Seven genomic subtypes of chronic fatigue syndrome/myalgic encephalomyelitis: a detailed analysis of gene networks and clinical phenotypes

J R Kerr, B Burke, R Petty, J Gough, D Fear, D L Mattey, J S Axford, A G Dalgleish and D J Nutt

J. Clin. Pathol. 2008;61;730-739; originally published online 5 Dec 2007; doi:10.1136/jcp.2007.053553

Seven genomic subtypes of chronic fatigue syndrome/myalgic encephalomyelitis: a detailed analysis of gene networks and clinical phenotypes

J R Kerr,[1,2] B Burke,[1] R Petty,[1] J Gough,[1,2] D Fear,[3] D L Mattey,[4] J S Axford,[1,2] A G Dalgleish,[1] D J Nutt[5]

[1] Department of Cellular and Molecular Medicine, St George's University of London, London, UK; [2] Sir Joseph Hotung Centre for Musculoskeletal Disorders, St George's Healthcare Trust and St George's University of London, London, UK; [3] Department of Asthma, Allergy and Respiratory Sciences, King's College London, London, UK; [4] Staffordshire Rheumatology Centre, Stoke on Trent, UK; [5] Psychopharmacology Unit, Department of Community Based Medicine, University of Bristol, Bristol, UK

Correspondence to: Dr J Kerr, St George's University of London, Cranmer Terrace, London SW17 0RE, UK; jkerr@sgul.ac.uk

BB and RP contributed equally to this work

Accepted 13 November 2007
Published Online First 5 December 2007

ABSTRACT

Aim: Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is a multisystem disease, the pathogenesis of which remains undetermined. The authors have recently reported a study of gene expression that identified differential expression of 88 human genes in patients with CFS/ME. Clustering of quantitative PCR (qPCR) data from patients with CFS/ME revealed seven distinct subtypes with distinct differences in Medical Outcomes Survey Short Form-36 scores, clinical phenotypes and severity.

Methods: In this study, for each CFS/ME subtype, those genes whose expression differed significantly from that of normal blood donors were identified, and then gene interactions, disease associations and molecular and cellular functions of those gene sets were determined. Genomic analysis was then related to clinical data for each CFS/ME subtype.

Results: Genomic analysis revealed some common (neurological, haematological, cancer) and some distinct (metabolic, endocrine, cardiovascular, immunological, inflammatory) disease associations among the subtypes. Subtypes 1, 2 and 7 were the most severe, and subtype 3 was the mildest. Clinical features of each subtype were as follows: subtype 1 (cognitive, musculoskeletal, sleep, anxiety/depression); subtype 2 (musculoskeletal, pain, anxiety/depression); subtype 3 (mild); subtype 4 (cognitive); subtype 5 (musculoskeletal, gastrointestinal); subtype 6 (postexertional); subtype 7 (pain, infectious, musculoskeletal, sleep, neurological, gastrointestinal, neurocognitive, anxiety/depression).

Conclusion: It was particularly interesting that in the seven genomically derived subtypes there were distinct clinical syndromes, and that those which were most severe were also those with anxiety/depression, as would be expected in a disease with a biological basis.

Chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) is a disease characterised by severe and debilitating fatigue, sleep abnormalities, impaired memory and concentration, and musculoskeletal pain.[1] In the western world, the population prevalence is estimated to be of the order of 0.5%.[2 3] Research studies have identified various features relevant to the pathogenesis of CFS/ME such as viral infection, immune abnormalities and immune activation, exposure to toxins, chemicals and pesticides, stress, hypotension, lymphocyte abnormalities and neuroendocrine dysfunction. However, the precise underlying disease mechanisms and means by which these abnormalities inter-relate in patients with CFS/ME remain to be clarified.[4 5]

We have previously described a study of gene expression in peripheral blood from 25 patients with CFS/ME, diagnosed according to the Centers for Disease Control (CDC) diagnostic criteria, and 50 normal blood donors using the Affymetrix U133+2 microarray. Genes showing differential expression were further analysed using quantitative PCR (qPCR) in 55 patients with CFS/ME and 75 normal blood donors. Differential expression was confirmed for 88 genes, 85 of which were upregulated and three downregulated. Highly represented functions were haematological disease and function, immunological disease and function, cancer, cell death, immune response and infection. Clustering of qPCR data from patients with CFS/ME revealed seven distinct subtypes with distinct differences in Medical Outcomes Survey Short Form-36 (SF-36) scores, clinical phenotypes and severity.[6]

In this study, we determined for each CFS subtype, the fold difference of each of the 88 CFS-associated genes compared with normal persons. Using a fold-difference cut-off of $\geqslant 1.5$, we then determined those genes that are differentially expressed in each CFS subtype. For each subtype, we report respective gene functions/pathways, gene interactions and disease associations, and relate these to the clinical phenotype details.

METHODS

Subjects and clinical characterisation

Analyses in this paper are based upon clinical and genomic data from patients with CFS/ME whose blood was used for qPCR confirmation of microarray data, as previously reported.[6] In total, 55 such patients were enrolled from clinics in Dorset, Bristol and London, UK, and New York City, USA (one patient from Leicester, UK, was managed by a clinic in London). These cases were diagnosed according to the Centers for Disease Control (CDC) diagnostic criteria for CFS/ME.[1] Patients with psychiatric disease were excluded using the Minnesota International Neuropsychiatric Interview, thus ensuring that none of our patients was suffering from major psychiatric disease or abuse of alcohol or other drugs. In addition, patients who smoked in the previous year, or were currently taking (or were within 3 months of taking) antibiotics, steroids or antidepressants were excluded from the study.

For all enrolled subjects, according to the recommendations of the International CFS Study Group,[7] severity of physical and mental fatigue was

assessed using the Chalder Fatigue Scale,[8] level of disability was assessed using the Medical Outcomes Survey SF-36, accompanying symptoms were characterised using the Somatic and Psychological Health Report, sleep abnormalities were assessed using the Pittsburgh sleep questionnaire, and assessment of type and severity of pain was performed using the McGill pain questionnaire. For the patients with CFS/ME, neurocognitive testing was performed using the spatial span (SSP) and verbal recognition memory (VRM) modules of the Cantab software (Cambridge Cognition, Cambridge, UK), which showed abnormal results in CFS/ME.[6 9]

For each patient with CFS, the severity of particular symptoms and level of function were taken from the questionnaires described above. Then for each CFS subtype, which was derived by clustering of qPCR data as previously described,[6] mean values for each symptom and score were calculated and compared between the subtypes. Analysis of variance (ANOVA) was used to determine the significance of differences in individual SF-36 domain scores between CFS subtypes.

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

Determination of differential expression of human genes in each CFS/ME subtype

The threshold cycle (Ct) for each test gene in each sample was compared with a calibrator sample to calculate a ΔCt value. ΔCt values were then normalised to the Ct value for an endogenous control gene, glyceraldehyde-3-phosphate dehydrogenase, in respective samples to give the ΔΔCt values. Relative quantity (RQ) values ($2^{-\Delta\Delta Ct}$) for each mRNA of interest were then calculated. Samples showing a difference between minimum and maximum RQ values of ⩾100 (indicating poor replicate concordance) were excluded. The t test was used to compare the RQ values for patients with CFS/ME with the RQ values of the controls. Genes whose mean RQ values differed between the groups (at $p \leqslant 0.05$) were included in our CFS/ME-associated gene signature.[6] RQ values for all 88 CFS/ME-associated genes were normalised and clustered using Genesis software.[10]

For each CFS subtype, mean RQ values were calculated. Then, for each gene, the mean RQ value for each CFS subtype was divided by the mean RQ value of the normal blood donors, to provide fold-difference values for each CFS subtype. For each subtype, genes were included for analysis assuming they showed fold-difference values (mean RQ in CFS subtype/mean RQ in normal) in qPCR experiments of ⩾1.5. Thus an individual gene list was generated for each CFS subtype within the 88-gene signature for CFS. Mean fold-difference values were clustered using Cluster version 2.11 software (without normalisation) and visualised using Treeview version 1.60 software.[11]

Analysis of gene function and interaction in each CFS/ME subtype

Each of these subtype-specific gene lists was analysed for gene function and interaction using Ingenuity Pathways Analysis (IPA) software (Ingenuity, Redwood City, California, USA) in order to link CFS/ME-subtype-associated genes into networks based on recognised interactions, and to discern the top associated diseases and disorders, molecular and cellular functions, associated physiological system development and function and canonical pathways.

RESULTS

Subjects and clinical characterisation

Clinical and genomic data from a total of 55 patients with CFS/ME fulfilling CDC diagnostic criteria were used for this study. Of these, 19 were male, and 36 were female, with an overall mean age of 41.6 years and a mean duration of disease of 3.2 years. Additional clinical details are provided elsewhere.[6] 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. Normal blood donors were used as a comparison group and clinical data for these are available elsewhere.[6]

Genomic CFS/ME subtypes

As previously reported, clustering of qPCR data revealed the presence of seven genomic CFS subtypes with distinct profiles of gene expression within the 88-gene CFS gene signature.[6] Fold-difference values (mean RQ in CFS/mean RQ in normal) for all patients with CFS and for each CFS subtype are shown in table 1. For each subtype, genes with fold-difference values of ⩾1.5 were noted and used in further analysis. This resulted in the following numbers of differentially-expressed genes in each subtype: 58 (CFS subtype 1), 70 (CFS subtype 2), 48 (CFS subtype 3), 27 (CFS subtype 4), 66 (CFS subtype 5), 69 (CFS subtype 6) and 71 (CFS subtype 7). In table 1, genes without values are those for which there were missing data for particular subtypes.

Analysis of gene function in each CFS/ME subtype

Using IPA software, the gene list for each CFS subtype was analysed to determine the most important associated diseases and disorders, molecular and cellular functions, associated physiological system development and function and canonical pathways. The results of this analysis are shown in table 2. As regards disease associations, analysis revealed some common (neurological, haematological, cancer) and some distinct (metabolic, endocrine, cardiovascular, immunological, inflammatory) disease associations among the subtypes. This was also true for cellular and molecular functions, and physiological system development and function analyses (table 2). As regards the canonical pathways implicated in each subtype, there was more variation between subtypes than for the previous analyses, probably because these assignments are based on fewer genes per pathway. Interleukin (IL)6 signalling was implicated in subtypes 1, 2, 5 and 6; B cell receptor signalling was implicated in subtypes 4 and 6; oestrogen receptor signalling was implicated in subtype 7; ephrin receptor signalling was implicated in subtypes 1, 2 and 7; and insulin receptor signalling was implicated in subtypes 3, 4 and 6 (table 2).

Analysis of gene interaction in each CFS/ME subtype

Gene interaction was assessed for each subtype using IPA software. For each subtype, this analysis generated between two and five large networks (arbitrarily defined as containing eight or more CFS-associated genes) based on published gene interactions (data not shown) and a variable number of smaller networks and single genes for which interactions were not known. For each subtype, all networks, large and small, were

Table 1 Fold-difference values for 88 genes in 55 patients with CFS: as a group, and in each of seven CFS subtypes

Gene symbol	mRNA accession no.	Taqman PCR assay ID‡	PCR*	Two-tail p value for PCR*	Subtype 1	2	3	4	5	6	7
ABCD4*	NM_020323	Hs00245340_m1	2.08	0.028	2.04		1.66	1.46		4.48	2.13
ACTR3	NM_005721	Hs00828586_m1	1.42	0.0042	3.77	2.44	1.93	0.93	1.86	0.87	1.76
AKAP10	NM_007202	Hs00183673_m1	1.54	0.0011	2.47	1.96	1.12	0.99	2.27	2.08	1.93
ANAPC11*	NM_016476	Hs00212858_m1	3.32	0.00033	4.33		2.26	2.02	5.57	3.27	4.26
ANAPC5	NM_016237	Hs00212120_m1	2.36	0.00016	3.88	3.64		1.69	3.56	3.88	0.83
APP†	NM_201413	Hs00169098_m1	2.5	4.33×10^{-9}	2.45†	3.11†	1.73†	1.77†	3.94†	2.87†	4.16†
ARL4C	NM_005737	Hs00255039_s1	2.96	8.90×10^{-6}			3.50	1.83	4.61	3.89	3.27
ARPC5	NM_005717	Hs00271722_m1	3.23	6.82×10^{-8}	9.03	5.54	5.27	2.25	3.59	1.72	5.81
ARSD	NM_001669	Hs00534692_m1	1.98	0.001		2.30	2.08	1.27		1.85	2.67
ATP6V1C1	NM_001695	Hs00184625_m1	2.03	0.00029	3.29	3.06	4.74	1.18	2.68	2.47	1.94
BCOR	NM_017745	Hs00372369_m1	1.6	0.0098	1.43	1.53	0.84	0.95	3.66	1.96	1.74
BMP2K	NM_198892	Hs00214079_m1	1.3	0.014	0.77	1.41	1.17	0.99	1.48	1.54	0.85
BRMS1*	NM_015399	Hs00363036_m1	2.68	0.0014		3.16	1.00	1.37		4.35	2.57
CD2BP2*	NM_006110	Hs00272036_m1	1.8	5.35×10^{-6}	2.19	1.76	2.37	1.34	1.63	2.78	3.45
CD47	NM_198793	Hs00179953_m1	2.2	0.00013	4.37	3.24	1.30	1.24	3.32	1.91	1.62
CEP350	NM_014810	Hs00402774_m1	2.02	0.0048	3.13	6.76	2.28	1.27	2.80	2.03	1.20
CITED2	NM_006079	Hs00366696_m1	2.39	4.45×10^{-6}		2.64	0.75	2.16		2.87	2.23
CMTM6	NM_017801	Hs00215083_m1	1.41	0.012	3.84	2.48	2.38	0.84	1.99	0.89	1.63
CREBBP	NM_004380	Hs00231733_m1	1.43	0.016			0.93	1.29	1.50	1.21	2.68
CRK	NM_016823	Hs00180418_m1	2.51	1.11×10^{-5}	3.65	4.39		1.79	2.31	3.63	3.59
CTBP1	NM_001328	Hs00179922_m1	1.45	0.062			0.86	0.94		1.35	2.49
CXCR4†	NM_003467	Hs00607978_s1	1.67	7.80×10^{-5}			1.12	1.15		2.18†	2.41†
EBI2	NM_004951	Hs00270639_s1	3.44	0.0012		11.68	2.91	1.47	6.18	4.89	4.38
EGR1	NM_001955	Hs00152928_m1	2.82	0.015	1.18	7.71	0.94	1.73	2.01	2.97	1.92
EGR3	NM_004421	Hs00231780_m1	1.92	0.017	1.14		1.28	2.60	3.46		1.61
EIF2B4*	NM_172195	Hs00248984_m1	2.06	0.025			3.68	1.55		1.10	2.86
EIF3S10	NM_003750	Hs00186707_m1	3.58	0.0029			1.98	3.78	5.64	2.60	5.78
EIF4G1*	NM_198241	Hs00191933_m1	3.05	0.0033	3.25	8.57	1.50	1.83		3.71	5.84
EIF4G3	NM_003760	Hs00186804_m1	1.67	1.37×10^{-5}		3.07	0.94	1.48	2.28	1.72	2.03
ETS1	NM_005238	Hs00901425_m1	2.11	1.00×10^{-5}	2.79	3.51	0.98	1.54	1.64	2.00	2.77
FAM126B	NM_173822	Hs00545158_m1	1.64	0.0034	4.64	1.80	2.23	0.95	2.79	1.35	1.01
FNTA†	NM_002027	Hs00357739_m1	2.18	3.82×10^{-6}	4.80†	3.32†	2.08†	1.49	2.48†	2.22†	1.49
GABARAPL1*	NM_031412	Hs00744468_s1	5.64	6.10×10^{-5}	12.43	13.97	8.31	2.49	5.02	4.75	4.37
GABPA	NM_002031	Hs00745591_s1	8.06	3.00×10^{-4}	6.59	23.28	15.56	2.92	3.36	9.27	3.44
GCN1L1	NM_006836	Hs00412445_m1	2.05	0.00052	2.03	2.00	1.54	1.38	2.41	4.26	3.22
GLTSCR2	NM_015710	Hs00414236_m1	1.24	0.026	1.97	1.85	1.14	0.77	1.30	1.41	2.20
GNAS	NM_080425	Hs00255603_m1	1.7	1.09×10^{-7}	2.14	3.16	1.73	1.27	1.67	1.46	1.87
GSN*	NM_198252	Hs00609276_m1	2.93	0.00017	4.39	7.07	3.51	1.42	3.79	2.36	5.90
GTF2A2	NM_004492	Hs00362112_m1	1.79	0.03		1.42	5.57	1.16		1.65	1.90
HIF1A	NM_001530	Hs00153153_m1	0.81	0.016		1.24	1.35	0.86		0.59	1.07
IFNAR1	NM_000629	Hs00265057_m1	1.76	0.00073	3.90	3.22	2.21	1.05	1.64	2.90	1.54
IL10RA*	NM_001558	Hs00387004_m1	1.73	9.87×10^{-6}	1.27	1.54	0.76	1.31	1.31	2.87	2.29
IL6R	NM_000565	Hs00794121_m1	1.19	0.06	5.42	3.11	3.32	2.94	3.10	3.03	4.08
IL6ST†	NM_002184	Hs00174360_m1	1.8	0.002	2.87†	1.52†	0.77	1.22	1.85†	1.71†	5.06†
IL7R	NM_002185	Hs00233682_m1	0.82	0.032	1.16	1.34	0.98	0.57	1.44	0.66	0.64
JAK1	NM_002227	Hs00233820_m1	1.91	1.86×10^{-8}	1.73	3.19	1.26	1.40	1.94	2.12	2.40
KHSRP*	NM_003685	Hs00269352_m1	1.67	0.00026	1.43	2.13	1.25	1.17	1.89	2.49	2.73
MAPK9	NM_139070	Hs00177102_m1	1.4	0.045	1.74	2.36		1.08	2.34	2.12	
METTL3	NM_019852	Hs00219820_m1	2.06	0.0001	3.94	2.95	0.76	1.38	1.97	3.14	2.11
MRPL23*	NM_021134	Hs00221699_m1	2.06	0.001	5.41	2.10	0.70	1.48	1.98	3.04	3.05
MRPS6	NM_032476	Hs00606808_m1	1.53	0.025		2.53	2.59	1.25	1.33	1.62	0.86
MRRF	NM_138777	Hs00751845_s1	8.91	0.0004	28.27	38.40	13.24	2.85	5.85	9.96	9.11
MSN*	NM_002444	Hs00792607_mH	1.33	0.0016	1.56	2.16	1.76	1.03	1.02	1.19	1.87
MTMR6	NM_004685	Hs00395064_m1	1.71	0.0025	5.07	2.10	2.38	1.23	2.17	0.59	1.89
NFKB1	NM_003998	Hs00231653_m1	1.59	4.04×10^{-5}		2.26	1.14	1.34		1.60	2.56
NHLH1	NM_005589	Hs00271582_s1	11.51	7.00×10^{-4}	3.87	47.19	20.19	5.34	3.90	15.62	5.32
NR1D2	NM_005126	Hs00233309_m1	2.44	0.00076	3.17	2.45	2.66	1.08		4.03	2.00
NTE*	NM_006702	Hs00198648_m1	1.7	0.04	4.40	8.11	0.97	0.94	1.55	1.33	2.24
NUFIP2	NM_020772	Hs00325168_m1	1.5	0.00036	2.05	1.94	0.74	1.02	2.44	1.97	1.45
PAPOLA	NM_032632	Hs00413685_m1	1.32	0.00194	1.90	1.93	1.41	0.88	1.99	1.03	1.40

Continued

Table 1 Continued

Gene symbol	mRNA accession no.	Taqman PCR assay ID‡	PCR*	Two-tail p value for PCR*	Subtype 1	2	3	4	5	6	7
PDCD2*	NM_002598	Hs00751277_sH	6.76	0.0096	26.46	24.28	17.55	2.84	3.78	8.81	5.38
PDCD6	NM_013232	Hs00737034_m1	1.74	0.00019	1.99	1.41	1.94	1.47	2.20	2.10	2.62
PEX16*	NM_004813	Hs00191337_m1	1.74	0.0034	2.58	2.46		1.15	2.08	2.13	2.33
PGM2	NM_018290	Hs00217619_m1	2.17	1.68×10^{-6}	4.35	3.49	2.76	1.28	3.04	2.26	2.41
PIK3R1	NM_181523	Hs00236128_m1	0.68	0.025			1.61	0.50	0.25	0.46	1.33
PKN1*	NM_213560	Hs00177028_m1	1.56	9.40×10^{-5}	2.30	2.11	0.95	1.14	1.97	1.73	2.60
POLR2G*	NM_002696	Hs00275738_m1	2.58	0.0078	3.38	3.04	2.79	1.23	2.56	8.07	2.63
PPP2R5C	NM_002719	Hs00604902_m1	1.38	0.022	2.59	1.63	1.03	0.95	2.15	1.85	1.14
PRKAA1	NM_006251	Hs01562315_m1	1.72	0.00052	3.27	1.98	1.43	1.36	2.23	1.68	0.97
PRKAR1A	NM_002734	Hs00267597_m1	2.63	2.91×10^{-8}	4.81	5.67	4.73	1.64	2.69	1.54	4.06
PUM2	NM_015317	Hs00209677_m1	1.39	0.00064			1.35	1.32	1.29	1.36	1.81
RAP2C	NM_021183	Hs00221801_m1	2.1	0.015	4.97	2.07		1.11	2.24	3.09	0.53
REPIN1	NM_013400	Hs00274221_s1	3.62	6.00×10^{-6}	3.86	7.29	2.66	1.67	1.93	4.76	2.19
RNF141	NM_16422	Hs00212656_m1	2.37	1.62×10^{-6}	3.79	2.86	3.16	1.85	3.56	2.26	1.02
SELENBP1	NM_003944	Hs00187625_m1	1.92	0.002	3.87	1.86	1.28		2.48	2.31	
SFXN1	NM_022754	Hs00224259_m1	1.6	0.022	4.86	1.89	1.57	1.21	3.50	0.96	1.41
SHPRH	NM_173082	Hs00542737_m1	1.77	0.05	0.92	2.85	0.98	2.21	1.71	1.23	
SNAP23	NM_003825	Hs00187075_m1	2.02	0.00018	5.60	1.82	3.11	1.32	3.37	2.02	1.99
SORL1	NM_003105	Hs00268342_m1	1.54	4.10×10^{-8}	1.31	1.89	1.48	1.26	1.51	1.74	3.09
SOS1	NM_005633	Hs00362308_m1	2.31	0.002	4.49	2.69	1.28	1.34	3.91	3.37	1.46
TAF11	NM_005643	Hs00194573_m1	1.87	0.05	2.32	4.97	0.86	1.23	2.71	1.93	2.60
TCF3	NM_003200	Hs00413032_m1	1.44	0.023			1.44	1.00		1.34	2.99
TDP1	NM_018319	Hs00217832_m1	1.67	0.0099	3.25	1.35	1.53	1.25	1.80	2.60	2.14
TNFRSF1A†	NM_001065	Hs00533560_m1	1.37	0.016		2.04†	1.13	1.32	0.72	1.48	1.46
UBTF	NM_014233	Hs00610729_g1	2.26	0.024		6.23	2.23	1.29		1.09	4.27
USP38	NM_032557	Hs00261419_m1	1.71	0.0021	4.07	2.08	0.96	1.10	2.03	2.05	2.07
WAPAL	NM_015045	Hs00386162_m1	1.69	0.027	3.27	3.64	3.35	1.07	2.00	1.16	0.91
WDR26	NM_025160	Hs00228535_m1	2.62	0.00012	5.62	6.11	2.84	1.80	2.33	1.20	3.19

For each subtype, genes with fold-difference values of ⩾1.5 up- or downregulated, as compared with normals, were noted and used in further analysis. Genes without values for the subtypes are those for which there was missing data for one or more subtypes.
*Kerr *et al* (2008).[6]
†Genes targeted by existing drugs, and those chronic fatigue syndrome subtypes in which there may be a rationale for a trial of a particular drug (see Results).
‡Assays were pre-designed by Applied Biosystems.

combined into a single network, indicating genes found to be upregulated and downregulated and then stratified to show the subcellular location of each (fig 1A–G).

Clinical features of each CFS subtype

Number of patients, mean age and male:female ratio for each subtype were as follows: subtype 1 (2, 27 years, 1:1), subtype 2 (5, 49 years, 4:1), subtype 3 (2, 32 years, 0:2), subtype 4 (19, 44.3 years, 8:11), subtype 5 (7, 51 years, 0:7), subtype 6 (14, 41.1 years, 6:8), subtype 7 (3, 47 years, 0:3). Subtypes 3, 5 and 7 were made up of females only, subtype 2 was predominantly male, and the remainder were mixed; age differences were less clearly demarcated.

Mean questionnaire scores for each subtype are shown in fig 2(A,B). Clinical symptom severity for each subtype is shown in fig 2(C). CFS subtypes 1 and 7 were the most severe, followed sequentially by subtypes 2, 4, 5 and 6/3. ANOVA testing revealed significant differences between groups for the SF-36 total score ($p = 0.016$), social functioning ($p = 0.03$), and emotional role ($p = 0.003$), while the difference between groups approached significance for general health ($p = 0.08$) and mental health ($p = 0.08$). After adjusting for multiple comparisons, significant associations were found between specific groups and clinical phenotypes. Subtype 7 had most pain, lowest SF-36 scores (along with subtype 1), and most severe individual symptoms including swollen glands, sore throat and headaches. Subtype 1 had the worst cognition and mental health score, and poor sleep despite having 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. Subtype 2 had marked postexertional malaise, muscle pain and joint pain but poor mental health (fig 2A–C).

Summary clinical features of each subtype were as follows: subtype 1 (cognitive, musculoskeletal, sleep, anxiety/depression); subtype 2 (musculoskeletal, pain, anxiety/depression); subtype 3 (mild); subtype 4 (cognitive); subtype 5 (musculoskeletal, gastrointestinal); subtype 6 (postexertional); subtype 7 (pain, infectious, musculoskeletal, sleep, neurological, gastrointestinal, neurocognitive, anxiety/depression). It is particularly interesting that in these genomically derived subtypes, there were distinct clinical syndromes and that those that were most severe were also those with anxiety/depression, as would be expected in a disease with a biological basis.

As regards subtype associations with geographical location, subtypes 4 and 6 were predominant in Dorset, subtype 4 was predominant in London and New York, and subtype 5 was predominant in Bristol (fig 2D).

Table 2 Disease associations, molecular and cellular functions, physiological systems and canonical pathway associations for differentially expressed genes in 55 patients with chronic fatigue syndrome

	Combined*	Subtype 1	Subtype 2	Subtype 3	Subtype 4	Subtype 5	Subtype 6	Subtype 7
Diseases	Haematological (22)	Neurological (18)	Neurological (21)	Neurological (16)	Neurological (11)	Neurological (21)	Neurological (25)	Neurological (24)
	Immunological (14)	Haematological (14)	Haematological (18)	Haematological (10)	Haematological (10)	Haematological (18)	Haematological (20)	Haematological (23)
	Cancer (31)	Cancer (21)	Inflammatory (11)	Cancer (15)	Immunological (7)	Immunological (13)	Immunological (17)	Immunological (17)
	Dermatological (3)	Cardiovascular (6)	Cancer (26)	Endocrine system (5)	Cancer (14)	Cancer (26)	Inflammatory (14)	Cancer (27)
	Endocrine system (9)	Inflammatory (6)	Immunological (15)	Metabolic (4)	Organismal injury (8)	Inflammatory (26)	Cancer (26)	Inflammatory (12)
Molecular and cellular functions	Cellular development (26)	Cell signalling (26)	Cell signalling (31)	Gene expression (16)	Cellular development (10)	Cell death (30)	Cell signalling (33)	Cell signalling (33)
	Cell death (33)	Cellular assembly and organisation (11)	Cellular growth and proliferation (24)	Cell death (15)	Cell death (13)	Cell death (25)	Cellular growth and proliferation (26)	Cell death (27)
	Gene expression (31)	Cellular compromise (6)	Gene expression (27)	Cell signalling (16)	Cell signalling (14)	Cellular movement (17)	Cellular development (20)	Cellular movement (20)
	Cellular growth and proliferation (31)	Cell death (20)	Cellular movement (18)	Cellular movement (10)	Cellular growth and proliferation (12)	Gene expression (25)	Gene expression (27)	Cellular development (21)
	Cellular assembly and organisation (15)	Cellular growth and proliferation (18)	Cell death (25)	Protein synthesis (6)	Gene expression (10)	Cellular development (18)	Cell death (25)	Gene expression (27)
Physiological system development and function	Haematological system development and function (22)	Nervous system development and function (15)	Immune response (17)	Haematological system development and function (7)	Haematological system development and function (8)	Organismal survival (13)	Haematological system development and function (17)	Tissue morphology (18)
	Immune and lymphatic system development and function (18)	Organismal survival (9)	Organismal survival (12)	Immune response (8)	Tissue morphology (10)	Tissue morphology (14)	Immune and lymphatic system development and function (16)	Haematological system development and function (16)
	Tissue morphology (18)	Haematological system development and function (11)	Haematological system development and function (15)	Organismal survival (8)	Organismal survival (9)	Haematological system development and function (16)	Tissue morphology (16)	Immune and lymphatic system development and function (16)
	Organismal survival (17)	Immune response (13)	Tissue morphology (13)	Tissue morphology (9)	Immune and lymphatic system development and function (6)	Immune and lymphatic system development and function (13)	Organismal survival (15)	Organismal survival (16)
	Immune response (20)	Tissue morphology (9)	Nervous system development and function (15)	Nervous system development and function (12)	Immune response (8)	Immune response (17)	Immune response (19)	Immune response (19)
Top canonical pathways	Interferon signalling (2)	Actin cytoskeleton signalling (6)	IL6 signalling (5)	Axonal guidance signalling (5)	Insulin receptor signalling (3)	Acute-phase response signalling (6)	Acute-phase response signalling (6)	Ephrin receptor signalling (5)
	IL6 signalling (6)	Ephrin receptor signalling (5)	Actin cytoskeleton signalling (6)	Actin cytoskeleton signalling (4)	B cell receptor signalling (3)	IL6 signalling (5)	IL6 signalling (5)	Axonal guidance signalling (6)
	EGF signalling (3)	ERK/MAPK signalling (5)	Acute-phase response signalling (5)	Hepatic fibrosis/stellate cell activation (3)	Acute-phase response signalling (3)	ERK/MAPK signalling (6)	ERK/MAPK signalling (6)	Oestrogen receptor signalling (4)
	IL2 signalling (3)	IL6 signalling (4)	Ephrin receptor signalling (5)	Insulin receptor signalling (3)	ERK/MAPK signalling (3)	Axonal guidance signalling (7)	Insulin receptor signalling (5)	Actin cytoskeleton signalling (5)
	PPAR signalling (5)	Axonal guidance signalling (6)	ERK/MAPK signalling (5)	Acute-phase response signalling	Actin cytoskeleton signalling (3)	Actin cytoskeleton signalling (6)	B cell receptor signalling (5)	Hepatic fibrosis and stellate cell activation (4)

Results are presented for the group as a whole and for each of the seven genomically derived subtypes.
*Kerr *et al* (2008).[6]
IL, interleukin; ERK, extracellular signal regulated kinase; MAPK, mitogen-activated protein kinase; PPAR, peroxisome proliferator-activated receptor; EGF, epidermal growth factor.

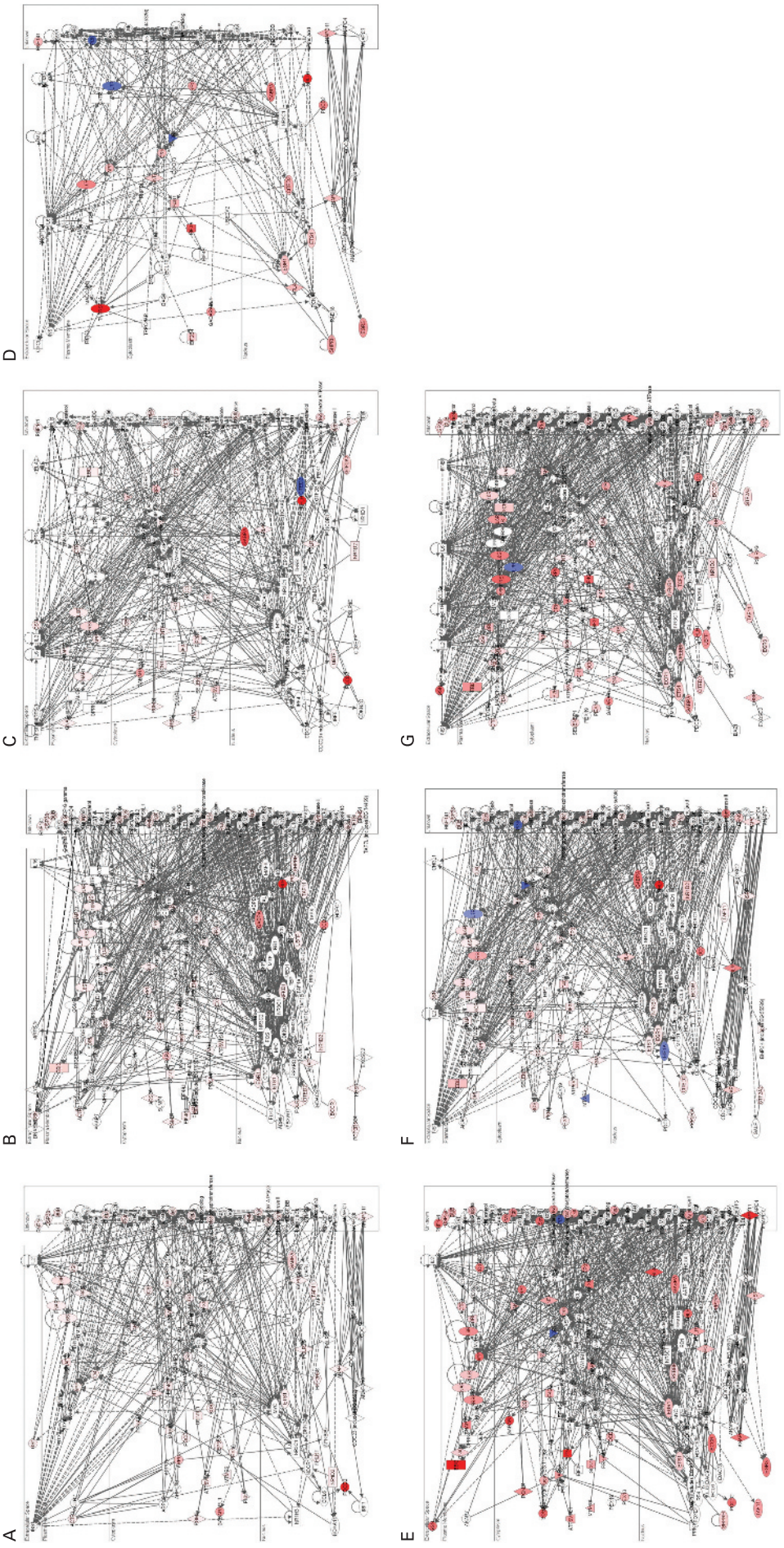

Figure 1 (A–G) Combined gene networks for chronic fatigue/myalgic encephalomyelitis (CFS/ME) subtypes 1–7, respectively. Networks are stratified by subcellular location. Genes are coloured according to expression levels determined in the present study: red, upregulation; blue; downregulation; colour intensity reflects the magnitude of the fold-difference between CFS/ME and normal subjects.[6] Networks were generated using Ingenuity Pathways Analysis.

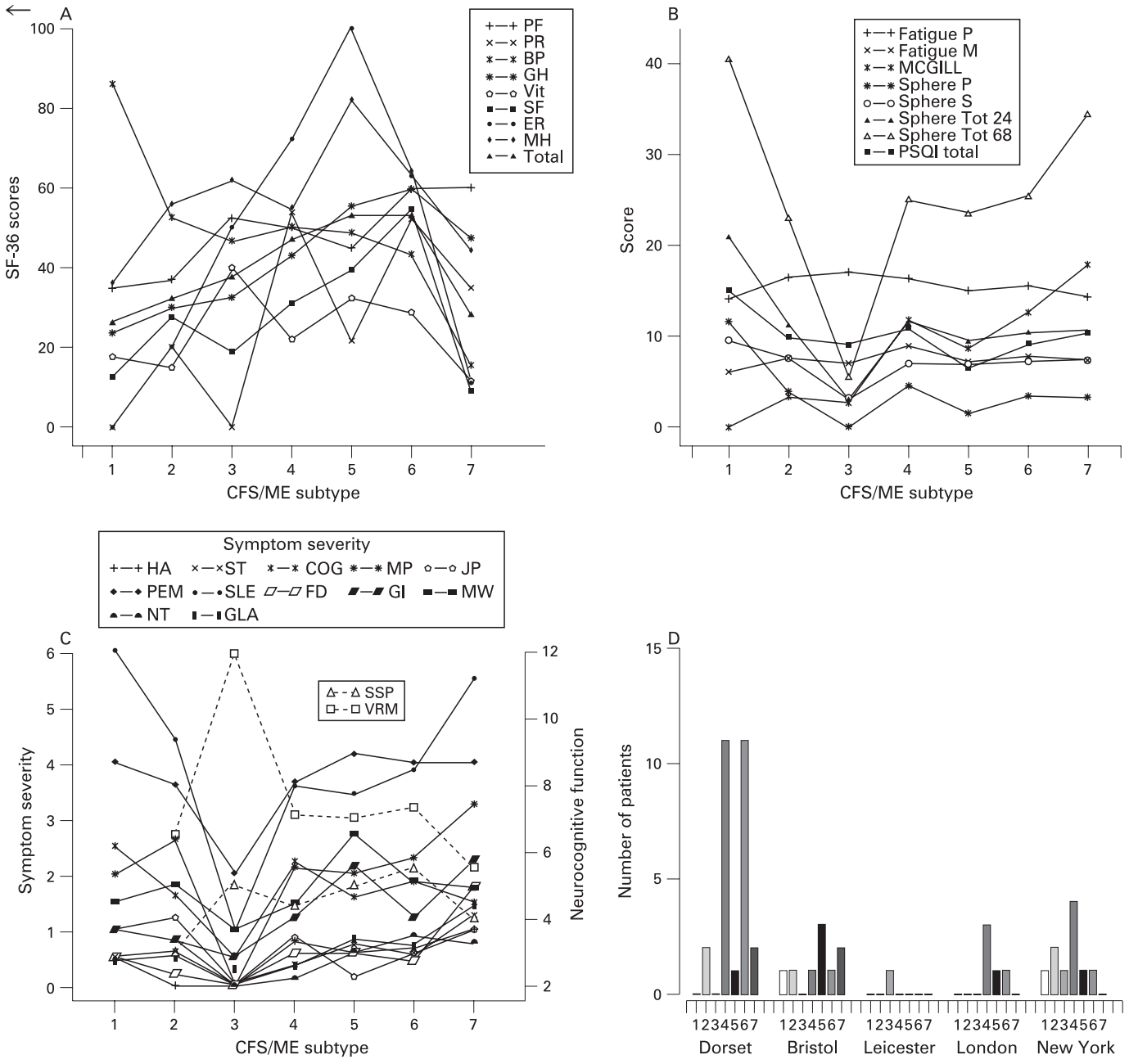

Figure 2 (A) Medical Outcomes Survey Short Form-36 domain and total scores for each chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) subtype; physical function, physical role (RP), bodily pain (BP), general health (GH), vitality (VIT), social functioning (SF), emotional role (RE), mental health (MH) and total score (Total). (B) Scores for the clinical questionnaires, Chalder Fatigue Scale (physical fatigue (P) and mental fatigue (M)), McGill pain questionnaire, Sphere (psychological (P), somatic (S), total/24, total/68), and Pittsburgh sleep quality index (PSQI). (C) Scores indicating occurrence and severity of 11 clinical symptoms and results of neurocognitive testing 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), verbal recognition memory (VRM). (D) Histogram showing the numbers of patients with CFS/ME of each subtype occurring in each of the five geographical locations. (A), (C) and (D) are reproduced from Kerr *et al*.[6]

CFS/ME associated genes that are specifically targeted by existing drugs

Within the CFS gene signature, there were five human genes that are known to be targeted by one or more existing drugs designed or intended for use in other diseases. Based on the expression levels of these five genes, these drugs may be predicted to be beneficial for particular CFS subtypes. These genes, corresponding drugs and CFS subtypes are as follows: APP (AAB-001; subtypes 1, 2, 3, 4, 5, 6 and 7); CXCR4 (JM1300; subtypes 5 and 6); FNTA (lonafarnib, tipifarnib; subtypes 1, 2, 3, 5 and 6); IL6ST (tocilizumab; subtypes 1, 2, 5, 6 and 7); TNF (golimumab, adalimumab, etanercept, certolizumab pegol, infliximab; subtype 2).

DISCUSSION

This study follows on from our paper describing differential expression of 88 human genes in patients with CFS[6] and its purpose is to expand upon the brief description of the genomic

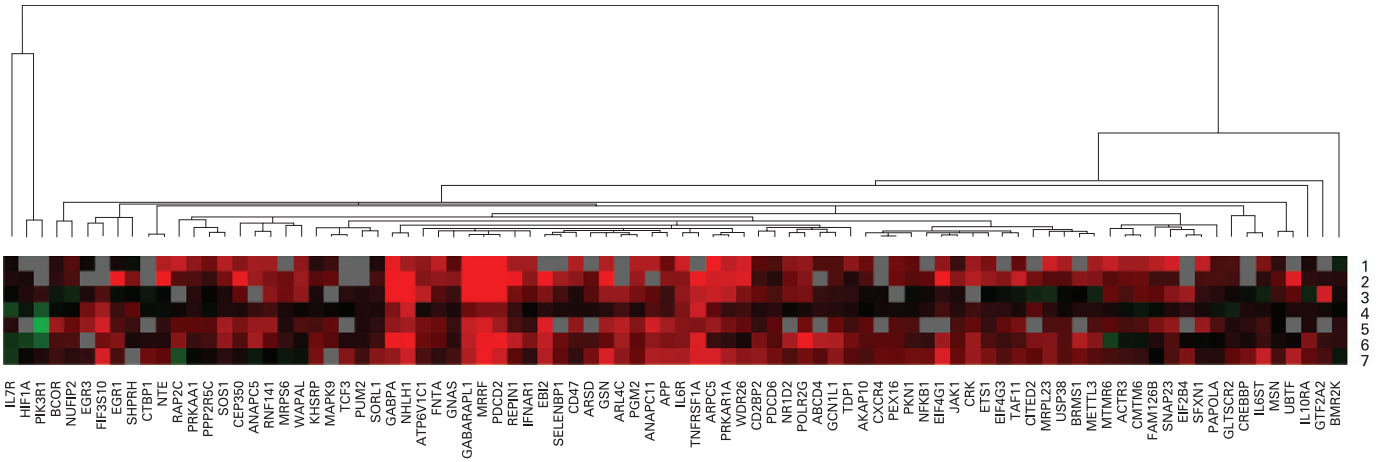

Figure 3 Clustering of logarithm (base 2) fold-difference values for patients of each of the subtypes compared with normal blood donors for 88 genes. Red, upregulation; green, downregulation. This figure is reproduced from Kerr *et al* (2008).[6]

and phenotypic aspects of the CFS subtypes given in the earlier paper.

It has long been recognised that subtypes of CFS/ME exist, and it has been believed that these subtypes may, at least in part, reflect particular aetiological factors.[12] A symptom-based approach has had some success in identifying musculoskeletal, inflammatory and neurological subtypes[13]; however, those groups had only minor differences in overall functional severity in contrast to those of the present study.

It is intriguing that within our 88-gene signature, there are several genes with links to various aetiological triggering factors. For example, virus infection (EIF4G1, EBI2) and organophosphate exposure (neuropathy target esterase (NTE)). EIF4G1 is an eukaryotic translation initiation factor that is bound and cleaved by a range of viruses, including enteroviruses, which both trigger and persistently infect patients with CFS.[14 15] Whistler and colleagues have also reported upregulation of EIF4G1 transcript variant 5 (the same variant as we report) in patients with CFS who have rapid (?triggered by virus infection) as compared with insidious onset.[16] EIF4G1 is a component of the protein complex EIF4F, which is crucial in translation.[17] These viruses divert EIF4G1 from its utilisation by the cellular machinery to facilitate production of viral proteins.[17] EIF4G1 is upregulated in CFS subtypes 1, 2, 3, 4, 6 and 7 (table 1; fig 3 and fig 1A,B,C,D,F,G).

Various CFS-associated genes identified have been previously shown to be upregulated in Epstein-Barr virus (EBV) infection, namely NFKB1, EGR1, ETS1, GABPA, CREBBP, CXCR4, EBI2, HIF1A, JAK1, IL6R, IL7R and PIK3R1. This is very interesting as EBV is a recognised trigger of CFS and is known to reactivate upon stress.[18] However, it is difficult to draw conclusions as to the inter-relationship of these genes in the different subtypes (fig 3). The EBV transcription factor BRLF1 was found to be over-represented in the original CFS gene signature; however, this was not tested by PCR.[6] 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 the BRLF1 gene has been identified as being over-represented in the transcription factor analysis, and that specific IgG to the Zebra protein (BZLF1 gene product) has been reported previously in patients with CFS/ME.[19]

EBI2 is a gene that is upregulated 200-fold in EBV-infected cells[20] and is upregulated in subtypes 2, 3, 5, 6 and 7, but not in the normal controls.[6] One subject with EBI2 upregulation was a 26-year-old female whose CFS had been triggered by laboratory-documented EBV infection, and who had a chronic course with detectable EBV replication in blood for several years after the acute phase. This suggests the possibility that EBI2 may be a surrogate marker for ongoing EBV replication in patients with CFS, although this remains to be clarified. If this is true, then this would be very useful to inform the decision as to which patients with CFS should be treated with valganciclovir, which has been shown to be beneficial in CFS.[21]

Three patients had markedly raised levels of NTE, while all normal controls had uniformly low levels; CFS subtypes with significantly raised NTE levels were 1, 2, 5 and 7, of which subtypes 1, 2 and 7 were the most severely affected subtypes. We have previously documented upregulation of NTE in CFS.[22] NTE is the primary site of action of organophosphate (OP) compounds such as sarin, which causes axonal degeneration and paralysis resulting from inactivation of its serine esterase activity,[23] and in the adult chicken nervous system, OP-modified NTE initiates neurodegeneration. Exposure to OP compounds may trigger CFS/ME[24] and Gulf War illness.[25]

IL10RA is a gene that is critical for T cell activation and immune system homeostasis, as polymorphisms in it have been shown to be associated with development of lymphoma, chronic obstructive pulmonary disease, autoimmunity, severity of hepatitis C infection, and multiple sclerosis.[26–30] In the present study IL10RA was upregulated,[6] although we have previously found it to be downregulated in patients with CFS.[22] There were more subjects in our pilot study[19] who were bed bound than in the present study,[6] and IL10RA levels appear to be a marker of severity in CFS (as they are closely correlated with SF-36 general health score), with lower levels reflecting increasing severity (data not shown).

It is interesting that disease associations identified in the various subtypes are mostly those that are already recognised in CFS. However, for any one disease association, there are important variations between the subtypes. For example, for "neurological disease", which applies to all subtypes, the number of genes in this category varies from subtype to subtype (table 2). Assuming that differential expression of these

Take-home messages

- Expression patterns of 88 chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) associated genes revealed seven distinct CFS/ME subtypes with distinct differences in clinical symptoms and severity.
- It was particularly interesting that those CFS/ME subtypes that were most severe were also those with anxiety/depression, as would be expected in a disease with a biological basis.

genes reflects, at least in part, the pathogenesis of CFS, the gene contribution to each disease association presumably affects the final phenotype and risk of complications (eg, lymphoma).[31 32] It is also interesting that these genomically derived subtypes represent distinct clinical syndromes, and that those that were most severe were also those with anxiety/depression, as would be expected in a disease with a biological basis.

Oestrogen receptor signalling is implicated in CFS subtype 7. Interestingly, it has previously been reported that patients with CFS exhibit a downregulation of oestrogen receptor β.[33 34] Oestrogen is an immunomodulator that has multiple effects on the immune system and on other hormones, which can themselves affect the immune response.[35]

Following repeat testing, and confirmation of these findings, it will be important to find a means by which we can determine the subtype of individual patients with CFS. For the purpose of subtype diagnosis, use of an 88-gene qPCR-derived signature is cumbersome and so it will be important to determine the most predictive genes within this signature, whose up- or down-regulation reliably predicts subtype status. Using this approach, and depending on further research, we may then be able to use a shortlist of 10–20 CFS-associated genes to subtype individual patients in clinical settings.

We believe these 88 genes to reflect real biological features of these patients with CFS, and this is supported by the fact that differential expression of 16 of these genes has been reported previously by our group.[22] If these findings are confirmed, there are various options for clinical trials using existing therapies that have been shown to be safe, based on targeting of key genes in patients of different CFS subtypes, namely IL6ST, TNF, CXCR4, APP and FNTA. Interestingly, one anti-tumour necrosis factor drug (etanercept) has already been trialled using an 8-week regimen in six patients with CFS, with reported clinical benefit in fatigue and pain in all subjects. Although this has not been published as a paper, the results were presented by Kristin Lamprecht and colleagues from Minnesota at the International Association for CFS meeting in Seattle in 2001.[36] Unfortunately, this was not followed up because the Peterson group moved out of CFS research around this time (P Peterson, personal communication).

In conclusion, we report in detail the genomic and phenotypic differences in seven genomically defined subtypes of CFS. Further work is required to validate these findings, and this work is underway in our laboratory.

Funding: We thank Sir Joseph Hotung for funding of the salary of JK, and the CFS Research Foundation, Hertfordshire, UK, for generous funding of this project.

Competing interests: None.

REFERENCES

1. **Fukuda K,** Straus SE, Hickie I, *et al*. 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. **CFS/ME Working Group**. A report of the CFS/ME Working Group. Department of Health, 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. **Kerr JR,** Petty R, Burke B, *et al*. Differentially expressed genes in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) patients reveal seven subtypes with distinct clinical phenotypes. *J Infect Dis* 2008;**197**:1–4.
7. **Reeves WC,** Lloyd A, Vernon SD, *et al*. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. *BMC Health Serv Res* 2003;**3**:25.
8. **Chalder T,** Berelowitz G, Pawlikowska T, *et al*. Development of a fatigue scale. *J Psychosom Res* 1993;**37**:147–53.
9. **Joyce E,** Blumenthal S, Wessely S. Memory, attention, and executive function in chronic fatigue syndrome. *J Neurol Neurosurg Psychiatry* 1996;**60**:495–503.
10. **Sturn A,** Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics* 2002;**18**:207–8.
11. **Eisen MB,** Spellman PT, Brown PO, *et al*. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;**95**:14863–8.
12. **Jason LA,** Corradi K, Torres-Harding S, *et al*. Chronic fatigue syndrome: the need for subtypes. *Neuropsychol Rev* 2005;**15**:29–58.
13. **Janal MN,** Ciccone DS, Natelson BH. Sub-typing CFS patients on the basis of "minor" symptoms. *Biol Psychol* 2006;**73**:124–31.
14. **Chia JK.** The role of enterovirus in chronic fatigue syndrome. *J Clin Pathol* 2005;**58**:1126–32.
15. **Chia JK,** Chia AY. Chronic fatigue syndrome is associated with chronic enterovirus infection of the stomach. *J Clin Pathol* 2008;**61**:43–8.
16. **Whistler T,** Unger ER, Nisenbaum R, *et al*. Integration of gene expression, clinical, and epidemiologic data to characterize chronic fatigue syndrome. *J Transl Med* 2003;**1**:10–8.
17. **Byrd MP,** Zamora M, Lloyd RE. Generation of multiple isoforms of eukaryotic translation initiation factor 4GI by use of alternate translation initiation codons. *Mol Cell Biol* 2002;**22**:4499–511.
18. **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.
19. **Sairenji T,** Yamanishi K, Tachibana Y, *et al*. Antibody responses to Epstein-Barr virus, human herpesvirus 6 and human herpesvirus 7 in patients with chronic fatigue syndrome. *Intervirology* 1995;**38**:269–73.
20. **Birkenbach M,** Josefsen K, Yalamanchili R, *et al*. Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors. *J Virol* 2003;**67**:2209–20.
21. **Kogelnik AM,** Loomis K, Hoegh-Petersen M, *et al*. Use of valganciclovir in patients with elevated antibody titers against human herpesvirus-6 (HHV-6) and Epstein-Barr virus (EBV) who were experiencing central nervous system dysfunction including long-standing fatigue. *J Clin Virol* 2006;**37**(Suppl 1):S33–8.
22. **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.
23. **Glynn P.** NTE: one target protein for different toxic syndromes with distinct mechanisms? *Bioessays* 2003;**25**:742–5.
24. **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.
25. **Proctor SP,** Heaton KJ, Heeren T, *et al*. Effects of sarin and cyclosarin exposure during the 1991 Gulf War on neurobehavioral functioning in US army veterans. *Neurotoxicology* 2006;**27**:931–9.
26. **He JQ,** Shumansky K, Zhang X, *et al*. Polymorphisms of interleukin-10 and its receptor and lung function in COPD. *Eur Respir J* 2007;**29**:1120–6.
27. **Nieters A,** Beckmann L, Deeg E, *et al*. Gene polymorphisms in Toll-like receptors, interleukin-10, and interleukin-10 receptor alpha and lymphoma risk. *Genes Immun* 2006;**7**:615–24.
28. **Qi ZM,** Wang J, Sun ZR, *et al*. Polymorphism of the mouse gene for the interleukin 10 receptor alpha chain Il10ra and its association with the autoimmune phenotype. *Immunogenetics* 2005;**57**:697–702.
29. **Saito T,** Ji G, Shinzawa H, *et al*. Genetic variations in humans associated with differences in the course of hepatitis C. *Biochem Biophys Res Commun* 2004;**317**:335–41.
30. **Heggarty S,** Sawcer S, Hawkins S, *et al*. A genome wide scan for association with multiple sclerosis in a N. Irish case control population. *J Neuroimmunol* 2003;**143**:93–6.
31. **Levine PH,** Fears TR, Cummings P, *et al*. Cancer and a fatiguing illness in Northern Nevada – a causal hypothesis. *Ann Epidemiol* 1998;**8**:245–9.

32. **Levine PH,** Atherton M, Fears T, *et al*. 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.
33. **Grans H,** Nilsson M, Dahlman-Wright K, *et al*. Reduced levels of oestrogen receptor beta mRNA in Swedish patients with chronic fatigue syndrome. *J Clin Pathol* 2007;**60**:195–8.
34. **Grans H,** Nilsson P, Evengard B. Gene expression profiling in the chronic fatigue syndrome. *J Intern Med* 2005;**258**:388–90.
35. **Lang TJ.** Estrogen as an immunomodulator. *Clin Immunol* 2004;**113**:224–30.
36. **Vallings R.** American Association for Chronic Fatigue Syndrome (AACFS), 5th international research, clinical and patient conference, 25–29 January 2001, Seattle, USA. http://www.cfs-news.org/aacfs-01.htm (accessed 6 February 2008).