Gene Expression Subtypes in Patients with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis

Jonathan R. Kerr,[1] Robert Petty,[1,a] Beverley Burke,[1,a] John Gough,[1] David Fear,[2] Lindsey I. Sinclair,[6] Derek L. Mattey,[7] Selwyn C. M. Richards,[8] Jane Montgomery,[8] Don A. Baldwin,[10] Paul Kellam,[3] Tim J. Harrison,[4] George E. Griffin,[1] Janice Main,[5] Derek Enlander,[11] David J. Nutt,[6] and Stephen T. Holgate[9]

[1]Department of Cellular & Molecular Medicine, St. George's University of London, [2]Department of Asthma, Allergy and Respiratory Sciences, King's College London, Departments of [3]Infection and [4]Medicine, Windeyer Institute of Medical Sciences, University College London, and [5]Department of Infectious Diseases and General Medicine, Imperial College London, St. Mary's Hospital, London, [6]Psychopharmacology Unit, Department of Community Based Medicine, University of Bristol, Bristol, [7]Staffordshire Rheumatology Centre, Stoke on Trent, [8]Dorset CFS Service, Poole Hospital, Dorset, and [9]MRC Department of Immunopharmacology, University of Southampton, Southampton General Hospital, Southampton, United Kingdom; [10]Penn Microarray Facility, University of Pennsylvania, Philadelphia; and [11]New York ME/CFS Service, New York, New York

Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is a multisystem disease, the pathogenesis of which remains undetermined. We set out to determine the precise abnormalities of gene expression in the blood of patients with CFS/ME. We analyzed gene expression in peripheral blood from 25 patients with CFS/ME diagnosed according to the Centers for Disease Control and Prevention diagnostic criteria and 50 healthy blood donors, using a microarray with a cutoff fold difference of expression of $\geqslant$2.5. Genes showing differential expression were further analyzed in 55 patients with CFS/ME and 75 healthy blood donors, using quantitative polymerase chain reaction. Differential expression was confirmed for 88 genes; 85 were upregulated, and 3 were downregulated. Highly represented functions were hematological disease and function, immunological disease and function, cancer, cell death, immune response, and infection. Clustering of quantitative polymerase chain reaction data from patients with CFS/ME revealed 7 subtypes with distinct differences in Medical Outcomes Survey Short Form-36 scores, clinical phenotypes, and severity.

Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is characterized by severe and debilitating fatigue, abnormal sleep behavior, impaired memory and concentration, and musculoskeletal pain [1]. In developed nations, the population prevalence is estimated to be 0.5% [2, 3]. Studies have identified various features relevant to the pathogenesis of CFS/ME, such as viral infection; abnormal immune function; exposure to toxins, chemicals, and pesticides; stress; hypotension; abnormal lymphocyte levels; and neuroendocrine dysfunction. However, the precise underlying mechanisms of disease and the means by which they interrelate in patients with CFS/ME remain to be clarified [4, 5].

Various studies have analyzed gene expression in the peripheral blood of patients with CFS/ME, and in each study genes associated with immunity and defense were prominent [6–17]. Unfortunately, in several of these studies quantitative polymerase chain reaction (PCR) confirmation was not performed, so the results may be unreliable [6, 8–10, 14–17]. The genes identified in the studies that used PCR confirmation suggest that CFS/ME has a complex pathogenesis [7, 11–13]. However, because none of these studies were comprehensive in terms of the number of human genes examined, our knowledge of the precise metabolic pathways involved in CFS/ME remains incomplete.

The goal of this study was to determine the precise abnormalities of gene expression in the blood of patients with CFS/ME. We enrolled patients with a rigorously defined CFS/ME phenotype and compared them to healthy blood donors by means of a microarray that represented the entire human genome and by

Received 14 July 2007; accepted 24 September 2007; electronically published 19 March 2008.

Potential conflicts of interest: none reported.

Presented in part: 2007 Conference of the International Association for Chronic Fatigue Syndrome/Myalgic Encephalomyelitis, 10–14 January 2007, Fort Lauderdale, Florida.

Financial support: Chronic Fatigue Syndrome Research Foundation, Hertfordshire, United Kingdom.

[a] R.P. and B.B. made equal contributions to this article.

Reprints or correspondence: Dr. Jonathan R. Kerr, St. George's University of London, Cranmer Terrace, Jenner Wing, Rm. 2.267, London SW17 0RE, United Kingdom (jkerr@sgul.ac.uk).

The Journal of Infectious Diseases 2008; 197:1171–84

0022-1899/2008/19708-0014$15.00
DOI: 10.1086/533453

quantitative PCR confirmation. Using this approach, we identified differential expression of 88 human genes in patients with CFS/ME. Among these genes, highly represented functions were hematological disease and function, immunological disease and function, cancer, cell death, immune response, and virus infection. Clustering of gene expression data revealed 7 CFS/ME subtypes with distinct clinical phenotypes and associated disease severity.

SUBJECTS, MATERIALS, AND METHODS

Subject enrollment, clinical characterization, and blood sampling. Twenty-five patients with CFS/ME from the Dorset CFS Service in southeast England were enrolled for the microarray study. Patients with CFS/ME whose blood was used for subsequent PCR studies comprised those in the microarray study along with an additional 30 patients from clinics in 3 United Kingdom cities (Dorset, Bristol, and London; 1 patient from Leicester was under the care of a clinic in London) and New York, New York. CFS/ME was diagnosed on the basis of Centers for Disease Control and Prevention (CDC) criteria [1]. Patients with psychiatric disease were excluded on the basis of findings of the Minnesota International Neuropsychiatric Interview, thus ensuring that no patients had major psychiatric disease or were abusing alcohol or drugs. In addition, patients who had smoked tobacco during the previous 12-month period and/or had taken antibiotics, steroids, or antidepressants during the previous 3-month period were excluded from the study.

Healthy blood donors were used as a comparison group for both the microarray and real-time PCR studies. For the microarray study, 50 healthy blood donors were enrolled from the Dorset National Blood Service (NBS) and matched to the CFS/ME group at a ratio of 2:1 on the basis of age, sex, and geographical location. For subsequent PCR studies, the comparison group comprised the donors involved in the microarray study plus an additional 25 donors, all enrolled from the NBS. Restrictions imposed by the NBS on persons who are allowed to donate blood are outlined elsewhere [13]. Blood donors were excluded from the study if they had smoked tobacco during the previous 12-month period and/or had taken antibiotics, steroids, or antidepressants during the previous 3-month period.

In accordance with the recommendations of the International Chronic Fatigue Syndrome Study Group [18], the following questionnaires were completed for all enrolled subjects (patients and control donors): the Chalder Fatigue Scale [19], to assess the severity of physical and mental fatigue; the Medical Outcomes Survey Short Form-36 (SF-36), to determine the level of disability; the Somatic and Psychological Health Report, to characterize accompanying symptoms; the Pittsburgh Sleep Questionnaire, to assess abnormal sleep behaviors; and the McGill Pain Questionnaire, to assess the type and severity of pain. For patients with CFS/ME, neurocognitive testing was performed using the Spatial Span (SSP) and Verbal Recognition Memory (VRM) modules of Cantab software (Cambridge Cognition). In another study, the SSP test yielded abnormal findings for patients with CFS/ME [20], and we found similar abnormalities in patients with CFS/ME enrolled in this study.

Patients and control subjects gave informed written consent, in accordance with guidance of the Wandsworth Research Ethics Committee (approval number 05/Q0803/137). For patients in New York, approval of the local institutional review board was obtained. Human experimentation guidelines of the US Department of Health and Human Services were followed in this study.

Fifteen-milliliter blood samples from patients with CFS/ME and healthy blood donors (as part of routine blood donation) were collected into PAXgene tubes (PreAnalytix), and total RNA was extracted using the PAXgene blood RNA kit (PreAnalytix), according to the instructions of the manufacturer. RNA quality and amount were confirmed by microspectrophotometry (Nanodrop). Total RNA samples used in this study had an absorbance ratio (A260/280) of 1.9–2.0.

Microarray analysis. RNA specimens were shipped as ethanol precipitates to the Penn Microarray Facility (Philadelphia, PA), where mass and qualitative assays were repeated, using the Nanodrop spectrophotometer and Agilent Bioanalyzer, respectively. Bioanalyzer traces indicated that intact ribosomal bands as well as various background and partial degradation bands typical of blood RNA were present in all samples. All microarray analyses were conducted with the GeneChip human genome U133+2 microarray (Affymetrix) and the One-Cycle target labeling and control reagents kit (Affymetrix), in accordance with the manufacturer's recommendations. The average cRNA yield was 48 μg for samples from the control group and 51 μg for samples from patients with CFS/ME. Samples from 10 patients with CFS/ME were randomly selected as technical replicates for hybridization to a second GeneChip, to assess concordance. Microarray Suite 5.0 (Affymetrix) was used to quantitate expression levels for targeted genes; default values provided by the manufacturer were applied to all analysis parameters. A weighted mean value of probe fluorescence (corrected for nonspecific signals by subtracting the mismatch probe value) was calculated using the Tukey 1-step biweight estimate. This signal value, a relative measure of the expression level, was computed for each assayed gene. Global scaling was performed to allow comparison of gene signals across multiple microarrays.

Microarray data were normalized in Excel 2003 (Microsoft) and imported into GeneSpring 7.3 (Agilent Technologies) within a genome consisting of the entirety of the probes on the microarray minus the probes that were specific for >1 gene (i.e., those with the suffix "_s_at" or "_x_at"); this effectively excluded probes whose specificity could not be assigned with certainty to any one gene and reduced the genome to 39,174 probes. Data were analyzed by use of a class comparison experiment to obtain a list of gene probes that showed differences in expression between test and control groups with a fold difference cutoff of 2.5 and a P value of $\leq$.05 (an arbitrary level designed to select for 5% of the probes whose values

Table 1. Demographic and clinical characteristics and results of diagnostic tests for patients with chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) and healthy blood donors involved in microarray and real-time polymerase chain reaction (PCR) studies.

	Microarray study		Real-time PCR study	
Characteristic	CFS/ME group (n = 25)	Healthy group (n = 50)	CFS/ME group (n = 55)	Healthy group (n = 75)
Female sex	19	38	36	51
Age, years, mean	43.2	44.1	41.6	43.3
Disease duration, years, mean	3.33	NA	3.17	NA
Symptoms/signs				
Headache	11	1	26	3
Sore throat	11	0	27	0
Poor memory/concentration	20	2	30	3
Muscle pain	18	0	37	0
Muscle weakness	12	0	36	0
Joint pain	20	1	41	2
Postexertional malaise	23	0	47	0
Sleep problem	22	2	44	4
Gastrointestinal problems	15	1	35	3
Fainting/dizziness	11	0	25	0
Numbness/tingling	10	0	24	0
Tender lymphadenopathy	12	0	27	0
Test, score, mean				
Chalder Fatigue Scale				
Physical	15.69	5.84	16.13	7.41
Mental	8.1	3.38	8.05	4.25
McGill Pain Questionnaire	16.11	1.08	15.28	1.22
SPHERE	12.17	1.42	11.25	1.76
SF-36	49.7	83.87	46.45	84.96
Pittsburgh Sleep Quality Index	10.31	4.36	10.22	4.40

NOTE. NA, not applicable; SF-36, Medical Outcomes Survey Short Form-36; SPHERE, Somatic and Psychological Health Report.

showed the greatest differences between the CFS/ME and control groups). Probe values for each of these genes were reviewed, and a gene was included only if the majority of probe values were in agreement with the mean fold difference of the probe that originally flagged the gene.

Quantitative PCR. Quantitative PCR (Applied Biosystems) with TaqMan primers and probes was used to confirm the significance of genes identified in array experiments by the comparative method, using custom 384-well low-density arrays (LDAs) and the ABI Prism 7900HT instrument (Applied Biosystems), with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as the endogenous control gene. Experiments were performed in triplicate, using a protocol described elsewhere [13]. Data were displayed using SDS software (version 2.2 [ABI]), and discordant data between replicates were omitted. Results for each LDA were calculated and loaded into SDS software (version 2.2, Enterprise Edition), to facilitate analysis of data involving up to 80–100 LDAs per experiment.

The threshold cycle (Ct) for the test gene in each sample was compared to that of a calibrator sample to calculate a ΔCt value. ΔCt values were then normalized to the Ct value for GAPDH in respective samples to give $\Delta\Delta$Ct values. Relative quantities of each mRNA of interest (RQ; defined as $2^{-\Delta\Delta Ct}$) were then calculated. Samples showing a difference of $\geqslant$100 between minimum and maximum RQs (which is indicative of poor replicate concordance) were excluded. The t test was used to compare RQs for the patients with CFS/ME with RQs for the control subjects. Genes with mean RQs that differed significantly (defined as a P value of $\leqslant$.05) between the groups in the same direction as in the microarray were included in our CFS/ME-associated gene signature.

Transcription factor binding sites. Promoter sequences for each of the 83 genes identified after real-time PCR analysis were extracted from the Mammalian Promoter Database at Cold Spring Harbor Laboratory (Cold Spring Harbor, NY; available

Table 2. Characteristics of differentially expressed genes in chronic fatigue syndrome/myalgic encephalomyelitis.

Gene symbol	Gene name	GenBank accession no.	TaqMan assay identification no.[a]	Microarray fold difference	Real-time PCR	
					Fold difference	2-tailed *P*
ABCD4[b]	ATP-binding cassette, subfamily D (ALD), member 4	NM_020323	Hs00245340_m1	NA	2.08	.028
ACTR3	ARP3 actin-related protein 3 homolog (yeast)	NM_005721	Hs00828586_m1	4.55	1.42	.0042
AKAP10	A-kinase (PRKA) anchor protein 10	NM_007202	Hs00183673_m1	12.5	1.54	.0011
ANAPC11[b]	APC11 anaphase promoting complex subunit 11 homolog (yeast)	NM_016476	Hs00212858_m1	6.3	3.32	.00033
ANAPC5	Anaphase promoting complex subunit 5	NM_016237	Hs00212120_m1	8.25	2.36	.00016
APP	Amyloid β (A4) precursor protein	NM_201413	Hs00169098_m1	3.1	2.5	4.33×10^{-9}
ARL4C	ADP-ribosylation factor–like 4C	NM_005737	Hs00255039_s1	2.23	2.96	8.90×10^{-6}
ARPC5	Actin-related protein 2/3 complex, subunit 5	NM_005717	Hs00271722_m1	4.8	3.23	6.82×10^{-8}
ARSD	Arylsulfatase D	NM_001669	Hs00534692_m1	6.97	1.98	.001
ATP6V1C1	ATPase, H^+ transporting, lysosomal V1 subunit C1, 42 kDa	NM_001695	Hs00184625_m1	3.8	2.03	.00029
BCOR	BCL6 corepressor	NM_017745	Hs00372369_m1	2.95	1.6	.0098
BMP2K	BMP2 inducible kinase	NM_198892	Hs00214079_m1	1.67	1.3	.014
BRMS1[b]	Breast cancer metastasis suppressor 1	NM_015399	Hs00363036_m1	4	2.68	.0014
CD2BP2[b]	CD2 (cytoplasmic tail) binding protein 2	NM_006110	Hs00272036_m1	3.6	1.8	5.35×10^{-6}
CD47	CD47 molecule	NM_198793	Hs00179953_m1	1.62	2.2	.00013
CEP350	Centrosomal protein, 350 kDa	NM_014810	Hs00402774_m1	2.05	2.02	.0048
CITED2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy terminal domain, 2	NM_006079	Hs00366696_m1	7.05	2.39	4.45×10^{-6}
CMTM6	CKLF-like MARVEL transmembrane domain containing 6	NM_017801	Hs00215083_m1	1.6	1.41	.012
CREBBP	CREB binding protein (Rubinstein-Taybi syndrome)	NM_004380	Hs00231733_m1	2.37	1.43	.016
CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	NM_016823	Hs00180418_m1	2.2	2.51	1.11×10^{-5}
CTBP1	C-terminal binding protein 1	NM_001328	Hs00179922_m1	2.85	1.45	.062
CXCR4	C-terminal binding protein 1	NM_003467	Hs00607978_s1	2.7	1.67	7.80×10^{-5}
EBI2	Epstein-Barr virus–induced gene 2 (lymphocyte-specific GPCR)	NM_004951	Hs00270639_s1	1.3	3.44	.0012
EIF2B4[b]	Eukaryotic translation initiation factor 2B, subunit 4 δ, 67 kDa	NM_172195	Hs00248984_m1		2.06	.025
EIF3S10	Eukaryotic translation initiation factor 3, subunit 10 θ, 150/170kDa	NM_003750	Hs00186707_m1	1.7	3.58	.0029
EIF4G1[b]	Eukaryotic translation initiation factor 4A, isoform 2	NM_198241	Hs00191933_m1		3.05	.0033
EIF4G3	Eukaryotic translation initiation factor 4 γ, 3	NM_003760	Hs00186804_m1	1.4	1.67	1.37×10^{-5}
FAM126B	Family with sequence similarity 126, member B	NM_173822	Hs00545158_m1	2.35	1.64	.0034
FNTA	Farnesyltransferase, CAAX box, α	NM_002027	Hs00357739_m1	2.6	2.18	3.82×10^{-6}
GABARAPL1[b]	GABA(A) receptor associated protein–like 1	NM_031412	Hs00744468_s1	1.65	5.64	6.10×10^{-5}
GCN1L1	GCN1 general control of amino acid synthesis 1–like 1 (yeast)	NM_006836	Hs00412445_m1	2.45	2.05	.00052
GLTSCR2	Glioma tumor suppressor candidate region gene 2	NM_015710	Hs00414236_m1	3.7	1.24	.026

(continued)

Table 2. ***(Continued.)***

Gene symbol	Gene name	GenBank accession no.	TaqMan assay identification no.[a]	Microarray fold difference	Real-time PCR	
					Fold difference	2-tailed *P*
GNAS	GNAS complex locus	NM_080425	Hs00255603_m1	1.7	1.7	1.09×10^{-7}
GSN[b]	Gelsolin (amyloidosis, Finnish type)	NM_198252	Hs00609276_m1	3.62	2.93	.00017
GTF2A2	General transcription factor IIA, 2, 12 kDa	NM_004492	Hs00362112_m1	1.25	1.79	.03
HIF1A	Hypoxia-inducible factor 1, α subunit	NM_001530	Hs00153153_m1	0.6	0.81	.016
IFNAR1	Interferon (-α, -β, and -ω) receptor 1	NM_000629	Hs00265057_m1	2.2	1.76	.00073
IL10RA[b]	Interleukin-10 receptor, α	NM_001558	Hs00387004_m1	1.7	1.73	9.87×10^{-6}
IL6R	Interleukin-6 receptor	NM_000565	Hs00794121_m1	1.45	1.19	.06
IL6ST	Interleukin-6 signal transducer	NM_002184	Hs00174360_m1	2.12	1.8	.002
IL7R	Interleukin-7 receptor	NM_002185	Hs00233682_m1	0.45	0.82	.032
JAK1	Janus kinase 1 (a protein tyrosine kinase)	NM_002227	Hs00233820_m1	1.85	1.91	1.86×10^{-8}
KHSRP[b]	KH-type splicing regulatory protein (FUSE binding protein 2)	NM_003685	Hs00269352_m1	6.5	1.67	.00026
MAPK9	Mitogen-activated protein kinase 9	NM_139070	Hs00177102_m1	2.75	1.4	.045
METTL3	Methyltransferase-like 3	NM_019852	Hs00219820_m1	2.55	2.06	.0001
MRPL23[b]	Mitochondrial ribosomal protein L23	NM_021134	Hs00221699_m1	3.4	2.06	.001
MRPS6	Mitochondrial ribosomal protein S6	NM_032476	Hs00606808_m1	1.47	1.53	.025
MRRF	Mitochondrial ribosome recycling factor	NM_138777	Hs00751845_s1	4.6	8.91	.0004
MSN[c]	Moesin	NM_002444	Hs00792607_mH	2.1	1.33	.0016
MTMR6	Myotubularin-related protein 6	NM_004685	Hs00395064_m1	2.95	1.71	.0025
NFKB1	Nuclear factor of κ light polypeptide gene enhancer in B-cells 1 (p105)	NM_003998	Hs00231653_m1	2.8	1.59	4.04×10^{-5}
NR1D2	Nuclear receptor subfamily 1, group D, member 2	NM_005126	Hs00233309_m1	2.1	2.44	.00076
NTE[b]	Neuropathy target esterase	NM_006702	Hs00198648_m1	3.5	1.7	.04
NUFIP2	Nuclear fragile X mental retardation protein interacting protein 2	NM_020772	Hs00325168_m1	5.6	1.5	.00036
PAPOLA	Poly(A) polymerase α	NM_032632	Hs00413685_m1	1.74	1.32	.00194
PDCD2[b]	Programmed cell death 2	NM_002598	Hs00751277_sH	2.55	6.76	.0096
PDCD6	Programmed cell death 6	NM_013232	Hs00737034_m1	2.6	1.74	.00019
PEX16[b]	Peroxisomal biogenesis factor 16	NM_004813	Hs00191337_m1	4.2	1.74	.0034
PGM2	Phosphoglucomutase 2	NM_018290	Hs00217619_m1	2.05	2.17	1.68×10^{-6}
PIK3R1	Phosphoinositide-3-kinase, regulatory subunit 1 (p85 α)	NM_181523	Hs00236128_m1	0.45	0.68	.025
PKN1[b]	Protein kinase N1	NM_213560	Hs00177028_m1	1.7	1.56	9.40×10^{-5}
POLR2G[b]	Polymerase (RNA) II (DNA-directed) polypeptide G	NM_002696	Hs00275738_m1	2.3	2.58	.0078
PPP2R5C	Protein phosphatase 2, regulatory subunit B (B56), γ isoform	NM_002719	Hs00604902_m1	1.83	1.38	.022
PRKAA1	Protein kinase, AMP-activated, α 1 catalytic subunit	NM_006251	Hs01562315_m1	2.22	1.72	.00052
PRKAR1A	Protein kinase, cAMP-dependent, regulatory, type I, α	NM_002734	Hs00267597_m1	1.3	2.63	2.91×10^{-8}
PUM2	Pumilio homolog 2 *(Drosophila)*	NM_015317	Hs00209677_m1	2.7	1.39	.00064
RAP2C	RAP2C, member of RAS oncogene family	NM_021183	Hs00221801_m1	1.85	2.1	.015
RNF141	Ring finger protein 141	NM_16422	Hs00212656_m1	2.17	2.37	1.62×10^{-6}
SELENBP1	Selenium binding protein 1	NM_003944	Hs00187625_m1	2.8	1.92	.002
SFXN1	Sideroflexin 1	NM_022754	Hs00224259_m1	3	1.6	.022

(continued)

Table 2. ***(Continued.)***

Gene symbol	Gene name	GenBank accession no.	TaqMan assay identification no.[a]	Microarray fold difference	Real-time PCR	
					Fold difference	2-tailed *P*
SHPRH	SNF2 histone-linker PHD RING helicase	NM_173082	Hs00542737_m1	3.35	1.77	.05
SNAP23	Synaptosomal-associated protein, 23 kDa	NM_003825	Hs00187075_m1	1.9	2.02	.00018
SORL1	Sortilin-related receptor, L(DLR class) A repeats–containing	NM_003105	Hs00268342_m1	1.57	1.54	4.10×10^{-8}
SOS1	Son of sevenless homolog 1 *(Drosophila)*	NM_005633	Hs00362308_m1	1.6	2.31	.002
TAF11	TAF11 RNA polymerase II, TATA box–binding protein (TBP)–associated factor, 28 kDa	NM_005643	Hs00194573_m1	3.1	1.87	.05
TCF3	Transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	NM_003200	Hs00413032_m1	1.8	1.44	.023
TDP1	Tyrosyl-DNA phosphodiesterase 1	NM_018319	Hs00217832_m1	2.1	1.67	.0099
TNFRSF1A	Tumor necrosis factor receptor superfamily, member 1A	NM_001065	Hs00533560_m1	1.46	1.37	.016
UBTF	Upstream binding transcription factor, RNA polymerase I	NM_014233	Hs00610729_g1	1.95	2.26	.024
USP38	Ubiquitin-specific peptidase 38	NM_032557	Hs00261419_m1	3.3	1.71	.0021
WAPAL	Wings apart–like homolog *(Drosophila)*	NM_015045	Hs00386162_m1	1.97	1.69	.027
WDR26	WD repeat domain 26	NM_025160	Hs00228535_m1	2.12	2.62	.00012
Transcription factor						
NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B cells 1 (p105)	NM_003998	Hs00231653_m1	2.8	1.59	4.04×10^{-5}
NHLH1	Nescient helix loop helix 1	NM_005589	Hs00271582_s1	NA	11.51	7.00×10^{-4}
REPIN1	Replication initiator 1	NM_013400	Hs00274221_s1	NA	3.62	6.00×10^{-6}
GABPA	GA binding protein transcription factor, α subunit, 60 kDa	NM_002031	Hs00745591_s1	NA	8.06	3.00×10^{-4}
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	NM_005238	Hs00901425_m1	NA	2.11	1.00×10^{-5}
EGR1	Early growth response 1	NM_001955	Hs00152928_m1	NA	2.82	.015
EGR3	Early growth response 3	NM_004421	Hs00231780_m1	NA	1.92	.017

NOTE. A graphic representation of these data appears in figure 1. Fold differences are defined as the relative quantity of mRNA transcripts for patients with CFS/ME divided by the value for healthy control subjects. NA, not available; PCR, polymerase chain reaction.

[a] Assays were predesigned by Applied Biosystems.
[b] Genes found in pilot study [13].
[c] Genes found by means of differential display polymerase chain reaction [7].

at: http://rulai.cshl.edu/cgi-bin/CSHLmpd2/hspd.pl) and analyzed for overrepresentation of transcription factor binding sites, using tools from the Institute of Bioinformatics (Beijing, China; available at: http://www.bioinfo.tsinghua.edu.cn/~zhengjsh/OTFBS); this analysis is described elsewhere [21]. Human transcription factors identified by this approach were tested in the same samples, using available real-time PCR assays (ABI 7500 Fast PCR System [Applied Biosystems]) performed in accordance with the method described above but in a 96-well format.

Analysis of gene function and interaction. Mean RQs from TaqMan real-time PCR testing of the 88 CFS/ME-associated genes were analyzed with Ingenuity Pathways Analysis (IPA) (Ingenuity) to link CFS/ME-associated genes into networks based on recognized interactions. IPA was also performed to discern the top associated diseases and disorders, molecular and cellular functions, associated physiological system development and function, and canonical pathways on the basis of over-representation analysis, in which, for each pathway, the fraction of CFS/ME-associated genes involved in that pathway was compared to the fraction of human genes involved in that pathway. For each pathway, the probability of involvement of the respective number of CFS/

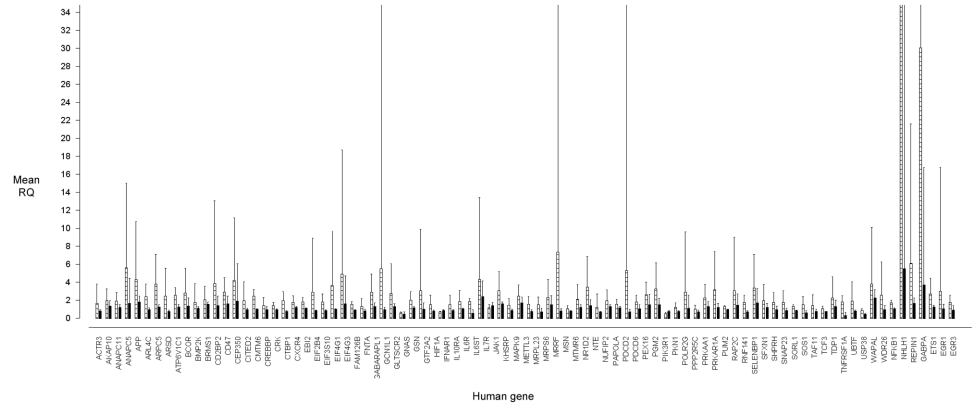

Figure 1. Mean relative quantity of mRNA transcripts (RQ) in patients with chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME; *hatched bars*) and healthy control subjects *(black bars)* for 88 genes that were differentially expressed in patients with CFS/ME. *Error bars* indicate variance from the mean in each case. The chart is truncated at an RQ of 34, to enhance the clarity of data for the greatest number of genes.

ME-associated genes was expressed as a *P* value; values of ⩽.05 were taken to be significant.

Clustering of quantitative PCR–generated gene values for patients with CFS/ME. RQs for all 88 CFS/ME-associated genes were normalized and clustered by use of Genesis software [22]. For each of the 7 CFS/ME subtypes identified by means of this approach, mean RQs were calculated for each gene and used to generate fold differences (defined as the value for patients with CFS/ME divided by the value for healthy control subjects) for each gene in each CFS/ME subtype. Mean fold differences for each gene in each CFS/ME subtype were then clustered with and without normalization/median centering by use of Cluster software, version 2.11, and visualized by use of Treeview software, version 1.60 [23].

RESULTS

Subjects and clinical characterization. Fifty-five patients with CFS/ME who fulfilled CDC diagnostic criteria and 82 healthy blood donors were enrolled in the study; 25 patients and 50 donors were evaluated in the microarray study, and 55 patients and 75 healthy donors were evaluated in the quantitative PCR study. Patient and clinical details are summarized in table 1. This study included several patients with CFS/ME whose disease was severe and necessitated bed rest for much of the day and patients who were able to attend an outpatient clinic (table 1).

Microarray analysis. Although overall RNA quality was similar between the 2 groups, as indicated by Bioanalyzer traces, cRNA amplification yields, and GAPDH integrity, the microarrays from patients with CFS/ME showed half as many genes that met the Microarray Suite 5.0 "Present" detection threshold, but the average level of expression of these genes was 1.9 times the average level in control samples. Technical microarray replicates showed a concordance correlation coefficient of 0.978. GeneSpring analysis of microarray data identified 1789 probes with a differential expression in patients with CFS/ME that was statistically significantly different from that in healthy controls at a fold difference cutoff of ⩾2.5. However, when all probe values

Table 3. Top gene networks associated with the chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) signature, according to Ingenuity Pathways Analysis.

Network[a]	No. of genes	Score[b]	Gene symbol
Hematological disease, cell development, hematological system development and function	17	33	*CD47, CRK, CXCR4, GNAS, HIF1A, IFNAR1, IL10RA, IL6R, IL6ST, IL7R, JAK1, PIK3R1, PKN11, POLR2G, PRKAA1, SOS1, TNFRSF1A*
Cell morphology, gene expression, cellular assembly and organization	17	21	*AKAP10, APP, BRMS1, CITED2, CREBBP, ETS1, GABPA, GSN, GTF2A2, MAPK9, MSN, NFKB1, PRKAR1A, SNAP23, SORL1, TCF3, UBTF*
Protein synthesis, cell cycle, gene expression	12	21	*ACTR3, ARPC5, CD2BP2, EIF3S10, EIF4G1, EIF4G3, FNTA, GLTSCR2, PAPOLA, PDCD6, PPP2R5C, TAF11*
Gene expression, cellular development, hematological system development and function	9	14	*ARL4C, BCOR, CEP350, CTBP1, EGR3, NHLH1, NR1D2, PNPLA6, SELENBP1*
Cell cycle, protein degradation, protein synthesis	6	9	*ANAPC5, ANAPC11, EBI2, EIF2B4, GCN1L1, PDCD2*

[a] See figure 2.
[b] Indicates significance of the network in terms of overrepresentation of network-eligible genes.

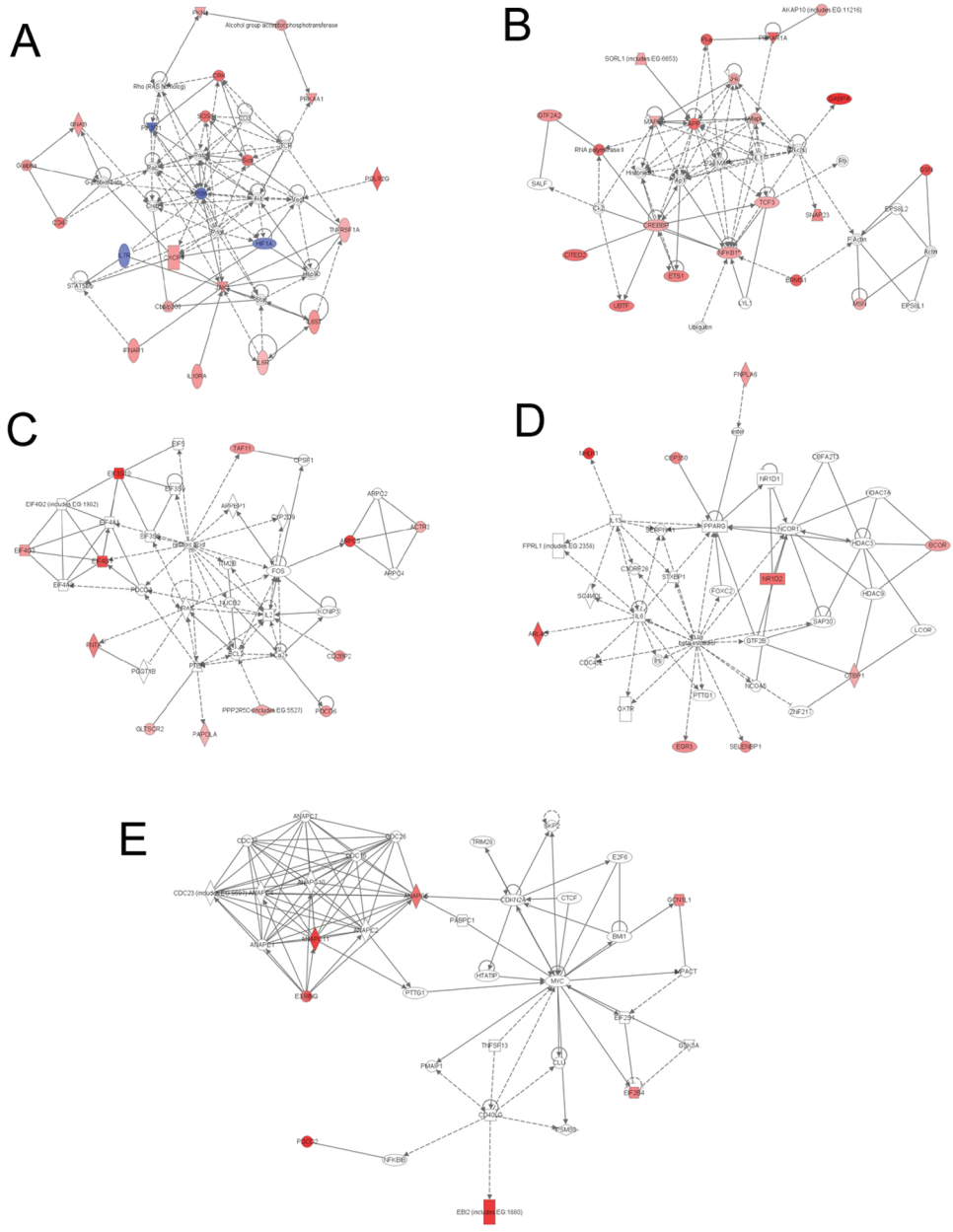

Figure 2. Five large and distinct gene networks *(A–E)* identified using Ingenuity Pathways Analysis (Ingenuity); functional themes of each are shown in tables 3–5. Genes are colored according to levels of expression determined in the present study: *red* denotes upregulation, and *blue* denotes downregulation. *Color intensity* reflects the magnitude of the fold difference between patients with CFS/ME and healthy subjects.

for each of these genes were reviewed, for only 182 genes were the majority of probe values in agreement with the mean fold difference of the probe that originally flagged that gene.

Quantitative PCR. Statistically significant differential expression with the same profile as that in gene arrays was confirmed for 82 of 182 genes tested, with upregulation of 79 genes and downregulation of 3 genes (*PIK3R1, IL7R,* and *HIF1A*) (table 2 and figure 1).

Transcription factors. Analysis of the promoter sequences of the 82 human genes specified above revealed over-representation of binding sites for the following transcription factor genes: *REPIN1, SP1, ETS1, GABPA, GTF3A, EGR1, EGR2, EGR3, NFKB1, NHLH1, EGR4, REST,* and the Epstein-Barr virus (EBV) R transactivator gene *BRLF1*. Of these 13 transcription factor genes, results of real-time PCR assays were available for all except *BRLF1; NFKB1* had already been tested because it was one of the original 82 genes. Real-time TaqMan PCR was used to test the remaining 11 genes in 25 patients with CFS/ME and 25 healthy controls, and significant upregulation in patients with CFS/ME was confirmed for 7 genes (*REPIN1, ETS1, GABPA, NFKB1, EGR1, EGR3,* and *NHLH1*) in patients with CFS/ME (table 2 and figure 1), suggesting that upregulation of these genes may be key to the overall gene signature ob-

Table 4. Top gene functions associated with the chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) signature, according to Ingenuity Pathways Analysis.

Gene function	No. of genes	*P*	Gene symbol
Diseases and disorders			
Hematological disease	22	1.37×10^{-9} to 1.04×10^{-2}	*APP, CD47, CRK, CTBP1, CXCR4, GNAS, HIF1A, IFNAR1, IL6ST, NFKB1, PIK3R1, TNFRSF1A, CREBBP, FNTA, IL6R, EGR3, GSN, IL7R, MAPK9, TCF3, MSN, PRKAR1A*
Immunological disease	14	1.27×10^{-6} to 1.04×10^{-2}	*CD47, CXCR4, EGR3, GNAS, GSN, IL6R, IL6ST, IL7R, MAPK9, NFKB1, PIK3R1, TNFRSF1A, ETS1, IFNAR1*
Cancer	31	2.19×10^{-5} to 8.81×10^{-3}	*CD47, CXCR4, GSN, CREBBP, CRK, EIF4G1, ETS1, MAPK9, NFKB1, PIK3R1, PRKAR1A, SOS1, UBTF, CRK, TCF3, IL6ST, GNAS, HIF1A, IFNAR1, IL6R, FNTA, TNFRSF1A, MSN, CTBP1, WAPAL, APP, JAK1, MTMR6, PPP2R5C, PRKAA1, BRMS1*
Dermatological diseases and conditions	3	2.19×10^{-5} to 9.48×10^{-4}	*CD47, CXCR4, GSN*
Endocrine system disorders	9	6.31×10^{-5} to 7.6×10^{-3}	*CREBBP, GNAS, IL6ST, MAPK9, NFKB1, PIK3R1, TNFRSF1A, APP, CXCR4*
Neurological disease	8	2.03×10^{-3} to 1.04×10^{-2}	*APP, NFKB1, CITED2, MAPK9, EGR3, GNAS, HIF1A, TNFRSF1A*
Nervous system development and function	12	7.31×10^{-5} to 9.41×10^{-3}	*APP, CD47, NFKB1, CREBBP, IL6R, IL6ST, CXCR4, EGR3, MAPK9, NHLH1, CITED2, EIF2B4*
Inflammatory disease	9	1.41×10^{-4} to 1.03×10^{-2}	*APP, CXCR4, IFNAR1, IL6R, IL6ST, MAPK9, NFKB1, PIK3R1, TNFRSF1A*
Infectious disease	5	4.54×10^{-3} to 9.26×10^{-3}	*IFNAR1, NFKB1, PIK3R1, APP, CXCR4*
Molecular and cellular functions			
Cell development	26	5.47×10^{-10} to 9.57×10^{-3}	*CREBBP, CXCR4, EGR3, ETS1, HIF1A, IL6R, IL6ST, IL7R, JAK1, NFKB1, PIK3R1, TCF3, TNFRSF1A, CD47, CRK, EIF2B4, GSN, SOS1, IFNAR1, IL10RA, APP, PNPLA6, PRKAR1A, SFXN1, FNTA, MSN*
Cell death	33	8.6×10^{-7} to 1.04×10^{-2}	*CD47, CXCR4, EGR3, GNAS, GSN, IL6R, IL6ST, IL7R, MAPK9, NFKB1, PIK3R1, TNFRSF1A, APP, CITED2, CREBBP, CRK, CTBP1, ETS1, HIF1A, MSN, PDCD2, PDCD6, PRKAA1, PRKAR1A, TCF3, MSN, FNTA, UBTF, WAPAL, PNPLA6, PPP2R5C, JAK1, MTMR6*
Gene expression	31	8.77×10^{-7} to 1.04×10^{-2}	*APP, CEP350, CITED2, CREBBP, EGR3, ETS1, GABPA, HIF1A, JAK1, NFKB1, PKN1, PPP2R5C, TAF11, TCF3, UBTF, CRK, PRKAR1A, TNFRSF1A, BCOR, EBI2, GTF2A2, IL6R, IL6ST, MAPK9, NR1D2, POLR2G, UBTF, RNF141, CD47, SOS1, CTBP1*
Cellular growth and proliferation	31	2.55×10^{-6} to 7.48×10^{-3}	*ANAPC5, CRK, CTBP1, CXCR4, EIF4G1, ETS1, GLTSCR2, GSN, HIF1A, IFNAR1, IL10RA, IL6R, IL6ST, IL7R, MAPK9, METTL3, NFKB1, PIK3R1, PKN1, PPP2R5C, PRKAR1A, SOS1, TCF3, TNFRSF1A, UBTF, CD47, CITED2, EGR3, APP, JAK1*
Cellular assembly and organization	15	3.53×10^{-6} to 9.99×10^{-3}	*GNAS, SNAP23, APP, CD47, MSN, GSN, EIF3S10, UBTF, CXCR4, ACTR3, ARPC5, TNFRSF1A, IL6R, IL6ST, CRK*
Physiological system development and function			
Hematological system development and function	22	5.08×10^{-9} to 1.04×10^{-2}	*CXCR4, TNFRSF1A, IL6R, IL6ST, APP, IL7R, CD47, NFKB1, PIK3R1, TCF3, CTBP1, ETS1, MAPK9, SFXN1, IFNAR1, JAK1, GNAS, IL10RA, EGR3, GSN, HIF1A, FNTA*
Immune and lymphatic system development and function	18	3.02×10^{-7} to 1.04×10^{-2}	*IL6ST, CXCR4, IL7R, NFKB1, CD47, PIK3R1, TCF3, IFNAR1, TNFRSF1A, CTBP1, IL10RA, EGR3, MAPK9, ETS1, CREBBP, JAK1, HIF1A, IL6R*
Tissue morphology	18	3.02×10^{-7} to 9.99×10^{-3}	*GNAS, APP, CITED2, IL6ST, NFKB1, ETS1, GSN, IL6R, TNFRSF1A, HIF1A, IL7R, CXCR4, CD47, PIK3R1, IFNAR1, TCF3, CREBBP, EGR3*
Organismal survival	17	6.27×10^{-7} to 1.17×10^{-6}	*APP, CITED2, CREBBP, CXCR4, EGR3, GNAS, HIF1A, IFNAR1, IL6ST, JAK1, MAPK9, NFKB1, NHLH1, PIK3R1, PNPLA6, TCF3, TNFRSF1A*
Immune response	20	4.93×10^{-5} to 1.04×10^{-2}	*IL6ST, IL6R, IL7R, IL10RA, CXCR4, TNFRSF1A, ETS1, IFNAR1, CD47, JAK1, APP, GNAS, PIK3R1, TCF3, NFKB1, CTBP1, EGR3, MAPK9, EBI2, GSN*

served in the CFS/ME group. This is supported by the fact that the fold differences for transcription factors were higher than those for most of the other genes. The transcription factor genes *EGR2*, *GTF3A*, and *SP1* were also upregulated in patients with CFS/ME; however, *P* values ranged from .18 to .09. Assays for *REST* and *EGR4* did not provide usable data for the cDNA concentrations used.

Analysis of gene function and interaction. Eighty-eight CFS/ME-associated genes were analyzed using IPA. The following 5 networks were revealed (table 3), all of which could

Table 5. Top canonical pathways associated with the chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) signature, according to Ingenuity Pathways Analysis.

Canonical pathway	Ratio	*P*	Gene symbols
Interferon signaling	2/29	8.69×10^{-3}	*IFNAR1, JAK1*
Interleukin (IL)–6 signaling	6/93	7.77×10^{-6}	*IL6R, IL6ST, MAPK9, NFKB1, TNFRSF1A, SOS1*
EGF signaling	3/47	1.88×10^{-3}	*JAK1, PIK3R1, SOS1*
IL-2 signaling	3/53	2.52×10^{-3}	*JAK1, PIK3R1, SOS1*
PPAR signaling	5/90	9.00×10^{-5}	*CITED2, CREBBP, NFKB1, TNFRSF1A, SOS1*
PDGF signaling	4/75	6.00×10^{-4}	*CRK, JAK1, PIK3R1, SOS1*
JAK-STAT signaling	3/59	3.44×10^{-3}	*JAK1, PIK3R1, SOS1*
GM-CSF signaling	3/62	3.79×10^{-3}	*ETS1, PIK3R1, SOS1*
Insulin receptor signaling	6/125	4.34×10^{-5}	*CRK, EIF2B4, JAK1, PIK3R1, PRKAR1A, SOS1*
IL-4 signaling	3/64	4.16×10^{-3}	*JAK1, PIK3R1, SOS1*
VEGF signaling	4/90	1.06×10^{-3}	*EIF2B4, HIF1A, PIK3R1, SOS1*
Estrogen receptor signaling	5/115	2.29×10^{-4}	*CREBBP, CTBP1, SOS1, TAF11, POLR2G*
IL-10 signaling	3/70	5.39×10^{-3}	*IL10RA, JAK1, NFKB1*
Amyloid processing	2/51	2.91×10^{-2}	*APP, PRKAR1A*
FGF signaling	3/84	9.34×10^{-3}	*CRK, PIK3R1, SOS1*
PI3K/AKT signaling	5/142	6.74×10^{-4}	*JAK1, PPP2R5C, PIK3R1, SOS1, NFKB1*
Neuregulin signaling	3/89	1.03×10^{-2}	*CRK, PIK3R1, SOS1*
B cell receptor signaling	5/149	9.9×10^{-4}	*ETS1, PIK3R1, SOS1, MAPK9, NFKB1*
IGF-1 signaling	3/90	9.97×10^{-3}	*PIK3R1, SOS1, PRKAR1A*
SAPK/JNK signaling	4/124	3.68×10^{-3}	*PIK3R1, SOS1, MAPK9, CRK*
Death receptor signaling	2/62	3.93×10^{-2}	*TNFRSF1A, NFKB1*
Cardiac β-adrenergic signaling	4/131	4.38×10^{-3}	*GNAS, PPP2R5C, AKAP10, PRKAR1A*
ERK/MAPK signaling	6/199	5.09×10^{-4}	*ETS1, PPP2R5C, PIK3R1, SOS1, CRK, PRKAR1A*
Fc-ε RI signaling	3/100	1.5×10^{-2}	*PIK3R1, SOS1, MAPK9*
Ephrin receptor signaling	6/206	4.95×10^{-4}	*GNAS, ACTR3, CXCR4, ARPC5, SOS1, CRK*
TGF-β signaling	3/104	1.42×10^{-2}	*SOS1, CREBBP, MAPK9*
T cell receptor signaling	3/104	1.63×10^{-2}	*PIK3R1, SOS1, NFKB1*
Actin cytoskeleton signaling	7/247	2.05×10^{-4}	*ACTR3, PIK3R1, ARPC5, SOS1, CRK, GSN, MSN*
NFKB signaling	4/142	5.87×10^{-3}	*TNFRSF1A, PIK3R1, CREBBP, NFKB1*
Neurotrophin/TRK signaling	2/72	4.79×10^{-2}	*PIK3R1, SOS1*
PTEN signaling	3/109	1.76×10^{-2}	*PIK3R1, SOS1, NFKB1*
Leukocyte extravasation signaling	5/192	3.22×10^{-3}	*CXCR4, PIK3R1, MAPK9, CRK, MSN*
Xenobiotic metabolism signaling	6/231	1.16×10^{-3}	*PPP2R5C, PIK3R1, CREBBP, MAPK9, NFKB1, CITED2*
G-protein coupled receptor signaling	5/200	3.75×10^{-3}	*GNAS, PIK3R1, SOS1, NFKB1, PRKAR1A*
Nicotinate and nicotinamide metabolism	3/129	2.13×10^{-2}	*PRKAA1, MAPK9, PKN1*
Inositol phosphate metabolism	4/177	4.11×10^{-2}	*PIK3R1, PRKAA1, MAPK9, PKN1*
Dopamine receptor signaling	2/91	8.2×10^{-2}	*PPP2R5C, PRKAR1A*
Huntington disease signaling	5/239	6.3×10^{-3}	*POLR2G, PIK3R1, SOS1, CREBBP, MAPK9*
Axonal guidance signaling	8/398	6.74×10^{-4}	*GNAS, ACTR3, CXCR4, PIK3R1, ARPC5, SOS1, CRK, PRKAR1A*
Integrin signaling	4/210	2.2×10^{-2}	*ACTR3, PIK3R1, ARPC5, SOS1*
cAMP-mediated signaling	3/159	4.82×10^{-2}	*GNAS, AKAP10, PRKAR1A*
Wnt/β-catenin signaling	3/162	4.75×10^{-2}	*PPP2R5C, CREBBP, TCF3*

NOTE. Ratios denote the total number of genes in the CFS/ME signature divided by the total number of genes in the human genome.

be linked together in a single network (data not shown): hematological disease, cell development, and hematological system development and function (figure 2*A*); cell morphology, gene expression, and cellular assembly and organization (figure 2*B*); protein synthesis, cell cycle, and gene expression (figure 2*C*); gene expression, cellular development, and hematological system development and function (figure 2*D*); and cell cycle, protein degradation, and protein synthesis (figure 2*E*).

Diseases and disorders, molecular and cellular functions, and physiological systems for these genes are shown in table 4. Prominent functional features involved immunity, inflammation, ap-

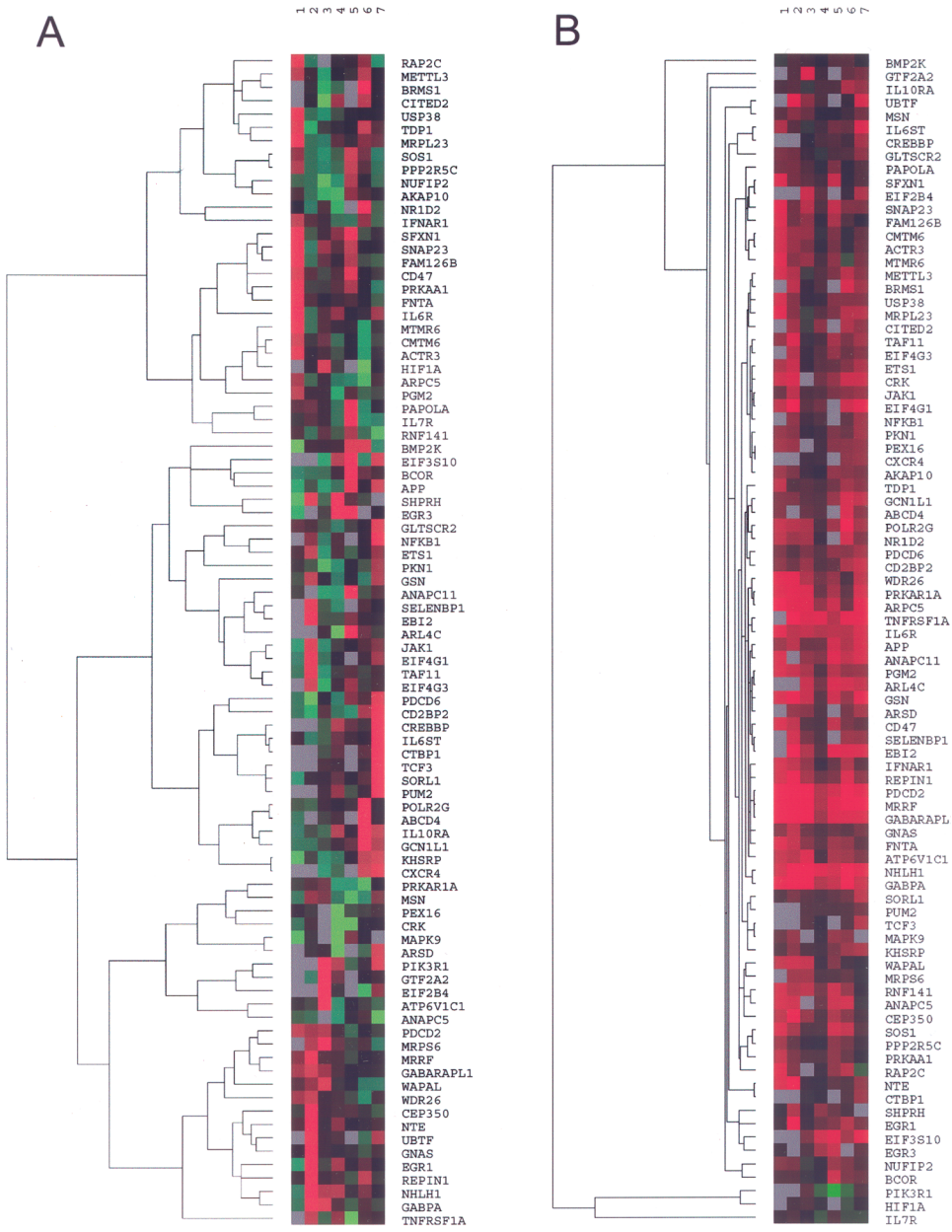

Figure 3. *A,* Clustering of normalized/median centered fold differences, calculated as the mean relative quantity of mRNA transcripts (RQ) in patients with chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) for each subtype divided by the mean RQ in healthy control subjects, for 88 genes associated with CFS/ME. *B,* Clustering of logarithm (base 2) fold differences, calculated as the mean RQ in patients with CFS/ME for each subtype divided by the mean RQ in healthy controls, for 88 genes associated with CFS/ME.

optosis, neurological disease and function, infectious disease, and cancer. Many canonical pathways were implicated by these genes (table 5); the most important themes were immunity, inflammation, apoptosis, and neurological disease and function.

Genomic CFS/ME subtypes. Clustering of RQs from patients with CFS/ME identified 7 subtypes consisting of 2, 5, 2, 19, 7, 14, and 3 patients. Clustering of mean fold differences (calculated as the mean subtype RQ of patients divided by the mean RQ for control subjects) for the 88 CFS/ME-associated genes revealed a distinct profile of gene expression in each subtype and clustering of genes with similar profiles of expression in the different subtypes (figure 3). Clustering of normalized/median centered fold differences for each gene in each subtype is shown in figure 3*A,* which highlights the distinct nature of the gene signature for each CFS/ME subtype. Clustering of logarithm (base 2) fold differences are shown in figure 3*B,* which emphasizes the predominance of upregulation for most genes in all CFS/ME subtypes.

Analysis of mean age and sex ratios for each subtype revealed that subtypes 3, 5, and 7 occurred only in females, subtype 2 was pre-

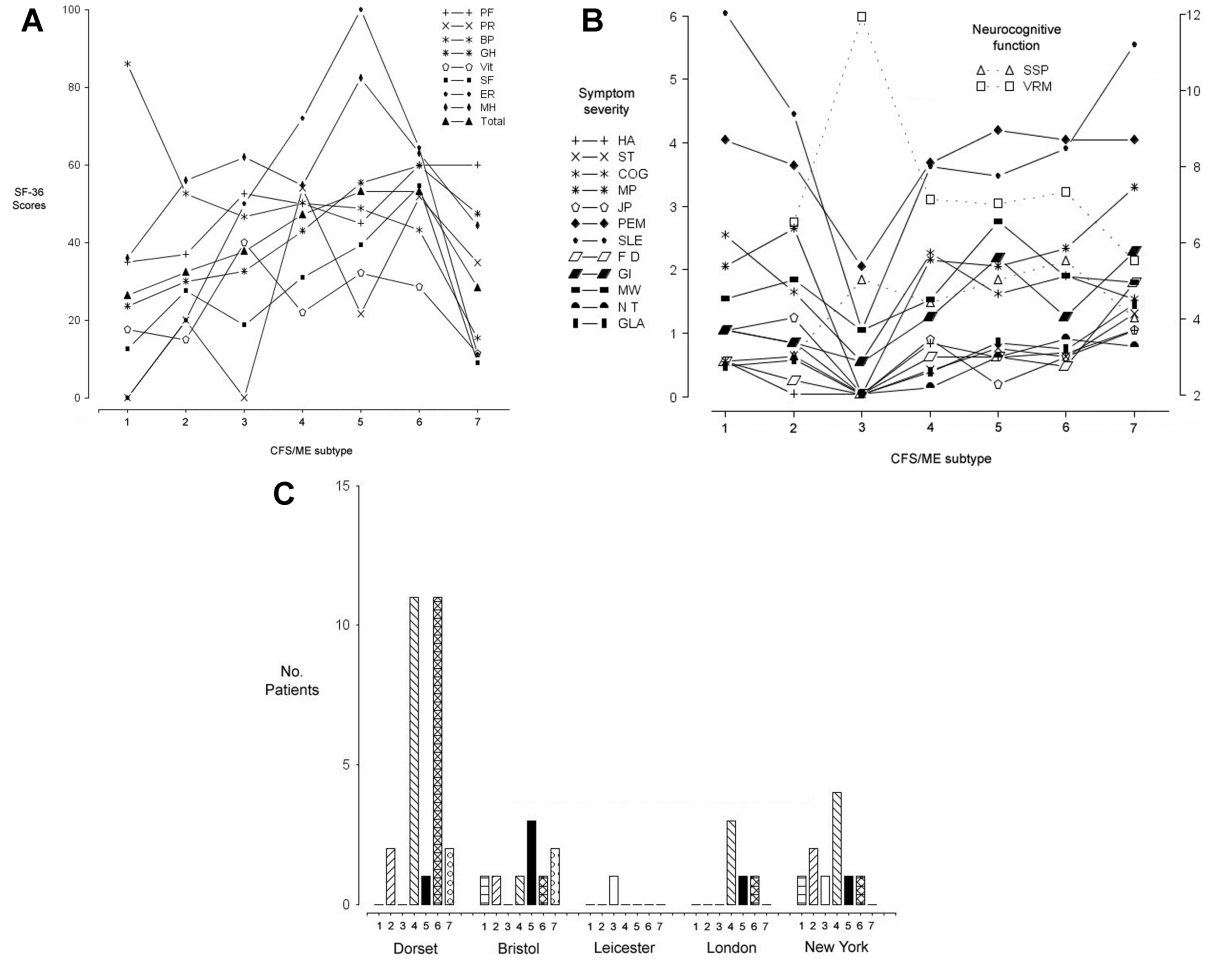

Figure 4. *A,* Medical Outcomes Survey Short Form-36 (SF-36) domain and total scores for each chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) subtype: physical function (PF), physical role (RP), bodily pain (BP), general health (GH), vitality (VIT), social functioning (SF), emotional role (RE), mental health (MH), and total score (Total). *B,* Scores indicating the occurrence and severity of the following 11 clinical symptoms and 2 neurocognitive functions for each CFS/ME subtype: headache (HA), sore throat (ST), swollen glands (GLA), cognitive defect (COG), muscle pain (MP), joint pain (JP), muscle weakness (MW), postexertional malaise (PEM), sleep problems (SLE), fainting/dizziness (F/D), gastrointestinal complaints (GI), numbness/tingling (N/T), spatial span (SSP), and verbal recognition memory (VRM). *C,* Histogram showing the number of patients with CFS/ME of each subtype, by geographical location.

dominantly male, and the remaining subtypes occurred in both males and females; age differences were less clearly demarcated. The clinical phenotype was distinct between subtypes; subtypes 1 and 7 were the most severe, followed sequentially by subtypes 2, 4, 5, 6, and 3. Analysis of variance testing revealed significant differences in the SF-36 total score ($P = .016$), social functioning ($P = .03$), and emotional role ($P = .003$) between the groups, whereas the difference between groups approached statistical significance for general health ($P = .08$) and mental health ($P = .08$).

After adjustment for multiple comparisons, significant associations were found between specific subtypes and clinical phenotypes. Subtype 7 had the most pain, the lowest SF-36 scores (along with subtype 1), and the most-severe individual symptoms, including swollen glands, sore throat, and headaches; subtype 1 had the worst cognition and mental health and poor sleep, despite being associated with the least pain; subtype 4 had moderate neurocognitive function and cognitive defects, combined with moderate levels of bodily pain and sleep problems; subtype 5 had the best mental health but poor neurocognitive function, gastrointestinal complaints, and the most marked muscle weakness and postexertional malaise; and subtype 2 had marked postexertional malaise, muscle pain, and joint pain but poor mental health (figures 4*A* and 4*B*).

Subtypes 4 and 6 were predominant in Dorset, subtype 4 was predominant in London and New York, and subtype 5 was predominant in Bristol (figure 4*C*).

DISCUSSION

In this article, we document the differential expression of 88 human genes in CFS/ME. We have confirmed involvement of all 16 genes reported in our pilot study [13] in the same direction as

previously reported, except *IL10RA*, which was downregulated in the previous study [13] but upregulated in the present study. This discrepancy remains unexplained, but because *IL10RA* is a critical gene required for T cell activation and since *IL10RA* levels correlate with measures of general health, we believe it to be particularly important.

The functions of these genes present a complex picture with links to several diseases and pathways (tables 3–5). Prominent themes that are well recognized in patients with CFS/ME are immunity, inflammation, and infection [5, 13]; cell death [13, 24]; cancer [25, 26]; and neurological disease [13, 27–29].

One of the most important viral triggers of CFS/ME is EBV, and this virus very likely plays an important role in perpetuation of disease, because it is reactivated by stress [30]. Within the gene signature identified in this study, the following 12 human genes that we found to be upregulated in patients with CFS/ME have been shown elsewhere to be upregulated, either directly or indirectly, by EBV infection: *NFKB1, EGR1, ETS1, GABPA, CREBBP, CXCR4,* and *EBI2* [31]; *HIF1A; JAK1; IL6R; IL7R;* and *PIK3R1*. A particularly interesting gene is *EBI2,* which was upregulated in 55% of patients with CFS/ME, one of whom was a 26-year-old woman with CFS/ME triggered by laboratory-documented EBV infection 10 years earlier. The EBV genes *BRLF1* and *BZLF1* mediate the switch from latent to lytic phases of EBV infection, and during this process they transactivate many human genes. It is interesting that *BRLF1* was identified as being overrepresented in the transcription factor analysis and that IgG specific to the Zebra protein (the *BZLF1* gene product) has been reported previously in patients with CFS/ME [32].

Enteroviruses are another very important viral trigger of CFS/ME [33]. Upregulation of *EIF4G1* transcript variant 5, which was found in this study and elsewhere [8, 13], is targeted during infection by various viruses, including enteroviruses, to subvert cellular machinery for the production of viral proteins.

The theme of neurological disease and function was prominent among the 88 CFS/ME-associated genes, almost all of which are expressed in multiple areas of the human brain (data not shown). Involvement of specific genes highlights the importance of neuregulin signaling, neurotrophin/TRK signaling, axonal guidance signaling, dopamine receptor signaling, and Huntington disease signaling (table 5). NTE was upregulated in blood of patients with CFS/ME, both in this study and our pilot study [13]. NTE is the primary site of action of organophosphate (OP) compounds, such as sarin, which cause axonal degeneration and paralysis due to inactivation of its serine esterase activity [34]. In the nervous system of adult chickens, OP-modified NTE initiates neurodegeneration. NTE probably regulates neuron-glial interactions during development and possibly also during adult life [35]. Exposure to OPs may trigger CFS/ME [36] and Gulf War illness [37]. Furthermore, the heterogeneity of NTE levels seen in patients with CFS/ME is consistent with a role for OPs in only a subset of patients with CFS/ME. In this study, 3 of 28 patients with CFS/ME for whom NTE data passed quality control analysis had high NTE levels, compared with uniformly low levels in healthy blood donors. This is in contrast to data from our pilot study, in which 10 of 17 patients with CFS/ME showed high NTE levels [13]. The discrepancy between the studies is interesting, given that in our pilot study all patients with individual PCR results for NTE were from Dorset, whereas in the present study 11 patients were from Dorset; 1 was from Leicester, United Kingdom; 5 were from Bristol; 5 were from London; and 6 were from New York.

Another interesting gene that affects neurological function is *EIF2B4,* mutations within which are associated with vanishing white matter disease [38]. This is interesting in view of the white matter hyperintensity lesions that have been documented in patients with CFS/ME [27–29].

We have identified involvement of several genes of the interleukin-6 (IL-6) signaling pathway, consistent with the findings of several previous studies [39, 40]. Although we found upregulation of the genes encoding IL-6R and IL-6ST, the 2 IL-6 receptors, we did not find evidence for upregulation of the gene encoding IL-6 itself, which may explain the inconsistent findings by different groups in this area.

It is intriguing that it is possible to identify CFS/ME subtypes on the basis of expression values for these 88 genes and even more so that these subtypes have distinct clinical phenotypes, with such marked differences in the occurrence of particular symptoms and their severity. It has been recognized for some time that subtypes of CFS/ME exist, and it has been thought that these subtypes may, at least in part, reflect particular etiological factors [41]. A symptom-based approach has had some success in identifying musculoskeletal, inflammatory, and neurological subtypes [42]; however, these groups had only minor differences between them in overall functional severity, in contrast to the subtypes in the present study. A more detailed analysis of clinical and molecular features of these subtypes will be presented in a separate article. Further work is urgently required to validate and develop these findings.

Acknowledgments

We thank Anne Faulkner and the CFS Research Foundation (Hertfordshire, United Kingdom), for generous funding of the project; Sir Joseph Hotung, for funding the salaries of J.R.K. and B.B; Dr. Frank Boulton, Ms. Julie Williams, Mr. Peter Rogers, Ms. Diana Carr, and the NBS teams of East Dorset, for their help in enrolling and collecting samples from healthy blood donors; Dr. Jon Sherlock (Applied Biosystems), for advice on TaqMan PCR; and all the patients with CFS/ME and blood donors for their participation. We also acknowledge the kind leadership of the late David A. J. Tyrrell [43], which was key in the early stages of this project.

References

1. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. Ann Intern Med **1994**; 121:953–9.

2. Department of Health. A report of the CFS/ME Working Group. January **2002**.
3. Papanicolaou DA, Amsterdam JD, Levine S, et al. Neuroendocrine aspects of chronic fatigue syndrome. Neuroimmunomodulation **2004**; 11: 65–74.
4. Komaroff AL, Buchwald D. Chronic fatigue syndrome: an update. Annu Rev Med **1998**; 49:1–13.
5. Devanur LD, Kerr JR. Chronic fatigue syndrome. J Clin Virol **2006**; 37: 139–50.
6. Vernon SD, Unger ER, Dimulescu IM, Rajeevan M, Reeves WC. Utility of the blood for gene expression profiling and biomarker discovery in chronic fatigue syndrome. Dis Markers **2002**; 18:193–9.
7. Powell R, Ren J, Lewith G, Barclay W, Holgate ST, Almond J. Identification of novel expressed sequences, up-regulated in the leucocytes of chronic fatigue syndrome patients. Clin Exp Allergy **2003**; 33:1450–6.
8. Whistler T, Unger ER, Nisenbaum R, Vernon SD. Integration of gene expression, clinical, and epidemiologic data to characterize chronic fatigue syndrome. J Transl Med **2003**; 1:10–8.
9. Whistler T, Jones JF, Unger ER, Vernon SD. Exercise responsive genes measured in peripheral blood of women with chronic fatigue syndrome and matched control subjects. BMC Physiol **2005**; 5:5.
10. Whistler T, Taylor R, Craddock RC, Broderick G, Klimas N, Unger ER. Gene expression correlates of unexplained fatigue. Pharmacogenomics **2006**; 7:395–405.
11. Grans H, Nilsson M, Dahlman-Wright K, Evengard B. Reduced levels of oestrogen receptor beta mRNA in Swedish patients with chronic fatigue syndrome. J Clin Pathol **2007**; 60:195–8.
12. Grans H, Nilsson P, Evengard B. Gene expression profiling in the chronic fatigue syndrome. J Intern Med **2005**; 258:388–90.
13. Kaushik N, Fear D, Richards SC, et al. Gene expression in peripheral blood mononuclear cells from patients with chronic fatigue syndrome. J Clin Pathol **2005**; 58:826–32.
14. Carmel L, Efroni S, White PD, Aslakson E, Vollmer-Conna U, Rajeevan MS. Gene expression profile of empirically delineated classes of unexplained chronic fatigue. Pharmacogenomics **2006**; 7:375–86.
15. Broderick G, Craddock RC, Whistler T, Taylor R, Klimas N, Unger ER. Identifying illness parameters in fatiguing syndromes using classical projection methods. Pharmacogenomics **2006**; 7:407–19.
16. Fang H, Xie Q, Boneva R, Fostel J, Perkins R, Tong W. Gene expression profile exploration of a large dataset on chronic fatigue syndrome. Pharmacogenomics **2006**; 7:429–40.
17. Fostel J, Boneva R, Lloyd A. Exploration of the gene expression correlates of chronic unexplained fatigue using factor analysis. Pharmacogenomics **2006**; 7:441–54.
18. Reeves WC, Lloyd A, Vernon SD, et al. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. International Chronic Fatigue Syndrome Study Group. BMC Health Serv Res **2003**; 3:25.
19. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. J Psychosom Res **1993**; 37:147–53.
20. Joyce E, Blumenthal S, Wessely S. Memory, attention, and executive function in chronic fatigue syndrome. J Neurol Neurosurg Psychiatry **1996**; 60:495–503.
21. Zheng Z, Wu J, Sun Z. An approach to identify over-represented cis-elements in related sequences. Nucleic Acids Res **2003**; 31:1995–2005.
22. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics **2002**; 18:207–8.
23. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A **1998**; 95:14863–8.
24. Kennedy G, Spence V, Underwood C, Belch JJ. Increased neutrophil apoptosis in chronic fatigue syndrome. J Clin Pathol **2004**; 57:891–3.
25. Levine PH, Fears TR, Cummings P, Hoover RN. Cancer and a fatiguing illness in Northern Nevada—a causal hypothesis. Ann Epidemiol **1998**; 8:245–9.
26. Levine PH, Atherton M, Fears T, Hoover R. An approach to studies of cancer subsequent to clusters of chronic fatigue syndrome: use of data from the Nevada State Cancer Registry. Clin Infect Dis **1994**; 18(Suppl 1):S49–53.
27. Lange G, DeLuca J, Maldjian JA, Lee H, Tiersky LA, Natelson BH. Brain MRI abnormalities exist in a subset of patients with chronic fatigue syndrome. J Neurol Sci **1999**; 171:3–7.
28. Greco A, Tannock C, Brostoff J, Costa DC. Brain MR in chronic fatigue syndrome. AJNR Am J Neuroradiol **1997**; 18:1265–9.
29. Schwartz RB, Garada BM, Komaroff AL, et al. Detection of intracranial abnormalities in patients with chronic fatigue syndrome: comparison of MR imaging and SPECT. AJR Am J Roentgenol **1994**; 162:935–41.
30. Glaser R, Padgett DA, Litsky ML, et al. Stress-associated changes in the steady-state expression of latent Epstein-Barr virus: implications for chronic fatigue syndrome and cancer. Brain Behav Immun **2005**; 19:91–103.
31. Birkenbach M, Josefsen K, Yalamanchili R, Lenoir G, Kieff E. Epstein-Barr virus–induced genes: first lymphocyte-specific G protein-coupled peptide receptors. J Virol **2003**; 67:2209–20.
32. Sairenji T, Yamanishi K, Tachibana Y, Bertoni G, Kurata T. Antibody responses to Epstein-Barr virus, human herpesvirus 6 and human herpesvirus 7 in patients with chronic fatigue syndrome. Intervirology **1995**; 38:269–73.
33. Chia JK. The role of enterovirus in chronic fatigue syndrome. J Clin Pathol **2005**; 58:1126–32.
34. Glynn P. NTE: one target protein for different toxic syndromes with distinct mechanisms? Bioessays **2003**; 25:742–5.
35. Glynn P. Neural development and neurodegeneration: two faces of neuropathy target esterase. Prog Neurobiol **2000**; 61:61–74.
36. Tahmaz N, Soutar A, Cherrie JW. Chronic fatigue and organophosphate pesticides in sheep farming: a retrospective study amongst people reporting to a UK pharmacovigilance scheme. Ann Occup Hyg **2003**; 47:261–7.
37. Proctor SP, Heaton KJ, Heeren T, White RF. Effects of sarin and cyclosarin exposure during the 1991 Gulf War on neurobehavioral functioning in US army veterans. Neurotoxicology **2006**; 27:931–9.
38. Leegwater PA, Pronk JC, van der Knaap MS. Leukoencephalopathy with vanishing white matter: from magnetic resonance imaging pattern to five genes. J Child Neurol **2003**; 18:639–45.
39. Chao CC, Janoff EN, Hu SX, et al. Altered cytokine release in peripheral blood mononuclear cell cultures from patients with the chronic fatigue syndrome. Cytokine **1991**; 3:292–8.
40. MacDonald KL, Osterholm MT, LeDell KH, et al. A case-control study to assess possible triggers and cofactors in chronic fatigue syndrome. Am J Med **1996**; 100:548–54.
41. Jason LA, Corradi K, Torres-Harding S, Taylor RR, King C. Chronic fatigue syndrome: the need for subtypes. Neuropsychol Rev **2005**; 15: 29–58.
42. Janal MN, Ciccone DS, Natelson BH. Sub-typing CFS patients on the basis of 'minor' symptoms. Biol Psychol **2006**; 73:124–31.
43. Kerr JR, Taylor-Robinson D. David Arthur John Tyrrell CBE: 19 June 1925–2 May 2005. Biogr Mem Fellows R Soc **2007**; 53:349–63.