STAT1 activation in association with *JAK2* exon 12 mutations

Anna L. Godfrey,[1,2,3] Edwin Chen,[1,3] Charles E. Massie,[1,3] Yvonne Silber,[1,3] Francesca Pagano,[1,3] Beatriz Bellosillo,[4] Paola Guglielmelli,[5] Claire N. Harrison,[6] John T. Reilly,[7] Frank Stegelmann,[8] Fontanet Bijou,[9] Eric Lippert,[10] Jean-Michel Boiron,[11] Konstanze Döhner,[8] Alessandro M. Vannucchi,[5] Carlos Besses,[4] and Anthony R. Green[1,2,3]

[1]Cambridge Institute for Medical Research and Wellcome Trust/MRC Stem Cell Institute, University of Cambridge, UK; [2]Department of Haematology, Addenbrooke's Hospital, UK; [3]Department of Haematology, University of Cambridge, Cambridge, UK; [4]Hematology Department, Hospital del Mar-IMIM, Barcelona, Spain; [5]Department of Experimental and Clinical Medicine, University of Florence and Azienda Ospedaliera Universitaria Careggi, Florence, Italy; [6]Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, UK; [7]Department of Haematology, Royal Hallamshire Hospital, Sheffield, UK; [8]University Hospital of Ulm, Department of Internal Medicine III, Germany; [9]Institut Bergonie, Department of Hematology, Bordeaux, France; [10]Laboratoire d'Hematologie, CHU de Bordeaux and Laboratorie Hematopoiese Leucemique et Cibles Thérapeutiques, INSERM U 1035, Universite Bordeaux Segalen, Bordeaux, France; [11]University of Bordeaux, Etablissement Français du Sang, France; [12] and Haematology Department, Hospital del Mar-IMIM, Barcelona, Spain

Correspondence: arg1000@cam.ac.uk

doi:10.3324/haematol.2015.128546

Supplementary information

Supplementary Table 1. Clinical details of patients used in this study

	***JAK2* exon 12 PV**	***JAK2V617F* ET**
Number of patients	13	7
Males, n (%)	5 (38)	3 (43)
Age at diagnosis, years (range)	38 (18-79)	63 (56-85)
Age at colony assay, years (range)	59 (33-83)	67 (60-85)
Hemoglobin at diagnosis, g/dl (range)	20.4 (17.2-22.8)	14.1 (12.4-15.7)
White blood cell count at diagnosis, x10^9/l (range)	7.5 (5.1-15.7)	9.9 (6.2-12.4)
Platelet count at diagnosis, x10^9/l	301 (90-597)	850 (541-1013)
Cytoreductive therapy at time of colony assay, n (%)	5 (38)	7 (100)

All patients met British Committee for Standards in Haematology diagnostic criteria[1, 2].
Of the *JAK2* exon 12-mutant PV patients, six were used for real-time PCR analyses and three for intracellular flow cytometry. Of the *JAK2V617F*-positive ET patients, six were used for real-time PCR analyses and two for intracellular flow cytometry.

Supplementary Table 2. Genes showing differential expression in *JAK2* exon 12-mutant and wild-type samples in microarray analysis

Cut-offs were <0.1 for q value and ≥1.25 for fold change. In the Name column, * denotes that at least 2 independent probes were identified for that gene in the microarray expression analysis. “+” in the V617F column denotes that the same gene was identified as being differentially expressed in a comparison of *JAK2V617F*-heterozygous and wild-type samples from 36 patients (20 PV, 16 ET; datasets published previously[3]). A statistical analysis comparing this dataset with the current microarray dataset was not performed, because of differences in the erythroid culture conditions and microarray platforms used for the two studies.

Genes upregulated in *JAK2* exon 12-mutant relative to wild-type samples

Symbol	Name	Fold change	q value	V617F
IFI27	interferon, alpha-inducible protein 27	2.72	0.046	+
CSAG3A	CSAG family, member 3	2.32	0.076	
FAM83A	family with sequence similarity 83, member A*	2.24	0.051	+
CSAG3B	CSAG family, member 2*	2.12	0.052	
IFI6	interferon, alpha-inducible protein 6	2.11	0.091	
WARS	tryptophanyl-tRNA synthetase	2.05	0.098	+
VAMP5	vesicle-associated membrane protein 5	1.94	0.078	
LOC646786	similar to Afadin (AF-6 protein) (withdrawn)	1.92	0.002	
HLA-B	major histocompatibility complex, class I, B	1.86	0.070	+
HLA-H	major histocompatibility complex, class I, H (pseudogene)*	1.86	0.041	
LOC653297	Similar to CSAG family, member 2 (withdrawn)	1.84	0.046	
HLA-F	major histocompatibility complex, class I, F	1.75	0.083	
VCL	vinculin*	1.71	0.046	+
HLA-E	major histocompatibility complex, class I, E	1.71	0.067	
IDO1	indoleamine 2,3-dioxygenase 1	1.65	0.098	
ESR2	estrogen receptor 2 (ER beta)*	1.65	0.041	+
C12ORF76	chromosome 12 open reading frame 76	1.63	0.098	
BST2	bone marrow stromal cell antigen 2	1.63	0.057	+
HCG4	HLA complex group 4 (non-protein coding)	1.57	0.078	
RAG1AP1	SLC50A1 (solute carrier family 50 (sugar transporter), member 1)	1.57	0.051	+
ADA	adenosine deaminase	1.57	0.067	
FBXO6	F-box protein 6	1.52	0.078	
STARD5	StAR-related lipid transfer (START) domain containing 5	1.47	0.051	
GDF15	growth differentiation factor 15	1.47	0.078	
CEBPB	CCAAT/enhancer binding protein (C/EBP), beta	1.46	0.098	+
HLA-G	major histocompatibility complex, class I, G	1.45	0.091	
MMD	monocyte to macrophage differentiation-associated	1.45	0.098	
STK19	serine/threonine kinase 19	1.44	0.098	+
GYPE	glycophorin E (MNS blood group)	1.44	0.098	
APOBEC3G	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G	1.41	0.098	

FCGR2A	Fc fragment of IgG, low affinity IIa, receptor	1.39	0.076	
ATP6V1B2	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B2	1.39	0.063	+
PARP3	poly (ADP-ribose) polymerase family, member 3	1.38	0.098	
CNOT10	CCR4-NOT transcription complex, subunit 10	1.38	0.078	
PLXNB2	plexin B2	1.36	0.065	
CASP1	caspase 1	1.36	0.089	
ITPRIP	inositol 1,4,5-trisphosphate receptor interacting protein	1.34	0.098	
MR1	major histocompatibility complex, class I-related	1.31	0.067	
CDC2L2	cyclin-dependent kinase 11A	1.31	0.090	+
LOC100132550	hypothetical protein LOC100132550	1.31	0.041	
TINF2	TERF1 (TRF1)-interacting nuclear factor 2	1.29	0.046	+
ZMAT5	zinc finger, matrin-type 5	1.29	0.098	
LMO4	LIM domain only 4	1.29	0.067	
MARCKSL1	MARCKS-like 1	1.29	0.097	
CISH	cytokine inducible SH2-containing protein	1.28	0.078	+
YIF1A	Yip1 interacting factor homolog A (S. cerevisiae)	1.27	0.098	+

Genes downregulated in *JAK2* exon 12-mutant relative to wild-type samples

Symbol	**Name**	**Fold change**	**q value**	**V617F**
TPSAB1	tryptase alpha/beta 1*	0.34	0.063	+
TPSB2	tryptase beta 2	0.37	0.067	
AKR1C2	aldo-keto reductase family 1, member C2	0.49	0.003	
PTRF	polymerase I and transcript release factor	0.51	0.041	
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog*	0.51	0.052	+
FAM129B	family with sequence similarity 129, member B*	0.52	0.041	
SH3PXD2A	SH3 and PX domains 2A	0.56	0.098	
SLC27A2	solute carrier family 27 (fatty acid transporter), member 2	0.58	0.052	+
HS.505676	SAP domain containing ribonucleoprotein	0.59	0.098	+
JUP	junction plakoglobin	0.60	0.098	
RNU1A3	RNA, U1 small nuclear 1	0.62	0.067	
RNU1-5	RNA, U1 small nuclear 5	0.62	0.092	
RNU1G2	RNA, U1 small nuclear 4	0.63	0.070	
TNS3	tensin 3	0.66	0.098	
SDPR	serum deprivation response	0.67	0.098	
TMEM51	transmembrane protein 51	0.68	0.046	
DCBLD2	discoidin, CUB and LCCL domain containing 2	0.72	0.046	
CD44	CD44 molecule (Indian blood group)	0.73	0.090	+
CLDN11	claudin 11	0.74	0.098	
ATP2B4	ATPase, Ca++ transporting, plasma membrane 4*	0.75	0.078	+

Supplementary Table 3. GSEA analysis of *JAK2* exon 12 mutant *vs* wild-type colonies

Analysis utilised gene sets from the MSigDB Chemical and Genetic Perturbations database[4]. FDR q-value cut-off was <0.1. NES = normalised enrichment score. "+" in the V617F column indicates that a gene set for exactly the same perturbation (e.g. IFN exposure) showed enrichment in a parallel analysis of *JAK2V617F*-heterozygous and wild-type samples from 36 patients (20 PV, 16 ET; datasets published previously[3]). A statistical analysis comparing this dataset with the current microarray dataset was not performed, because of differences in the erythroid culture conditions and microarray platforms used for the two studies.

Gene sets enriched in *JAK2* exon 12 mutant samples

Name	**Description**	**NES**	**Nom p-val**	**FDR q-val**	**V617F**
BENNETT_SLE_UP	Genes Significantly up-regulated in SLE patient Blood Mononuclear Cells	2.995	0.000	0.000	
CMV_8HRS_UP	Upregulated at 8hrs following infection of primary human foreskin fibroblasts with CMV	2.837	0.000	0.000	+
CMV_HCMV_TIMECOURSE_12HRS_UP	Up-regulated in fibroblasts following infection with human cytomegalovirus (at least 3-fold, with Affymetrix change call, in at least two consectutive timepoints), with maximum change at 12 hours	2.716	0.000	0.000	+
CMV-UV_HCMV_6HRS_UP	Up-regulated in fibroblasts at 6 hours following infection with UV-inactivated human cytomegalovirus	2.479	0.000	0.000	+
DER_IFNA_UP	Genes up-regulated by interferon-alpha in HT1080 (fibrosarcoma)	2.608	0.000	0.000	+
IFNA_HCMV_6HRS_UP	Up-regulated in fibroblasts at 6 hours following treatment with interferon-alpha	3.328	0.000	0.000	+
IFNA_UV-CMV_COMMON_HCMV_6HRS_UP	Up-regulated in fibroblasts at 6 hours following either infection with UV-inactivated CMV or interferon-alpha	2.927	0.000	0.000	+
IFNALPHA_HCC_UP	Upregulated by interferon alpha treatment in Hep3B hepatocellular carcinoma cells	2.558	0.000	0.000	+
IFNALPHA_NL_UP	Upregulated by interferon alpha treatment in normal primary hepatocytes	2.741	0.000	0.000	+
RADAEVA_IFNA_UP	Genes up-regulated by interferon-alpha in primary hepatocyte	3.054	0.000	0.000	+
SANA_IFNG_ENDOTHELIAL_UP	Genes up-regulated by interferon-gamma in colon,derm,iliac,aortic,lung endothelial cells	3.197	0.000	0.000	+
SANA_TNFA_ENDOTHELIAL_UP	Genes up-regulated by TNFA in colon,derm,iliac,aortic,lung endothelial cells	2.776	0.000	0.000	+
TAKEDA_NUP8_HOXA9_10D_UP	Effect of NUP98-HOXA9 on gene transcription at 10 d after transduction UP	2.622	0.000	0.000	
TAKEDA_NUP8_HOXA9_3D_UP	Effect of NUP98-HOXA9 on gene transcription at 3 d after transduction UP	2.791	0.000	0.000	
WIELAND_HEPATITIS_B_INDUCED	Genes induced in the liver during hepatitis B viral clearance in chimpanzees.	2.459	0.000	0.000	
TAKEDA_NUP8_HOXA9_8D_UP	Effect of NUP98-HOXA9 on gene transcription at 8 d after transduction UP	2.343	0.000	0.000	
IFNALPHA_NL_HCC_UP	Upregulated by interferon alpha treatment in both normal primary hepatocytes and Hep3B hepatocellular carcinoma cells	2.355	0.000	0.000	+
DER_IFNB_UP	Genes up-regulated by interferon-beta in HT1080 (fibrosarcoma)	2.365	0.000	0.000	+
IFN_ALPHA_UP	Upregulated 2-fold in HT1080 cells 6 hours following treatment with interferon alpha	2.394	0.000	0.000	+
UV-CMV_UNIQUE_HCMV_6HRS_UP	Up-regulated in fibroblasts at 6 hours following infection with UV-inactivated CMV, but not untreated CMV	2.418	0.000	0.000	+

ZHAN_MM_MOLECULAR_CLASSI_UP	123 genes differentiate CD-1 from CD-2 group upregulated	2.231	0.000	0.000	
BLEO_HUMAN_LYMPH_HIGH_4HRS_UP	Up-regulated at 4 hours following treatment of human lymphocytes (TK6) with a high dose of bleomycin	2.199	0.000	0.000	
REOVIRUS_HEK293_UP	Up-regulated at any timepoint up to 24 hours following infection of HEK293 cells with reovirus strain T3Abney	2.163	0.000	0.000	
DER_IFNG_UP	Genes up-regulated by interferon-gamma in HT1080 (fibrosarcoma)	2.155	0.000	0.000	+
SMITH_HTERT_UP	Genes upregulated by telomerase	2.135	0.000	0.001	
NF90_UP	Upregulated by ectopic expression of NF90 in GHOST(3)CXCR4 cells	2.137	0.000	0.001	
MMS_HUMAN_LYMPH_HIGH_24HRS_UP	Up-regulated at 24 hours following treatment of human lymphocytes (TK6) with a high dose of methyl methanesulfonate (MMS)	2.085	0.000	0.001	
PENG_LEUCINE_UP	Genes upregulated in response to leucine starvation	2.073	0.000	0.001	
GALE_FLT3ANDAPL_DN	The 19 probe sets differentially down-regulated in FLT3 ITD+ samples	2.016	0.000	0.003	+
RIBAVIRIN_RSV_UP	Up-regulation by infection of human pulmonary epithelian cells (A549) with respiratory syncytial virus (RSV) is enhanced by the anti-viral drug ribavirin	1.995	0.000	0.003	
BECKER_TAMOXIFEN_RESISTANT_UP	Genes that are reproducibly up-regulated in untreated or permanently tamoxifen-treated MaCa 3366/TAM (a tamoxifen-resistant, patient-derived mammary carcinoma xenograft) compared with untreated MaCa 3366 (a tamoxifen-sensitive, patient-derived mammary carcinoma xenograft). (fold change > 2)	1.978	0.002	0.004	
IFN_BETA_UP	Upregulated 2-fold in HT1080 cells 6 hours following treatment with interferon beta	1.968	0.000	0.004	+
HEMATOP_STEM_ALL_UP	Up-regulated in populations of human hematopoietic stem cells (CD34+/CD38-/Lin-) from bone marrow, umbilical cord blood, and peripheral blood stem-progenitor cells, compared to the stem cell-depleted population (CD34+/[CD38/Lin++])	1.972	0.000	0.004	
KANNAN_P53_UP	Target genes up regulated by p53	1.957	0.002	0.004	
ZHAN_MM_CD138_CD1_VS_REST	50 top ranked SAM-defined over-expressed genes in each subgroup_CD-1	1.958	0.000	0.004	
TAVOR_CEBP_DN	C/EBP down-regulated genes in KCL22 cells	1.949	0.004	0.005	
CMV_ALL_UP	Upregulated at any timepoint following infection of primary human foreskin fibroblasts with CMV	1.932	0.000	0.006	+
CMV_HCMV_6HRS_UP	Up-regulated in fibroblasts at 6 hours following infection with human cytomegalovirus (CMV)	1.916	0.000	0.006	+
HPV31_DN	Downregulated in normal human keratinocytes carrying episomal HPV31 DNA	1.900	0.002	0.007	
CANCERDRUGS_PROBCELL_UP	Up-regulated by at least two of four cancer drugs (cisplatin, camptothecin, methotrextate and/or paclitaxel) in pro-B cells (FL5.12)	1.872	0.004	0.010	+
AGUIRRE_PANCREAS_CHR8	Genes on chromosome 8 with copy-number-driven expression in pancreatic adenocarcinoma.	1.842	0.000	0.013	
IFN_ANY_UP	Upregulated 2-fold in HT1080 cells 6 hours following treatment with any of interferons alpha, beta and gamma	1.834	0.000	0.014	+
CMV_UV-CMV_COMMON_HCMV_6HRS_UP	Up-regulated in fibroblasts at 6 hours following infection with either human cytomegalovirus (CMV) or UV-inactivated CMV	1.828	0.002	0.014	+
TAKEDA_NUP8_HOXA9_16D_UP	Effect of NUP98-HOXA9 on gene transcription at 16 d after transduction UP	1.822	0.000	0.015	
TPA_SKIN_UP	Upregulated in murine dorsal skin cells 6 hours after treatment with the phorbol ester carcinogen TPA	1.820	0.000	0.015	
CAMPTOTHECIN_PROBCELL_UP	Up-regulated in pro-B cells (FL5.12) following treatment with camptothecin	1.779	0.012	0.022	+
CISPLATIN_PROBCELL_UP	Up-regulated in pro-B cells (FL5.12) following treatment with cisplatin	1.775	0.009	0.022	
IFN_GAMMA_UP	Upregulated 2-fold in HT1080 cells 6 hours following treatment with interferon gamma	1.767	0.006	0.023	+
VEGF_HUVEC_2HRS_UP	Up-regulated 2 hours after VEGF treatment in human umbilical vein endothelial cells	1.753	0.011	0.026	
CMV_HCMV_TIMECOURSE_ALL_UP	Up-regulated in fibroblasts following infection with human cytomegalovirus (at least 3-fold, with Affymetrix change call, in at least two consectutive timepoints)	1.744	0.000	0.027	+
BRENTANI_DEATH	Cancer related genes involved in cell death	1.699	0.005	0.039	
INNEREAR_UP	Genes prefentially expressed in human inner ear tissue (cochlea and vestibule), at least 10-fold higher from a mixture of 29 other tissues	1.694	0.018	0.040	
DSRNA_UP	Upregulated by dsRNA (polyI:C) in IFN-null GRE cells	1.689	0.009	0.040	
LEE_TCELLS4_UP	Transcripts enriched in more mature cells (SP4, CB4, and AB4) more than 3-fold, with average signal value differences of at least 100 between less mature (ITTP, DP) and more mature (SP4, CB4, and AB4) cells	1.683	0.002	0.042	

DAC_PANC_UP	Upregulated by DAC treatment in at least one of four pancreatic cancer cell lines, but not in normal (HPDE) cells	1.665	0.000	0.047	
SCHURINGA_STAT5A_UP	Differential Gene Expression in CB CD34 Cells Expressing STAT5A(1*6) Up-regulated	1.656	0.026	0.050	
PENG_GLUTAMINE_UP	Genes upregulated in response to glutamine starvation	1.653	0.000	0.050	
CMV_24HRS_UP	Upregulated at 24hrs following infection of primary human foreskin fibroblasts with CMV	1.644	0.012	0.052	+
GAY_YY1_UP	List of YY1 target genes identified in MEFs expressing ~25% of YY1 UP	1.626	0.002	0.059	+
ADIP_DIFF_CLUSTER1	Progressively downregulated over 24 hours during differentiation of 3T3-L1 fibroblasts into adipocytes (cluster 1)	1.618	0.024	0.059	+
ET743_HELA_UP	Upregulated by Et-743 in HeLa cells	1.619	0.007	0.060	+
PENG_RAPAMYCIN_UP	Genes upregulated in response to rapamycin starvation	1.620	0.000	0.060	
BRG1_SW13_UP	Up-regulated by transient expression of BRG1 at 24 hours in human, BRG1-lacking SW-13 cells	1.620	0.011	0.061	
LEE_MYC_E2F1_UP	Genes up-regulated in hepatoma tissue of Myc+E2f1 transgenic mice	1.593	0.025	0.067	
GENOTOXINS_ALL_24HRS_REG	Genes most consistently regulated in mouse lymphocytes (TK 3.7.2C) at 24 hours by all six genotoxins tested (cisplatin, methyl methanesulfonate, mitomycin C, taxol, hydroxyurea and etoposide)	1.596	0.028	0.067	+
BLEO_HUMAN_LYMPH_HIGH_24HRS_UP	Up-regulated at 24 hours following treatment of human lymphocytes (TK6) with a high dose of bleomycin	1.598	0.002	0.067	
UVC_TTD_4HR_UP	Up-regulated at 4 hours following treatment of XPB/TTD fibroblasts with 3 J/m^2 UVC	1.600	0.009	0.067	+
ZMPSTE24_KO_UP	Fifty genes most strongly up-regulated in liver tissue from mice deficient in the lamin-protease Zmpste24/Face1, versus wild-type controls	1.586	0.021	0.068	
CHEN_HOXA5_TARGETS_UP	The following tables are gene lists which included differentially expressed genes with at least two-fold changes (Sig log-average >1) in 4 out of 4 comparisons as described in the text Up-regulated	1.583	0.000	0.069	
GERY_CEBP_TARGETS	Complete list of differentially regulated C/EBP-target genes, sorted by P-value	1.587	0.005	0.069	
STRESS_ARSENIC_SPECIFIC_UP	Genes up-regulated 4 hours following arsenic treatment that discriminate arsenic from other stress agents	1.579	0.002	0.070	
FSH_OVARY_MCV152_UP	Up-regulated in ovarian epithelial cells (MCV152) 72 hours following FSH treatment, compared to untreated	1.568	0.011	0.075	+
UVC_TTD_ALL_UP	Up-regulated at any timepoint following treatment of XPB/TTD fibroblasts with 3 J/m^2 UVC	1.548	0.009	0.085	+
KRETZSCHMAR_IL6_DIFF	Shown are those probe sets that report at least a 15-fold expression change in response to IL-6 addition to INA-6 cells	1.528	0.002	0.098	+

Gene sets enriched in wild-type samples

Name	**Description**	**NES**	**Nom p-val**	**FDR q-val**	**V617F**
AGED_MOUSE_CORTEX_UP	Up-regulated in the cerebral cortex of aged (22 months) BALB/c mice, compared to young (2 months) controls	-2.195	0.000	0.001	+
IDX_TSA_UP_CLUSTER3	Strongly up-regulated at 16-24 hours during differentiation of 3T3-L1 fibroblasts into adipocytes with IDX (insulin, dexamethasone and isobutylxanthine), vs. fibroblasts treated with IDX + TSA to prevent differentiation (cluster 3)	-2.212	0.000	0.001	+
CHANG_SERUM_RESPONSE_UP	CSR (Serum Response) signature for activated genes (Stanford)	-2.030	0.000	0.009	+
TAKEDA_NUP8_HOXA9_10D_DN	Effect of NUP98-HOXA9 on gene transcription at 10 d after transduction Down	-2.020	0.000	0.010	
SERUM_FIBROBLAST_CORE_UP	Core group of genes consistently up-regulated following exposure to serum in a variety of human fibroblast cell lines (higher expression in activated cells, not cell-cycle dependent)	-2.041	0.000	0.010	+
ADIP_DIFF_CLUSTER4	Strongly upregulated at 16 hours during differentiation of 3T3-L1 fibroblasts into adipocytes (cluster 4)	-2.053	0.000	0.010	+
BRCA_PROGNOSIS_NEG	Genes whose expression is consistently negatively correlated with breast cancer outcomes - higher expression is associated with metastasis and poor prognosis	-1.960	0.000	0.012	+
UVC_HIGH_D4_DN	Progressively down-regulated through 12 hours following treatment of WS1 human skin fibroblasts with UVC at a high dose (50 J/m^2) (cluster d4)	-1.988	0.000	0.012	+
ZHAN_MULTIPLE_MYELOMA_VS_NORMAL_DN	The 50 most significantly down-regulated genes in MM in comparison with normal bone marrow PCs	-1.949	0.000	0.012	

PARK_MSCS_DIFF	List of differentially expressed genes between RhloLin2/loSca-11c-kit1 and RhhiLin2/loSca-11c-kit1 HSCs	-1.962	0.000	0.012	+
TAKEDA_NUP8_HOXA9_8D_DN	Effect of NUP98-HOXA9 on gene transcription at 8 d after transduction Down	-1.937	0.000	0.013	
SERUM_FIBROBLAST_CELLCYCLE	Cell-cycle dependent genes regulated following exposure to serum in a variety of human fibroblast cell lines	-1.950	0.000	0.013	+
NAKAJIMA_MCS_UP	Most increased transcripts in activated human and mouse MCs	-1.930	0.000	0.013	
P21_P53_ANY_DN	Down-regulated at any timepoint (4-24 hrs) following ectopic expression of p21 (CDKN1A) in OvCa cells, p53-dependent	-1.966	0.000	0.013	+
NAKAJIMA_MCSMBP_MAST	Top 50 most-increased mast cell specific transcripts	-1.989	0.000	0.014	
ADIP_VS_FIBRO_DN	Downregulated following 7-day differentiation of murine 3T3-L1 fibroblasts into adipocytes	-1.913	0.000	0.014	
KLEIN_PEL_UP	Genes downregulated in AIDS-related primary effusion lymphoma (PEL) cells compared to normal B cells and other tumor subtypes.	-1.969	0.000	0.015	
IGLESIAS_E2FMINUS_UP	Genes that increase in the absence of E2F1 and E2F2	-1.916	0.000	0.015	
TPA_SENS_LATE_UP	Upregulated by TPA at two consecutive timepoints between 12-48 hrs in sensitive HL-60 cells	-1.899	0.004	0.015	
HDACI_COLON_BUT12HRS_DN	Downregulated by butyrate at 12 hrs in SW260 colon carcinoma cells	-1.902	0.000	0.015	+
BREASTCA_TWO_CLASSES	Gene set that can be used to differentiate BRCA1-linked and BRCA2-linked breast cancers	-1.892	0.000	0.016	+
CREB_BRAIN_8WKS_UP	Up-regulated in the nucleus accumbens of mice after 8 weeks of induction of transgenic CREB	-1.855	0.000	0.026	
CANTHARIDIN_DN	Downregulated in HL-60 promyeloid leukemic cells after treatment with the cytotoxic drug cantharidin	-1.827	0.002	0.033	
ZHAN_MM_CD138_PR_VS_REST	50 top ranked SAM-defined over-expressed genes in each subgroup__PR	-1.817	0.002	0.034	+
DOX_RESIST_GASTRIC_UP	Upregulated in gastric cancer cell lines reistant to doxorubicin, compared to parent chemosensitive lines	-1.814	0.005	0.034	+
BECKER_TAMOXIFEN_RESISTANT_DN	Genes that are reproducibly down-regulated in untreated or permanently tamoxifen-treated MaCa 3366/TAM (a tamoxifen-resistant, patient-derived mammary carcinoma xenograft) compared with untreated MaCa 3366 (a tamoxifen-sensitive, patient-derived mammary carcinoma xenograft). (fold change > 2)	-1.820	0.000	0.035	
HDACI_COLON_BUT16HRS_DN	Downregulated by butyrate at 16 hrs in SW260 colon carcinoma cells	-1.805	0.000	0.035	+
SCHUMACHER_MYC_UP	Genes up-regulated by MYC in P493-6 (B-cell)	-1.810	0.000	0.035	+
RUTELLA_HEPATGFSNDCS_UP	Genes up-regulated by HGF treatments	-1.789	0.000	0.037	
OLDAGE_DN	Downregulated in fibroblasts from old individuals, compared to young	-1.795	0.000	0.037	+
POD1_KO_UP	Up-regulated in glomeruli isolated from Pod1 knockout mice, versus wild-type controls	-1.790	0.000	0.037	+
P21_P53_MIDDLE_DN	Down-regulated at intermediate timepoints (12-16 hrs) following ectopic expression of p21 (CDKN1A) in OvCa cells, p53-dependent	-1.776	0.004	0.040	+
BREASTCA_THREE_CLASSES	Gene set that can be used to differentiate BRCA1-linked, BRCA2-linked, and sporadic primary breast cancers	-1.758	0.007	0.044	
HIF1_TARGETS	Hif-1 (hypoxia-inducible factor 1) transcripional targets	-1.755	0.007	0.045	
MANALO_HYPOXIA_DN	Genes downregulated in human pulmonary endothelial cells under hypoxic conditions or after exposure to AdCA5, an adenovirus carrying constitutively active hypoxia-inducible factor 1 (HIF-1alpha).	-1.759	0.002	0.045	+
CMV_HCMV_TIMECOURSE_20HRS_DN	Down-regulated in fibroblasts following infection with human cytomegalovirus (at least 3-fold, with Affymetrix change call, in at least two consectutive timepoints), with maximum change at 20 hours	-1.737	0.009	0.045	+
JECHLINGER_EMT_DN	Genes downregulated for epithelial plasticity in tumor progression	-1.751	0.004	0.046	
RUIZ_TENASCIN_TARGETS	Tenascin-C target genes	-1.738	0.000	0.046	
MENSSEN_MYC_UP	Genes up-regulated by MYC in HUVEC (umbilical vein endothelial cell)	-1.762	0.008	0.046	+
HYPOXIA_REVIEW	Genes known to be induced by hypoxia	-1.760	0.004	0.046	
SCHRAETS_MLL_UP	Expression profile of Mll wild-type cells Top 40 list includes Genbank accession codes (GAC#), gene names (target gene), and log2 factors of differential gene expression	-1.727	0.006	0.046	
SANA_TNFA_ENDOTHELIAL_DN	Genes down-regulated by TNFA in colon,derm,iliac,aortic,lung endothelial cells	-1.744	0.002	0.046	+
EMT_DN	Down-regulated during the TGFbeta-induced epithelial-to-mesenchymal transition (EMT) of Ras-transformed mouse mammary epithelial (EpH4) cells (EMT is representative of late-stage tumor progression and metastasis)	-1.739	0.004	0.046	
KNUDSEN_PMNS_DN	Genes downregulated in PMNs upon migration to skin lesions	-1.728	0.000	0.047	+

HG_PROGERIA_DN	Downregulated in fibroblasts from patients with Hutchinson-Gilford progeria, compared to normal young individuals	-1.732	0.007	0.047	+
FLECHNER_KIDNEY_TRANSPLANT_REJECTION_DN	Genes downreglated in acute rejection transplanted kidney biopsies relative to well functioning transplanted kidney biopsies from stable, immunosuppressed recipients. (median FDR < 0.14% per comparison).	-1.746	0.000	0.047	+
INOS_ALL_DN	Downregulated following iNOS induction in hepatocytes (Tables 3-17)	-1.729	0.004	0.047	
E2F3_ONCOGENIC_SIGNATURE	Genes selected in supervised analyses to discriminate cells expressing E2F3 oncogene from control cells expressing GFP.	-1.740	0.002	0.047	+
BLEO_MOUSE_LYMPH_HIGH_24HRS_DN	Down-regulated at 24 hours following treatment of mouse lymphocytes (TK 3.7.2C) with a high dose of bleomycin	-1.711	0.007	0.050	
HDACI_COLON_BUT24HRS_DN	Downregulated by butyrate at 24 hrs in SW260 colon carcinoma cells	-1.712	0.002	0.050	+
HIPPOCAMPUS_DEVELOPMENT_POSTNATAL	Highly expressed in late postnatal mouse hippocampus (clusters 11 and 15)	-1.714	0.002	0.050	
UVB_NHEK3_C5	Regulated by UV-B light in normal human epidermal keratinocytes, cluster 5	-1.715	0.006	0.050	+
VANTVEER_BREAST_OUTCOME_GOOD_VS_POOR_DN	Poor prognosis marker genes in Breast Cancer (part of NKI-70) from Van't Veer et al 2002	-1.705	0.003	0.052	+
KNUDSEN_PMNS_UP	Genes up-regulated in PMNs upon migration to skin lesions	-1.697	0.000	0.052	
LE_MYELIN_UP	Genes upregulated in Egr2Lo/Lo mice (who bear mutations in the transcription factor Egr2 and in which peripheral nerve myelination is disrupted) whose expression is significantly altered after sciatic nerve injury.	-1.699	0.002	0.053	+
HADDAD_CD45CD7_PLUS_VS_MINUS_UP	Genes enriched in CD45RAhiCD7hi vs CD45RAintCD7- HPCs	-1.700	0.005	0.053	
SANA_IFNG_ENDOTHELIAL_DN	Genes down-regulated by interferon-gamma in colon,derm,iliac,aortic,lung endothelial cells	-1.692	0.003	0.053	+
IRITANI_ADPROX_DN	BEC-specific suppressed by AdProx-1	-1.693	0.011	0.053	
ZHAN_MM_MOLECULAR_CLASSI_DN	123 genes differentiate CD-1 from CD-2 group downregulated	-1.688	0.006	0.054	
CMV_IE86_UP	Upregulated by expression of cytomegalovirus IE86 protein in primary human fibroblasts	-1.682	0.002	0.056	+
PRMT5_KD_UP	Up-regulated by stable RNAi knock-down of PRMT5 in NIH 3T3 cells	-1.675	0.002	0.060	+
HADDAD_HSC_CD7_UP	Genes upregulated in human hematopoietic stem cells of the line CD45RA(hi) CD7+, which are biased toward developing into T lymphocytes or natural killer cells, versus CD45RA(int) CD7-.	-1.670	0.007	0.060	
TNFALPHA_4HRS_UP	Upregulated 4hrs after TNF-alpha treatment of HeLa cells	-1.672	0.011	0.060	
HTERT_UP	Upregulated in hTERT-immortalized fibroblasts vs. non-immortalized controls	-1.666	0.004	0.061	
ADIP_DIFF_CLUSTER5	Strongly upregulated at 24 hours during differentiation of 3T3-L1 fibroblasts into adipocytes (cluster 5)	-1.666	0.005	0.062	+
WANG_MLL_CBP_VS_GMP_DN	MLL-CBP alters the genetic programs of GMP The top 50 genes most closely associated with MLL-CBP expression Down-regulated	-1.660	0.011	0.063	+
ADIP_VS_PREADIP_DN	Downregulated in mature murine adipocytes (7 day differentiation) vs. preadipocytes (6 hr differentiation)	-1.647	0.009	0.070	
TPA_SENS_EARLY_UP	Upregulated by TPA at two consecutive timepoints between 15min-3hrs in sensitive HL-60 cells	-1.643	0.008	0.071	
TPA_SENS_MIDDLE_UP	Upregulated by TPA at two consecutive timepoints between 2-24hrs in sensitive HL-60 cells	-1.639	0.016	0.073	
TARTE_PLASMA_BLASTIC	Genes overexpressed in mature plasma cells isolated from tonsils (TPCs) and mature plasma cells isolated from bone marrow (BMPCs) as compared to polyclonal plasmablastic cells (PPCs).	-1.636	0.000	0.074	+
ZUCCHI_EPITHELIAL_UP	The 50 most upregulated genes in primary invasive breast dutcal carcinoma or metastatic breast carcinoma isolated from lymph nodes, as compared to normal mammary epithelium.	-1.633	0.009	0.074	
HDACI_COLON_SUL48HRS_DN	Downregulated by sulindac at 48 hrs in SW260 colon carcinoma cells	-1.628	0.009	0.076	+
VERNELL_PRB_CLSTR1	pRB pathway target genes CLUSTER 1 The listed genes were found regulated by pRB and p16 and one of the E2Fs (E2F1, E2F2, or E2F3) Cluster 1 genes are up-regulated by E2F and down-regulated by pRB and p16	-1.626	0.006	0.077	+
IDX_TSA_UP_CLUSTER4	Progressively up-regulated from 8-48 hours during differentiation of 3T3-L1 fibroblasts into adipocytes with IDX (insulin, dexamethasone and isobutylxanthine), vs. fibroblasts treated with IDX + TSA to prevent differentiation (cluster 4)	-1.623	0.014	0.078	+
TPA_RESIST_MIDDLE_UP	Upregulated by TPA at two consecutive timepoints between 2-24hrs in resistant HL-525 cells	-1.617	0.013	0.081	

CORDERO_KRAS_KD_VS_CONTROL_UP	Genes upregulated in kras knockdown vs control in a human cell line	-1.615	0.011	0.082	
TAKEDA_NUP8_HOXA9_6H_UP	Effect of NUP98-HOXA9 on gene transcription at 6 h after transfection UP	-1.612	0.024	0.082	
HESS_HOXAANMEIS1_UP	Genes upregulated in Hoxa9/Meis1 transduced cells vs control	-1.608	0.012	0.084	+
KENNY_WNT_UP	Genes up-regulated by Wnt in HC11 (mammary epithelial cells)	-1.593	0.010	0.092	
HESS_HOXAANMEIS1_DN	Genes downregulated in Hoxa9/Meis1 transduced cells vs control	-1.595	0.009	0.092	+
BLEO_MOUSE_LYMPH_LOW_24HRS_DN	Down-regulated at 24 hours following treatment of mouse lymphocytes (TK 3.7.2C) with a low dose of bleomycin	-1.588	0.026	0.095	
CARIES_PULP_UP	Up-regulated in pulpal tissue from extracted carious teeth (cavities), compared to tissue from extracted healthy teeth	-1.584	0.003	0.097	
ABRAHAM_MM_VS_AL_DN	Genes highly expressed in immunoglobulin light chain amyloidosis (AL) versus multiple myeloma (MM) in plasma cells.	-1.579	0.035	0.099	

Supplementary Methods

Patient recruitment

The study was approved by the Cambridge and Eastern Region Ethics Committee, patients gave written informed consent and research was carried out in accordance with the Declaration of Helsinki.

Patients were recruited from the following sites:

Addenbrooke's Hospital, Cambridge, UK
Guy's and St Thomas' Hospitals, London, UK
Royal Hallamshire Hospital, Sheffield, UK
Hospital del Mar-IMIM, Barcelona, Spain
Azienda Ospedaliera Universitaria Careggi, Florence, Italy
University Hospital of Ulm, Ulm, Germany
University Hospital of Bordeaux, Bordeaux, France

RNA microarrays and data analysis

Individual BFU-E colonies were resuspended in RLT lysis buffer (Qiagen, Maryland, USA) and a portion used for DNA extraction and genotyping by direct sequencing[5]. Pools of approximately 20 colonies of the same genotype underwent RNA extraction (Qiagen RNeasy kit, Qiagen, Maryland, USA). Quantification and RNA quality assessment (RIN values) were performed using an Agilent 2100 Bioanalyser (Agilent, Waldbronn, Germany). 50-100 ng total RNA were processed using the Illumina TotalPrep RNA amplification kit (Ambion, TX, USA), according to the manufacturer's instructions. Amplified cRNA was hybridised onto Illumina Human-12 v4.0 Expression BeadChips, chips were scanned using an Illumina BeadScanner, and fluorescence intensity data were extracted and quantified using Beadstudio.

Data were subjected to quality assessment, VST transformation and quantile normalisation using the lumi package in R. Datasets were filtered to include only probes that were expressed in over 50% of samples. Differential expression analyses were performed using the limma package in R[6], with genotype and patient code incorporated into the contrast matrices so that corresponding wild-type and mutant samples from each patient were analysed in a pairwise fashion. For identification of differentially expressed genes, probes with a q value <0.1 (false discovery rate <10%) and a fold change of at least 1.25 between the two genotypes were selected.

Gene set enrichment analysis[4] was performed to analyse enrichment of gene sets for chemical and genetic perturbations (CGP), using GSEA software (http://www.broadinstitute.org/gsea/index.jsp). Analyses were performed using a pre-ranked list of genes according to T values from the limma analysis, according to the developer's protocol. FDR cut-off of 10% (q<0.1) was used.

Microarray data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3946.

Real-time PCR

Eleven genes were selected for validation based on possible roles in hematopoiesis, cell cycle or other processes relevant to neoplasia: *ADA*[7], *CEBPB*[8], *ESR2*[9], *FAM83A*[10], *FBXO6*[11], *GDF15*[12], *LMO4*[13], *PARP3*[14], *STK19*[15], *TINF2*[16] and *KIT*[17].

cDNA was synthesised from pooled colony RNA using the SuperScript III First-Strand Synthesis System (Life Technologies, Paisley, UK). In order to analyse expression of a panel of genes from limited amounts of erythroid colony cDNA, the Fluidigm BioMark system[18] (Fluidigm, CA, USA) was used. All steps were performed according to the manufacturer's instructions. Intron-spanning TaqMan Gene Expression Assays (Applied Biosystems, CA, USA) were selected for each gene of interest (see table below). After 14-cycle specific target amplification (STA), a 48.48 Dynamic Array integrated fluidic circuit (IFC) chip (Fluidigm, CA, USA) was used for real-time PCR, with a BioMark HD Reader and BioMark Data Collection Software (Fluidigm, CA, USA). Reaction conditions were 2 minutes at 50°C, 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Data analysis was performed using Fluidigm Real-Time PCR Analysis software (Fluidigm, CA, USA). Expression levels of each test gene were normalised to levels of *HPRT1* as an internal control. Expression of genes in each *JAK2*-mutant sample, relative to expression in the respective wild-type sample, was calculated according to the ΔΔCt method. Each experiment was performed twice, including an independent STA reaction, and mean relative expression was calculated for each gene in each patient.

Gene	**Taqman gene expression assay**
ADA	Hs01110945_m1
CEBPB	Hs00270923_s1
ESR2	Hs01100353_m1
FAM83A	Hs00999384_m1
FBXO6	Hs00218350_m1
FOSL1	Hs04187685_m1
GDF15	Hs00171132_m1
HLA-E	Hs03045171_m1
HPRT1	Hs01003267_m1
IFI30	Hs00173838_m1
IFITM1	Hs01652522_g1
IRF1	Hs00971960_m1
IRF9	Hs00196051_m1
KIT	Hs00174029_m1
LMO4	Hs00232488_m1
PARP3	Hs00193946_m1
PSMB8	Hs00544757_g1
STK19	Hs00261086_m1
TAP1	Hs00388675_m1
TINF2	Hs01554305_g1

Flow cytometry using BFU-E colonies

For flow cytometry, colonies were resuspended in PBS for approximately 4 hours whilst a portion was used for genotyping by real-time PCR for *JAK2V617F*[5], or allele-specific PCR for *JAK2N542-E543del* using primers CTCCTCTTTGGAGCAATTCA (forward), CATCTAACACAAGGTTGGCATA (reverse, control) and TGACTTACAAATATCAAATCTCTGAT (reverse, mutation-specific).

Flow cytometry using erythroid colonies was performed as previously described[3], using either APC-conjugated anti-CD71 (BD Biosciences, NJ, USA) and PE-conjugated anti-GpA (Invitrogen, Paisley, UK) antibodies or Alexafluor 488-conjugated anti-pSTAT1 and PE-conjugated anti-pSTAT5 antibodies (BD Biosciences, NJ, USA).

Cell cultures, transfections and immunoblotting

JAK2-null gamma2A cells were cultured at 37°C and 5% CO_2 in DMEM (Sigma-Aldrich, Dorset, UK) with 10% fetal calf serum. Cells at approximately 50% confluency were transfected with expression constructs for *Jak2* wild-type, *V617F* and *K539L* using GeneJuice transfection reagent (Merck Millipore, Darmstadt, Germany). Western immunoblotting was performed on total cell lysates using the following antibodies: rabbit anti-JAK2, rabbit anti-p-Y701-STAT1, rabbit anti-STAT1 (all Cell Signaling, MA, USA); mouse anti-beta actin (Sigma-Aldrich, Dorset, UK); anti-mouse IgG-HRP and anti-rabbit IgG-HRP (Bio-Rad, CA, USA).

References

1. McMullin MF, Reilly JT, Campbell P, et al. Amendment to the guideline for diagnosis and investigation of polycythaemia/erythrocytosis. Br J Haematol. 2007;138(6):821-822.
2. Harrison CN, Bareford D, Butt N, et al. Guideline for investigation and management of adults and children presenting with a thrombocytosis. Br J Haematol. 2010;149(3):352-375.
3. Chen E, Beer PA, Godfrey AL, et al. Distinct clinical phenotypes associated with JAK2V617F reflect differential STAT1 signaling. Cancer Cell. 2010;18(5):524-535.
4. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-15550.
5. Godfrey AL, Chen E, Pagano F, et al. JAK2V617F homozygosity arises commonly and recurrently in PV and ET, but PV is characterized by expansion of a dominant homozygous subclone. Blood. 2012;120(13):2704-2707.
6. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology. 2004;Vol 3(No 1, Article 3.
7. Sokolic R, Maric I, Kesserwan C, et al. Myeloid dysplasia and bone marrow hypocellularity in adenosine deaminase-deficient severe combined immune deficiency. Blood. 2011;118(10):2688-2694.
8. Li H, Gade P, Xiao W, Kalvakolanu DV. The interferon signaling network and transcription factor C/EBP-beta. Cell Mol Immunol. 2007;4(6):407-418.
9. Schroeder C, Gibson L, Nordstrom C, Beug H. The estrogen receptor cooperates with the TGF alpha receptor (c-erbB) in regulation of chicken erythroid progenitor self-renewal. Embo J. 1993;12(3):951-960.
10. Macaulay IC, Tijssen MR, Thijssen-Timmer DC, et al. Comparative gene expression profiling of in vitro differentiated megakaryocytes and erythroblasts identifies novel activatory and inhibitory platelet membrane proteins. Blood. 2007;109(8):3260-3269.

11. Zhang YW, Brognard J, Coughlin C, et al. The F box protein Fbx6 regulates Chk1 stability and cellular sensitivity to replication stress. Mol Cell. 2009;35(4):442-453.

12. Ramirez JM, Schaad O, Durual S, et al. Growth differentiation factor 15 production is necessary for normal erythroid differentiation and is increased in refractory anaemia with ring-sideroblasts. Br J Haematol. 2009;144(2):251-262.

13. Meier N, Krpic S, Rodriguez P, et al. Novel binding partners of Ldb1 are required for haematopoietic development. Development. 2006;133(24):4913-4923.

14. Boehler C, Gauthier LR, Mortusewicz O, et al. Poly(ADP-ribose) polymerase 3 (PARP3), a newcomer in cellular response to DNA damage and mitotic progression. Proc Natl Acad Sci U S A. 2011;108(7):2783-2788.

15. Gomez-Escobar N, Chou CF, Lin WW, Hsieh SL, Campbell RD. The G11 gene located in the major histocompatibility complex encodes a novel nuclear serine/threonine protein kinase. J Biol Chem. 1998;273(47):30954-30960.

16. Kim SH, Kaminker P, Campisi J. TIN2, a new regulator of telomere length in human cells. Nat Genet. 1999;23(4):405-412.

17. Ratajczak MZ, Luger SM, DeRiel K, Abrahm J, Calabretta B, Gewirtz AM. Role of the KIT protooncogene in normal and malignant human hematopoiesis. Proc Natl Acad Sci U S A. 1992;89(5):1710-1714.

18. Spurgeon SL, Jones RC, Ramakrishnan R. High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. PLoS One. 2008;3(2):e1662.