MM	WES: 16 WGS: 23 WES/WGS: 1	*DIS3#, FAM46C, LRRK2, BRAF•†, IRF4†,* 11 NFκB pathway genes	[103]

Summarized information of whole-exome (WES) and whole-genome sequencing (WGS) studies of selected publications with highlighted novel mutated genes indicated. ccRCC = clear cell renal cell carcinoma, HNSCC = head-and-neck squamous cell carcinoma, HGS-OvCa = high grade serous ovarian carcinoma, HCC = hepatocellular carcinoma, AML = acute myeloid leukemia, MDS = myelodysplastic syndromes, CLL = chronic lymphocytic leukemia, DLBCL = diffuse large B cell lymphoma, MM = multiple myeloma, MSI = microsatelite instable, MSS = microsatelite stable. Predicted or reported consequences of mutations are shown: •hotspot identified, †activating, or likely activating #inactivating, or likely inactivating.

•hotspot identified

†activating, or likely activating

#inactivating, or likely inactivating

Supplementary FIG. 1 | Recurrent and significantly mutated genes identified from analysis of whole-exome and whole-genome sequencing studies of solid tumours. Recurrently mutated genes (Recurrent Mut.) and significantly mutated genes (SMGs) demonstrated to possess a higher mutation rate than the calculated background mutation rate (BMR) are shown for whole-exome/whole-genome sequencing (WES/WGS) studies for a number of solid tumour types where available. Approaches to identify SMGs (Sig. approaches) and False Discovery Rate (FDR), a statistical method used to correct for multiple hypothesis testing indicating the expected proportion of "false discoveries", are shown where available.

Supplementary FIG. 2 | Recurrent and significantly mutated genes identified from analysis of whole-exome and whole-genome sequencing studies of haematologic malignancies. Recurrently mutated genes (Recurrent Mut.) and significantly mutated genes (SMGs) demonstrated to possess a higher mutation rate than the calculated background mutation rate (BMR) are shown for whole-exome/whole-genome sequencing (WES/WGS) studies for a number of haematologic malignancies where available. Approaches to identify SMGs (Sig. approaches) and False Discovery Rate (FDR), a statistical method used to correct for multiple hypothesis testing indicating the expected proportion of "false discoveries", are shown where available.

Type of file: figure
Label: Supplementary Figure 1
Filename: Supplementary Fig 1.pdf

Glioblastoma (GBM)

Parsons *et al.*, 2008 (REF. 6)

SMG	Incidence
TP53	35%
EGFR	14%
PTEN	26%
NF1	15%
RB1	8%
IDH1	11%
PIK3CA	10%
PIK3R1	8%

Sig. Approach:
Described in:
Sjoblom *et al.*, 2006;
Parmigiani *et al.*, 2007

Paediatric GBM

Schwartzentruber *et al.*, 2012 (REF. 118)

Recurrent Mut.
TP53
ATRX
DAXX
H3F3A
NF1
IDH1
PDGFRA

Sig. Approach:
Knowledge based

Clear-cell Renal Cell Carcinoma (ccRCC)

Sato *et al.*, 2013 (REF. 55)

SMG
VHL
PBRM1
BAP1
TCEB1
SETD2
FPGT
MUDENG
KEAP1
TET2
MUC4
MLLT10
MSGN1
KRT32
M6PR
RPL14
GRB7
TP53
CSMD3
DNHD1
PIK3CA
NLRP12
VMO1
OR4C13
KCNMA1
LMAN2L
MTOR
ZNF536
YIPF3

Sig. Approach:
Unique to study. (Note: Calculated BMR taking into consideration DNA replication time as described in Hellman *et al.*, 2005; and Stamatoyannopoulos *et al.*, 2009)
(FDR<0. 05)

TCGA,2013 (REF. 19)

SMG
VHL
PBRM1
SETD2
KDM5C
PTEN
BAP1
MTOR
TP53
PIK3CA
MSR1
TXNIP
TCEB1
NFE2L2
BTNL3
SLITRK6
RHEB
ARID1A
NPNT
CCNB2

Sig. Approach:
Mutsig (FDR<0.1)

Head and Neck Squamous Cell Carcinoma (HNSCC)

Stransky *et al.*, 2011 (REF. 57)

SMG
TP53
CDKN2A
PRDM9
CASP8
FAT1
CSMD3
COL22A1
NOTCH1
PTEN
SYNE1
HRAS
RIMS2
USH2A
C19orf26
MAGEC1
CLSTN2
KCTD8
THSD7A
PIK3CA
OR4C15
ALLC
PRB4
CDH10
GRM8
NETO2
VN1R4
OR5L2
PPP1R9A
MAFA
LRFN5
MED1
ADCY8
ZFHX4
MLL2
SI
ZNF804B
LILRB1
PKHD1L1
CNTNAP2

Sig. Approach:
MutSig (FDR<0.1)

Agrawal *et al.*, 2011 (REF. 56)

Recurrent Mut.	Incidence
TP53	47%
NOTCH1	15%
CDKN2A	9%
PIK3CA	6%
FBXW7	5%
HRAS	4%

Sig. Approach:
Frequency

High-grade serous ovarian adenocarcinoma (HGS-OvCa)

TCGA, 2011 (REF. 25)

SMG
TP53
BRCA1
CSMD3
NF1
CDK12
FAT3
GABRA6
BRCA2
RB1

Sig. Approach:
MuSiC and MutSig (FDR <0.15)

Melanoma

Hodis *et al.*, 2012 (REF. 43)

SMG	Incidence
BRAF	63%
NRAS	26%
TP53	19%
PTEN	12%
P16INK4a	19%
PPP6C	9%
RAC1	5%
MAP2K1	5%
SNX31	7%
TACC1	7%
STK19	4%
ARID2	9%

Sig. Approach:
InVEx (FDR ≤0.2)

Krauthammer *et al.*, 2012 (REF. 62)

SMG
BRAF
NRAS
DCC
TNC
TP53
PTPRK
PPP6C
TLR4
CD163L1
GRM3
NPAP1
SLC15A2
RAC1
MAGEC1
JAKMIP2

Sig. Approach:
Modified protocol described in Ding *et al.*, 2008 (FDR <0.05)

Lung Cancer

Non-small-cell lung cancer (NSCLC)

Adenocarcinoma (AC)

Imielinski *et al.* 2012 (REF. 64)

SMG
TP53
KRAS
EGFR
STK11
KEAP1
ATM
NF1
SMARCA4
ARID1A
BRAF
RBM10
SETD2
PIK3CA
CBL
FBXW7
PPP2R1A
RB1
SMAD4
CTNNB1
U2AF1
KIAA0427
PTEN
BRD3
FGFR3
GOPC

Sig. Approach:
InVEx (FDR<0.25)

Squamous cell carcinoma (SCC)

TCGA, 2012 (REF. 26)

SMG	Incidence
TP53	81%
CDKN2A	15%
PTEN	8%
PIK3CA	16%
KEAP1	12%
MLL2	20%
HLA-A	3%
NFE2L2	15%
NOTCH1	8%
RB1	7%

Sig. Approach:
MutSig (FDR<0.1)

Small-cell lung cancer (SCLC)

Rudin *et al.*, 2012 (REF. 65)

SMG
TP53
COL22A1
RB1
ELAVL2
RASSF8
CNTNAP2
BCLAF1
GRM8
KIF21A
GRIK3
C17orf108
RUNX1T1
PLSCR4
CDYL
RIMS2
ZDBF2
KHSRP
SATB2
COL4A2
DIP2C
TMEM132D
ADCY1

Sig. Approach:
Described in Kan *et al.*, 2010
(FDR ≤0. 1)

Prostate Cancer

Barbieri *et al.*, 2012 (REF. 68)

SMG	Incidence
SPOP	13%
TP53	6%
PTEN	4%
FOXA1	4%
CDKN1B	3%
ZNF595	4%
THSD7B	5%
MED12	5%
NIPA2	3%
PIK3CA	4%
C14orf49	5%
SCN11A	5%

Sig. Approach:
MutSig (FDR<0.1)

Grasso *et al.* 2012 (REF. 69)

SMG	Incidence
TP53	40%
AR	9%
ZFHX3	14%
RB1	10%
PTEN	10%
APC	10%
MLL2	9%
OR5L1	3%
CDK12	5%

Sig. Approach:
MutSig (FDR ≤ 0.1)

Colorectal Cancer

TCGA, 2012 (REF. 22)

Hypermutated

SMG	Incidence
ACVR2A	63%
APC	51%
TGFBR2	51%
BRAF	46%
MSH3	40%
MSH6	40%
MYO1B	31%
TCF7L2	31%
CASP8	29%
CDC27	29%
FZD3	29%
MIER3	29%
TCERG1	29%
MAP7	26%
PTPN12	26%

Sig. Approach:
MutSig (FDR<0.1) and **manual curation** of expressed genes

Non-hypermutated

SMG	Incidence
APC	81%
TP53	60%
KRAS	43%
TTN	31%
PIK3CA	18%
FBXW7	11%
SMAD4	10%
NRAS	9%
TCF7L2	9%
FAM123B	7%
SMAD2	6%
CTNNB1	5%
KIAA1804	4%
SOX9	4%
ACVR1B	4%
GPC6	4%
EDNRB	3%

Sig. Approach:
MutSig (FDR<0.1) and **manual curation** of expressed genes

Seshagiri *et al.*, 2012 (REF. 70)

MSS Samples

SMG	Incidence
KRAS	55%
TP53	53%
APC	26%
PIK3CA	26%
SMAD4	25%
FBXW7	11%
CSMD1	19%
NRXN1	15%
TCF7L2	8%
DNAH5	15%
MRVI1	8%
ZNF208	8%
TRPS1	9%
OR10A7	6%
DMD	9%
KIF2B	8%
ATM	8%
FAM5C	8%
EVC2	9%
OR2W3	6%
IL13RA1	6%
TMPRSS11A	6%
C1orf170	4%

Sig. Approach:
Described in Kan *et al.*, 2010 (FDR<0. 0.05)

Gastric Cancer

Zang *et al.*, 2011 (REF. 35)

SMG
TP53
DBR1
RIT2
CCNL1
APC
ATF5
PKHD1
ADAM18
HTR1E
ARID1A
NFAT5
PPP1R1B
ZNF193
CERK
OR4C46
OR4C15
PIK3CA

Sig. Approach:
MuSiC and described in Greenman *et al.*, 2006 (FDR<0.2)

Wang *et al.*, 2011 (REF. 34)

SMG
TP53
PTEN
ARID1A
RPL22
TTK
FMN2
SPRR2B
PTN
ACVR2A
PMS2L3
DNAH7
TTN
FSCB
CTNNB1
SEMA3E
MCHR1
SPANXN2
METTL3
EIF3A
EPB41L3

Sig. Approach:
Described in Kan *et al.*, 2010 (FDR ≤ 0.2)

Breast Cancer

Ellis *et al.*, 2012 (REF. 73)

SMG
MAP3K1
PIK3CA
TP53
GATA3
CDH1
TBX3
ATR
RUNX1
ENSG00000212670
RB1
LDLRAP1
STMN2
MYH9
MLL3
CDKN1B
AGTR2
SF3B1
CBFB

Sig. Approach:
MuSiC (FDR<0.26)

Stephens et al. , 2012 (REF. 31)

Recurrent Mut.
TP53
PIK3CA
GATA3
CDH1
CDKN2A
MAP2K4
MAP3K1
NCOR1
PTEN
MLL3
AKT1
ARID1B
SF3B1
ARID1A
CASP8
NF1
RB1
TBX3
MAP3K13
AKT2
APC
ARID2
ASXL1
BAP1
BRCA1
BRCA2
CDKN1B
KRAS
MLL2
SETD2
SMAD4
SMARCD1
STK11

Sig. Approach:
Described in Greenman *et al.*, 2006 and prior biological knowledge

TCGA, 2012 (REF. 23)

SMG
TP53
PIK3CA
GATA3
MAP3K1
MLL3
CDH1
MAP2K4
RUNX1
PTEN
TBX3
PIK3R1
AKT1
CBFB
TBL1XR1
NCOR1
CTCF
ZFP36L1
GPS2
SF3B1
CDKN1B
USH2A
RPGR
RB1
AFF2
NF1
PTPN22
RYR2
PTPRD
OR6A2
HIST1H2BC
GPR32
CLEC19A
CCND3
SEPT13
DCAF4L2

Sig. Approach:
MuSiC (FDR<0.05)

Banerji *et al.*, 2012 (REF. 72)

SMG	Incidence
TP53	37%
PIK3CA	36%
AKT	6%
CBFB	4%
GATA3	4%
MAP3K1	3%

Sig. Approach:
MutSig (FDR<0.1)

Shah *et al.*, 2012 (REF. 75)

SMG
PIK3CA
RB1
TP53
PTEN
MYO3A
GH1

Sig. Approach:
Described in Youn *et al.*, 2011 (FDR q < 0.1)

Pancreatic Ductal Adenocarcinoma (PDAC)

Biankin *et al.*, 2012 (ICGC) (REF. 21)

SMG
KRAS
TP53
CDKN2A
SMAD4
MLL3
TGFBR2
ARID1A
ARID2
EPC1
ATM
SF3B1
ZIM2
MAP2K4
NALCN
SLC16A4
MAGEA6

Sig. Approach:
MuSiC (FDR<0.1)

O-viverrini-related Cholangiocarcinoma (CCA)

Ong *et al.*, 2012 (REF. 39)

SMG
TP53
SMAD4
KRAS
RNF43
NDC80
ROBO2
GNAS

Sig. Approach:
MuSiC (FDR<0.3)

Hepatocellular Carcinoma (HCC)

Li *et al.* ,2011 (REF. 37)

Recurrent Mut.
ARID2
CTNNB1
TP53
DMXL1
NLRP1

Sig. Approach:
Frequency

Huang *et al.*, 2012 (REF. 38)

SMG
TP53
ARID1A
SAM9DL
ARID2
C18orf34
AMPH
ELMO1
DSE
SPAG17
CSMD3

Sig. Approach:
Unique to study

Guichard *et al.*, 2002 (REF. 36)

Recurrent Mut.
CTNNB1
TP53
ARID1A
AXIN1
RPS6KA3
CDKN2A
NFE2L2
ARID2
IRF2
HNF1A
APC
IL6ST
PIK3CA
KRAS

Sig. Approach:
Frequency and copy number integration

Fujimoto *et al.*, 2012 (REF. 27)

SMG
TP53
ERRFI1
ZIC3
CTNNB1
GXYLT1
OTOP1
ALB
ATM
ZNF226
USP25
WWP1
IGSF10
ARID1A
UBR3
BAZ2B

Sig. Approach:
Unique to study (FDR≤0.05)

Type of file: figure
Label: Supplementary Figure 2
Filename: Supplementary Fig 2.pdf

Acute Myeloid Leukemia (AML)

TCGA, 2013 (REF. 33)

SMG
DNMT3A
NPM1
FLT3
TET2
RUNX1
IDH2
IDH1
CEBPA
TP53
NRAS
WT1
KIT
PTPN11
KRAS
U2AF1
SMC1A
SMC3
PHF6
FAM5C
STAG2
RAD21
EZH2
HNRNPK

Sig. Approach:
MuSiC (FDR < 0.05)

Myelodysplastic Syndromes (MDS)

Yoshida *et a*l., 2011 (REF. 89)

MDS/CMML/AML-MRC

Recurrent Mut.
TET2
U2AF1
ZRSR2
NRAS
SRSF2
ASXL1
BCOR
DNMT3A
EZH2
KRAS
RUNX1
TP53
ATRX
CBL
FLT3
IDH1
IDH2
SF3B1
STAG2

Sig. Approach:
Knowledge based

Chronic Lymphocytic Leukemia (CLL)

Wang *et al.*, 2011 (REF. 92)

SMG	Incidence
TP53	15%
SF3B1	15%
MYD88	10%
ATM	9%
FBXW7	4%
NOTCH1	4%
ZMYM3	4%
DDX3X	3%
MAPK1	3%

Sig. Approach:
MutSig (FDR≤0.1)

Puente *et al.*, 2011 (ICGC) (REF. 29)

	IGHV: unmutated	mutated
Recurrent Mut.	Incidence	Incidence
NOTCH1	20.4%	7%
MYD88	0.8%	5.6%
XPO1	4.6%	0%
KLHL6	0%	4.5%

Sig. Approach:
Frequency

Quesada *et al.*, 2012 (ICGC) (REF. 30)

	IGHV: unmutated	mutated
Recurrent Mut.	Incidence	Incidence
NOTCH1	2.8%	10.1%
SF3B1	7.9%	20.5%
POT1	0%	11.1%
CHD2	8.3%	0%
LRP1B	5.0%	4.4%

Sig. Approach:
Frequency

Fabbri *et al.*, 2011 (REF. 94)

Recurrent Mut.	Incidence
NOTCH1	15%
TP53	8%
PLEKHG5	4%
TGM7	4%
BIRC3	4%

Sig. Approach:
Frequency

Diffuse Large B Cell Lymphoma (DLBCL)

Lohr *et al.*, 2012 (REF. 99)

SMG	Incidence
PCLO	35%
PIM1	31%
MLL2	29%
TP53	24%
TNFRSF14	22%
GNA13	20%
CARD11	20%
MEF2B	18%
CD79B	16%
BTG1	16%
HIST1H1C	14%
EZH2	14%
TMSL3	12%
CD58	10%

Sig. Approach:
MutSig (FDR≤0.1)

Pasqualucci *et al.*, 2011 (REF. 101)

Recurrent Mut.	Recurrent Mut.
MLL2	TMEM30A
CREBBP	EP300
TP53	DPYD
MYOM2	DSC3
TNFAIP3	KLF2
PIM1	MED12L
CD36	MTMR8
B2M	PMS1
CD79B	TSC22D1
PRDM1	OFD1
CARD11	TRAF3
BCL2	BCL2L10
MEF2B	BRSK1
MYD88	CAMTA1
ANKLE2	CYTSB
KDM2B	DUSP9
HNF1B	FBXO31
MYC	GRB2
CD58	HMGB1
EZH2	MYO1G
ADAMTSL3	PPP2R5A
AKAP8	RASGEF1A
C12orf35	RGAG1
CCND3	SERPINA1
DCHS1	SMARCA1
DUSP27	ZNF521
MAGEC3	ZWILCH
TLL2	NOTCH1

Sig. Approach:
Frequency

Paediatric Acute Lymphoblastic Leukemia (ALL)

Zhang *et al.*, 2012 (REF. 97)

ETP ALL		Non-ETP ALL	
Recurrent Mut.	Incidence	Recurrent Mut.	Incidence
ETV6	33%	CDKN2A	81%
WT1	28%	NOTCH1	43%
CDKN2A	25%	PTEN	26%
PHF6	25%	PHF6	21%
DNM2	13%	FBXW7	14%
NRAS	17%	WT1	12%
SUZ12	11%	EZH2	12%
EZH2	16%	ETV6	10%
NOTCH1	16%	DNM2	10%
RUNX1	16%	NRAS	10%
FLT3	14%	CTCF	10%
EED	13%	EED	7%
IKZF1	13%	BCL11B	7%
JAK3	11%	ECT2L	7%
NF1	11%	RUNX1	5%
GATA3	9%	NF1	5%
PTEN	9%	IL7R	5%
CTCF	8%	RELN	5%
ECT2L	8%	SUZ12	2%
IL7R	8%	IKZF1	2%
JAK1	8%	JAK1	2%
SETD2	8%		
RELN	6%		
SH2B3	6%		
BCL11B	5%		
EP300	5%		
FBXW7	5%		
PTPN11	5%		
KRAS	3%		
BRAF	2%		
DCLRE1C	2%		
HIST1H1B	2%		
HNRNPA1	2%		
HNRNPR	2%		

Sig. Approach:
Frequency

Multiple Myeloma (MM)

Chapman *et al.*, 2011 (REF. 103)

SMG	Incidence
KRAS	26%
NRAS	24%
FAM46C	13%
DIS3	11%
TP53	8%
CCND1	5%
PNRC1	5%
ALOX12B	8%
HLA-A	5%
MAGED1	5%

Sig. Approach:
MutSig (FDR≤0.1)