Figure S1. Characterization of somatic mutations in DLBCL. (A) Number and proportion of mutation classes with different functional impact in DLBCL. (B) Somatic mutation signatures in DLBCL. (C) Difference in mutation densities in MMR-mutant and MMR-wildtype DLBCL samples. Mb, million bases; DLBCL, diffuse large B cell lymphoma; MMR, mismatch-repair pathways; Ins, insertion; Del, deletion.

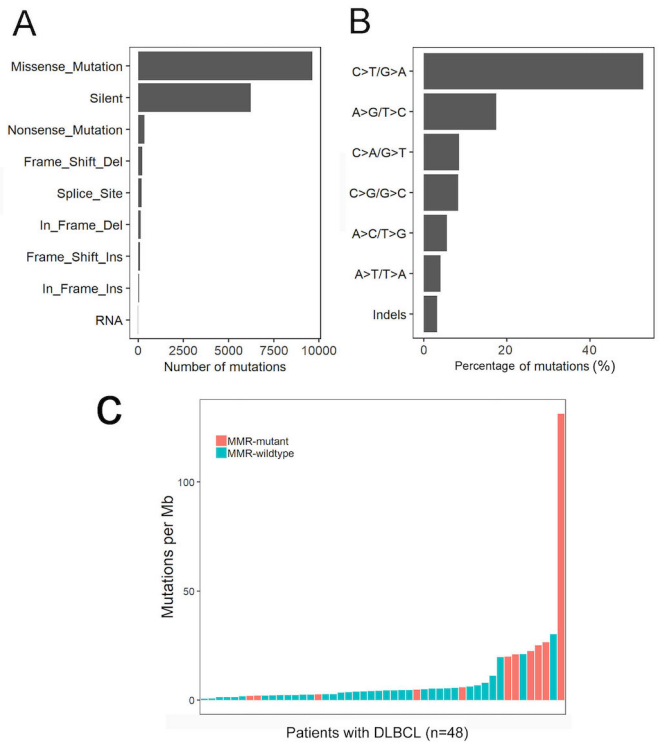

Figure S2. Frequencies of copy number variations in patients with diffuse large B cell lymphoma. Frequency of copy number variations was computed by divided by the number of patients with significant focal copy number variations by the total number of patients.

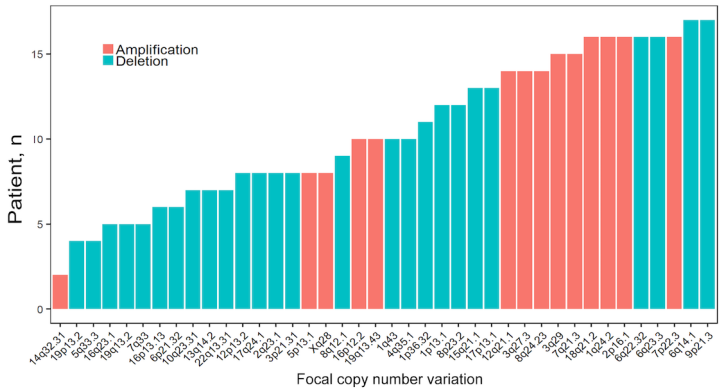

Figure S3. Unsupervised hierarchical clustering of the 20 driver genes with the most frequent gene-level copy-number alterations. The gene-level copy-number alterations of the top 20 driver genes were clustered using the heatmap function in R. Dark red, red, white, blue and dark blue represent high level amplification, amplification, copy-neutral, deletion and high-level deletion, respectively.

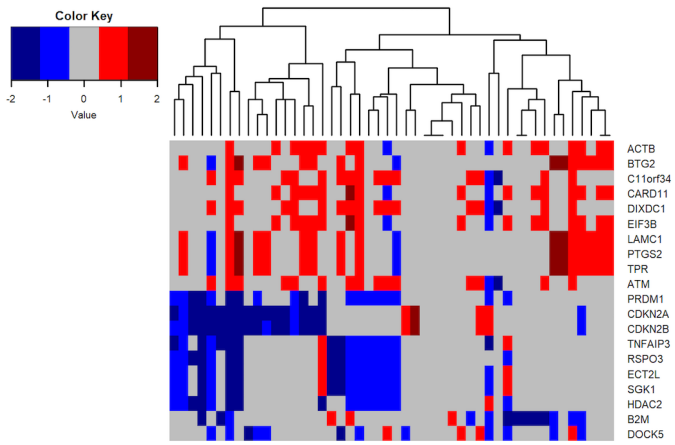

Table SI. Driver genes detected using OncodriveFM in patients with diffuse large B cell lymphoma.

Gene	P-value	Q-value
ZNF814	<0.001	<0.001
HLA-C	<0.001	<0.001
CD79B	<0.001	<0.001
RHPN2	<0.001	<0.001
MYD88	<0.001	<0.001
EZH2	<0.001	0.02
PIM1	0.05	0.03
MLL2	<0.001	<0.001
PDE4DIP	0.01	0.04
SOCS1	0.05	0.03
IRF4	0.03	0.01
CARD11	0.04	0.01
TNFAIP3	<0.001	0.01
P2RY8	<0.001	0.03
HLA-A	<0.001	<0.001
CIITA	<0.001	0.01
PRDM1	<0.001	<0.001
ARID1A	<0.001	<0.001
TP53	<0.001	0.01
STAT6	0.04	0.01
MSN	0.01	0.01
PTEN	0.02	0.01
BTG2	0.02	0.01
B2M	<0.001	<0.001
POLQ	0.01	0.04
ACTB	<0.001	<0.001
ACTG1	0.01	<0.001
FLNC	<0.001	0.03
NFKBIE	<0.001	<0.001
PCDH7	0.01	0.01
UBE2A	0.01	0.01
CSNK2A1	<0.001	<0.001
PTPN6	<0.001	<0.001
FLNA	0.02	0.03
DDX3X	0.12	<0.001
HLA-B	0.01	<0.001
DHX33	0.03	0.03
CGN	0.01	0.02
VANGL1	<0.001	0.01
HFE	0.01	<0.001
ATP6V0A1	0.01	<0.001
POLD3	0.01	0.03
TNPO1	0.04	0.01
CSNK1A1	0.02	<0.001
DGCR8	0.02	0.03
LYN	0.02	<0.001
ZMYM4	0.05	0.01

Driver genes were predicted by oncodriveFM (v0.0.1) (https://www.intogen.org). OncodriveFM compares the actual functional impact with a null distribution model generated by 1,000,000 permutations and computes the raw P-values by calculating the fraction of permutated functional scores greater than the actual one. Benjamini-Hochberg false discovery rate procedure was used to calculate the q values for genes in the oncodriveFM analysis.

Table SII. Driver genes detected by OncodriveCluster in patients with diffuse large B cell lymphoma.

Gene	P-value	Q-value
ZNF814	<0.001	<0.001
HLA-C	<0.001	0.01
CD79B	<0.001	<0.001
RHPN2	<0.001	<0.001
MYD88	0.01	0.03
EZH2	<0.001	<0.001
KDR	<0.001	0.01
MLH1	<0.001	<0.001
MCM8	<0.001	<0.001
HLA-DRB1	0.01	0.02
UBXN11	<0.001	0.01
SCYL1	<0.001	0.01

Driver genes were predicted by oncodriveCLUST v.0.4.1 (https://www.intogen.org). OncodriveCLUST performed the following steps to calculate P-values for genes. First, protein affecting mutations of each gene across a cohort of tumors are evaluated looking for those protein residues having a number of mutations barely expected by chance. Second, these positions are thereafter grouped to form mutation clusters. Third, each cluster is scored with a figure proportional to the percentage of the gene mutations that are enclosed within that cluster and inversely related to its length. The gene clustering score is obtained as the sum of the scores of all clusters (if any) found in that gene. Finally, each gene clustering score is compared with the background model to obtain a P-value. Background model is obtained performing the same steps than above but assessing only coding silent mutations. Benjamini-Hochberg false discovery rate procedure was used to calculate the q values for genes in the oncodriveCLUST analysis.

Table SIII. Driver genes detected by iCAGES in patients with diffuse large B cell lymphoma.

Gene	iCAGES
TP53	1.00
EGFR	0.98
CDKN1A	0.98
CREBBP	0.98
MYC	0.97
RAF1	0.97
STAT3	0.96
BRCA1	0.96
GRB2	0.95
MAP2K1	0.94
MAPK8	0.93
PDGFRB	0.92
CDH1	0.91
STAT1	0.91
PIK3CG	0.90
PTGS2	0.88
CDKN2A	0.88
ATM	0.87
PIK3CD	0.85
IKBKB	0.85
RHOA	0.85
TGFB2	0.84
FAS	0.84
PLCG2	0.83
HDAC2	0.82
PRKCD	0.81
AKT2	0.81
CDKN2B	0.81
CD44	0.81
MLH1	0.80
GATA2	0.79
SOCS1	0.79
TNF	0.78
ITPR1	0.77
PTCH1	0.76
ITGA5	0.76
CHEK2	0.74
PPARG	0.74
FGF5	0.74
PTPN6	0.73
IL6	0.72
FLT1	0.72
FN1	0.71
SOS2	0.71
GNAI2	0.70
GAB1	0.70
GNAQ	0.69
LYN	0.69
MMP2	0.69
KITLG	0.68
NFKB2	0.68
SMARCA4	0.67
COL4A1	0.67
ADCY3	0.67
ITGB4	0.66
TLR4	0.66
UBB	0.66
ITPR2	0.65
ABCB1	0.65
KAT2B	0.65
ACTB	0.64
SMAD7	0.64
E2F2	0.63
INHBA	0.63
CAV1	0.63
ADCY6	0.62
PPP2CA	0.62
WNT5A	0.61
TPR	0.61
XPO1	0.61
CCND3	0.60
AXIN2	0.60
BMP2	0.60
FZD1	0.59
RPTOR	0.59
TCF7L1	0.59
TLR2	0.58
PCK2	0.58
SMO	0.58
SKP1	0.57
CXCR4	0.56
LEPR	0.56
RAPGEF3	0.56
MAP2K3	0.56
NGFR	0.56
TLN1	0.56
JAK3	0.56
PLCB4	0.56
RET	0.56
SHH	0.55
STAT6	0.55
PPP1CA	0.55
WNT10A	0.55
PGR	0.55
CBLC	0.54
MAP2K4	0.54
ATR	0.54
LAMC1	0.53
VAV3	0.53
BIRC2	0.53
JUP	0.53
NOTCH3	0.52
NCOA2	0.52
CUL1	0.52
GATA3	0.51
NF1	0.51
IL6ST	0.50
F2R	0.50
FZD7	0.50

The driver genes were predicted by Integrated Cancer Genome Score (iCAGES, https://github.com/WGLab/icages).

Table SIV. Driver genes detected using Driver Genes and Proteins in patients with diffuse large B cell lymphoma.

Gene	P-value	Q-value
AGA	0.01	1.00
APMAP	0.05	1.00
BTG2	<0.001	1.00
C11orf34	0.03	1.00
C1orf87	<0.001	1.00
CARS2	0.02	1.00
CBX3	<0.001	<0.001
CYYR1	0.03	1.00
DIXDC1	0.05	1.00
DLGAP1	<0.001	1.00
DOCK5	0.04	1.00
ECT2L	0.04	1.00
EIF3B	0.03	1.00
ETV6	0.04	1.00
FAIM	0.02	1.00
FAM161A	0.01	1.00
FAS	<0.001	0.50
FGGY	<0.001	1.00
GTF2A1	0.04	1.00
IGLL5	<0.001	<0.001
IKBKB	0.04	1.00
IZUMO3	0.01	1.00
KCNH8	0.04	1.00
KLF2	0.01	1.00
KLRC4	<0.001	1.00
L3MBTL4	0.05	1.00
LALBA	0.02	1.00
LARP1B	0.03	1.00
MMEL1	0.04	1.00
NOTUM	<0.001	0.69
NTN4	0.01	1.00
NUP153	0.02	1.00
PABPC1	<0.001	0.45
PDCL3	0.03	1.00
PFN1	0.01	1.00
PPP2R2B	0.02	1.00
RBMS3	<0.001	1.00
RECQL4	<0.001	<0.001
RIF1	0.03	1.00
RIMS4	<0.001	0.51
RRAS2	0.04	1.00
RSPO3	0.02	1.00
SFTPD	<0.001	1.00
SGK1	0.03	1.00
SLC16A8	0.03	1.00
SPECC1	0.03	1.00
TECRL	0.03	1.00
TGIF2	0.02	1.00
TMEM168	0.01	1.00
TMPO	0.04	1.00
TMSB4X	0.02	1.00
TNF	<0.001	0.74
TP53	0.01	1.00
TYRO3	0.02	1.00
UBE2A	<0.001	0.02
UBXN2B	0.03	1.00
UNC13B	0.01	1.00
ZBTB17	0.01	1.00
ZNF675	0.01	1.00

The driver genes were predicted by Driver Genes and Proteins (DrGaP, https://code.google.com/archive/p/drgap/). DrGaP integrates biological knowledge of the mutational process in tumors, including the length of protein-coding regions, transcript isoforms, variation in mutation types, differences in background mutation rates, redundancy of the genetic code and multiple mutations in one gene. DrGaP use a Poisson process to model the random nature of somatic mutations, a Bayesian model to estimate background mutation rates and a likelihood ratio test to test the significance of driver genes and pathways. The Benjamini-Hochberg method is used to control false discovery rate (FDR) of P-values.

Table SV. Mutation frequencies of 208 driver genes in patients with diffuse large B cell lymphoma (n=48).

Driver gene	Number of patients with mutations	Mutation frequency
IGLL5	20	0.42
MLL2	17	0.35
BTG2	16	0.33
B2M	13	0.27
PIM1	12	0.25
CARD11	10	0.21
ATM	9	0.19
EGFR	8	0.17
P2RY8	8	0.17
STAT3	8	0.17
TNFAIP3	8	0.17
HLA-C	7	0.15
IRF4	7	0.15
MYD88	7	0.15
PDE4DIP	7	0.15
RHPN2	7	0.15
ZNF814	7	0.15
AGA	6	0.13
CREBBP	6	0.13
DOCK5	6	0.13
KDR	6	0.13
TMSB4X	6	0.13
ARID1A	5	0.10
CD79B	5	0.10
CIITA	5	0.10
FAS	5	0.10
FLNC	5	0.10
HLA-A	5	0.10
MYC	5	0.10
NF1	5	0.10
PCDH7	5	0.10
POLQ	5	0.10
SGK1	5	0.10
SOCS1	5	0.10
TLN1	5	0.10
TP53	5	0.10
ACTG1	4	0.08
DDX3X	4	0.08
FLNA	4	0.08
FN1	4	0.08
HLA-B	4	0.08
HLA-DRB1	4	0.08
IL6ST	4	0.08
KLF2	4	0.08
NFKBIE	4	0.08
PABPC1	4	0.08
PGR	4	0.08
PLCG2	4	0.08
PRDM1	4	0.08
RECQL4	4	0.08
RET	4	0.08
TNF	4	0.08
TPR	4	0.08
TYRO3	4	0.08

Table SV. Continued.

Driver gene	Number of patients with mutations	Mutation frequency
UBE2A	4	0.08
UBXN11	4	0.08
UNC13B	4	0.08
ACTB	3	0.06
APMAP	3	0.06
BRCA1	3	0.06
CBX3	3	0.06
CSNK2A1	3	0.06
CXCR4	3	0.06
DHX33	3	0.06
EZH2	3	0.06
FAM161A	3	0.06
GATA3	3	0.06
IKBKB	3	0.06
ITGB4	3	0.06
ITPR1	3	0.06
MAP2K1	3	0.06
MCM8	3	0.06
MLH1	3	0.06
NFKB2	3	0.06
PTCH1	3	0.06
PTPN6	3	0.06
RAF1	3	0.06
RAPGEF3	3	0.06
RIF1	3	0.06
SCYL1	3	0.06
SLC16A8	3	0.06
SMARCA4	3	0.06
SMO	3	0.06
STAT6	3	0.06
TLR4	3	0.06
UBXN2B	3	0.06
VANGL1	3	0.06
ZMYM4	3	0.06
ABCB1	2	0.04
ATP6V0A1	2	0.04
ATR	2	0.04
CCND3	2	0.04
CD44	2	0.04
CDH1	2	0.04
CDKN1A	2	0.04
CGN	2	0.04
CHEK2	2	0.04
CSNK1A1	2	0.04
CYYR1	2	0.04
DGCR8	2	0.04
DLGAP1	2	0.04
ECT2L	2	0.04
FAIM	2	0.04
FLT1	2	0.04
FZD1	2	0.04
GATA2	2	0.04
GNAI2	2	0.04
HFE	2	0.04
JAK3	2	0.04

Table SV. Continued.

Driver gene	Number of patients with mutations	Mutation frequency
KITLG	2	0.04
KLRC4	2	0.04
LEPR	2	0.04
LYN	2	0.04
MAP2K3	2	0.04
MMEL1	2	0.04
MSN	2	0.04
NGFR	2	0.04
NOTCH3	2	0.04
NUP153	2	0.04
PDGFRB	2	0.04
PFN1	2	0.04
PLCB4	2	0.04
POLD3	2	0.04
PPP2R2B	2	0.04
PTEN	2	0.04
RHOA	2	0.04
RIMS4	2	0.04
RPTOR	2	0.04
SOS2	2	0.04
TECRL	2	0.04
TLR2	2	0.04
TNPO1	2	0.04
XPO1	2	0.04
ADCY3	1	0.02
ADCY6	1	0.02
AKT2	1	0.02
AXIN2	1	0.02
BIRC2	1	0.02
BMP2	1	0.02
C11orf34	1	0.02
C1orf87	1	0.02
CARS2	1	0.02
CAV1	1	0.02
CBLC	1	0.02
CDKN2A	1	0.02
CDKN2B	1	0.02
COL4A1	1	0.02
CUL1	1	0.02
DIXDC1	1	0.02
E2F2	1	0.02
EIF3B	1	0.02
ETV6	1	0.02
F2R	1	0.02
FGF5	1	0.02
FGGY	1	0.02
FZD7	1	0.02
GAB1	1	0.02
GNAQ	1	0.02
GRB2	1	0.02
GTF2A1	1	0.02
HDAC2	1	0.02
IL6	1	0.02
INHBA	1	0.02
ITGA5	1	0.02

Table SV. Continued.

Driver gene	Number of patients with mutations	Mutation frequency
ITPR2	1	0.02
IZUMO3	1	0.02
JUP	1	0.02
KAT2B	1	0.02
KCNH8	1	0.02
L3MBTL4	1	0.02
LALBA	1	0.02
LAMC1	1	0.02
LARP1B	1	0.02
MAP2K4	1	0.02
MAPK8	1	0.02
MMP2	1	0.02
NCOA2	1	0.02
NOTUM	1	0.02
NTN4	1	0.02
PCK2	1	0.02
PDCL3	1	0.02
PIK3CD	1	0.02
PIK3CG	1	0.02
PPARG	1	0.02
PPP1CA	1	0.02
PPP2CA	1	0.02
PRKCD	1	0.02
PTGS2	1	0.02
RBMS3	1	0.02
RRAS2	1	0.02
RSPO3	1	0.02
SFTPD	1	0.02
SHH	1	0.02
SKP1	1	0.02
SMAD7	1	0.02
SPECC1	1	0.02
STAT1	1	0.02
TCF7L1	1	0.02
TGFB2	1	0.02
TGIF2	1	0.02
TMEM168	1	0.02
TMPO	1	0.02
UBB	1	0.02
VAV3	1	0.02
WNT10A	1	0.02
WNT5A	1	0.02
ZBTB17	1	0.02
ZNF675	1	0.02

Mutation frequency was calculated by dividing the number of patients with mutations by 48 patients.

Table SVI. Significant GO terms that are enriched for driver genes in patients with diffuse large B cell lymphoma.

GO terms	Gene count	P-values	Benjamini adjusted P-values
Stimulatory C-type lectin receptor signaling pathway	10	4.70×10^{-6}	4.10×10^{-4}
Fc-ε receptor signaling pathway	11	5.80×10^{-5}	3.30×10^{-3}
T cell receptor signaling pathway	10	7.40×10^{-5}	3.80×10^{-3}
Stress-activated MAPK cascade	4	1.90×10^{-3}	4.10×10^{-2}
Replicative senescence	6	1.70×10^{-7}	3.00×10^{-5}
Cellular response to DNA damage stimulus	13	8.00×10^{-6}	6.30×10^{-4}
Regulation of signal transduction by p53 class mediator	8	7.30×10^{-4}	2.00×10^{-2}
DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	4	8.40×10^{-4}	2.20×10^{-2}
DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	6	8.70×10^{-4}	2.20×10^{-2}
Cellular response to gamma radiation	4	1.90×10^{-3}	4.10×10^{-2}
Positive regulation of NF-κB transcription factor activity	10	3.20×10^{-5}	2.20×10^{-3}
Positive regulation of NF-κB import into nucleus	5	1.00×10^{-4}	5.10×10^{-3}
Positive regulation of interleukin-6 production	6	1.90×10^{-4}	7.90×10^{-3}
Defense response to Gram-positive bacterium	7	5.40×10^{-4}	1.60×10^{-2}
Lipopolysaccharide-mediated signaling pathway	5	5.50×10^{-4}	1.60×10^{-2}
Positive regulation of chemokine production	4	1.00×10^{-3}	2.50×10^{-2}
Phosphatidylinositol-mediated signaling	10	5.10×10^{-6}	4.30×10^{-4}
Phosphatidylinositol phosphorylation	8	1.30×10^{-4}	6.00×10^{-3}
Phosphatidylinositol-3-phosphate biosynthetic process	6	2.90×10^{-4}	1.00×10^{-2}
Regulation of phosphatidylinositol 3-kinase signaling	7	3.40×10^{-4}	1.20×10^{-2}
I-kappaB kinase/NF-kappaB signaling	7	8.00×10^{-5}	4.10×10^{-3}
Regulation of tumor necrosis factor-mediated signaling pathway	5	4.30×10^{-4}	1.40×10^{-2}
Positive regulation of interleukin-6 production	6	1.90×10^{-4}	7.90×10^{-3}
Cellular response to lipopolysaccharide	7	2.40×10^{-3}	4.90×10^{-2}
Interferon-gamma-mediated signaling pathway	10	1.70×10^{-7}	3.10×10^{-5}
Antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent	4	1.30×10^{-4}	6.10×10^{-3}
Antigen processing and presentation of endogenous peptide antigen via major histocompatibility complex class I via Estrogen receptor pathway, tracheal antimicrobial peptide -independent	3	4.20×10^{-4}	1.40×10^{-2}
Antigen processing and presentation of peptide antigen via MHC class I	5	4.30×10^{-4}	1.40×10^{-2}
Interleukin-6-mediated signaling pathway	4	1.30×10^{-4}	6.10×10^{-3}
Vascular endothelial growth factor receptor signaling pathway	6	1.70×10^{-3}	3.70×10^{-2}
Ras protein signal transduction	6	1.50×10^{-3}	3.50×10^{-2}
Extracellular matrix organization	11	1.30×10^{-4}	5.90×10^{-3}
Stress-activated mitogen-activated protein kinase cascade	4	1.90×10^{-3}	4.10×10^{-2}
Renal system development	4	5.60×10^{-4}	1.60×10^{-2}
Negative regulation of fat cell differentiation	5	1.60×10^{-3}	3.60×10^{-2}

In order to characterize the functional enrichment of all driver genes, GO (1) controlled vocabulary that can be applied to all eukaryotes even as knowledge of gene and protein roles in cells is accumulating and changing. To this end, three independent ontologies accessible on the World-Wide Web (http://www.geneontology.org biological process term analysis (1) was performed with The Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) (2). (1) Ashburner M, Ball CA, Blake JA, *et al*: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25-29, 2000. (2) Huang DW, Sherman BT and Lempicki RA: Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. 37: 1-13, 2009.

Table SVII. Significant KEGG pathways that are enriched for driver genes.

KEGG pathway	Gene count	P-value	Benjamini adjusted P-value
Hepatitis B	26	9.10×10^{-16}	4.40×10^{-14}
Endometrial cancer	15	4.50×10^{-12}	1.10×10^{-10}
Forkhead box, sub-group O signaling pathway	21	1.40×10^{-11}	2.70×10^{-10}
Chronic myeloid leukemia	16	4.20×10^{-11}	6.90×10^{-10}
Glioma	15	1.20×10^{-10}	1.70×10^{-9}
B cell receptor signaling pathway	15	2.80×10^{-10}	3.70×10^{-9}
Prostate cancer	16	8.40×10^{-10}	9.80×10^{-9}
Pancreatic cancer	14	1.50×10^{-9}	1.60×10^{-8}
Fc epsilon RI signaling pathway	14	2.70×10^{-9}	2.50×10^{-8}
Acute myeloid leukemia	13	2.90×10^{-9}	2.50×10^{-8}
Melanoma	14	4.70×10^{-9}	3.80×10^{-8}
Erb-b2 receptor tyrosine kinase signaling pathway	15	6.90×10^{-9}	4.90×10^{-8}
Colorectal cancer	13	9.80×10^{-9}	6.80×10^{-8}
Chemokine signaling pathway	20	3.10×10^{-8}	1.90×10^{-7}
Non-small cell lung cancer	12	3.50×10^{-8}	2.10×10^{-7}
Neurotrophin signaling pathway	16	6.70×10^{-8}	3.60×10^{-7}
Ras signaling pathway	21	1.50×10^{-7}	7.60×10^{-7}
Bladder cancer	10	2.40×10^{-7}	1.20×10^{-6}
Hepatitis C	16	2.70×10^{-7}	1.30×10^{-6}
T cell receptor signaling pathway	14	3.30×10^{-7}	1.50×10^{-6}
Central carbon metabolism in cancer	11	1.40×10^{-6}	5.30×10^{-6}
Prolactin signaling pathway	11	3.80×10^{-6}	1.30×10^{-5}
Osteoclast differentiation	14	7.30×10^{-6}	2.50×10^{-5}
Cathelicidin antimicrobial peptide signaling pathway	17	9.30×10^{-6}	3.10×10^{-5}
Insulin signaling pathway	14	1.30×10^{-5}	4.20×10^{-5}
Renal cell carcinoma	10	1.50×10^{-5}	4.90×10^{-5}
Progesterone-mediated oocyte maturation	10	1.40×10^{-4}	3.90×10^{-4}
Natural killer cell mediated cytotoxicity	11	4.20×10^{-4}	1.10×10^{-3}
Vascular endothelial growth factor signaling pathway	8	4.20×10^{-4}	1.10×10^{-3}
Choline metabolism in cancer	10	4.40×10^{-4}	1.10×10^{-3}
Fc γ R-mediated phagocytosis	9	5.90×10^{-4}	1.40×10^{-3}
Cyclic GMP- Protein Kinase G signaling pathway	12	8.80×10^{-4}	2.00×10^{-3}
Chagas disease (American trypanosomiasis)	18	1.20×10^{-10}	1.80×10^{-9}
Influenza A	21	1.70×10^{-9}	1.70×10^{-8}
Toll-like receptor signaling pathway	14	6.60×10^{-7}	2.70×10^{-6}
Tumor necrosis factor signaling pathway	14	7.30×10^{-7}	2.90×10^{-6}
Insulin resistance	12	3.00×10^{-5}	9.20×10^{-5}
Apoptosis	9	7.00×10^{-5}	2.00×10^{-4}
Type II diabetes mellitus	7	6.90×10^{-4}	1.60×10^{-3}
Adipocytokine signaling pathway	8	9.80×10^{-4}	2.20×10^{-3}
Mammalian target of rapamycin signaling pathway	7	1.90×10^{-3}	4.00×10^{-3}
Non-alcoholic fatty liver disease	9	2.10×10^{-2}	3.60×10^{-2}
Cell cycle	16	1.00×10^{-7}	5.50×10^{-7}
p53 signaling pathway	9	1.20×10^{-4}	3.40×10^{-4}
Toxoplasmosis	17	2.50×10^{-9}	2.40×10^{-8}
Tuberculosis	15	4.40×10^{-5}	1.30×10^{-4}
Inflammatory bowel disease (IBD)	9	8.80×10^{-5}	2.50×10^{-4}
Leishmaniasis	9	1.90×10^{-4}	4.90×10^{-4}
Pertussis	8	1.50×10^{-3}	3.20×10^{-3}
Rheumatoid arthritis	8	3.70×10^{-3}	7.10×10^{-3}
Malaria	6	4.80×10^{-3}	9.00×10^{-3}
Legionellosis	6	7.30×10^{-3}	1.30×10^{-2}
Estrogen signaling pathway	17	5.10×10^{-10}	6.30×10^{-9}
GnRH signaling pathway	16	1.40×10^{-9}	1.50×10^{-8}
Platelet activation	18	4.20×10^{-9}	3.50×10^{-8}
Gap junction	13	5.60×10^{-7}	2.30×10^{-6}

Table SVII. Continued.

KEGG pathway	Gene count	P-value	Benjamini adjusted P-value
Oxytocin signaling pathway	16	1.30×10^{-6}	4.90×10^{-6}
Inflammatory mediator regulation of TRP channels	13	1.80×10^{-6}	6.70×10^{-6}
Long-term depression	9	5.50×10^{-5}	1.60×10^{-4}
Long-term potentiation	9	1.10×10^{-4}	3.10×10^{-4}
Cholinergic synapse	11	1.90×10^{-4}	5.10×10^{-4}
Vascular smooth muscle contraction	11	3.00×10^{-4}	7.80×10^{-4}
Cyclic GMP- Protein Kinase G signaling pathway	12	8.80×10^{-4}	2.00×10^{-3}
Oocyte meiosis	10	8.80×10^{-4}	2.00×10^{-3}
Retrograde endocannabinoid signaling	9	2.00×10^{-3}	4.10×10^{-3}
Dopaminergic synapse	10	2.40×10^{-3}	4.90×10^{-3}
Serotonergic synapse	9	3.60×10^{-3}	7.00×10^{-3}
Adrenergic signaling in cardiomyocytes	10	4.00×10^{-3}	7.60×10^{-3}
Gastric acid secretion	7	6.00×10^{-3}	1.10×10^{-2}
Renin secretion	6	1.50×10^{-2}	2.60×10^{-2}
Glutamatergic synapse	8	1.50×10^{-2}	2.60×10^{-2}
Pancreatic secretion	7	1.90×10^{-2}	3.20×10^{-2}
Thyroid hormone synthesis	6	2.10×10^{-2}	3.60×10^{-2}
Glucagon signaling pathway	7	2.50×10^{-2}	4.10×10^{-2}
Graft-versus-host disease	7	8.20×10^{-5}	2.40×10^{-4}
Allograft rejection	6	1.40×10^{-3}	3.10×10^{-3}
Antigen processing and presentation	8	1.60×10^{-3}	3.40×10^{-3}
Viral myocarditis	7	1.70×10^{-3}	3.70×10^{-3}
Type I diabetes mellitus	6	2.40×10^{-3}	4.90×10^{-3}
Autoimmune thyroid disease	5	3.00×10^{-2}	4.80×10^{-2}
Vascular endothelial growth factor signaling pathway	8	4.20×10^{-4}	1.10×10^{-3}
Regulation of lipolysis in adipocytes	7	1.60×10^{-3}	3.40×10^{-3}
Pathogenic Escherichia coli infection	5	2.80×10^{-2}	4.60×10^{-2}
Dilated cardiomyopathy	8	2.80×10^{-3}	5.70×10^{-3}
Hypertrophic cardiomyopathy	7	8.30×10^{-3}	1.50×10^{-2}
Arrhythmogenic right ventricular cardiomyopathy	6	1.80×10^{-2}	3.10×10^{-2}
Hedgehog signaling pathway	4	2.20×10^{-2}	3.80×10^{-2}
Phosphatidylinositol signaling system	7	2.40×10^{-2}	3.90×10^{-2}

In order to characterize the functional enrichment of all driver genes, KEGG (1) controlled vocabulary that can be applied to all eukaryotes even as knowledge of gene and protein roles in cells is accumulating and changing. To this end, three independent ontologies accessible on the World-Wide Web (http://www.geneontology.org pathway analysis (1) was performed with The Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) (2). (1) Ogata H, Goto S, Sato K, Fujibuchi W, Bono H and Kanehisa M: KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 27: 29-34, 1999. (2) Huang DW, Sherman BT and Lempicki RA: Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. 37: 1-13, 2009.

Table SVIII. Linear regression model analysis between overall survival in patients with diffuse large B cell lymphoma and clinicopathological features.

Clinical features	T-value	P-value
Age, years	-0.1	0.91
Clinical stage	0.73	0.47
Sex	0.58	0.56
Radiation therapy	-1.12	0.27
Ethnicity	-1.55	0.13

P<0.05 was predefined statistically significant. Clinical stage was evaluated based on the Ann Arbor staging system (1). (1) Carbone PP, Kaplan HS, Musshoff K, Smithers DW and Tubiana M: Report of the Committee on Hodgkin's Disease Staging Classification. Cancer Res 31: 1860LP-1861, 1971.