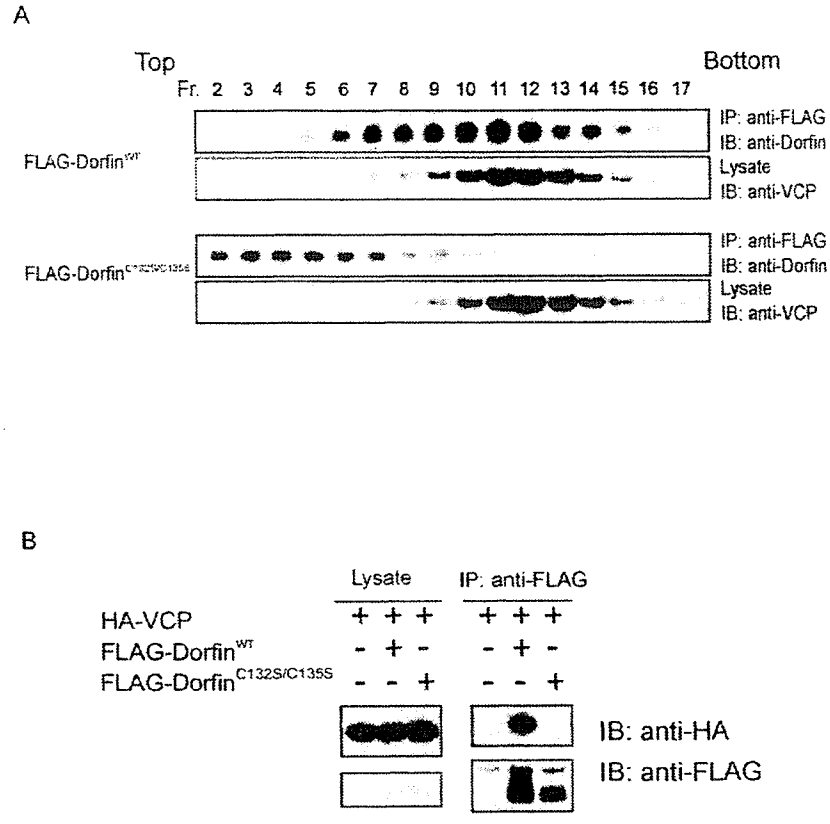

FIG. 2. **Loss of physical interaction between Dorfin$^{C132S/C135S}$ and VCP.** *A*, transfected Dorfin^{WT}, but not Dorfin$^{C132S/C135S}$ (*Dorfin$^{C132S-C135S}$*), forms a high M_r complex. Lysate of HEK293 cells overexpressed with FLAG-Dorfin^{WT} or FLAG-Dorfin$^{C132S/C135S}$ was fractionated by 10–40% glycerol gradient centrifugation. The selected fractions (*Fr.*), fractions 2–17, were subjected to immunoprecipitation (*IP*) using anti-FLAG (M2) antibody. Immunoblotting (*IB*) with anti-Dorfin antibody revealed that exogenous FLAG-Dorfin^{WT} formed a high molecular weight complex, whose peak was at fraction 11, whereas FLAG-Dorfin$^{C132S/C135S}$ migrated in fractions of smaller M_r (around fraction 7). Ten percent of the fractionated samples were shown as "lysate." *B*, Dorfin^{WT} can interact with VCP, but Dorfin$^{C132S/C135S}$ cannot. FLAG-Dorfin^{WT} or FLAG-Dorfin$^{C132S/C135S}$ and HA-VCP were co-expressed in HEK293 cells. FLAG-mock vector was used as a negative control. The amounts of HA-VCP in 10% of the lysate used are shown (*Lysate*); the rest was subjected to immunoprecipitation with anti-FLAG (M2) antibody. Following immunoblotting with anti-HA (12CA5) antibody revealed that HA-VCP was co-immunoprecipitated with FLAG-Dorfin^{WT} but not with FLAG-Dorfin$^{C132S/C135S}$.

associated complexes. These peptide data identified nine proteins as candidates for Dorfin-associated proteins. One of these identified proteins was VCP that has been proposed to have multiple functions, such as membrane fusion or endoplasmic reticulum-associated degradation (ERAD) (18–22). In the next step, we examined the relationship between Dorfin and VCP, because the latter has been reported to be linked to various aspects of neurodegeneration (15).

Dorfin Interacts with VCP in Vivo—To verify the interaction between Dorfin and VCP, FLAG-Dorfin and HA-VCP were transiently overexpressed in HEK 293 cells. Immunological analyses revealed that HA-VCP was co-immunoprecipitated with FLAG-Dorfin but not with FLAG-mock (Fig. 1*A*), confirming their physical interactions in the cells. To determine whether endogenous Dorfin forms a complex, the lysate from mouse brain homogenate was fractioned by glycerol density gradient centrifugation. Each fraction was immunoblotted with anti-Dorfin antibody. The majority of endogenous Dorfin was co-sedimented with VCP around a size of 400–600 kDa, although endogenous Parkin, which is another RING-IBR type E3 ligase (12), existed in the fractions of much lighter molecular weight (M_r) (Fig. 1*B*, *top panels*). Moreover, Dorfin was sedimented in the fractions of 400–600 kDa in other tissues, such as the liver, kidney, and muscle of mouse, and various cultured cells including Neuro2a, HeLa, and HEK293 cells (Fig. 1*B*, *bottom panels*). To determine whether endogenous Dorfin interacts with VCP, immunoprecipitation using polyclonal anti-Dorfin antibody (Dorfin-30) was performed on the fractions shown in Fig. 1*B*, *top panels*. Endogenous VCP was co-immunoprecipitated with endogenous Dorfin in the fractions of high M_r (fractions (*Fr.*) 13 and 14). No apparent band was observed when precipitated with rabbit IgG (Fig. 1*C*).

Mutations of RING Finger Domain of Dorfin Results in Loss of Dorfin-VCP Interactions—Next, we examined whether transfected Dorfin (FLAG-Dorfin^{WT}) and its RING mutant (FLAG-Dorfin$^{C132S/C135S}$), in which the two Cys residues at positions 132 and 135 within the RING finger domain were substituted for Ser residues, form a complex. The results showed overexpression of FLAG-Dorfin^{WT} in high molecular fractions (*Fr.* in Fig. 2), whose peak was between fractions 10 and 12, whereas overexpressed FLAG-Dorfin$^{C132S/C135S}$ did not consist of high molecular weight complex. Overexpression of FLAG-Dorfin^{WT} or FLAG-Dorfin$^{C132S/C135S}$ did not change the sedimentation pattern of VCP (Fig. 2*A*). Furthermore, immunoprecipitation analysis showed that FLAG-Dorfin^{WT}, but not FLAG-Dorfin$^{C132S/C135S}$, could interact with HA-VCP in HEK293 cells (Fig. 2*B*).

Dorfin Interacts with VCP in Vitro—To confirm the direct

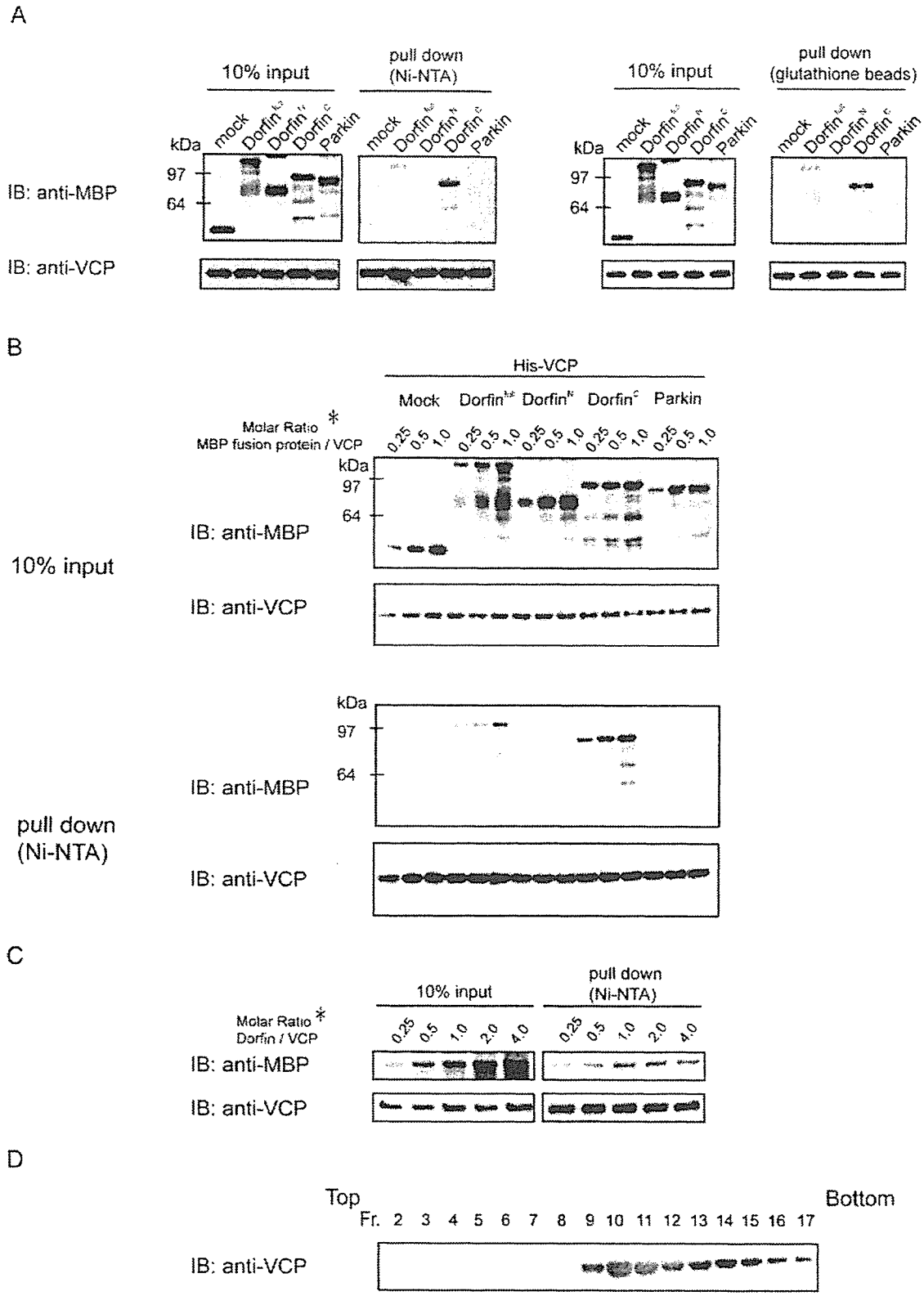

FIG. 3. ***In vitro* interaction between Dorfin and VCP.** *A*, recombinant His- or GST-VCP protein was incubated with MBP-mock, MBP-Dorfin^{full}, MBP-Dorfin^{N}, MBP-Dorfin^{C}, and MBP-Parkin proteins *in vitro*. Two μg of His- or GST-VCP proteins and MBP fusion proteins at similar molar concentrations to VCP proteins were used for the assays. The amounts of MBP fusion and GST fusion Dorfin derivatives and His-VCP in 10% of the samples used are shown (*10% input*). *NTA*, nitrilotriacetic acid. *IB*, immunoblot. *B*, 2 μg of His-VCP was incubated with MBP-mock,

binding between Dorfin and VCP and to determine the exact portion of Dorfin that interacts with VCP *in vitro*, we performed pull-down assays using recombinant proteins. Recombinant MBP-Dorfin or its deletion mutants (*i.e.* MBP-Dorfin^{N} and MBP-Dorfin^{C}) and the same molar of recombinant His-VCP or GST-VCP were mixed and incubated for 1 h at 4 °C. MBP-mock protein was used as a negative control in these experiments. A small portion of MBP-Dorfin^{full} or Dorfin^{C} (C-terminal substrate-recognizing domain) bound to both His-VCP and GST-VCP, whereas MBP-mock, MBP-Dorfin^{N} (N-terminal RING-IBR domain), and MBP-Parkin did not bind to His-VCP or GST-VCP (Fig. 3*A*). We next determined the number of Dorfins that bind one hexamer of VCP. To investigate this issue, we incubated His-VCP with increasing amounts of MBP-Dorfin^{full}, MBP-Dorfin^{N}, MBP-Dorfin^{C}, MBP-mock, or MBP-Parkin. As shown in Fig. 3*B*, the amount of binding portion of MBP-Dorfin^{full} and -Dorfin^{C} pulled down with His-VCP was not saturated below the even molar ratio. The pull-down experiments using excess amounts of MBP-Dorfin^{full} revealed that MBP-Dorfin^{full} was saturated at the even molar ratio (Fig. 3*C*). As reported previously (15), recombinant His-VCP sedimented in high molecular weight fractions, indicating that it formed a hexamer *in vitro* (Fig. 3*D*). These findings indicated that six Dorfin molecules were likely bind to a VCP complex *in vitro*.

Subcellular Localization of Dorfin and VCP in HEK293 Cells—In previous studies, we showed that exogenous and endogenous Dorfin resided perinuclearly and was colocalized with Vimentin in cultured cells treated with a proteasome inhibitor (4). The staining patterns of Dorfin were indistinguishable from those of the aggresome, namely a pericentriolar, membrane-free, cytoplasmic inclusion containing misfolded ubiquitylated proteins packed in a cage of intermediate filaments (4). VCP immunostaining was also observed throughout aggresomes in cultured neuronal cells when induced by treatment with a proteasome inhibitor (15). In order to examine the subcellular localization of Dorfin and VCP, GFP-Dorfin and HA-VCP were co-expressed in HEK293 cells. Without proteasome treatment, GFP-Dorfin-expressing cells showed granular fluorescence in the cytosol, and the HA-VCP-expressing cells showed diffuse and uniform cytoplasmic staining (Fig. 4*A*). Treatment with MG132 (1 μM, 16 h) resulted in accumulation of both GFP-Dorfin and HA-VCP and perinuclear colocalization as a clear large protein aggregate that mimics aggresomes (Fig. 4*B*).

Colocalization of Dorfin and VCP in the Affected Neurons of ALS and PD—In previous studies, immunostaining of Dorfin and VCP was independently noted in LBs of PD, and the peripheral staining pattern of both proteins in LBs was similar (7, 23). To confirm the immunoreactivities of Dorfin and VCP in the affected neurons in ALS and PD, we performed a double-labeling immunofluorescence study using a rabbit polyclonal anti-Dorfin antibody (Dorfin-41) and a mouse monoclonal VCP antibody on the postmortem samples of ALS and PD. In the ALS spinal cords, both proteins were colocalized in the LB-like inclusions (Fig. 5, *A–F*). The margin of LBs in PD was intensely immunostained for Dorfin and VCP, and merged images confirmed their strong colocalization (Fig. 5, *G–L*). Dorfin and VCP were also positive in Lewy neurites in the affected neurons of PD (Fig. 5, *M–O*).

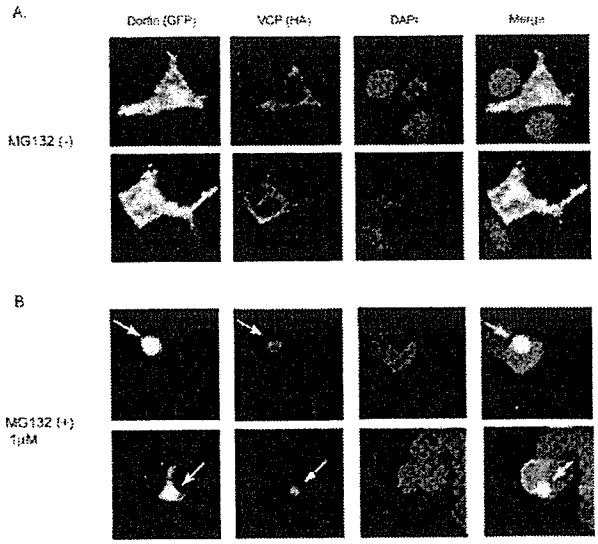

FIG. 4. **Subcellular localization of GFP-Dorfin and HA-VCP in HEK293 cells treated or untreated with a proteasome inhibitor.** GFP-Dorfin and HA-VCP were co-expressed transiently in HEK 293 cells. Cells were treated with (*B*) or without (*A*) 1 μM MG132 for 16 h. HA-VCP was stained with anti-monoclonal HA antibody (12CA5). Nuclei were stained with 4′,6-diamidino-2-phenylindole (*DAPI*). Without the treatment of MG132, GFP-Dorfin was spread through the cytosol, and it appeared like small aggregations. HA-VCP was also seen mainly in the cytosol and partly colocalized with GFP-Dorfin (*A*). After treatment with 1 μM MG132 for 16 h, both GFP-Dorfin and HA-VCP showed perinuclear accumulation and colocalization and appeared as clear large protein aggregates (*B*; *arrows*).

Dorfin Ubiquitylates Mutant SOD1 in Vivo—Unlike the wild-type form, mutant SOD1 proteins are rapidly degraded by the ubiquitin-proteasome system. Consistent with our previous results (5), SOD1^{G93A} and SOD1^{G85R} were polyubiquitylated, and co-expression with FLAG-Dorfin^{WT} enhanced polyubiquitylation of these mutant SOD1s compared with co-expression with FLAG-BAP, a negative control construct (Fig. 6*A*). Boiling with 1% SDS-containing buffer did not change the level of ubiquitylated mutant SOD1, indicating that mutant SOD1 itself was ubiquitylated by Dorfin (Fig. 6*B*). We also performed the same *in vivo* ubiquitylation assay using Neuro2a cells to examine for E3 activity of Dorfin in neuronal cells. The enhanced polyubiquitylation of these mutant SOD1s by Dorfin was observed in Neuro2a cells as well as in HEK293 cells (Fig. 6*C*). FLAG-Dorfin$^{C132S/C135S}$ did not enhance polyubiquitylation of mutant SOD1s, indicating that this RING finger mutant form was functionally inactive (Fig. 6*D*).

VCP^{K524A} Suppresses the E3 Activity of Dorfin—VCP has two ATPase binding domains (D1 and D2). A D2 domain mutant, VCP^{K524A}, induces cytoplasmic vacuoles, which mimics vacuole formation seen in the affected neurons in various neurodegenerative diseases (11, 15). The D2 domain represents the major ATPase activity and is essential for VCP function (11). The ATPase activity of VCP^{K524A} is much lower than that of VCP^{WT}, and VCP^{K524A} caused accumulation of polyubiquitylated proteins in the nuclear and membrane fractions together with elevation of ER stress marker proteins due to ERAD

MBP-Dorfin^{full}, MBP-Dorfin^{N}, MBP-Dorfin^{C}, and MBP-Parkin with increasing amounts (molar ratio to VCP: 0.25, 0.5, and 1.0). The amounts of MBP fusion Dorfin derivatives and His-VCP in 10% of the samples used are shown (10% input). *C*, 2 μg of His-VCP was incubated with MBP-Dorfin^{full} with increasing amounts (molar ratio to VCP: 0.25, 0.5, 1, 2, and 4). The amounts of MBP-Dorfin^{full} and His-VCP in 10% of the samples used are shown (*10% input*). *D*, His-VCP protein (0.5 μg) was fractionated by 10–40% glycerol gradient centrifugation followed by separation into 30 fractions using a fraction collector. Immunoblotting using anti-VCP antibody was performed on the selected fractions (fractions 2–17). *, The molar ratio was calculated by the amount of VCP monomers, not VCP complexes.

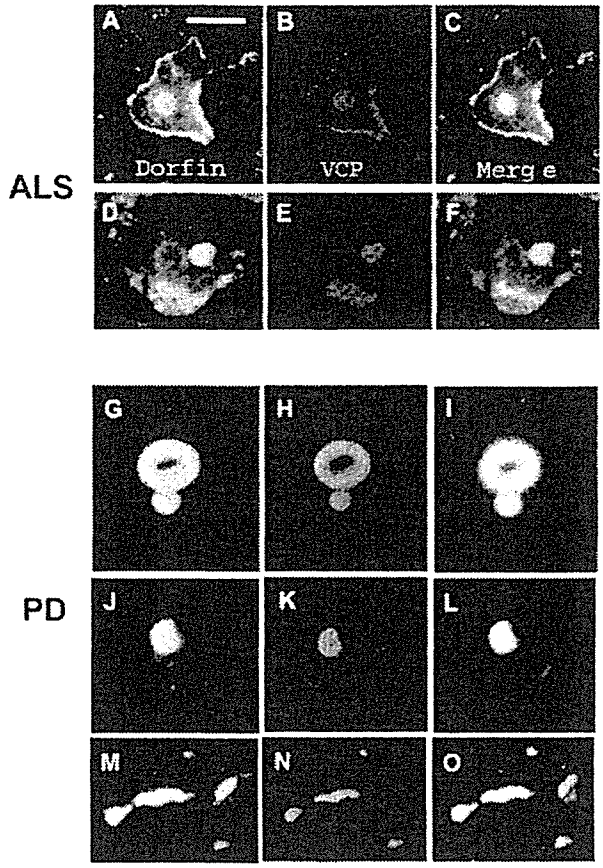

FIG. 5. **Colocalization of Dorfin-41 immunoreactivity with VCP in neuronal inclusions in ALS and PD.** Sections were doubly labeled with anti-Dorfin-41 antiserum and monoclonal VCP antibody and analyzed with a laser-scanning confocal microscope. The *left panels* (*green*) correspond to Dorfin, *middle panels* (*red*) correspond to VCP, and *right panels* correspond to merged images; structures in *yellow* indicate colocalization. Colocalization of Dorfin and VCP is seen in LB-like inclusions in motor neurons of the spinal cord of ALS (*A–F*). Dorfin is also colocalized with VCP in the margin of LBs (*G–I*), premature LBs (*J–L*), and Lewy neurites (*M–O*) in the nigral neurons of PD. *Scale bars*, 20 μm (*A–L*) and 10 μm (*M–O*).

inhibition, whereas its expression level, localization, and complex formation were indistinguishable from those of VCP^{WT} (11). In order to examine the functional effect of VCP on Dorfin, VCP^{WT}, VCP^{K524A}, or LacZ was co-expressed with $SOD1^{G85R}$, FLAG-Dorfin, and HA-Ub in HEK293 cells. Co-expression with VCP^{K524A} showed a marked decline of polyubiquitylation of $SOD1^{G85R}$ compared with co-expression with VCP^{WT} or LacZ (Fig. 7*A*, *top* and *middle*). Since Dorfin physically interacts with mutant SOD1s (5), we next investigated whether this decline of polyubiquitylation of $SOD1^{G85R}$ was mediated by reduced affinity between $SOD1^{G85R}$ and Dorfin. Immunoprecipitation by anti-FLAG antibody showed that VCP^{K524A} did not change affinity between $SOD1^{G85R}$ and Dorfin (Fig. 7*A*, *bottom*). Neither VCP^{WT} nor VCP^{K524A} changed the level of polyubiquitylation protein in the total lysate (Fig. 7*B*). To clarify whether this negative effect of VCP^{K524A} is specific for Dorfin, we assessed the autoubiquitylation of FLAG-Parkin in the presence of VCP^{WT}, VCP^{K524A}, or LacZ. Co-expression of VCP^{K524A} did not decrease autoubiquitylation of FLAG-Parkin compared with co-expression of LacZ or VCP^{WT} (Fig. 7*C*). We performed the same experiments using Neuro2a cells to see whether VCP^{K524A} suppress the E3 activity of Dorfin in neuronal cells. The marked decline of polyubiquitylation of $SOD1^{G85R}$ by VCP^{K524A} expression was also seen in Neuro2a cells (Fig. 7*D*).

DISCUSSION

UBIs in the affected neurons are histopathological hallmarks in various neurodegenerative disorders (8). Dorfin is an E3 ligase, which can ubiquitylate mutant SOD1s and synphilin-1 (5, 24). These substrates and Dorfin were identified in UBIs in various neurodegenerative diseases, such as LB-like inclusions in ALS and LBs in PD and dementia with Lewy bodies (7). This finding suggests that Dorfin may play a crucial role in the process of generating inclusions in the affected neurons. In the present study, we identified VCP as one of the Dorfin-associated proteins using mass spectrometry, and VCP-Dorfin physical interaction was confirmed by an immunoprecipitation experiment using FLAG-Dorfin and HA-VCP overexpressed in HEK293 cells (Fig. 1*A*). VCP is an essential and highly conserved protein of the AAA-ATPase family, which is considered to have diverse cellular functions, such as membrane fusion (25–27), nuclear trafficking (28), cell proliferation (29, 30), and the ERAD pathway (18–22). Many reports have implied that VCP is involved in the pathogenesis of various neuromuscular diseases. VCP has been implicated as a factor that modifies the progress of polyglutamine-induced neuronal cell death (15). In addition, histopathological studies revealed positive staining for VCP in UBIs in PD and ALS with dementia (23). VCP is also associated with MJD protein/atxin-3, in which abnormal expansion of polyglutamine tracts causes Machado-Joseph disease/spinocerebellar ataxia type 3 (31). VCP is also required for the degradation of ataxin-3 in collaboration with E4B/Ufd2a, a ubiquitin chain assembly factor (E4) (32). Recent studies have indicated that missense mutations in the VCP gene cause inclusion body myopathy associated with Paget's disease of bone and frontotemporal dementia, which is characterized by the presence of vacuoles in the cytoplasm in muscle fibers (33).

Our results showed that endogenous Dorfin formed a 400–600-kDa complex in various tissues and various cultured cells (Fig. 1*B*). Dorfin is a ~91-kDa protein; therefore, this high M_r complex should include Dorfin-associated proteins, although the possibility that Dorfin itself oligomerizes in the cell cannot be excluded. Glycerol gradient centrifugation analysis and immunoprecipitation experiments in the present study showed that endogenous Dorfin interacted with endogenous VCP in a complex of approximately 600 kDa, possibly including a Dorfin molecule and a hexametric form of VCP (Fig. 1*C*).

The first RING mutant of Dorfin, in which Cys at positions 132 and 135 changed to Ser, was prepared. This mutant Dorfin, $Dorfin^{C132S/C135S}$, could not ubiquitylate mutant SOD1s (Fig. 6*D*). Glycerol gradient centrifugation analysis revealed that $Dorfin^{C132S/C135S}$ did not form a high M_r complex, whereas exogenous wild type Dorfin ($Dorfin^{WT}$) formed a high M_r complex similar to endogenous Dorfin (Fig. 2*A*). Furthermore, an immunoprecipitation experiment using $Dorfin^{WT}$ and $Dorfin^{C132S/C135S}$ revealed that $Dorfin^{WT}$ could interact with VCP, whereas $Dorfin^{C132S/C135S}$ could not (Fig. 2*B*).

Our *in vitro* study using recombinant proteins showed that full-length ($MBP\text{-}Dorfin^{full}$) and the C terminus of Dorfin ($MBP\text{-}Dorfin^{C}$) directly interacted with VCP, whereas the $MBP\text{-}Dorfin^{N}$ mutant, containing the entire RING finger domain (amino acid residues 1–367), did not bind to VCP (Fig. 3*A*). This finding was unexpected, since *in vivo* binding analysis suggested that Dorfin could interact with VCP at the RING finger domain. It is plausible that certain structural changes in $Dorfin^{C132S/C135S}$ might render the C-terminal VCP-binding portion incapable of accessing VCP molecules. This may explain the result that $Dorfin^{C132S/C135S}$ did not form a high M_r complex.