

Identification of Marker Genes for Differential Diagnosis of Chronic Fatigue Syndrome

Takuya Saiki,¹ Tomoko Kawai,² Kyoko Morita,² Masayuki Ohta,³ Toshiro Saito,³ Kazuhito Rokutan,² and Nobutaro Ban¹

¹Department of General Medicine, Nagoya University Hospital, Nagoya, Japan; ²Department of Stress Science, Institute of Health Biosciences, University of Tokushima Graduate School, Tokushima, Japan; and ³Life Science Group, Hitachi, Ltd, Saitama, Japan

Chronic fatigue syndrome (CFS) is a clinically defined condition characterized by long-lasting disabling fatigue. Because of the unknown mechanism underlying this syndrome, there still is no specific biomarker for objective assessment of the pathological fatigue. We have compared gene expression profiles in peripheral blood between 11 drug-free patients with CFS and age- and sex-matched healthy subjects using a custom microarray carrying complementary DNA probes for 1,467 stress-responsive genes. We identified 12 genes whose mRNA levels were changed significantly in CFS patients. Of these 12 genes, quantitative real-time PCR validated the changes in 9 genes encoding granzyme in activated T or natural killer cells (*GZMA*), energy regulators (*ATP5J2*, *COX5B*, and *DBI*), proteasome subunits (*PSMA3* and *PSMA4*), putative protein kinase c inhibitor (*HINT*), GTPase (*ARHC*), and signal transducers and activators of transcription 5A (*STAT5A*). Next, we performed the same microarray analysis on 3 additional CFS patients and 20 other patients with the chief complaint of long-lasting fatigue related to other disorders (non-CFS patients) and found that the relative mRNA expression of 9 genes classified 79% (11/14) of CFS and 85% (17/20) of the non-CFS patients. Finally, real-time PCR measurements of the levels of the 9 involved mRNAs were done in another group of 18 CFS and 12 non-CFS patients. The expression pattern correctly classified 94% (17/18) of CFS and 92% (11/12) of non-CFS patients. Our results suggest that the defined gene cluster (9 genes) may be useful for detecting pathological responses in CFS patients and for differential diagnosis of this syndrome.

Online address: <http://www.molmed.org>
doi: 10.2119/2007-00059.Saiki

INTRODUCTION

Chronic fatigue syndrome (CFS) is a clinically defined condition characterized by long-lasting disabling fatigue, resulting in severe impairment in daily functioning and associated symptoms such as memory and concentration difficulties, muscle aches, sleep disturbances, and headache (1). CFS poses a diagnostic challenge because of the unknown mechanism underlying this syndrome and the difficulty in making an objective assessment of pathological fatigue.

Several groups have been searching for reliable biomarkers for diagnosing CFS and have shown altered gene expression profiles in peripheral blood leukocyte populations, which can distinguish the majority of CFS cases (2–5). Recently, a data-intensive analysis has been conducted successfully by the Wichita CFS project (6). In the 2-day in-hospital study, gene expression levels of 20,000 genes in isolated peripheral blood mononuclear cells were analyzed to identify biologically and clinically meaningful signatures of gene expression rele-

vant to classification, diagnosis, and treatment of CFS (6).

Peripheral leukocytes express receptors for stress mediators, such as hormones, neurotransmitters, growth factors, and cytokines. Also, leukocytes produce a number of mediators, including cytokines, some of which can activate the hypothalamus-pituitary-adrenal (HPA) axis (7). Leukocytes may be potential targets for mediators eliciting pathological responses associated with stress-related disorders. CFS has been hypothesized to involve an abnormal response to various stressful experiences such as infection, overwork, or psychological stresses resulting in immunologic dysfunction, dysregulation of the HPA axis, and dysautonomia (8–11). At the same time, psychological and sociocultural factors, when present in patients with CFS, also influence the severity of the illness and treatment outcome (8–10). In

Address correspondence and reprint requests to Kazuhito Rokutan, Department of Stress Science, Institute of Health Biosciences, Graduate School of Medicine, The University of Tokushima Graduate School, 3-18-15 Kuramoto-cho, Tokushima 770-8503, Japan. Phone: + 81-88-633-9007; Fax: + 81-88-633-9008; E-mail: rokutan16@basic.med.tokushima-u.ac.jp. Submitted May 31, 2007; Accepted for publication June 12, 2008; Epub (www.molmed.org) ahead of print June 18, 2008.

fact, CFS is accompanied frequently by psychiatric disorders such as mood disorders, and the clinical manifestations of these two conditions partly overlap. Therefore, it is important that physicians are able to make the differential diagnosis between CFS and mood disorders, particularly major depression. However, at present, we have no reliable laboratory tool linking or separating these two disease states (10).

We developed a custom cDNA microarray specifically designed to measure mRNA levels of 1,467 stress-responsive genes in blood (12). Using this microarray, a whole-blood RNA collection system, and real-time PCR, we have identified a cluster of nine genes in blood as marker genes useful for differential diagnosis of CFS.

MATERIALS AND METHODS

Subjects

The present study was approved by the institutional review boards of the Nagoya University School of Medicine. After the experimental procedures were fully explained, written informed consent was obtained from all patients. All procedures were in accordance with the institutional guidelines and the Helsinki Declaration. Patients were recruited from a series of patients referred to the Department of General Medicine, Nagoya University Hospital, Nagoya, Japan. Initially, 11 patients with CFS (four males and seven females; aged 33.4 ± 9.4 years) were selected according to the Centers for Disease Control and Prevention criteria for CFS (1). Next, for a discriminating analysis of CFS versus non-CFS patients, 3 patients with CFS and 20 patients who presented with the chief complaint of general fatigue related to other disorders (non-CFS patients) were enrolled additionally in microarray analysis. Finally, 18 CFS and 12 non-CFS patients also were enrolled in quantitative real-time PCR assay for checking the validity of differential diagnosis. We obtained clinical information concerning current disability, duration of illness, number and nature of

accompanying symptoms (Table 1), the clinical data on blood chemistry, and complete blood cell counts (CBC) (supplementary Table 1) by standard laboratory tests. To confirm the diagnosis, all CFS and non-CFS patients underwent a psychiatric evaluation by a psychiatrist accustomed to confirming CFS patients' diagnoses. Age- and sex-matched healthy volunteers were recruited randomly to each experiment as controls. The controls were free of medication and underwent comprehensive medical examination for past and current health problems. Three months prior to enrollment in this study, all patients were removed from medications being taken.

Measures

RNA preparation, amplification, and hybridization. Venous blood (5 or 10 mL) was taken from patients and healthy volunteers under fasting conditions before lunch. Whole blood was poured directly into the PAXgene Blood RNA tube (Becton Dickinson, Franklin Lakes, NJ, USA). Total RNA was extracted from the whole blood mixture using a PAXgene Blood RNA kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Contaminating DNA was removed using an RNase-free DNase kit included in the PAXgene Blood RNA kit (Qiagen). The quality of the purified RNA and its applicability for microarray analysis was assessed by the Agilent 2100 Bioanalyzer using an RNA 6000 Nano Labchip kit (Agilent Technologies, Palo Alto, CA, USA). Quality of RNA was considered to be acceptable when the RIN value was > 8.0 . All RNA samples fulfilled this criterion. The labeling of RNA was done by an indirect aminoallyl labeling methodology. Five μg of total RNA was first reverse transcribed with oligo dT primer conjugating T7 sequence. The yield of first strand cDNA complementary to poly(A) RNA was amplified by using a MEGascript T7 *in vitro* RNA transcription kit (Applied Biosystems, Foster City, CA, USA). Amplified RNA (6 μg) was reverse transcribed by using random hexamer and aminoallyl-

dUTP. The synthesized cDNA was labeled by reaction with a dye (NHS-ester Cy5 or Cy3; Amersham Biosciences, Piscataway, NJ, USA). Cy5-cDNAs prepared from each patient were mixed with the equivalent amount of Cy3-cDNAs from the respective healthy subject, and the mixture was applied to the cDNA microarray. Hybridization was performed at 62°C for 12 h. After washing, fluorescence intensity at each spot was assayed using a scanner (ScanArray 5000; GSI-Lumonics, Billerica, MA, USA).

Microarray analysis. The construction of our microarray already has been described (12). To minimize non-specific hybridization reactions, mainly with hemoglobin RNAs, we selected 1,467 genes whose mRNAs were confirmed to be detectable in whole blood RNA samples by reverse transcriptase-PCR. The genes carried on our cDNA microarray are categorized into stress hormones, neurotransmitters, cytokines, growth factors, receptors, signal transduction molecules, transcription factors, heat shock proteins, growth- or apoptosis-associated factors, metabolic enzymes, and others (see supplementary Table 2).

Signal intensities of Cy5 and Cy3 were quantified and analyzed by subtracting backgrounds, using the QuantArray software (GSI-Lumonics). The global normalization was performed by scaling the Cy3 signal intensities to the median Cy5: Cy3 ratio. The normalized values for duplicate cDNA probes were averaged. Then, we selected 1,072 genes having fluorescence intensities higher than a cut-off value of 300 in either Cy5 or Cy3 conditions among all samples. The relative expression values (Cy5: Cy3) for 1,072 genes were subjected to hierarchical clustering using the GeneSpring 7.3 software (Agilent). Average linkage and cosine θ were used for clustering algorithm and the calculations of distance metric, respectively. After the Cy5: Cy3 ratios of 1,072 genes were transformed to logarithm base 2, statistical significance between CFS patients and age- and sex-matched controls was examined by the paired *t* test statistic using the

Table 1. Clinical features of patients

Patient Number	Age	Gender	Duration of illness (years)	Clinical diagnosis	Symptoms according to the revised CDC criteria for CFS										
					Unexplained fatigue	Impairment in memory	Sore throat	Tender nodes	Muscle pain	Multijoint pain	Headaches	Un-refreshing sleep	Post-exertional malaise		
1	41	F	12	CFS	+	+	-	-	-	+	+	+	+	+	
2	25	M	2	CFS + Adjustment disorder	+	+	-	-	-	-	-	-	+	+	
3	35	M	3	CFS	+	-	-	+	-	-	+	+	+	+	
4	30	F	5	CFS	+	+	+	+	+	+	+	+	+	+	
5	27	F	1	CFS	+	+	-	-	+	+	+	+	+	+	
6	25	F	7	CFS + Mood disorder	+	+	+	-	-	+	+	+	+	+	
7	26	F	3	CFS	+	+	-	-	+	+	+	+	+	+	
8	46	F	13	CFS	+	+	-	-	-	+	+	+	+	+	
9	29	M	2	CFS + Mood disorder	+	+	-	+	-	-	-	+	+	+	
10	30	M	1	CFS	+	+	-	-	+	-	-	-	+	+	
11	53	F	5	CFS	+	+	-	+	-	-	-	+	+	+	
Additionally enrolled patients for microarray analysis															
12	28	F	2	CFS + Adjustment disorder	+	+	-	+	-	-	-	+	+	-	
13	24	F	8	CFS	+	+	+	+	-	-	-	-	+	+	
14	25	M	5	CFS + Mood disorder	+	+	+	-	+	+	+	+	+	+	
15	30	F	3	Mood disorder	+	+	+	-	+	+	+	+	+	+	
16	24	F	3	Somatoform disorder	+	-	+	+	-	+	+	+	-	+	
17	45	F	12	Personality disorder	+	+	-	-	-	-	-	-	+	-	
18	38	F	3	Mood disorder	+	-	-	-	-	+	+	+	-	-	
19	39	F	1	Adjustment disorder	+	+	+	-	+	+	+	+	+	-	
20	25	F	0.5	Eosinophilia	+	-	-	+	+	+	+	+	+	-	
21	42	M	13	Diabetes mellitus type II + hyper-CK-emia	+	-	-	-	-	-	-	-	+	-	
22	34	M	0.5	Mood disorder	+	+	-	-	+	-	-	-	+	+	
23	15	M	1	Sleep disorder	+	+	-	+	-	-	-	-	+	+	
24	42	F	8	Mood disorder	+	+	-	+	-	+	+	+	-	-	
25	35	F	5	Anxiety disorder	+	+	-	-	-	-	+	+	-	-	
26	60	F	5	Mood disorder	+	-	+	-	-	+	+	+	-	-	
27	23	F	1	Mood disorder	+	+	+	-	-	+	+	+	-	-	
28	41	F	2	Mood disorder	+	-	+	+	+	-	-	+	-	-	
29	30	M	4	Adjustment disorder	+	+	-	-	+	+	+	+	+	+	
30	17	F	1	Adjustment disorder	+	-	+	-	+	+	+	-	-	-	
31	46	M	7	Unexplained fatigue	+	+	+	-	+	+	+	+	+	+	
32	31	F	4	Unexplained fatigue	+	+	+	+	+	+	+	+	+	+	
33	23	F	0.5	Unexplained fatigue	+	-	+	+	+	-	-	+	+	+	
34	33	M	3	Adjustment disorder	+	+	-	-	-	-	-	+	+	+	

Continued

Table 1. Continued

		Additionally enrolled patients for quantitative real-time PCR																			
35	45	F	0.5	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
36	13	F	0.75	CFS + Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
37	28	F	7	CFS + Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
38	31	F	2.5	CFS	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
39	24	F	4	CFS	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
40	25	M	2	CFS + Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
41	25	F	3	CFS + Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
42	35	M	10	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
43	27	F	10	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
44	29	M	3	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
45	41	F	15	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
46	45	M	7	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
47	46	F	0.8	CFS + Secondary depression	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
48	35	M	1.1	CFS + Major depression	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
49	53	M	6	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
50	26	M	4	CFS + Secondary depression	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
51	38	M	6	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
52	26	F	1	CFS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
53	37	F	1	Major depression	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
54	48	F	1.5	Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
55	31	F	5	Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
56	28	M	5	Major depression	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
57	26	M	2	Unexplained fatigue	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
58	72	F	5	Unexplained fatigue	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
59	41	F	4	Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
60	38	F	9	Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
61	33	F	7	Adjustment disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
62	32	M	2.5	Hypo-cerebrospinal fluid	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
63	28	F	4	Mood disorder	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
64	37	M	1	Major depression + Sleep apnea	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

CK, creatine kinase.

Cyber-T stats program written in the R stats language (see <http://cybert.microarray.ics.uci.edu/help/index.html>) (13). Statistical significance was defined as a Bonferroni-corrected *P* value of < 0.05, after the problem of a multiple test was addressed.

Quantitative real-time PCR. cDNA was prepared from total RNA (0.5 µg) using oligo dT primer according to the instructions of SuperScript II reverse transcriptase kit (Invitrogen, Carlsbad, CA, USA). The mRNA levels of ten target genes based on GenBank accession numbers (Table 2) were analyzed by quantitative real-time PCR using pre-designed, gene-specific TaqMan probes and primer sets (search the batch ID for each gene in Table 2 at <https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=ABGEBatchSearch>) and the ABI-PRISM 7500 sequence detection system (Applied Biosystems, Foster City, CA, USA). Appropriate pre-designed TaqMan probe and primer sets for detecting specific mRNA types of *COX7C* and *HSPA2* were not available because of their gene structure (Applied Biosystems). Each PCR reaction was performed according to the protocol of TaqMan Universal PCR Mastermix (Applied Biosystems), and data were analyzed using SDS 2.2 software (Applied Biosystems). A no template control and a no RT step control also were run for every reaction to see that the amplification was not off genomic DNA. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as an endogenous quantity control, and quantity values were normalized to *GAPDH* mRNA expression. After the relative ratio of each mRNA between CFS patients and control subjects was calculated, the unpaired *t* test was performed to compare the relative ratio for each gene in the microarray (*n* = 11) versus quantitative real time PCR results (*n* = 11). Finally, we also examined relative mRNA ratios of the target genes in newly enrolled 18 CFS and 12 non-CFS patients using age- and sex-matched healthy subjects as controls by quantitative real-time PCR.

Table 2. List of TaqMan probes and primer sets

Gene symbol	GenBank accession number	Batch ID	Slope in standard curve (Ct/Quantity)
<i>GZMA</i>	M18737	Hs00196206_m1	-3.58
<i>ATP5J2</i>	AF047436	Hs00934710_m1	-3.30
<i>DBI</i>	BC006466	Hs01554584_m1	-3.69
<i>HINT</i>	U51004	Hs00602163_m1	-3.19
<i>KIAA0194</i>	D83778	Hs00412357_m1	-3.77
<i>COX5B</i>	NM_001862	Hs00426948_m1	-3.24
<i>ARHC</i>	L25081	Hs00237129_m1	-3.70
<i>PSMA4</i>	BC005361	Hs00160566_m1	-3.88
<i>PSMA3</i>	BC005265	Hs00541061_m1	-3.09
<i>STAT5A</i>	L41142	Hs00559643_m1	-3.29
<i>GAPDH</i>	M33197	Hs99999905_m1	-3.29

All supplementary materials are available online at molmed.org.

RESULTS

Clinical Features of CFS Patients

The initially enrolled 11 patients with CFS were comprised of four males and seven females whose median age was 33.4 ± 9.4 years. The clinical features of the 11 patients are shown in Table 1. All patients were complaining of debilitating fatigue lasting for longer than one year, and the median duration of fatigue was 4.9 years. Five patients (patients 3, 6, 7, 9, and 10) reported some linkage between infectious episodes and onset of their symptoms, but none of them were positive for serum antibody against *Coxiella burnetti* (Q fever) that could trigger CFS-like symptoms (14), human herpes viruses type 6 and 7, Epstein-Barr virus, or cytomegalovirus (data not shown). Biochemical data also demonstrated the absence of current active infection (data not shown). CBC of one patient (patient 6) was not measured at the time of sampling for microarray analysis, since the medical record from another hospital described no abnormalities in CBC and leukocyte populations. CBC data and white blood cell differential counts of the other patients were normal (supplementary Table 1). All 11 patients showed normal BMI values (20.3 ± 2.2, mean ± SD) (see supplementary Table 1). Three CFS patients (patients 2, 6, and 9) had past histories of major depression or adjust-

ment disorder. We consulted the psychiatrist and confirmed that these disorders were not active during the experimental period.

Gene Expression Profile in Patients with CFS

Hierarchical cluster analysis of 1,072 genes whose fluorescence intensities were higher than a cut-off value of 300 in either Cy5 or Cy3 conditions among all samples suggested the presence of genes that commonly changed among all patients, compared with age- and sex-matched controls (data not shown). The statistic analysis of 1,072 genes (paired *t* test; Bonferroni [experiment-wide false positive rate] adjusted *P* value = 0.05) identified 12 genes whose mRNA levels were changed significantly in CFS patients compared with the healthy controls (Figure 1). Several upregulated genes were categorized into regulators of energy metabolism; a mitochondrial ATP subunit (*ATP5J2*; *f* subunit of the F0 complex), nuclear-encoded subunits of the mitochondrial cytochrome *c* (*COX7C* and *COX5B*), and an intracellular acyl-coenzyme A transporter (*DBI*). The significantly upregulated genes also contained a cytotoxic T lymphocyte- and natural killer cell-specific serine protease, granzyme A (*GZMA*), a member of the ras homolog gene family (*ARHC*), and proteasome subunits (*PSMA3* and *PSMA4*). The mRNA levels of two other genes encoding a protein with unknown function (*KIAA0194*), and a putative pro-

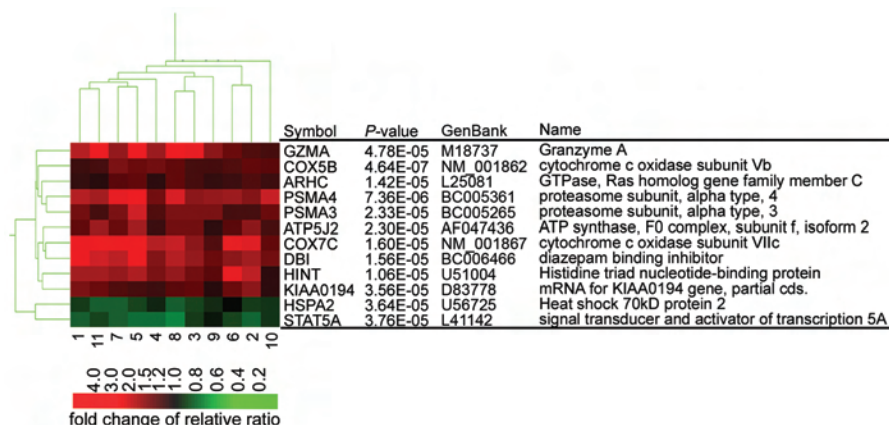


Figure 1. Identification of 12 CFS-associated genes in patients with CFS. The expressions of 1,072 genes were compared between 11 patients with CFS and age- and sex-matched healthy controls, and 12 common genes were identified that were significantly changed in CFS patients according to the paired *t* test using the cyber-T stats program (Bonferroni wide false positive rate = 0.05). Heatmap values represent the relative ratios of gene expressions between CFS and age- and sex-matched healthy controls. Hierarchical clustering was performed by the results of average linkage and distance metric of cosine θ . Gene symbols, *P* values, GenBank accession numbers, and names of the 12 genes are listed. *P* values are based on paired *t* tests. Patient's numbers indicated at the bottom of column are corresponding to the numbers in Table 1.

tein kinase C inhibitor of the HINT family (*HINT*) also were elevated significantly. In contrast, CFS patients showed decreased mRNA levels of heat shock 70-KDa protein 2 (*HSPA2*) and a member of the signal transducer and activator of transcription (STAT) family of transcription factors (*STAT5A*).

Quantitative Real-Time PCR

Although we applied the indirect labeling method to reduce the extent of cyanine dye bias in our microarray assay, there may be some concerns regarding the possibility of the genes exhibiting some dye-bias in dual-labeled spotted cDNA microarrays. Therefore, we performed TaqMan real-time PCR to validate the microarray data. The mRNA levels of two genes (*COX7C* and *HSPA2*) were not measured, since appropriate TaqMan probes for these genes were not available. As for the rest of genes, we confirmed that each PCR reaction had a similar efficiency of reaction when we checked the slope in a standard curve for each PCR reaction using the same total RNA from whole blood as standard

(Table 2). Among the 10 mRNA levels measured, real-time PCR did not demonstrate any significant change in the *KIAA0194* mRNA level, while the other nine mRNA levels in CFS patients were confirmed to be changed significantly (Figure 2).

Microarray Analysis of CFS and Non-CFS Patients with Prolonged Fatigue

To test whether gene expression profiling could be useful for differential diagnosis of CFS, 3 CFS patients and 20 patients who presented with the chief complaint of general fatigue related to other disorders (non-CFS patients) were enrolled in this study additionally. Relative gene expression levels of the CFS and non-CFS patients also were measured by the dual-labeled cDNA microarray using age- and sex-matched healthy subjects as controls. RNA samples from the newly added patients (3 CFS and 20 non-CFS) and age- and sex-matched healthy controls were labeled Cy5 and Cy3, respectively. All 20 non-CFS patients complained of abnormal fatigue lasting for more than 6 months, while

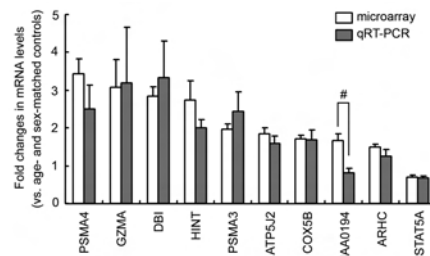


Figure 2. Relative expression of 10 marker genes in patients with CFS by microarray and real-time PCR. RNA prepared from 11 patients with CFS and age- and sex-matched healthy controls was subjected to TaqMan real-time PCR as described in the method section. Of the 12 genes, 10 mRNA levels were measured and then normalized to *GAPDH* mRNA expression. After the relative ratios of 10 mRNAs between CFS patients and control subjects were calculated, they were compared between microarray (empty bars) and real time PCR results (solid bars). Values are mean fold changes \pm SD (*n* = 11). #*P* < 0.05 by the paired *t* test.

their clinical features did not completely meet the CDC criteria for CFS (Table 1). First, we compared gene expression profiles of 1,072 genes in 14 CFS patients, including 3 additionally enrolled patients, and 20 non-CFS patients. Hierarchical cluster analysis of the relative mRNA levels of 1,072 genes showed that gene expression patterns could be classified roughly into CFS and non-CFS patterns, but it was difficult to draw a margin between the two patterns (data not shown).

Next, we tested whether the changes in nine genes, whose expressions were confirmed to be changed significantly between 11 CFS patients and healthy subjects by both microarray (see Figure 1) and quantitative real-time PCR (see Figure 2), could exclude non-CFS patients. As shown in Figure 3, the hierarchical clustering of the expression of nine genes classified 34 patients into two groups (A and B) or 3 groups (A, B1, and B2). Group A branches contained 13 CFS patients and 3 non-CFS patients (mood disorder, 30-year-old female; somatoform disorder, 24-year-old female; personality

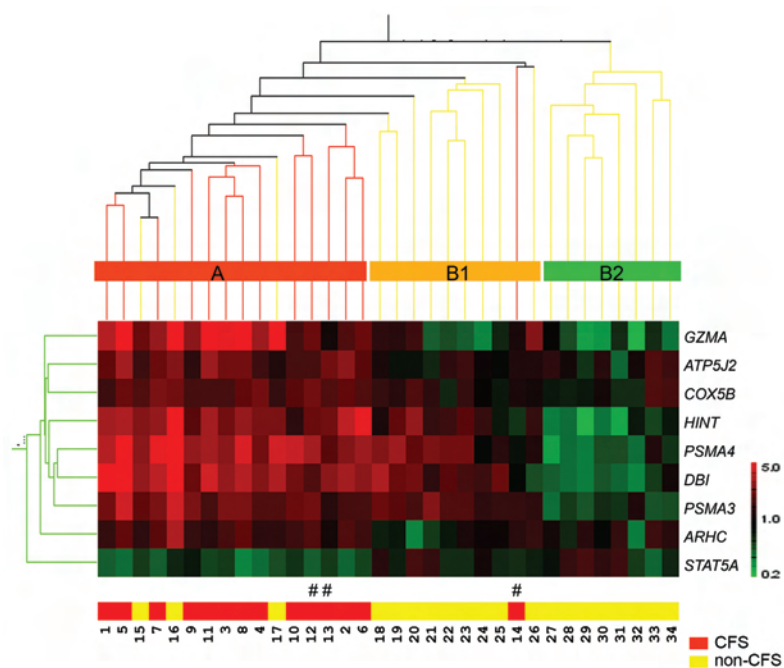


Figure 3. Expression of nine marker genes in CFS and non-CFS cases. RNA was prepared from whole blood of 3 additionally enrolled patients with CFS and 20 patients with the chief complaint of general fatigue related to other disorders, and subjected to the microarray analysis. The expression patterns of nine genes in these patients were compared with those of the initially enrolled 11 patients with CFS by hierarchical cluster analysis. CFS and non-CFS cases are indicated as red and yellow bars, respectively. Patient's numbers indicated at the bottom of columns are corresponding to the numbers in Table 1. #Three newly enrolled patients with CFS.

disorder, 45-year-old female). Among 10 branches of group B1, only 1 CFS patient was included. All branches of group B2 were composed of non-CFS patients. Thus, the cluster analysis of relative mRNA levels of nine genes measured by the microarray suggested that the nine marker genes might be useful for differential diagnosis of CFS.

Use of Nine Marker Genes for Differential Diagnosis of CFS

Finally, we tested whether the nine marker genes could be useful for differential diagnosis of CFS. To correctly assess this issue, we omitted the 11 patients in whom we had identified the nine genes. A total of 18 newly enrolled CFS patients and 12 non-CFS patients (Table 1) were subjected to quantitative real time RT-PCR analysis using *GAPDH* mRNA as an endogenous quantity control. Age- and sex-matched healthy subjects (total 30

subjects) also were used as controls of individual patients. As shown in Figure 4A, the expression levels of six genes (*PSMA4*, *PSMA3*, *HINT1*, *DBI*, *GZMA*, and *ATHC*) out of nine genes in 18 CFS patients were significantly different from that in 12 non-CFS patients. The hierarchical clustering of the expression of nine genes classified 30 patients into two groups (*a* and *b*) (Figure 4B). Group *a* branches contained 17 CFS patients and 1 non-CFS patient (major depression, 37 F). Among 12 branches of group *b*, 1 case of CFS was included. Thus, the expression pattern of nine genes could distinguish the majority of our CFS patients from non-CFS patients.

DISCUSSION

The microarray or differential display approach has been used to examine the CFS-specific gene expression in peripheral blood mononuclear cells (2–5). Two

sources are commonly used for preparation of RNA, whole blood, or its leukocyte populations (15). Because of advantages and disadvantages associated with both systems (16–20), at present there is no consensus regarding the optimal technique for isolation of RNA from peripheral blood. A whole-blood RNA collection system is appealing, particularly in clinical settings, since the RNA isolation method is easy to use and reduces operator time and sample volume. In addition, this system reduces the risk of exposure of laboratory personnel to biohazards relative to the risk involved in isolation of leukocyte populations.

Using RNA from whole blood, we show here that both microarray analysis and real-time PCR identify nine genes whose mRNA expression are significantly different in 11 patients with CFS, compared with age- and sex-matched healthy controls. Although the individual genes identified as CFS-related genes did not overlap with those identified in other studies (2–5), most them could be categorized into distinct clusters, including host defense, energy metabolism, or small G protein-dependent signal transduction (2–5). The significance of our study can be considered from three different perspectives.

First, the identified genes are informative in considering the pathophysiology of CFS. The upregulated *GZMA* encodes a T cell- and natural killer cell-specific serine protease that functions as a common component necessary for lysis of target cells by these cytotoxic cells. The proteasome subunits *PSMA3* and *PSMA4* also were upregulated. The proteasome is the central proteolytic system that also plays an important role in the major histocompatibility complex-class I antigen processing. Previous studies identified genes involved in T cell activation (2–5). Our findings also suggest that patients with CFS may have altered immunity, such as that involved in anti-viral defense. As reported in other studies (3,5), we also have identified genes encoding molecules catalyzing oxidative phosphorylation in mitochondria (*COX5B* and

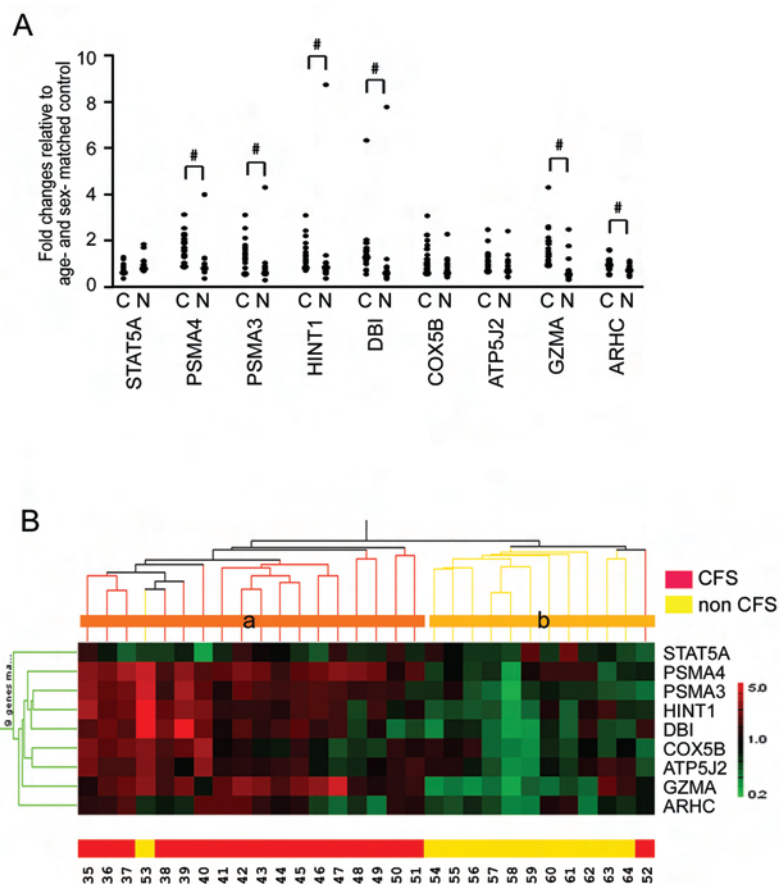


Figure 4. Differentiation of CFS and non-CFS cases by expression levels of nine marker genes. (A) The expressions of nine marker genes were validated by real-time PCR, and their mean values were compared in 18 newly enrolled CFS patients (indicated as “C”) with 12 non-CFS patients (indicated as “N”). Values are mean fold changes \pm SD. # $P < 0.05$ by the unpaired t test. (B) The expression levels of nine genes from 18 CFS and 12 non-CFS patients were differentiated using hierarchical cluster analysis by the results of average linkage and distance metric of cosine θ . CFS and non-CFS patients are indicated as red and yellow bars, respectively. Patient’s numbers indicated at the bottom of columns are corresponding to the numbers in Table 1.

ATP5J2). *COX5B* and *ATP5J2* encode a cytochrome *c* oxidase subunit and a subunit of the mitochondrial proton channel, respectively. Although we were unable to measure the mRNA level of another cytochrome *c* oxidase subunit (*COX7C*) by real time PCR, these nuclear-encoded subunits (*COX5B* and *COX7C*) function in the regulation and assembly of the cytochrome *c* oxidase complex and mitochondrial ATPase. In addition, our CFS patients had significantly increased *DBI* mRNA levels. The diazepam binding inhibitor (*DBI*) is known as a GABA receptor modulator or acyl-coenzyme A (acyl-

CoA) binding protein (ACBP). ACBP binds thiol esters of long fatty acids and coenzyme A in a one-to-one binding mode with high specificity and affinity. This molecule is suggested to act as an intracellular acyl-CoA transporter and to form a pool of ACBP-acyl-CoA complex that is an important intermediate in lipid synthesis and fatty acid degradation that participates in regulating intermediary metabolism and gene expression. The increased mRNA expression of *DBI*, *COX5B*, and *ATP5J2* strongly suggests abnormalities in energy metabolism in our CFS patients.

We also found that the *STAT5A* mRNA level was decreased significantly in CFS patients. The protein encoded by *STAT5A* is a member of the STAT family of transcription factors. STAT-5 mediates the signal transduction triggered by various cell ligands, such as IL2, IL4, colony-stimulating factor 1, and growth hormones. Adult growth hormone deficiency (AGHD) is a CFS-like disorder characterized by fatigue, tiredness, and myalgia; replacement therapy with human growth hormone improves these symptoms (21). Growth hormone activates STAT1, 3, 5A, and 5B in different cell systems (22). Webb *et al.* reported that STAT-5 isoform, but not STAT-1 or STAT3, were increased markedly in skeletal muscles in patients with AGHD and suggested that the STAT5 signal transduction pathway in skeletal muscle might be abnormal in AGHD (21). The decreased expression of *STAT5A* mRNA in peripheral blood cells from CFS patients suggests that the abnormality in STAT5 signaling might be associated with symptoms of CFS.

In the Wichita study directed by CDC (6), fatigue-associated gene expression patterns in isolated blood mononuclear cells were identified by several groups sharing the same data sets. Most of the groups in that study did not divide subjects into CFS and non-CFS cases by CDC classification but focused instead on fatigue itself and accompanying symptoms for elucidation of fatigue-associated genes. It was confirmed that 9 of 16 genes reported by Kaushik *et al.* as differentially expressed genes in CFS (5) also were included among fatigue-associated genes measured by quantitative trait analysis (QTA) in the Wichita study (23). Our study also revealed that two genes, *STAT5A* and *COX5B*, were categorized in the same pathways as *STAT5B* and *COX7A2*, which were identified as fatigue-associated genes according to QTA. *STAT5A* and *COX5B* belong to the Jak-STAT signaling pathway and oxidative phosphorylation pathway, respectively. Furthermore, Fang *et al.* in the

Wichita study succeeded in separating CFS from non-CFS patients based on expression profiles of 24 genes that were differentially expressed in subjects between those who received high scores and low scores in both multidimensional fatigue inventory scores and the Zung depression scale (24). The reported 24 genes included *ACBD6* that encodes one of the acyl-CoA-binding proteins. Another homologue of the acyl-CoA-binding proteins, *DBI* (*ACBD1*), was up-regulated in our CFS cases. The Wichita study was a population-based study, while our data are based on a clinical cohort. Despite the fact that we used a different RNA preparation method and a different microarray platform, there was a significant overlap between our results and those of the Wichita study.

Second, most of the non-CFS patients in our study were psychiatric disorders, which usually presents a challenge for the clinicians to differentiate. Thus, the present study may provide a potential tool to clinicians who see chronically fatigued patients in daily practice with no objective marker for CFS.

Third, the expression pattern of the nine marker genes did not distinguish all of our clinically diagnosed CFS cases, since they include heterogeneous populations. However, identification of a group of CFS patients having this unique expression pattern will be useful for future treatment studies.

With regard to the limitations of our study, although we could not detect any common abnormality in the CBC and leukocyte subpopulations of our CFS cases, the use of whole blood RNA could not rule out heterogeneity of the cell population and the potential diversity of the cell-specific responses. The second concern is that microarrays carrying different gene probes may pick up different groups of genes. Although our microarray carries cDNA probes for 1,467 genes that have been confirmed to be actually detectable by reverse transcription PCR, more genome-wide examinations in larger numbers of cases may reveal additional marker genes for CFS. However,

we also found that the expression pattern of nine genes measured by the microarray could classify 79% (11/14) of CFS and 85% (17/20) of non-CFS patients. Finally, real-time PCR measurement of the nine mRNA levels in another group of subjects (18 CFS and 12 non-CFS patients) classified 94% (17/18) of CFS and 92% (11/12) of non-CFS patients. A clinical trial of a larger number of CFS and non-CFS patients with long-lasting fatigue is now under way.

ACKNOWLEDGMENTS

The authors thank Tetsushi Yoshikawa (Fujita Healthscience University) for his measurements of the titers of HHV6 and 7, Motohiko Ogawa (National Infectious Disease Research Center), and Yumiko Furuya (Kanagawa Health Research Center) for their measurements of the titer of *C. burnetii* for Q-fever, and Mitsuo Kondo (Kondo Mental Clinic) for psychiatric evaluation of the patients. The authors also thank Gyoshin Ko and Teruko Mizutani for their technical assistance. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (16390199 to NB) and a grant from Japan Science and Technology Agency (to KR).

REFERENCES

1. Fukuda K, *et al.* (1994) The chronic fatigue syndrome: a comprehensive approach to its definition and study. International chronic fatigue syndrome study group. *Ann. Intern. Med.* 121:953–9.
2. Vernon SD, Unger ER, Dimulescu IM, Rajeevan M, Reeves WC. (2002) Utility of the blood for gene expression profiling and biomarker discovery in chronic fatigue syndrome. *Dis. Markers* 18:193–9.
3. Powell R, *et al.* (2003) Identification of novel expressed sequences, up-regulated in the leukocytes of chronic fatigue syndrome patients. *Clin. Exp. Allergy* 33:1450–6.
4. Whistler T, Unger ER, Nisenbaum R, Vernon S. (2003) Integration of gene expression, clinical, and epidemiologic data to characterize chronic fatigue syndrome. *J. Transl. Med.* 1:10.
5. Kaushik N, *et al.* (2005) Gene expression in peripheral blood mononuclear cells from patients with chronic fatigue syndrome. *J. Clin. Pathol.* 58:826–32.
6. Vernon SD, Reeves WC. (2006) The challenge of integrating disparate high-content data: epidemi-

- ological, clinical and laboratory data collected during an in-hospital study of chronic fatigue syndrome. *Pharmacogenomics* 7:345–54.
7. Chesnokova V, Melmed S. (2002) Neuro-immuno-endocrine modulation of the hypothalamic-pituitary-adrenal (HPA) axis by gp130 signaling molecules. *Endocrinology* 143:1571–74.
8. Salit IE. (1997) Precipitating factors for the chronic fatigue syndrome. *J. Psychiat. Res.* 31:59–65.
9. Cleare AJ. (2003) The neuroendocrinology of chronic fatigue syndrome. *Endocr. Rev.* 24:236–52.
10. Prins JB, van der Meer JWM, Bleijenberg G. (2006) Chronic fatigue syndrome. *Lancet* 367:346–55.
11. Chaudhuri A, Behan PO. (2004) Fatigue in neurological disorders. *Lancet* 363:978–88.
12. Morita K, *et al.* (2005) Expression analysis of psychological stress-associated genes in peripheral blood leukocytes. *Neurosci. Lett.* 381:57–62.
13. Baldi P, Long AD. (2001) A Bayesian framework for the analysis of microarray expression data: regularized *t* test and statistical inferences of gene changes. *Bioinformatics* 17:509–519
14. Wildman MJ, *et al.* (2002) Chronic fatigue following infection by *Coxiella burnetii* (Q fever): ten-year follow-up of the 1989 UK outbreak cohort. *QJM.* 95:527–38.
15. Fan H, Hegde PS. (2005) The transcriptome in blood: challenges and solutions for robust expression profiling. *Curr. Mol. Med.* 5:3–10.
16. Feezor RJ, *et al.* (2004) Whole blood and leukocyte RNA isolation for gene expression analyses. *Physiol. Genomics* 19:247–54.
17. Gnatenko DV, *et al.* (2003) Transcript profiling of human platelets using microarray and serial analysis of gene expression. *Blood* 101:2285–93.
18. Rainen L, *et al.* (2002) Stabilization of mRNA expression in whole blood samples. *Clin. Chem.* 48:1883–90.
19. Debey S, *et al.* (2004) Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. *Pharmacogenomics J.* 4:193–207.
20. Thach DC, *et al.* (2003) Assessment of two methods for handling blood in collection tubes with RNA stabilizing agent for surveillance of gene expression profiles with high density microarrays. *J. Immunol. Methods* 283:269–79.
21. Webb SM, *et al.* (2003) Neuromuscular dysfunction in adult growth hormone deficiency. *Clin. Endocrinol. (Oxf).* 59:450–8.
22. Herrington J, Smit LS, Schwartz J, Carter-Su C. (2000) The role of STAT proteins in growth hormone signaling. *Oncogene* 19:2585–97.
23. Whistler T, *et al.* (2006) Gene expression correlates of unexplained fatigue. *Pharmacogenomics* 7:395–405.
24. Fang H, *et al.* (2006) Gene expression profile exploration of a large dataset on chronic fatigue syndrome. *Pharmacogenomics* 7:429–40.