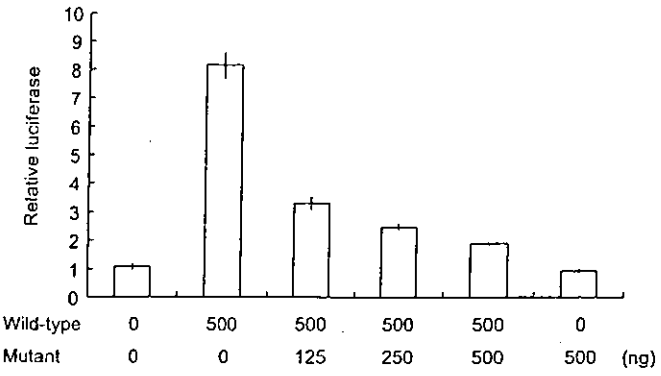

Fig. 1. Functional analysis of the identified HNF-1β mutation. The wild-type expression construct was prepared using human HNF-1β cDNA and vector pCMV6b. The L95fsinsAGCT mutation was introduced in the sequence by site-directed mutagenesis and cloned into pCMV6b. The reporter construct pGL3-RA was prepared by cloning the rat albumin gene promoter into the firefly luciferase reporter vector pGL3-Basic (Promega, Madison, Wis., USA). HepG2 cells were transfected with various amounts of the reporter and test plasmid constructs. The transactivation activity of the wild-type and mutant proteins was measured using the Dual-Luciferase Reporter Assay System (Promega) as described previously [8]

ic features, wild-type and mutant proteins were transiently expressed in cultured hepatic cells. Luciferase reporter assay revealed L95fsinsAGCT to be a null function mutation (Fig. 1). However, when co-expressed with wild-type, the mutation behaved as strong dominant negative regulation; an equimolar amount of the mutant reduced wild-type activity by ~85%. However, whether the dominant negative mechanism influences the diversity of phenotypes seen in this patient is unknown.

Since molecular cooperation with other diabetogenes might increase the severity of symptoms or hasten onset, other MODY genes were also screened for mutations in this patient, resulting in the identification of one missense mutation, D78H, in the BETA2/NeuroD1 (MODY6) gene; the allele frequency in the general population is 0.002 (three in 1330 alleles). Since the mutation was transmitted from the non-diabetic mother, this mutation alone should not cause diabetes. However, because HNF-1 directly regulates BETA2/NeuroD1 gene transcription, some synergistic effect on increased morbidity is possible.

This report of a solitary case owing to a new mutation expands our clinical knowledge on HNF-1β deficiency. Accumulation of additional cases and identification of the molecular participants in the transcriptional events are required to clarify this distinct syndrome of varied phenotypes.

Acknowledgements. We thank S. Oike, Y. Yaginuma, I. Uda and T. Takahashi for excellent technical services. This study was supported by Grants-in-Aid for Scientific Research on Priority Areas from the Japan Ministry of Science, Education, Sports, Culture and Technology, by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science, by Comprehensive Research on Aging and Health and Research on Human Genome, and Tissue Engineering Food Biotechnology from the Japan Ministry of Health, Labor and Welfare, by the Yamanouchi Foundation for Research, and by Core Research for Evolutional Science and Technology (CREST) from Japan Science and Technology Corporation (JST).

N. Shihara, Y. Horikawa
Laboratory of Molecular Genetics, Department of Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, Maebashi, Gunma 371-8512, Japan

T. Onishi, M. Ono, K. Kashimada
Department of Pediatrics and Developmental Biology, Graduate School of Medicine, Tokyo Medical and Dental University, Japan

J. Takeda
Department of Endocrinology, Diabetes and Rheumatology, Division of Bioregulatory Medicine, Gifu University School of Medicine, Gifu, Japan

References

1. Horikawa Y, Iwasaki N, Hara M et al. (1997) Mutation in hepatocyte nuclear factor-1β gene (*TCF2*) associated with MODY. Nat Genet 17:384–385
2. Nishigori H, Yamada S, Kohama T et al. (1998) Frameshift mutation, A263fsinsGG, in the hepatocyte nuclear factor-1beta gene associated with diabetes and renal dysfunction. Diabetes 47:1354–1355
3. Bingham C, Bulman MP, Ellard S et al. (2001) Mutations in the hepatocyte nuclear factor-1beta gene are associated with familial hypoplastic glomerulocystic kidney disease. Am J Hum Genet 68:219–224
4. Bingham C, Ellard S, van't Hoff WG et al. (2003) Atypical familial juvenile hyperuricemic nephropathy associated with a hepatocyte nuclear factor-1beta gene mutation. Kidney Int 63:1645–1651
5. Linder TH, Njolstad PR, Horikawa Y, Bostad L, Bell GI, Sovik O et al. (1999) A novel syndrome of diabetes mellitus, renal dysfunction and genital malformation associated with a partial deletion of the pseudo-POU domain of hepatocyte nuclear factor-1beta. Hum Mol Genet 11:2001–2008
6. Iwasaki N, Okabe I, Momoi MY, Ohashi H, Ogata M, Iwamoto Y (2001) Splice site mutation in the hepatocyte nuclear factor-1 beta gene, IVS2nt+1G>A, associated with maturity-onset diabetes of the young, renal dysfunction and bicornuate uterus. Diabetologia 44:387–388
7. Takeda A, Ohgushi H, Mizusawa Y et al. (1999) Dysplastic glomerulocystic kidney. Clin Exp Nephrol 1:51–53
8. Tomura H, Nishigori H, Sho K, Yamagata K, Inoue I, Takeda J (1999) Loss-of-function and dominant negative mechanisms associated with hepatocyte nuclear factor-1β mutations in familial Type 2 diabetes mellitus. J Biol Chem 274:12975–12978

Y. Horikawa (✉)
Laboratory of Molecular Genetics,
Department of Cell Biology,
Institute for Molecular and Cellular Regulation,
Gunma University, 3-39-15, Showa-machi, Maebashi,
Gunma 371-8512, Japan
E-mail: yhorikaw@showa.gunma-u.ac.jp
Tel.: +81-27-2208831, Fax: +81-27-2208889

Abbreviations: HNF, hepatocyte nuclear factor · MRI, magnetic resonance imaging · FJHN, familial juvenile hyperuricaemic nephropathy

Available online at www.sciencedirect.com

SCIENCE DIRECT

Molecular Brain Research 129 (2004) 20–32

MOLECULAR BRAIN RESEARCH

www.elsevier.com/locate/molbrainres

Research report

Expression profile of mRNAs from rat hippocampus and its application to microarray

Takeshi Tanaka[a], Yukio Horikawa[b,c,*], Takanori Kawamoto[a], Noriko Kabe-Sakurai[a], Jun Takeda[b,c,d], Masahiko Mikuni[a]

[a]*Department of Psychiatry and Human Behavior, Gunma University, Graduate School of Medicine, Gunma, Japan*
[b]*Laboratory of Molecular Genetics, Department of Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, Gunma, Japan*
[c]*Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), Kawaguchi, Japan*
[d]*Division of Bioregulatory Medicine, Department of Endocrinology, Diabetes and Rheumatology, Gifu University School of Medicine, Gifu, Japan*

Accepted 17 June 2004
Available online 23 July 2004

Abstract

Stress refers to physiological or psychological stimuli that disrupt homeostasis and induce pathophysiological conditions due to maladaptive response, sometimes resulting in mental disorders including depression and post-traumatic stress disorder. Severe stress has been shown to induce neuronal atrophy and apoptosis, especially in the hippocampus, which is thought to be a region of the brain important in stress-related disorders. We have analyzed gene expression in rat hippocampus comprehensively to clarify the molecular mechanism of stress-related disorders. In the present study, we identified and catalogued 13,660 partial complementary DNA sequences (expressed sequence tags (ESTs)) of randomly selected clones from a cDNA library of rat hippocampus. Sequence analysis showed that these clones cluster into 7173 non-redundant sequences comprising 1794 clusters and 5379 singletons. As a result of nucleotide and peptide database search, 2594 were found to represent known rat sequences. Of the remaining 4579 genes, 599 non-redundant ESTs represent rat homologs of genes identified in other species or new members of structurally related families. In addition, we illustrate the use of these clone sets by constructing a cDNA microarray focused on genes categorized into 'cell/organism defense'. These ESTs and our own microarray thus provide an improved genomic source for molecular studies of animal models of stress-related disorders.

Theme: Cellular and molecular biology
Topic: Gene structure and function: general

Keywords: Hippocampus; Stress; Expressed sequence tags (ESTs); cDNA library; Microarray

1. Introduction

The hippocampus is not only crucial in learning and memory but also is especially vulnerable to stress. This region of the brain also is involved in feedback regulation of the hypothalamus–pituitary–adrenal axis, dysfunction of which is associated with depression [12,30]. The effects of chronic stress on brain function via CRF, ACTH, and glucocorticoids may trigger some of the pathophysiological changes in brain function related to depression and other stress-related disorders. Glucocorticoids are known to influence most brain regions, but have particularly dramatic effects on limbic structures such as hippocampus and amygdala [24]. Recent studies suggest that stress-induced atrophy and loss of hippocampal neurons may contribute to the pathophysiology of depression [6,20]. Interestingly, hippocampal volume is decreased in patients with stress-related disorders, including depression and

* Corresponding author. Laboratory of Molecular Genetics, Department of Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15 Showa-machi, Maebashi, Gunma 371-8512, Japan. Tel.: +81 27 220 8831; fax: +81 27 220 8889.
E-mail address: yhorikaw@showa.gunma-u.ac.jp (Y. Horikawa).

0169-328X/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbrainres.2004.06.017

post-traumatic stress disorder [24,25]. Furthermore, the hippocampus is one of only a few brain regions where the production of neurons occurs throughout the lifetime of animals, including human [7]. Furthermore, hippocampal neurogenesis is influenced by various environmental factors and stimuli [11,21,29]. For example, both acute and chronic stress cause a decrease in cell proliferation [8].

These findings indicate that cell death, neurogenesis, and the more dramatic changes induced by chronic stress occur in hippocampus together with stress-related disorders. To compare gene expression in the hippocampus in normal and an animal model of mental disorder, we analyzed gene expression in this region of the brain by large-scale sequencing of randomly selected clones from the cDNA library to generate expressed sequence tags (ESTs).

We also illustrate one use of these clone sets by constructing a cDNA microarray focused on genes categorized into 'cell/organism defense'. These nonredundant hippocampus clone sets and our own microarray promise to become a useful tool for molecular studies of animal models of stress-related disorders.

2. Materials and methods

2.1. cDNA sequencing

A non-unidirectional cDNA library with inserts larger than ~400 bp, which was constructed using mRNAs from adult rat hippocampus and Lambda ZAP® II vector system, was purchased from commercial company (Stratagene, La Jolla, CA, USA). Plasmid DNA were prepared as described previously [28]. Briefly, the non-unidirectional cDNA library was excised in vivo from the λ phage into phagemid DNA using the ExAssist® helper phage (Stratagene). Phagemid particles were transfected into *Escherichia coli* SOLR (Stratagene) and plated on LB plates containing ampicillin to generate plasmid forms. The colonies were randomly selected from the plates and plasmid DNAs were extracted using the Biomek 2000 miniprep systems (Beckman, Fullerton, CA, USA). The inserts of the cDNA clones were sequenced from both ends. DNA sequencing was performed using an ABI PRISM BigDye Terminator Cycle Sequencing FS Ready Reaction Kit® (Applied Biosystems, Foster, CA, USA). The sequencing reaction products were analyzed by an Applied Biosystems DNA sequencer model 377. Quality assessment and base trimming of each sequence were performed using PE Sequencing Analysis 3.3 software (Applied Biosystems). Contaminated vector sequences were removed using Assembly LIGN® (copyright by Oxford Molecular Group). Sequences containing less than 1% ambiguous bases longer than 200 bp were counted as good sequences.

2.2. Database analysis of rat hippocampus ESTs

We analyzed ~13,867 ESTs from rat hippocampus with non-redundant nucleotide and peptide sequences extracted in silico from GenBank databases at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/). We first removed tracks of ambiguous residues from the obtained ESTs and masked the highly repetitive sequences by RepeatMasker (http://ftp.genome.washington.edu/RM/RepeatMasker.html). The resultant sequences were subjected to a BLAST search against a merged database containing daily updates of rat sequences from GenBank. The program BLASTN [2] was used to compare the sequences at the level of the nucleic acids. If a query EST sequence shared more than 95% sequence identity without masked and ambiguous nucleotides and showed a score of more than 365 with any other sequences in the database, it was grouped with the query. If there was at least one sequence in common, groups were merged into a single cluster. An EST sequence that did not belong to any of the clusters is a singleton. To assemble the sequences that belonged in each cluster, we applied the Labo Server® system to make contigs (World Fusion, Tokyo, Japan). The EST clones without any match to known genes in the nucleotide database were retrieved by the BLASTX program [2], which was used to conceptually translate the sequence in all six reading frames and compare the sequences with those in the peptide database at NCBI (http://www.ncbi.nlm.nih.gov/).

Role categories and subcategories were chosen to encompass a broad view of rat cell biology. Although many categorization schemes might be considered equally valid, we have attempted to group together proteins that share similar functional characteristics or cellular roles rather than by a strict biochemical classification. Roles were assigned according to the known or putative involvement of a gene or protein in a cellular process or pathway as opposed to participation in a specific binding or catalysis function on which Gene Ontology (GO) annotations are based.

We used a seventh broad category, unclassified, for proteins and genes of unknown role or which could not be assigned with confidence based on searches of the literature [1]. The EST clones matching known genes (excluding repetitive elements and probable microbial contaminant sequences) were catalogued into seven general categories (cell division, cell signaling/cell communication, cell structure/motility, cell/organism defense, gene/protein expression, metabolism and unclassified) and subcategorized according to specific function based on the putative functions of the known genes using the Genome Directory (http://www.tigr.org/tdb/hgi.html), UniGene, Entrez and PubMed at the NCBI. Two subcategories were included in cell structure/motility, namely, contractile proteins and vesicular transport [1,13].

Table 1
Summary of rat hippocampus ESTs

	Known genes	Unknown genes	Total
Cluster	1282	512	1794
Singleton	1312	4067	5379
Total	2594	4579	7173

2.3. Animals and treatment

Adult Sprague–Dawley rats (Charles River, Yokohama, Japan) were sacrificed by decapitation, and hippocampus were quickly dissected on an ice plate, immediately frozen with liquid nitrogen and stored at −80 °C until RNA isolation. All procedures were performed in accordance with our institutional guidelines after obtaining the permission of the Laboratory Animal Committee of Gunma University.

2.4. Construction of an original cDNA microarray

For future investigation of the genotype of stress responses in the nerve system, 115 clones related in 'cell/organism defense' were selected from the collected ESTs. Clones were amplified by PCR using ExTaq® (TaKaRa Shuzou, Kyoto, Japan) in a 50 μl reaction mixture and PCR was performed 12 times for each clone. Amplification was performed as follows: 3 min at 94 °C for initial denaturation, 35 cycles of 94 °C denaturing for 30 s, 60 °C annealing for 30 s and 72 °C extension for 1 min, followed by a final extension at 72 °C for 10 min. The quality and quantity of purified PCR product was confirmed using 1.2% agarose gel electrophoresis. One hundred and four of 115 clones that gave a single band then were used to construct an original cDNA microarray. Purified PCR products of each clone were resuspended in 3×SSC so that concentrations of nucleotide would be about 1 μg/μl. cDNA solutions were spotted onto poly-L-lysine-coated microarray slides (Matsunami Glass, Japan) using a capillary pen styled arrayer (OmniGrid™). cDNA spotted slides were then exposed to 120 mJ of 254 nm light to crosslink DNA on slides. Lambda phage DNA were spotted as negative controls, and GAPDH and 18S rRNA were used as positive controls.

Table 2
Redundancy of nucleotide sequences from the cDNA clones

Redundancy	No. of groups	Percentage
1	5379	75.0
2	953	13.3
3	348	4.9
4	165	2.3
5	85	1.2
6–10	158	2.2
11–20	62	0.9
21–50	22	0.3
51–100	0	0.0
>100	2	0.0

2.5. Hybridization and analysis

Total RNA was extracted from hippocampus using Qiagen RNeasy RNA extraction Kit (Qiagen, Valencia, CA, USA). We confirmed extraction of a high yield of intact total RNA by 1.2% formaldehyde agarose gel. The cDNA probes were generated by RNA reverse transcription under BD PowerScript Reverse Transcriptase (Clontech, Palo Alto, CA, USA) with a modified oligo (dT) primer (the BD SMART CDS Primer IIA, Clontech). cDNA probes then were labeled with a modified indirect labeling protocol using BD Atlas SMART Fluorescent Probe Amplification Kit® (Clontech). Briefly, primary aliphatic amino groups are incorporated through primer extension using a dNTP mix, which includes the dTTP analog, aminoallyl-dUTP. The aminoallyl-dUTP-labeled cDNA probes then are labeled with Cy3 dye (Amersham Biosciences, Piscataway, NJ, USA). In preparation for hybridization, the cDNA pellets were resuspended in 25 μl sterile deionized water. The probes then were mixed with 20 μg poly dA, 20 μg tRNA and 20 μg mouse Cot1 DNA, and finally resuspended in 50 μl of 3.4×SSC/0.5% SDS. The probe was incubated at 95 °C for 5 min, transferred to a prehybridized glass array and incubated for 18 h in a hybridization chamber (KakenGeneqs, Chiba, Japan) at 65 °C. After the hybridization, glass arrays were washed three times with agitation in the following solutions: 2×SSC/0.1% SDS for 2 min, 1×SSC/0.1% SDS for 2 min and 0.2×SSC/0.1% SDS for 2 min at room temperature. Arrays then were dried by centrifugation in a slide rack for 2 min at 800 rpm. All slides were scanned immediately using a ScanArray®Lite (PerkinElmer, Boston, MA, USA). Image analysis was performed with QuantArray (PerkinElmer) and background intensities were determined by the median pixel values.

3. Results

3.1. Characterization of rat hippocampus ESTs

A total of ~15,000 random clones from a non-unidirectional cDNA library were partially sequenced from the 3′- and 5′-end to generate 13,660 sequences with good quality. Such large-scale sequencing generally provides highly redundant ESTs that can be aligned and assembled for a set of unique genes. After 985 repetitive (7.1%) sequences and 323 mitochondrial (2.3%) DNAs were removed, the remaining ESTs were assembled into non-redundant sequence groups. The clustering analysis generated 7173 non-redundant sequences comprising 1794 groups of sequences and 5379 singletons (Table 1). Of these, 2594 were known genes. Relative frequencies of the ESTs for each gene reflect the average level of expression of the corresponding mRNAs in the pooled tissues. Since groups with redundancy of 1–5 times accounted for 96.6% of the groups, our massive sequencing was clearly effective in

Table 3
List of highly redundant cDNA clones

Redundancy	Gene products	Cellular function
119	myelin basic protein	cell structure/motility
111	proteolipid protein	cell structure/motility
50	synaptic vesicle glycoprotein 2 b	vesicular transport
47	hydroxy-δ-5-steroid dehydrogenase, 3β- and steroid δ-isomerase 1	lipid
44	myelin-associated oligodendrocytic basic protein	cell structure/motility
42	polyubiquitin	posttranslation modification/targeting
39	SNAP25 interacting protein	vesicular transport
38	calmodulin 1	effectors/modulators
35	glial fibrillary acidic protein	cytoskeletal
35	heat shock protein 8	stress response
34	eukaryotic translation elongation factor 1α1	translation factors
33	β-spectrin 3	cytoskeletal
32	calcium/calmodulin-dependent protein kinase IIα subunit	protein modification
32	SPARC-like 1	extracellular matrix
28	kinesin family member 5C	microtubule-associated proteins/motors
27	amyloidogenic glycoprotein (rAG), cognate of human A4 amyloid precursor protein	cell adhesion
26	development-related protein	unclassified
26	glutamine synthetase 1	amino acid
24	ATPase, Na^+K^+ transporting, α2	transport
24	heat shock protein 1, α	stress response
23	microtubule-associated protein 2	microtubule-associated proteins/motors
22	ATPase Na^+/K^+ transporting β1 polypeptide	transport
21	heat shock protein 90	stress response
21	stearoyl-coenzyme A desaturase 2	lipid
19	calmodulin 3	effectors/modulators
19	ribonucleotide reductase M2 subunit	nucleotide
18	myelin and lymphocyte protein	unclassified
18	neurochondrin	unclassified
18	reticulon 3	unclassified
18	syntaxin binding protein 1	vesicular transport
17	α-spectrin 2	cytoskeletal
17	cadherin 22	cell adhesion
17	glutamate oxaloacetate transaminase 1	amino acid
17	*Rattus norvegicus* clone RP31-464J4 strain Brown Norway	unclassified
17	tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein, ζ polypeptide	protein modification
16	aldolase C, fructose-biphosphate	sugar/glycolysis
16	α-tubulin	cytoskeletal
16	glutamate receptor, ionotropic, 2	receptors
16	prion protein	transcription factors
16	protein carrying the RING-H2 sequence motif	posttranslation modification/targeting
16	protein tyrosine phosphatase, receptor type, D	receptors
15	adaptor-related protein complex 2, μ1 subunit	vesicular transport
15	carboxypeptidase E	protein turnover
15	dynamin 1	cytoskeletal
15	neural visinin-like Ca^{2+}-binding protein type 2	effectors/modulators
15	neuronal pentraxin receptor	receptors
15	solute carrier family 1, member 3	channels/transport
14	ATPase, H^+ transporting, lysosomal (vacuolar proton pump), β 56/58 kDa, isoform 2	transport
14	chimerin (chimaerin) 1	intracellular transducers
14	DEAD (Asp–Glu–Ala–Asp) box polypeptide 5	RNA processing
14	myelin-associated glycoprotein	cell structure/motility
14	*N*-ethylmaleimide sensitive factor	carrier proteins/ membrane transport
14	prosaposin	unclassified
14	triosephosphate isomerase 1	sugar/glycolysis
13	growth arrest specific 7	cell cycle
13	protein tyrosine kinase 2β	protein modification
13	SNRPN upstream reading frame	unclassified
13	system N1 Na^+ and H^+-coupled glutamine transporter	channels/transport
13	tumor differentially expressed 1	unclassified
12	ankyrin 3 (G)	cytoskeletal
12	brain Ntab mRNA sequence	unclassified
12	C1-13 gene product	unclassified
12	eukaryotic translation elongation factor 2	translation factors
12	hippocalcin	effectors/modulators
12	nasal embryonic LHRH factor	unclassified
12	*Rattus norvegicus* clone RP31-422M21 strain Brown Norway	unclassified
12	S100 protein, β polypeptide	effectors/modulators
12	similar to RIKEN cDNA 1700001E04	unclassified
12	synaptotagmin 1	effectors/modulators
12	tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein, γ polypeptide	protein modification
12	v-raf-1 murine leukemia viral oncogene homolog 1	cell division
11	adaptor-related protein complex 3, β2 subunit	vesicular transport
11	amyloid β (A4) precursor-like protein 1	protein turnover
11	ATPase, Na^+K^+ transporting, α3 subunit	transport
11 •	C1q-like	unclassified
11	cytoplasmic FMR1 interacting protein 2	unclassified
11	diacylglycerol kinase ζ	lipid
11	glycoprotein m6b	cell structure/motility

(continued on next page)

Table 3 (*continued*)

Redundancy	Gene products	Cellular function
11	inositol 1,4,5-triphosphate receptor 1	receptors
11	mitogen-activated protein kinase 8 interacting protein 3	protein modification
11	nel-like 2 homolog (chicken)	unclassified
11	neurexin 1	cell adhesion
11	nucleolar protein 3 (apoptosis repressor with CARD domain)	unclassified
11	similar to expressed sequence C85658	unclassified
11	thymus cell antigen 1, θ	immunology
11	tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein, θ polypeptide	protein modification
10	adenomatosis polyposis coli	cell division
10	synaptosomal-associated protein	cell division
10	ATP/GTP binding protein 1	intracellular transducers
10	cyclic nucleotide phosphodiesterase 1	metabolism
10	nuerabin 2	cytoskeletal
10	neurofilament 3, medium	cytoskeletal
10	dystonin	extracellular matrix
10	limbic system-associated membrane protein	immunology
10	bruno-like 4, RNA binding protein (Drosophila)	RNA processing
10	tripartite motif protein 3	transcription factors
10	similar to ORF2 consensus sequence encoding endonuclease and reverse transcriptase minus RNaseH	transcription factors
10	protein phosphatase 2a, catalytic subunit, α isoform	protein modification
10	phosphofructokinase, platelet	sugar/glycolysis
10	Nogo-A	unclassified
10	sperm membrane protein (YWK-II)	unclassified

identifying a larger number of non-redundant mRNAs expressed at moderate levels (Table 2). Approximately 2.2% of the ESTs were identified 6–10 times. One hundred and one abundant sequences identified more than nine times (1.4%) are shown in Table 3. Of these, myelin basic protein (118 times) and brain myelin proteolipid protein (PLP) (111 times), the major extrinsic myelin protein and the major integral myelin membrane protein, respectively, are most abundant in this library.

3.2. Expression profile of known genes in rat hippocampus

The ESTs showing identity or high similarity to known genes were classified into seven major categories on the basis of putative general functions of the protein encoded, as described previously (categories; 'cell division', 'cell signaling/communication', 'cell structure/motility', 'cell/organism defense', 'gene/protein expression', 'metabolism' and 'unclassified') [1,13]. In total, 2594 known genes are represented in the classified data set (supplement at http://imcr.showa.gunma-u.ac.jp/lab/genetics/RHippocampus.zip). In concordance with the results in ESTs from brain observed by Adams et al., the largest category of genes was 'cell signaling and communication' except for the category 'unclassified' (618 genes, 23.8%) (Fig. 1). Successively smaller categories were 'gene/protein expression' (19.1%), 'metabolism' (13.8%), 'cell structure/motility' (9.5%), 'cell division' (4.8%) and 'cell/organism defense' (4.4%). ESTs lacking sufficient information to be classified constituted the remainder (24.5%). To further analyze the molecular complexity, each major category was subdivided according to the putative specific functions of the proteins (supplement at http://imcr.showa.gunma-u.ac.jp/lab/genetics/RHippocampus.zip). For example, the largest category, 'cell signaling and communication', was subdivided into eight subgroups (Fig. 2). Of these, 'protein modification' includes the largest number of non-redundant genes (145 genes) and ESTs for that function are also identified most frequently (429 ESTs for 145 different proteins).

3.3. Rat homologs of known genes and new members of gene families

In this study, 63.8% of the non-redundant ESTs did not match any of the known genes in the nucleotide database. To identify novel rat genes encoding proteins structurally related to the known proteins, we performed BLASTX search in the peptide databases using 4579 ESTs, with *P*-value of 10^{-10} and score of 60 as the cut off for significant similarity. Five hundred and ninety-nine non-redundant ESTs match this condition and, of these, 169 ESTs represent rat homologs of genes identified in mouse or new members of structurally related families in rat (Table 4). Of these, the proteins similar to NEDD-4 protein, retrovirus-related POL polyproteins and zinc finger proteins were most abundant. Functional analyses of the proteins identified through this approach should clarify the role of new members of structurally related families in hippocampus. The remaining ESTs (3980 genes) were not related to any other sequences in the databases. As found in similar large-scale cDNA sequencing studies carried out in other tissues, about 50% of the clones appear to be derived from genes that have not previously been described.

3.4. Construction of an original cDNA microarray

In this study, we illustrate one use of these clone sets by constructing a cDNA microarray focused on genes categorized into 'cell/organism defense' for use in further molecular studies of animal models of stress-related disorders. The hybridization pattern of normal adult rat hippocampal cDNA by our own microarray is shown in Fig. 3A and B. The 104 clones, 2 positive and 1 negative controls are spotted on the glass 10 times each. (Table 5). A

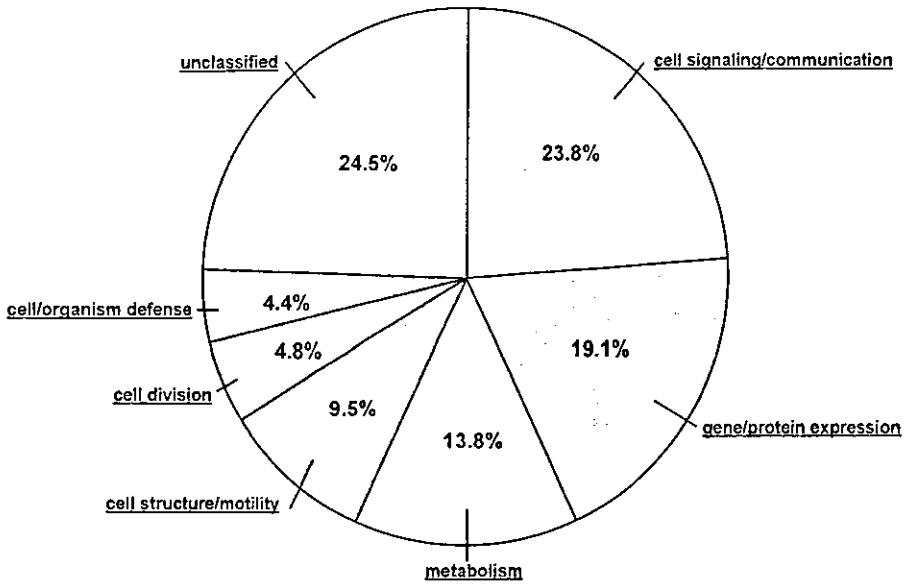

Fig. 1. Functional distribution of known genes in rat hippocampus. ESTs showing identity or high similarity to known genes were classified into seven major categories on the basis of the putative general functions of the protein encoded.

number of heat shock proteins (HSPs) and stress inducible proteins are certainly expressed also in normal rat hippocampus. As shown in Fig. 3 C, there was significant correlation between the frequencies of observed ESTs and the signal intensities of the spots (r=0.713).

4. Discussion

Expression profiling using serial analysis of gene expression (SAGE) tags and ESTs is a potent method for identifying and characterizing both known and novel genes in a given tissue. Over the past few years, cDNA libraries have been prepared from many tissues and cell lines, from which a large number of SAGE tags and ESTs have been studied. An expression profile of 30,000 genes in rat hippocampus using the SAGE method has been reported previously [5]. While SAGE analysis is unique in its ability to quantify gene expression in a given tissue with extremely high throughput, there are several limitations for the analysis of the data. For example, SAGE generates tags from the most 3′ -*Nla*III restriction site, but only on those mRNAs that have the site. Therefore, SAGE may under-perform because specific transcripts may be missed due to the absence of a recognition site for the anchoring enzyme or GC-content bias [17]. In addition, tag to gene

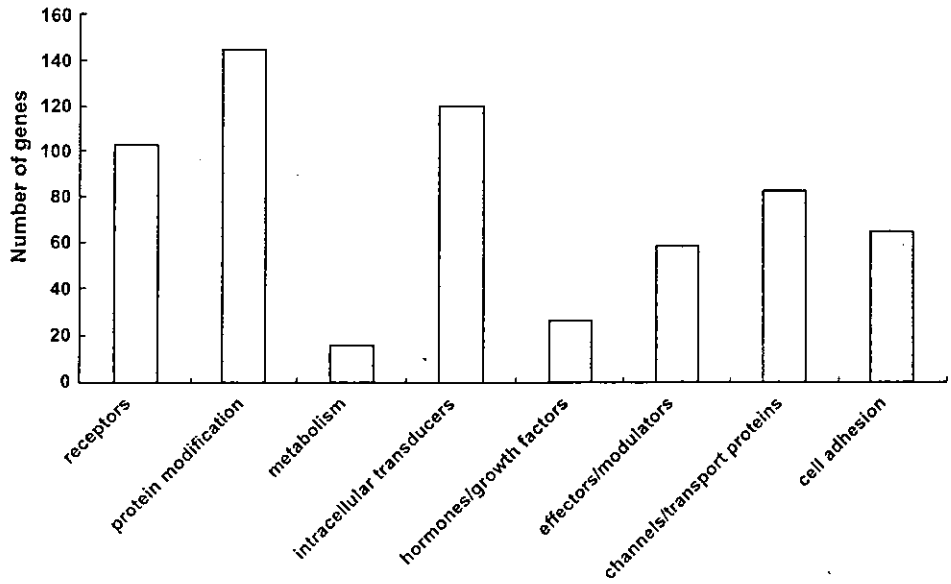

Fig. 2. Subclasses of the cell signaling/communication category. The largest category in Fig. 1 was subdivided into eight subgroups. Of these, the protein modification subgroup contains the largest number of non-redundant genes.

Table 4
Homologs of known genes and new members of gene families

Gene name	Species	Score	*P*-value	No.
14-3-3 protein eta	*Mus musculus*	85	4.00E−17	1
60S ribosomal protein L10	*Mus musculus*	58	1.00E−08	1
60S ribosomal protein L14	*Rattus norvegicus*	88	7.00E−26	1
60S ribosomal protein L34	*Mus musculus*	106	3.00E−23	1
Acetyl-CoA acetyltransferase, mitochondrial precursor	*Rattus norvegicus*	73	3.00E−13	1
Acyl-coenzyme A oxidase 1, peroxisomal	*Rattus norvegicus*	104	4.00E−23	1
AF-10 protein	*Mus musculus*	125	4.00E−29	1
Alcohol dehydrogenase class III	*Rattus norvegicus*	219	2.00E−60	1
α-Actinin 3	*Mus musculus*	199	3.00E−51	1
Amine oxidase	*Rattus norvegicus*	78	3.00E−15	1
Amyloid-like protein 1 precursor	*Mus musculus*	150	5.00E−37	1
Armadillo repeat protein deleted in velo-cardio-facial syndrome homolog	*Mus musculus*	116	3.00E−26	1
Armadillo repeat protein deleted in velo-cardio-facial syndrome homolog	*Mus musculus*	114	1.00E−25	1
Armadillo repeat protein deleted in velo-cardio-facial syndrome homolog	*Mus musculus*	51	7.00E−07	1
ATP synthase A chain	*Mus musculus*	56	1.00E−08	1
ATP-dependent DNA helicase II, 70-kDa subunit	*Mus musculus*	70	2.00E−12	1
Basement membrane-specific heparan sulfate proteoglycan core protein precursor	*Mus musculus*	87	1.00E−17	1
BCL2/adenovirus E1B 19-kDa protein-interacting protein 2	*Mus musculus*	134	6.00E−32	1
β-Chimerin	*Rattus norvegicus*	109	3.00E−24	1
β-Secretase precursor	*Rattus norvegicus*	99	6.00E−21	1
BRCA1-associated RING domain protein 1	*Rattus norvegicus*	77	2.00E−14	1
C-Rel proto-oncogene protein	*Mus musculus*	138	9.00E−38	1
Calcium-binding mitochondrial carrier protein Aralar2	*Mus musculus*	67	9.00E−12	1
cAMP-dependent protein kinase type I-β regulatory chain	*Rattus norvegicus*	130	2.00E−30	3
cAMP-dependent protein kinase type II-α regulatory chain	*Rattus norvegicus*	58	3.00E−17	2
Carbonic anhydrase XIV precursor	*Mus musculus*	57	2.00E−08	1
Carboxypeptidase H precursor	*Rattus norvegicus*	173	2.00E−43	1
Cathepsin B precursor	*Rattus norvegicus*	97	6.00E−21	1
Chloride channel protein 6	*Mus musculus*	260	9.00E−70	1
Chromobox protein homolog 6	*Mus musculus*	192	4.00E−50	1
Cofilin, muscle isoform	*Mus musculus*	52	2.00E−07	2
Cyclic-AMP-dependent transcription factor ATF-5	*Mus musculus*	69	5.00E−12	1
Cytochrome *B*	*Rattus norvegicus*	70	2.00E−12	1
Cytochrome *c* oxidase polypeptide II	*Mus musculus*	104	7.00E−23	1
Cytohesin 2	*Mus musculus*	64	2.00E−10	1
Density-regulated protein	*Mus musculus*	115	6.00E−26	1
Destrin	*Mus musculus*	145	2.00E−35	1
Developmentally regulated GTP-binding protein 1	*Mus musculus*	150	9.00E−37	1
DGCR6 protein	*Mus musculus*	84	6.00E−17	1
Disks large-associated protein 1	*Rattus norvegicus*	103	6.00E−23	1
DNA binding protein URE-B1	*Rattus norvegicus*	79	5.00E−15	1
DNA-binding protein SATB1	*Mus musculus*	62	4.00E−10	1
DnaJ homolog subfamily C member 4	*Mus musculus*	85	2.00E−17	1
Dual specificity protein phosphatase 8	*Mus musculus*	161	5.00E−40	1
Ectoderm-neural cortex-1 protein	*Mus musculus*	68	5.00E−16	1
Ectoderm-neural cortex-1 protein	*Mus musculus*	83	2.00E−16	1
Ectonucleotide pyrophosphatase/ phosphodiesterase 1	*Mus musculus*	55	8.00E−08	2
Elongation factor 2	*Rattus norvegicus*	99	5.00E−21	1
Enhancer of zeste homolog 1	*Mus musculus*	62	6.00E−10	1
Exostosin-1	*Mus musculus*	104	6.00E−23	1
FK506-binding protein precursor	*Mus musculus*	67	2.00E−11	1
Focal adhesion kinase 1	*Rattus norvegicus*	102	1.00E−22	1
Galactocerebrosidase precursor	*Mus musculus*	90	7.00E−19	1

Table 4 (*continued*)

Gene name	Species	Score	*P*-value	No.
Glucose-6-phosphate isomerase	*Mus musculus*	159	3.00E−39	1
Glutamate receptor 1 precursor	*Mus musculus*	116	3.00E−26	2
Glutamate receptor, ionotropic kainate 5 precursor	*Rattus norvegicus*	56	2.00E−08	1
Guanine nucleotide exchange factor DBS	*Rattus norvegicus*	89	5.00E−18	1
Guanine nucleotide releasing protein (GNRP) (P140 Ras-GRF)	*Rattus norvegicus*	100	7.00E−22	1
Guanine nucleotide-binding protein β subunit 5	*Mus musculus*	79	2.00E−15	3
Guanine nucleotide-binding protein G(q), α subunit	*Rattus norvegicus*	59	2.00E−09	2
Guanine nucleotide-binding protein G(S), α subunit	*Mus musculus*	99	1.00E−21	1
Histidine-rich membrane protein Ke4	*Mus musculus*	85	4.00E−17	1
Histone deacetylase 6	*Mus musculus*	188	4.00E−48	2
Importin α-3 subunit	*Mus musculus*	78	5.00E−15	1
Inhibitor of nuclear factor κ-B kinase α subunit	*Mus musculus*	85	9.00E−28	1
Integral membrane protein 2B	*Mus musculus*	74	1.00E−13	1
Integrin α-6 precursor	*Mus musculus*	72	9.00E−13	2
Inter-α-trypsin inhibitor heavy chain H3 precursor	*Rattus norvegicus*	102	4.00E−22	1
Interferon-α/β receptor α chain precursor	*Mus musculus*	65	3.00E−11	1
Kinesin-like protein KIF3A	*Mus musculus*	74	1.00E−13	1
Lamin B3	*Mus musculus*	125	7.00E−29	1
Latent transforming growth factor β binding protein 1 precursor	*Rattus norvegicus*	95	9.00E−20	1
Leukocyte tyrosine kinase receptor precursor	*Mus musculus*	78	9.00E−15	1
LIM/homeobox protein Lhx6.1	*Mus musculus*	80	1.00E−15	1
Low molecular weight phosphotyrosine protein phosphatase ACP1/ACP2	*Rattus norvegicus*	97	2.00E−20	1
Lysosomal α-mannosidase precursor	*Mus musculus*	136	2.00E−32	1
Lysosomal α-mannosidase precursor	*Mus musculus*	110	1.00E−32	1

Table 4 (*continued*)

Gene name	Species	Score	*P*-value	No.
Methionyl-tRNA formyltransferase, mitochondrial precursor	*Mus musculus*	87	2.00E−17	1
Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial precursor	*Rattus norvegicus*	84	1.00E−16	1
Microtubule-associated protein 1A	*Mus musculus*	73	1.00E−13	1
Microtubule-associated protein 4	*Mus musculus*	95	6.00E−20	2
Mitochondrial trifunctional enzyme α subunit precursor	*Rattus norvegicus*	112	3.00E−25	1
Mitogen-activated protein kinase 7	*Mus musculus*	125	4.00E−29	1
Myelin and lymphocyte protein	*Rattus norvegicus*	58	1.00E−08	1
Myelin basic protein S	*Rattus norvegicus*	90	2.00E−18	1
Myotubularin-related protein 3	*Mus musculus*	50	4.00E−07	1
NADPH/adrenodoxin oxidoreductase, mitochondrial precursor	*Rattus norvegicus*	72	5.00E−13	1
NEDD-4 protein	*Mus musculus*	56	3.00E−08	1
NEDD-4 protein	*Mus musculus*	54	1.00E−07	1
NEDD-4 protein	*Mus musculus*	54	1.00E−07	1
NEDD-4 protein	*Mus musculus*	54	1.00E−07	1
NEDD-4 protein	*Mus musculus*	53	4.00E−07	1
NEDD-4 protein	*Mus musculus*	53	2.00E−07	1
NEDD-4 protein	*Mus musculus*	51	8.00E−07	1
NEDD-4 protein	*Mus musculus*	56	4.00E−08	1
NEDD-4 protein	*Mus musculus*	54	5.00E−08	1
NEDD-4 protein	*Mus musculus*	53	1.00E−07	1
Neighbor of A-kinase anchoring protein 95	*Mus musculus*	88	1.00E−17	1
Neural Wiskott-Aldrich syndrome protein	*Rattus norvegicus*	125	1.00E−29	1
Neuroendocrine convertase 3 precursor	*Mus musculus*	70	3.00E−12	2
Neuronal membrane glycoprotein M6-A	*Mus musculus*	56	1.00E−08	2
Neuronal-specific septin 3	*Mus musculus*	55	1.00E−07	3
NGFI-A binding protein 1	*Rattus norvegicus*	92	7.00E−19	1
Nidogen-2 precursor	*Mus musculus*	50	2.00E−12	2
NK-tumor recognition protein	*Mus musculus*	120	2.00E−27	1
Nucleolin	*Rattus norvegicus*	57	8.00E−09	2
Numb-like protein	*Mus musculus*	213	2.00E−55	1
Peroxisomal targeting signal 2 receptor	*Mus musculus*	59	7.00E−09	1

(*continued on next page*)

Table 4 (*continued*)

Gene name	Species	Score	*P*-value	No.
Phosphatidylinositol-glycan-specific phospholipase D1 precursor	*Mus musculus*	127	1.00E−29	1
Phospholipase D2	*Rattus norvegicus*	58	1.00E−08	1
Phospholipid hydroperoxide glutathione peroxidase, mitochondrial precursor	*Rattus norvegicus*	114	6.00E−26	1
Polyadenylate-binding protein 1	*Mus musculus*	61	3.00E−10	1
Potential phospholipid-transporting ATPase IIA	*Mus musculus*	54	4.00E−08	1
Pristanoyl-CoA oxidase	*Rattus norvegicus*	82	2.00E−16	1
Probable calcium-binding protein Dd112	*Mus musculus*	58	1.00E−08	1
Probable cation-transporting ATPase 1	*Mus musculus*	130	4.00E−31	1
Prostaglandin F2-α receptor regulatory protein precursor	*Rattus norvegicus*	74	2.00E−13	1
Protein kinase C, γ type	*Mus musculus*	129	5.00E−30	1
Proto-oncogene tyrosine-protein kinase MER precursor	*Rattus norvegicus*	68	1.00E−11	1
Protocadherin 3 precursor	*Rattus norvegicus*	120	9.00E−28	1
Putative protein C21orf62 homolog	*Mus musculus*	185	6.00E−47	3
Ras-related protein Rab-1B	*Rattus norvegicus*	113	1.00E−25	1
Regulator of G-protein signaling 5	*Rattus norvegicus*	97	9.00E−21	1
Retrovirus-related POL polyprotein	*Mus musculus*	146	1.00E−35	1
Retrovirus-related POL polyprotein	*Mus musculus*	52	4.00E−07	1
Retrovirus-related POL polyprotein	*Mus musculus*	134	1.00E−31	1
Retrovirus-related POL polyprotein	*Mus musculus*	94	2.00E−19	1
Retrovirus-related POL polyprotein	*Mus musculus*	75	3.00E−14	1
Retrovirus-related POL polyprotein	*Mus musculus*	65	7.00E−11	1
Retrovirus-related POL polyprotein	*Mus musculus*	60	2.00E−09	1
Retrovirus-related POL polyprotein	*Mus musculus*	59	3.00E−15	2
Retrovirus-related POL polyprotein	*Mus musculus*	56	5.00E−08	1
RING finger protein 27	*Mus musculus*	147	9.00E−36	2
RING finger protein 4	*Rattus norvegicus*	65	3.00E−11	1

Table 4 (*continued*)

Gene name	Species	Score	*P*-value	No.
Semaphorin 4D precursor	*Mus musculus*	91	2.00E−18	1
Semaphorin 5A precursor	*Mus musculus*	121	2.00E−28	1
Semaphorin 6B precursor	*Rattus norvegicus*	61	2.00E−09	1
Septin 2	*Mus musculus*	87	3.00E−17	1
Serine/threonine protein kinase 25	*Mus musculus*	70	9.00E−13	7
Serine/threonine-protein kinase 19	*Mus musculus*	114	9.00E−26	1
Single-minded homolog 2	*Mus musculus*	68	9.00E−12	1
Sodium/calcium exchanger 2 precursor	*Rattus norvegicus*	73	1.00E−13	1
SOX-13 protein	*Mus musculus*	149	2.00E−36	1
Splicing factor 3B subunit 1	*Mus musculus*	123	7.00E−29	1
SSXT protein	*Mus musculus*	71	7.00E−13	1
Surfeit locus protein 6	*Mus musculus*	58	2.00E−10	1
Synaptojanin 2	*Rattus norvegicus*	70	4.00E−12	1
T-cell receptor α chain V region 2B4 precursor	*Mus musculus*	80	3.00E−15	3
T-complex protein 1, δ subunit	*Mus musculus*	53	2.00E−07	1
TLM protein	*Mus musculus*	62	8.00E−10	5
Transcription factor 17	*Mus musculus*	86	7.00E−17	1
Ubiquinol-cytochrome *C* reductase complex core protein I, mitochondrial precursor	*Mus musculus*	205	5.00E−53	1
Ubiquitin carboxyl-terminal hydrolase 2	*Mus musculus*	127	3.00E−30	2
Uridine kinase	*Mus musculus*	127	7.00E−30	1
Voltage-gated potassium channel protein Kv3.1	*Mus musculus*	109	1.00E−24	1
VPS26 protein homolog	*Mus musculus*	134	1.00E−31	1
VPS26 protein homolog	*Mus musculus*	74	5.00E−14	1
Werner syndrome helicase homolog	*Mus musculus*	159	3.00E−39	1
X inactive specific transcript protein	*Mus musculus*	49	5.00E−07	1
Zinc finger homeobox protein 1b	*Mus musculus*	137	3.00E−33	1
Zinc finger protein 27	*Mus musculus*	57	1.00E−08	1
Zinc finger protein 37	*Mus musculus*	86	3.00E−17	1
Zinc finger protein 37	*Mus musculus*	77	1.00E−14	1
Zinc finger protein 46	*Mus musculus*	135	2.00E−32	1
Zinc finger protein 60	*Mus musculus*	136	2.00E−32	1
Zinc finger protein 90	*Mus musculus*	94	2.00E−19	1
Zinc finger protein 92	*Mus musculus*	219	3.00E−59	1
Zinc-finger protein RFP	*Mus musculus*	84	1.00E−16	1

mapping is not completely definitive, as some tags correspond to several genes. Furthermore, incorrect tag counts can arise from incomplete digestion or alternative

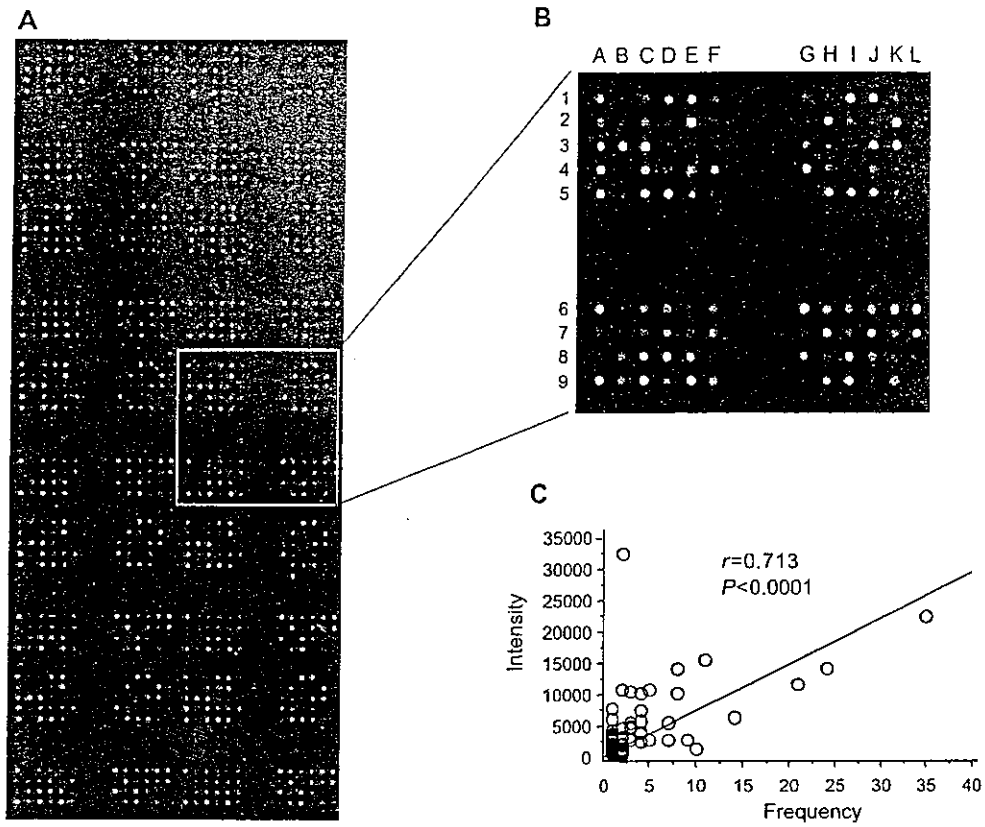

Fig. 3. (A) Hybridization pattern of normal adult rat hippocampal cDNA. The 104 EST clones, 2 positive and 1 negative controls are spotted on the glass 10 times each. (B) Zoom up figure of one sub-array. There are 107 spots in each sub-array. (C) The correlation between the frequencies of observed ESTs and the signal intensities of the spots. There is significant correlation between the frequencies and the signal intensities of the ESTs (r=0.713).

polyadenylation, giving rise to multiple tags derived from a single transcript.

However, there has been no report of large-scale generation of ESTs from rat hippocampus. EST analysis has certain advantages over other methods such as SAGE for examining the transcript repertoire of tissues. In particular, EST sequences that cover regions of the coding sequence can reveal variant transcripts and splice forms, many of which have functional significance.

In this study, we describe a collection of 13,660 hippocampus-related ESTs representing 7173 different transcripts. With respect to overall EST distribution (i.e. known gene matches), the results in rat hippocampus in this study differ somewhat from those obtained from other tissues. The largest category of function has been reported to be gene/protein expression in studies of ESTs in tissues other than brain [1,13,15]. The largest number of genes obtained from the rat hippocampus cDNA library encoded proteins related to 'cell signaling and communication', as in earlier EST study of brain [1]. The functional categorization of known genes reflects general differences in gene expression between different tissues, and may reflect tissue-specific function. For example, calmodulin and CAMK II, genes involved in 'protein modification' of 'cell signaling/communication' were identified most frequently in this research, and play an important role in the early stages of LTP (Long-term potentiation) in hippocampus [9,19,26]. Genes of mitogen-activated protein kinase signaling that belong to the core signaling pathways involved in memory storage [16,18,27] also were found frequently. Tyrosine 3-monooxgenase/tryptophan 5-monooxgenase activation proteins were found in this library abundantly, and play a role in the regulation of serotonin and biosynthesis of brain noradrenaline, reuptake of which is inhibited by antidepressant drugs [14]. In this study, there was also a larger portion of unclassified genes in hippocampus (24.5%) due, at least in part, to the large numbers of hypothetical proteins generated by recent high throughput genome sequencing efforts.

Although the rat genome project has been released [10], a significant fraction of the genes are hypothetical, revealed only by a computer prediction program. Thus, hippocampal transcripts provide a richer resource for analysis of novel genes related to known proteins. This catalog of expressed genes should facilitate the development of tissue-specific cDNA microarray technology.

Various kinds of stress induce the synthesis of stress proteins that protect cells from subsequent lethal stress. HSPs are ubiquitous cellular proteins with a highly conserved structure, mode of regulation, and function, indicating their important role in cellular functions. HSPs are induced by physical and chemical insults, and confer

Table 5
Gene list of microarray

Spot	Gene name
A1	β prime COP
A2	adenylate kinase 1
A3	HLA-B-associated transcript 1A
A4	heat shock protein 90 (rats, brain, mRNA, 2524 nt)
A5	calnexin
A6	branched chain aminotransferase 1, cytosolic
A7	P450 (cytochrome) oxidoreductase
A8	X-ray repair cross-complementing group 1 protein
A9	carbonic anhydrase 2
B1	oxygen-regulated protein
B2	neogenin
B3	transferrin
B4	RPT protein similar to yeast MRS2
B5	adenylate kinase 4
B6	nuclear factor of κ light polypeptide gene enhancer in B cells inhibitor-like 1
B7	branched chain aminotransferase 2, mitochondrial
B8	metallothionein 1, pseudogene A
B9	RT1 class Ib gene, H2-TL-like, grc region (N1)
C1	glutathione *S*-transferase omega 1
C2	carbonic anhydrase 11
C3	thymus cell antigen 1, θ
C4	coatomer protein complex, subunit γ1
C5	heat shock protein 60 (liver)
C6	tweety homolog 1 (Drosophila)
C7	excision repair cross-complementing rodent repair deficiency, complementation group 1
C8	calcium binding atopy-related autoantigen 1
C9	heat shock protein 8
D1	hemoglobin, α1
D2	odd Oz/ten-m homolog 4 (Drosophila)
D3	DnaJ (Hsp40) homolog, subfamily B, member 5
D4	endoplasmic reticulum protein 29
D5	Nsf: *N*-ethylmaleimide sensitive factor
D6	esterase 10
D7	acyl-coA oxidase
D8	coatomer protein complex, subunit γ2
D9	surfeit 1
E1	DnaJ-like protein
E2	glutathione *S*-transferase, μ type 3 (Yb3)
E3	activated leukocyte cell adhesion molecule
E4	α thalassemia/mental retardation syndrome X-linked homolog (human)
E5	GAPDH
E6	cleft lip and palate associated transmembrane protein 1
E7	heat shock 70 kDa protein 4
E8	heat shock 70 kDa protein 5
E9	peroxiredoxin 2
F1	cytochrome P450-like protein
F2	topoisomerase (DNA) III β
F3	thioredoxin domain containing 5
F4	superoxide dismutase 2
F5	negative control
F6	complement component factor h
F7	adenylate kinase 5
F8	similar to MHC class Ib RT1.S3
F9	limbic system-associated membrane protein
G1	epoxide hydrolase 1
G2	TAP binding protein
G3	Transthyretin (prealbumin, amyloidosis type I)
G4	hemoglobin β chain complex
G5	MHC class II-associated invariant chain
G6	glutathione *S*-transferase, μ1
G7	DnaJ (Hsp40) homolog, subfamily B, member 1
G8	cytochrome P450, subfamily 51
G9	immunoglobulin superfamily, member 8
H1	ferritin light chain 2
H2	glutathione transferase subunit 8
H3	mannose-binding protein associated serine protease-1
H4	NY-REN-18 antigen
H5	DnaJ (Hsp40) homolog, subfamily B, member 6
H6	T cell receptor γ locus, TCR γ1 and γ3 gene clusters
H7	coatomer protein complex, subunit β1
H8	glutamate cysteine ligase, modifier subunit
H9	Sacm21/RT1-A intergenic region, haplotype RT1n and partial RT1-A gene for MHC class I antigen
I1	heat shock protein 1, α
I2	RAD50 homolog (*S. cerevisiae*)
I3	germline MHC class I gene, complete cds
I4	acyl-coenzyme A oxidase 3, pristanoyl
I5	β-2-microglobulin
I6	DnaJ (Hsp40) homolog, subfamily C, member 5
I7	non MHC restricted killing associated
I8	suppression of tumorigenicity 13 (colon carcinoma) Hsp70-interacting protein
I9	ferritin, heavy polypeptide 1
J1	creatine kinase, brain
J2	HLA-B-associated transcript 3
J3	thioredoxin-like 2
J4	B-cell receptor-associated protein 37
J5	selenium-dependent glutathione peroxidase mRNA, complete cds
J6	thioredoxin 2
J7	glycoprotein Ib (platelet), β polypeptide
J8	topoisomerase (DNA) II β
J9	HLA-B associated transcript 2
K1	islet cell autoantigen 1, 69 kDa
K2	heat shock factor binding protein 1
K3	stress-induced-phosphoprotein 1 (Hsp70/Hsp90-organizing protein)
K4	C4 complement protein
K5	18S rRNA
K6	thioredoxin domain containing 1
K7	platelet-activating factor acetylhydrolase α2 subunit (PAF-AH α2)
K8	coagulation factor C homolog (*Limulus polyphemus*)
K9	adenylate kinase 3
L1	glutathione *S*-transferase, θ2
L2	RT1 class Ib gene (Aw2)
L3	proprotein convertase subtilisin/kexin type 3
L4	thioredoxin-like (32 kDa)
L5	blank
L6	glutathione *S*-transferase Yc1 subunit (rats, fetal liver, mRNA, 1052 nt)
L7	DnaJ (Hsp40) homolog, subfamily C, member 7
L8	ligase III, DNA, ATP-dependent
L9	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily e, member 1

cellular resistance to subsequent lethal stressors. For example, heat and ischemia are well known stimuli that induce the HSP70 family in the central nervous system [22]. In mammals, the HSP70 family is also stimulated by stress mediators such as adrenocorticotropic hormone and catecholamines [3]. Accordingly, expression of the HSP70 family may be associated with stress responses involving

the endocrine, nervous, and immune systems. Glucocorticoid levels also are increased in depressed patients [4] and glucocorticoid receptor function is regulated by HPSs [23]. Thus, further investigation of the relationship between HSPs and psycho-physiological stress in hippocampus should be fruitful. In the present study, for example, we constructed a cDNA microarray focused on genes categorized into 'cell/organism defense', including a number of stress inducible factors such as HSPs, for further molecular studies in animal models of stress-related disorders.

As shown in Fig. 3 C, there was significant correlation between the frequencies of observed ESTs categorized into 'cell/organism defense' and the signal intensities of the spots, suggesting that the profiling of transcripts by ESTs reflects the actual gene expression pattern well. These clone sets allow for the production of large numbers of cDNA microarrays at low cost, permitting the use of large numbers of replicates in gene expression profiling experiments, which should lead to increased data quality. In addition, because many of the cDNAs spotted on our microarrays are not contained on commercial platforms at present, they should provide a unique and useful tool for molecular studies of animal models of stress-related disorders.

Functional analysis of newly discovered genes through this approach might clarify the molecular mechanisms underlying the pathogenesis of stress-related disorders sufficiently to reveal novel therapeutic targets. Integrated information on hippocampus-specific functions and mapping of our ESTs on the human chromosome should complement genetic linkage studies and facilitate positional candidate cloning for the identification of genes of memory-, learning- and stress-related disorders in genetically defined regions.

Acknowledgments

This study was supported by Grants-in-Aid for Scientific Research A-C and for Scientific Research on Priority Areas (C) "Medical Genome Science" from the Japanese Ministry of Science, Education, Sports, Culture and Technology; for a Health and Labor Science Research Grant for Special Research from the Japanese Ministry of Health, Labor and Welfare; and for the Yamanouchi Foundation for Research on Metabolic Disorders, Japan Diabetes Foundation and Takeda Science Foundation.

References

[1] M.D. Adams, A.R. Kerlavage, R.D. Fleischmann, R.A. Fuldner, C.J. Bult, N.K. Lee, E.F. Kirkness, K.G. Weinstock, J.D. Gocayne, O. White, et al., Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence, Nature 377 (1995) 3–174 (Suppl.).

[2] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[3] M.J. Blake, D.J. Buckley, A.R. Buckley, Dopaminergic regulation of heat shock protein-70 expression in adrenal gland and aorta, Endocrinology 132 (1993) 1063–1070.

[4] B.J. Carroll, The dexamethasone suppression test for melancholia, Br. J. Psychiatry 140 (1982) 292–304.

[5] N.A. Datson, J. van der Perk, E.R. de Kloet, E. Vreugdenhil, Expression profile of 30,000 genes in rat hippocampus using SAGE, Hippocampus 11 (2001) 430–444.

[6] R.S. Duman, J. Malberg, J. Thome, Neural plasiticity to stress and antidepressant treatment, Biol. Psychiatry 46 (1999) 1181–1191.

[7] P.S. Eriksson, E. Perfilieva, T. Bjork-Eriksson, A.M. Alborn, C. Nordborg, D.A. Peterson, F.H. Gage, Neurogenesis in the adult human hippocampus, Nat. Med. 4 (1998) 1313–1317.

[8] E. Fuchs, G. Flugge, Stress, glucocorticoids and structural plasticity of the hippocampus, Neurosci. Biobehav. Rev. 23 (1998) 295–300.

[9] K. Fukunaga, L. Stoppini, E. Miyamoto, D. Muller, Long-term potentiation is associated with an increased activity of Ca^{2+}/calmodulin-dependent protein kinase II, J. Biol. Chem. 268 (1993) 7863–7867.

[10] R.A. Gibbs, G.M. Weinstock, M.L. Metzker, D.M. Muzny, E.J. Sodergren, S. Scherer, G. Scott, D. Steffen, K.C. Worley, P.E. Burch, et al., Genome sequence of the Brown Norway rat yields insights into mammalian evolution, Nature 428 (2004) 493–521.

[11] E. Gould, P. Tanapat, Stress and hippocampal neurogenesis, Biol. Psychiatry 46 (1999) 1472–1479.

[12] J.P. Herman, M.K. Schafer, E.A. Young, R. Thompson, J. Douglass, H. Akil, S.J. Watson, Evidence for hippocampal regulation of neuroendocrine neurons of the hypothalamo–pituitary–adrenocortical axis, J. Neurosci. 9 (1989) 3072–3082.

[13] D.M. Hwang, A.A. Dempsey, R.X. Wang, M. Rezvani, J.D. Barrans, et al., A genome-based resource for molecular cardiovascular medicine. Toward a compendium of cardiovascular genes, Circulation 96 (1997) 4146–4203.

[14] T. Ichimura, T. Isobe, T. Okuyama, T. Yamauchi, H. Fujisawa, Brain 14-3-3 protein is an activator protein that activates tryptophan 5-monooxygenase and tyrosine 3-monooxygenase in the presence of Ca^{2+}, calmodulin-dependent protein kinase II, FEBS Lett. 13 (1987) 79–82.

[15] L. Jin, H. Wang, T. Narita, R. Kikuno, O. Ohara, N. Shihara, T. Nishigori, Y. Horikawa, J. Takeda, Expression profile of mRNAs from human pancreatic islet tumors, J. Mol. Endocrinol. 31 (2003) 519–528.

[16] E.R. Kandel, The molecular biology of memory storage: a dialogue between genes and synapses, Science 294 (2001) 1030–1038.

[17] E.H. Margulies, S.L.R. Kardia, J.W. Innis, Identification and prevention of a GC content bias in SAGE libraries, Nucleic Acids Res. 29 (12) (2001) E60-0.

[18] A. Matynia, S.G. Angnostaras, A.J. Silva, Weaving the molecular and cognitive strands of memory, Neuron 32 (2001) 557–559.

[19] M. Mayford, J. Wang, E. Kandel, T.S. O'Dell, CAMK II regulates the frequency–response function of hippocampal synapses for the production of both LTD and LTP, Cell 81 (1995) 891–904.

[20] B.S. McEwen, The neurobiology of stress: from serendipity to clinical relevance, Brain Res. 886 (2000) 172–189.

[21] M. Nilsson, E. Perfilieva, U. Johansson, O. Orwar, P.S. Eriksson, Enriched environment increases neurogenesis in the adult rat dentate gyrus and improves spatial memory, J. Neurobiol. 39 (1999) 569–578.

[22] T.S. Nowak Jr., U. Bond, M.J. Schlesinger, Heat shock RNA levels in brain and other tissues after hyperthermia and transient ischemia, J. Neurochem. 54 (1990) 451–458.

[23] W.B. Pratt, D.O. Tort, Steroid receptor interactions with heat shock protein and immunophilin chaperones, Endocr. Rev. 18 (1997) 306–360.

[24] R.M. Sapolsky, Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders, Arch. Gen. Psychiatry 57 (2000) 925–935.

[25] Y.I. Sheline, P.W. Wang, M.H. Gado, J.G. Csernansky, M.W. Vannier, Hippocampal atrophy in recurrent major depression, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 3908–3913.
[26] A.C. Silva, C. Stevens, S. Tonegawa, Y. Wang, Deficient hippocampal long-term potentination in α-calcium-calmodulin kinase II mutant mice, Science 257 (1992) 201–206.
[27] J.D. Sweatt, The neuronal MAP kinase cascade: a biochemical signal integration system subserving synaptic plasticity and memory, J. Neurochem. 76 (2001) 1–10.
[28] J. Takeda, H. Yano, S. Eng, G.I. Bell, A molecular inventory of human pancreatic islets: sequence analysis of 1000 cDNA clones, Hum. Mol. Genet. 2 (1993) 1793–1798.
[29] H. van Praag, G. Kempermann, F.H. Gage, E. Gould, P. Tanapat, Running increase cell prolification and neurogenesis in the adult mouse dentate gyrus, Nat. Neurosci. 2 (1999) 266–270.
[30] E.A. Young, R.F. Haskett, V. Murphy-Weinberg, S.J. Watson, H. Akil, Loss of glucocorticoid fast feedback in depression, Arch. Gen. Psychiatry 48 (1991) 693–699.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Genetics and Metabolism 82 (2004) 238–245
Molecular Genetics and Metabolism
www.elsevier.com/locate/ymgme

Genetic variations in the WFS1 gene in Japanese with type 2 diabetes and bipolar disorder

Takanori Kawamoto,[a] Yukio Horikawa,[b,c,*] Takeshi Tanaka,[a] Noriko Kabe,[a] Jun Takeda,[b,c,d] and Masahiko Mikuni[a]

[a] *Department of Neuropsychiatry, Gunma University School of Medicine, Gunma, Japan*
[b] *Laboratory of Molecular Genetics, Department of Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, Gunma 371-8512, Japan*
[c] *Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), Kawaguchi, Japan*
[d] *Division of Bioregulatory Medicine, Department of Endocrinology, Diabetes and Rheumatology, Gifu University School of Medicine, Gifu, Japan*

Received 12 February 2004; received in revised form 23 April 2004; accepted 27 April 2004
Available online 11 June 2004

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Genetics and Metabolism 82 (2004) 238–245
Molecular Genetics and Metabolism
www.elsevier.com/locate/ymgme

Genetic variations in the WFS1 gene in Japanese with type 2 diabetes and bipolar disorder

Takanori Kawamoto,[a] Yukio Horikawa,[b,c,*] Takeshi Tanaka,[a] Noriko Kabe,[a] Jun Takeda,[b,c,d] and Masahiko Mikuni[a]

[a] *Department of Neuropsychiatry, Gunma University School of Medicine, Gunma, Japan*
[b] *Laboratory of Molecular Genetics, Department of Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, Gunma 371-8512, Japan*
[c] *Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), Kawaguchi, Japan*
[d] *Division of Bioregulatory Medicine, Department of Endocrinology, Diabetes and Rheumatology, Gifu University School of Medicine, Gifu, Japan*

Received 12 February 2004; received in revised form 23 April 2004; accepted 27 April 2004
Available online 11 June 2004

Abstract

Diabetic and psychiatric symptoms often appear in patients with Wolfram syndrome, and obligate carriers of *WFS1* have increased prevalence of type 2 diabetes and are more likely to require hospitalization for psychiatric illness including bipolar disorder. To identify the polymorphisms in Japanese, we examined a region of ~50 kb covering the entire *WFS1* gene, and evaluated the patterns of linkage disequilibrium. We found a total of 42 variations including 8 novel coding single nucleotide polymorphisms (A6T, A134A, N159N, T170T, E237K, R383C, V412L, and V503G), 14 novel non-coding polymorphisms, and 2 linkage disequilibrium blocks. We also performed association studies in patients with type 2 diabetes mellitus and patients with bipolar disorder. The haplotype comprising R456 and H611 was most associated with type 2 diabetes ($p = 0.013$) and the haplotype comprising g. −15503C/T and g. 16226G/A was most associated with bipolar disorder ($p = 0.006$), but neither reached significant difference after multiple adjustment. These genetic variations and linkage disequilibrium patterns in *WFS1* in Japanese should be useful in further investigation of genetic diversities of *WFS1* and various related disorders.

Keywords: Genetics; Mood disorder; Population study; Linkage disequilibrium; Single nucleotide polymorphism

Introduction

Wolfram syndrome, also known as DIDMOAD (diabetes insipidus, diabetes mellitus, optic atrophy, and deafness [OMIM 222300]), was first described by Wolfram and Wagener [1]. While only juvenile onset diabetes mellitus and progressive optic nerve atrophy are required for the diagnosis, many patients also develop diabetes insipidus, sensorineuronal hearing loss, ataxia, peripheral neuropathy, urinary tract atonia, and psychiatric illnesses such as psychosis, severe depression, and dementia [2].

Wolfram syndrome has been shown to have links with D4S432–D4S431 at chromosome 4p16 [3,4]. We earlier reported two siblings with Wolfram syndrome who demonstrated mood symptoms [5], and proceeded with a multi-institutional coordinated effort to discover the genetic etiology of the disease. Recently, mutations in the gene *WFS1/wolframin* were identified in patients with Wolfram syndrome [6,7]. The gene, which encodes a novel protein containing the predicted transmembrane domains, is expressed ubiquitously, with high expression in pancreatic islets and specific neurons (hippocampus CA1, amygdaloid areas, olfactory tubercles, and superficial layers of the allocortex). The subcellular localization of this protein has been determined to be primarily in the endoplasmic reticulum [8], but it's function is not established.

* Corresponding author. Fax: +81-27-220-8889.
E-mail address: yhorikaw@showa.gunma-u.ac.jp (Y. Horikawa).

1096-7192/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2004.04.012

The prevalence of this autosomal recessive syndrome was estimated as ~1/770,000 in the United Kingdom [9]. While Wolfram syndrome is rare, obligate carriers have increased prevalence of type 2 diabetes mellitus [9,10], and heterozygous carriers are reported to be 26-fold more likely to require hospitalization for psychiatric illness [11].

Bipolar disorder, also called manic-depression, is characterized by mood swings between states of depression and elation, and genetic predisposition is thought to be an important factor. A linkage study of a large Scottish family provided a maximum LOD score of 4.8 in the region D4S431–D4S403 [12]. Other groups have presented supportive linkage evidence with markers in this region [13,14]. The psychiatric phenotypes observed in carriers and patients with Wolfram syndrome and the location of *WFS1* in the region linked to bipolar disorder have suggested a role of the gene in the development of the disease.

Evidence of abnormal glucose metabolism in psychiatric patients has been accumulating since the early 20th century [15]. Some studies report an increased prevalence of diabetes in hospitalized manic-depressives [16–19]. Gavard et al. carried out an evaluation of 20 studies of the diabetic personality, and found an increased prevalence of depression. A relationship between psychiatric disorder and diabetes mellitus has been suggested by mutations in *WFS1* that affect both diabetic and psychiatric phenotypes. Indeed, we estimated the LOD score for susceptibility to type 2 diabetes in one of the Wolfram pedigrees (WS-1) [6] using additional family information, and found suggestive linkage (M. Mikuni et al. unpublished).

In this study, we examined all of the regions of *WFS1* in Japanese to detect single nucleotide polymorphisms (SNPs)[1] as genetic markers, and evaluated the pattern of linkage disequilibrium (LD) to provide information on population diversity in this gene. We also performed association studies in Japanese patients with type 2 diabetes mellitus and patients with bipolar disorder.

Materials and methods

Subjects

One hundred and ninety two patients with type 2 diabetes mellitus (male/female, 114/78; age, 62.0 ± 11.2 years; age at diagnosis, 49.8 ± 11.0 years; postprandial glucose, 168.5 ± 69.0 mg/dl; hemoglobin (Hb) A_{1C}, 6.7 ± 1.1%; body mass index (BMI), 23.9 ± 3.5 kg/m^2) and 192 controls (male/female, 74/118; age, 67.6 ± 5.8; HbA_{1C}, 4.9 ± 0.3%; BMI, 22.9 ± 2.7 kg/m^2) were examined. Patients were diagnosed with type 2 diabetes by medical records or by 75 g oral glucose tolerance test according to the criteria of the Japan Diabetes Society. Control subjects were recruited on the following criteria: 60 or more years of age, no past history of diagnosis of diabetes, HbA_{1C} less than 5.6%, and no diabetes in family members or second degree relatives. Eighteen patients with bipolar I disorders (male/female, 9/9; age, 46.7 ± 12.4) and 29 patients with bipolar II disorders (male/female, 16/13; age, 53.6 ± 12.3) were recruited from Gunma University Hospital and local hospitals in Gunma prefecture, met DSM-IV diagnostic criteria (Diagnostic and Statistical Manual of Mental Disorders, 4th edition) [20], and were assessed by trained clinicians on the basis of unstructured interviews supplemented by case-note reviews. Ninety-six Japanese random controls (male/female, 39/57; age, 68.8 ± 5.6) were examined for comparisons of genetic variations. The study was approved by the Ethics Committee of Gunma University, and included the written informed consent of each subject.

Detection of polymorphisms in WFS1

Genomic DNA was extracted from samples of whole blood using QIAamp DNA Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's instruction. Twelve of the random control samples (24 alleles) were used to detect single nucleotide polymorphisms (SNPs) in *WFS1*. Primers for PCR experiments were designed by Primer3 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) on the basis of the genomic contig sequence (GenBank Accession No. NT_006051) of the *WFS1* region. The mixture for the PCR was 20 μl in 10 ng template DNA, 0.5 mM each dNTP, 2.5 pmol each forward and reverse primer, 0.5U Ex*Taq* polymerase (Takara, Kyoto, Japan), and 2 μl of 10× PCR buffer. The reaction conditions were an initial denaturation step of 95 °C for 3 min, subsequent 40 cycle reactions at 94 °C for 30 s, 52–62 °C for 30 s, and 72 °C for 1 min, and a final extension step of 72 °C for 10 min. A 3 μl aliquot from each reaction was assayed on a 1% agarose gel to confirm the product, and the remainder was purified using MultiScreen Filtration System (MILLIPORE, Billerica, MA, USA) with Sephadex G-75 (Amersham Biosciences, Piscataway, NJ, USA). Each PCR product was subjected to cycle sequencing with BigDye terminator Cycle Sequencing FS (Applied Biosystems, Foster, CA, USA) using each forward and reverse primer. Reaction products were purified by ethanol precipitation, and sequenced by ABI PRISM 377 sequencer. Results were processed with Autoassembler, version 2.1 (Applied Biosystems, Foster, CA, USA) to compare sequences.

[1] *Abbreviations used:* bp, base pair; SNPs, single nucleotide polymorphisms; LD, linkage disequilibrium; PCR, polymerase chain reaction; cSNPs, coding single nucleotide polymorphisms.

Mutation screening and genotyping of frequent polymorphisms in WFS1

We examined the coding region of *WFS1* and genotyped sixteen frequent SNPs in the 47 bipolar patients and 96 control subjects. All exons were examined in the 192 type 2 diabetic patients and 192 controls.

Estimation of haplotype frequencies and evaluation of pattern of linkage disequilibrium

Haplotypes were inferred by the expectation-maximization method by Arlequin Software (http://anthro.unige.ch/arlequin). The coefficient for LD, D', and r^2 value was estimated by GOLD software (http://www.well.ox.ac.uk/asthma/GOLD).

Statistical analyses

Statistical difference in allele frequencies between bipolar disorder or diabetes and control groups was assessed by χ^2 test (including Fisher's test when one sample number was less than five for a corresponding 2×2 table). Statistical analysis was performed with StatView 5.0 software (SAS Institute, Cary, NC).

Results

Identification of polymorphisms in WFS1

Twelve of the random controls were examined to detect genetic variations in the entire region of *WFS1*, and a total of 42 polymorphisms were identified in this study as shown in Fig. 1 and Table 1. Comparing our data with the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/index.html), 22 of the SNPs are novel. The distribution of polymorphisms was approximately 1/1000 bp in the 49.2 kb of DNA examined.

Evaluation of the pattern of linkage disequilibrium

As shown in Fig. 2, 16 SNPs were used to define haplotypes and to evaluate the pattern of LD. The other SNPs were excluded because of the rarity of minor alleles. As shown in Fig. 2, there are two LD blocks in this region, one ranging from position g. −15503 to g. 14909 and the other from position g. 16226 to g. 25103. The two SNPs at position g. 16226 and g. 16568, and the four SNPs at position g. 19460, g. 20758, g. 23707, and g. 25103 are in complete linkage disequilibrium.

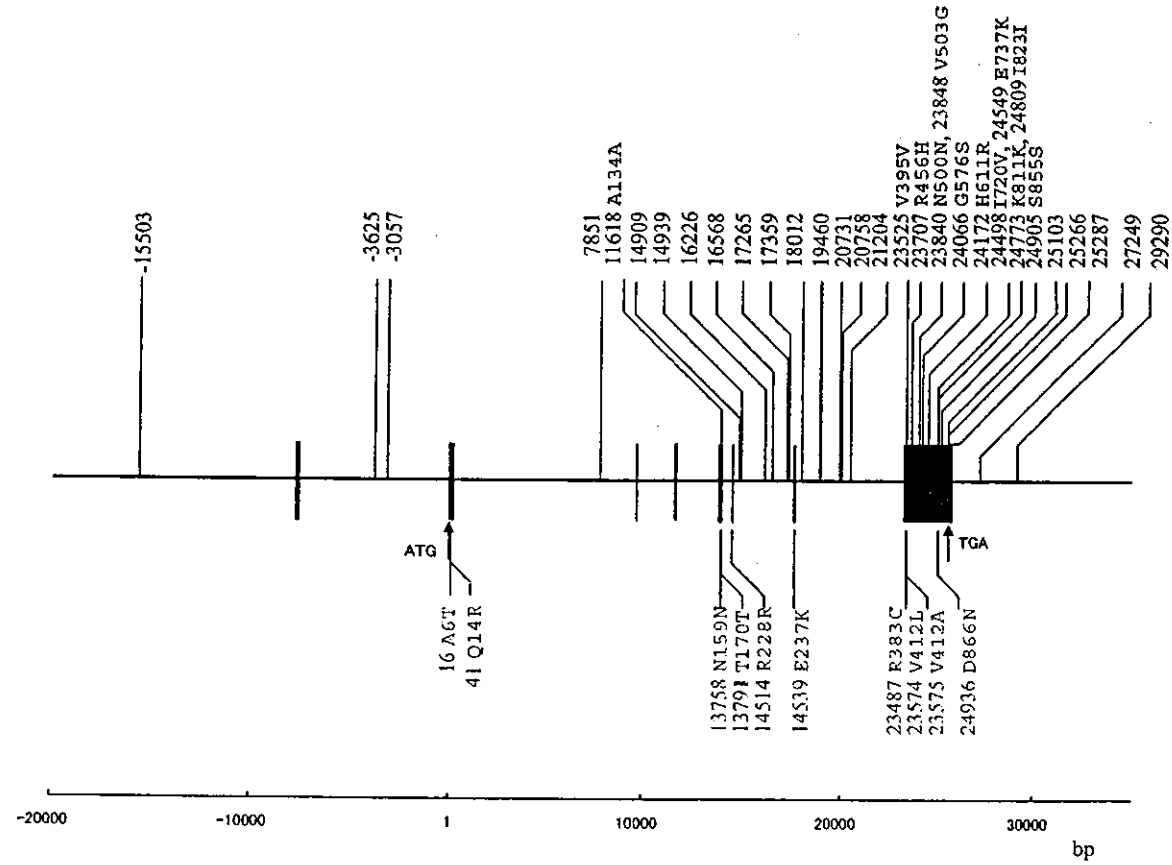

Fig. 1. Polymorphisms of *WFS1* identified in this study. The locations of the polymorphisms described in the text are shown. The nucleotide indicates the location of the SNP relative to the A of the ATG of the initiator Met of *WFS1* (GenBank No. NT_006051). The cSNPs shown in red are observed only in Type 2 diabetic patients. The cSNP shown in blue is observed only in patients with bipolar disorder.

Table 1
Polymorphisms identified in *WFS1* region in this study

Position genome	AA change	Variation	Location	Frequency of minor allele
−15503		C>T*	5′ flanking	0.42
−3625		C>T*	Intron 1	0.21
−3057		G>A	Intron 1	0.46
16	A6T	G>A*	Exon 2	—
41	Q14R	A>G	Exon 2	0.0026
7851		A>G	Intron 2	0.29
11618	A134A	G>A*	Exon 4	0.010
13758	N159N	C>T*	Exon 5	—
13791	T170T	C>G*	Exon 5	0.0079
14514	R228R	G>C	Exon 6	0.010
14539	E237K	G>A*	Exon 7	—
14909		G>A*	Intron 6	0.29
14939		T>C*	Intron 6	0.083
16226		G>A*	Intron 6	0.13
16568		G>A	Intron 6	0.13
17265		G>T*	Intron 6	0.13
17359		C>T*	Intron 6	0.042
18012		G>A*	Intron 7	0.13
19460		G>A*	Intron 7	0.13
20731		C>T	Intron 7	0.29
20758		T>C*	Intron 7	0.13
21204		delCTCA*	Intron 7	0.083
23487	R383C	C>T*	Exon 8	—
23525	V395V	T>C	Exon 8	0.010
23574	V412L	G>C*	Exon 8	0.0026
23575	V412A	T>C	Exon 8	0.0026
23707	R456H	G>A	Exon 8	0.078
23840	N500N	T>C	Exon 8	0.010
23848	V503G	T>G*	Exon 8	—
24066	G576S	G>A	Exon 8	0.12
24172	H611R	A>G	Exon 8	0.094
24498	I720V	A>G	Exon 8	0.063
24549	E737K	G>A	Exon 8	0.047
24773	K811K	A>G	Exon 8	0.010
24809	I823I	C>T	Exon 8	0.005
24905	S855S	G>A	Exon 8	0.010
24936	D866N	G>A	Exon 8	0.0052
25103		G>A*	3′ UTR	0.13
25266		G>A	3′ UTR	0.042
25287		GA	3′ UTR	0.042
27249		delCT*	3′ flanking	0.042
29290		C>T*	3′ flanking	0.13

The nucleotide indicates the location of the SNP relative to the A of the ATG of the initiator Met of *WFS1* (GenBank No. NT_006051). The frequencies of minor alleles of non-coding SNPs shown in this table are observed in random control samples. The frequencies of minor alleles of coding SNPs are observed in 192 non diabetic controls. Asterisk indicates a novel polymorphism.

Association study of genetic variations of WFS1 in patients with type-2 diabetes

All exons were examined in 192 type 2 diabetic patients. We found a total of 21 cSNPs, ten silent mutations and eleven missense mutations, of which seven are novel cSNPs (A6T, A134A, N159N, T170T, E237K, R383C, and V412L). As shown in Table 2, minor alleles H456 and R611were present more frequently in type 2 diabetic patients than in control subjects ($p = 0.091$ and $p = 0.050$, respectively). Because these two cSNPs are in strong linkage disequilibrium, as shown in Fig. 2, the haplotype defined by these SNPs was investigated for association with type 2 diabetes mellitus. The R456–H611 haplotype was less frequent in type 2 diabetic patients than in control subjects (Table 3, $p = 0.013$, $1 - \beta \approx 0.4$), but when we compared the two groups with and without this haplotype, there were no significant differences in age, BMI, fasting and postprandial glucose, or HbA_{1C} (data not shown).

Association study of genetic variations of WFS1 in patients with bipolar disorder

Mutation screening of *WFS1* in 47 patients with bipolar disorders revealed twelve coding SNPs. The allelic frequencies in patients and controls are shown in Table 4. One SNP (c. 402G>A, A134A) was located in exon 4