Supplementary materials for

Ubiquitinated PCNA drives USP1 synthetic lethality in cancer

Antoine Simoneau*, Justin L. Engel*, Madhavi Bandi*, Katherine Lazarides, Shangtao Liu, Samuel R. Meier, Ashley H. Choi, Hongxiang Zhang, Binzhang Shen, Lauren Martires, Deepali Gotur, Truc V. Pham, Fang Li, Lina Gu, Shanzhong Gong, Minjie Zhang, Erik Wilker, Xuewen Pan, Douglas A. Whittington, Scott Throner, John P. Maxwell, Yingnan Chen, Yi Yu, Alan Huang, Jannik N. Andersen, Tianshu Feng^

Correspondence to: pfeng@tangotx.com

SUPPLEMENTARY METHODS

Table I. Cell lines and culture conditions

Cell Line	Vendor	Cat#	Complete medium	Purchase date
HAP-1 USP1 knockout	Horizon discovery	HZGHC000455c007	DMEM + 10% FBS	06/19/2019
HAP-1 parental	Horizon discovery	C631	DMEM + 10% FBS	06/19/2019
UWB1.289	ATCC	CRL-2945	MGEM/1640+3%FBS	9/20/2017
UWB1.289+BRCA1	ATCC	CRL-2946	MGEM/1640+3%FBS+200ug/ml G418	9/20/2017
OVSAHO	JCRB	JCRB1046	RPMI 1640+10%FBS	12/15/2015
OVTOKO	JCRB	JCRB1048	RPMI 1640+10%FBS	12/15/2015
OV56	ECACC	96020759	DMEM:HAMS F12 (1:1) + 2mM Glutamine + 5% Foetal Bovine Serum (FBS) + 0.5 ug/ml hydrocortisone + 10ug/ml insulin	3/2/2016
KURAMOCHI	JCRB	JCRB0098	RPMI1640+10%FBS	12/7/2012
COV362	ECACC	7071910	DMEM + 2mM Glutamine + 10% FBS	12/24/2012
ES-2	ATCC	CRL-1978	McCoy's 5a +10%FBS	8/3/2012
SK-OV-3	ATCC	HTB-77	McCoy's 5a+10%FBS	9/27/2007
OVISE	JCRB	JCRB1043	RPMI1640+10%FBS	12/7/2012
Caov-3	ATCC	HTB-75	DMEM+10%FBS	
SUM149PT	Asterand		DMEM:HAMS F12 (1:1) + 2mM Glutamine + 5% (FBS) + 0.5 ug/ml hydrocortisone + 10ug/ml insulin	10/13/2017
MDA-MB-436	ATCC	HTB-130	Leibovitz's L-15 medium with 10 mcg/ml insulin, 16 mcg/ml glutathione, 90%; +10%FBS	11/17/2008
MCF-7	ATCC	HTB-22	Eagle's Minimum Essential Medium+0.01 mg/ml bovine insulin+ 10%FBS	9/27/2007
AU565	ATCC	CRL-2351	RPMI-1640 +10%FBS	8/3/2012
CAMA-1	ATCC	HTB-21	EMEM+10%FBS	8/3/2012
MCF 10A	ATCC	CRL-10317	MEGM + 100ng/mL Cholera toxin	
BT-549	ATCC	HTB-122	RPMI1640+10%FBS+0.023 IU/ml human insulin	12/20/2010
HCC1954	ATCC	CRL-2338	RPMI-1640+10%FBS	8/3/2012
HCC1395	ATCC	CRL-2324	RPMI1640+10%FBS	12/20/2010
NCI-H1693	ATCC	CRL-5887	RPMI-1640+5%FBS	8/3/2012
NCI-H292	ATCC	CRL-1848	RPMI1640+10%FBS	
NCI-H1792	ATCC	CRL-5895	RPMI1640+10%FBS	9/6/2010
NCI-H1299	ATCC	CRL-5803	RPMI1640+10%FBS	

JHOS-2	RIKEN	RCB1521	1:1 DMEM/HAM-F12+10% FBS + 0.1mM NEAA	11/30/2020
JHOS-4	RIKEN	RCB1678	1:1 DMEM/HAM-F12+10% FBS + 0.1mM NEAA	11/30/2020
Lenti-X	Takara Bio	632180	DMEM + 10% FBS	9/20/2017

Table II. Antibodies

Target	Vendor	Catalog#	Dilution
ATM	CST	2873S	1:1000
BRCA1, D-9	SCBT	sc-6954	1:1000
FANCD2	Abcam	ab108928	1:1000
Goat anti Rabbit AlexaFluor568	Thermofisher	A21244	1:500
Goat anti-Mouse IgG HRP	CST	#7076	1:5000
Goat anti-rabbit AlexaFluor488	Thermofisher	A11034	1:500
Goat anti-Rabbit IgG HRP	CST	#7074	1:1000
H3	Abcam	ab176842	1:10000
p-ATM S1981	Abcam	ab81292	1:1000
p-CHK1 S317	CST	2360S	1:1000
PCNA K164ub	CST	#13439	1:500 IF, 1:1000 WB
PCNA, PC-10	CST	#2586	1:500 IF, 1:1000 WB
p-RPA32 S33	Bethyl	A300-246A	1:1000
RAD18	Thermofisher	A301-340A	1:1000
RAD51	Abcam	ab133534	1:1000
RPA32	Bethyl	A300-244A	1:1000
UBE2K	CST	#3847S	1:1000
USP1	Proteintech	66069-1	1:1000
Vinculin	CST	#13901	1:1000
yH2AX	CST	9718S	1:400 IF, 1:1000 WB
β-actin	CST	#4970	1:5000

Table III. Guide sequences

Target	Sequence
NTC	GTACATGAAAAGGCTCTAGG
RAD18	ACAATAGATGATTTGCTG
UBE2K	CAATGACAATAATACCGTG

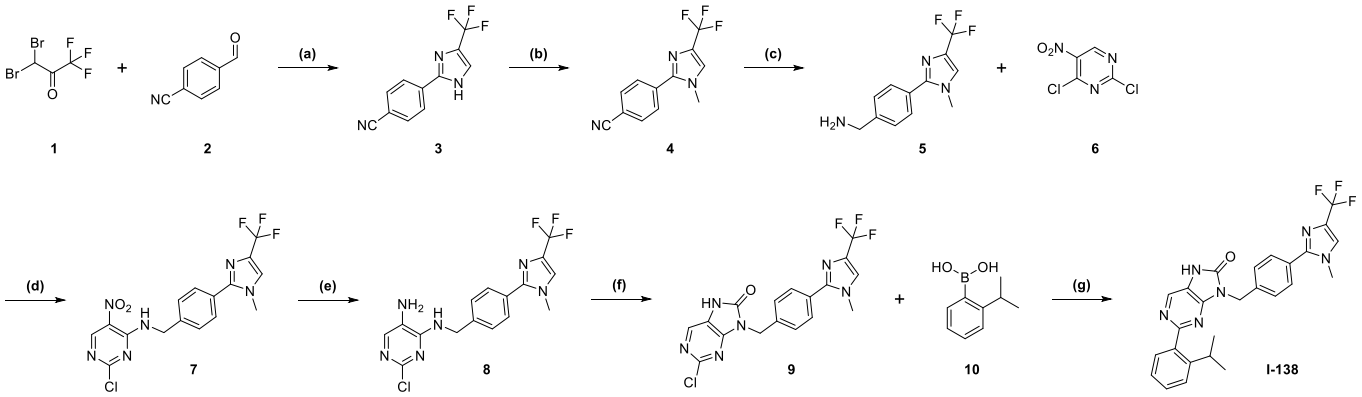
Synthesis of I-138: 2-(2-isopropylphenyl)-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one

The synthesis of I-138 (2-(2-isopropylphenyl)-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one) was performed as described below in a similar manner as previously described in US patent application US 2017/07145012.

Abbreviations

CDI, Carbonyldiimidazole
DCM, Dichloromethane
DIEA, *N,N*-Diisopropylethylamine
DMF, *N,N*-dimethylformamide
DMSO, Dimethylsulfoxide
EtOAc, Ethyl acetate
EtOH, Ethanol
HPLC, High pressure liquid chromatography
LCMS or LC-MS, Liquid chromatography mass spectrometry
MeCN, Acetonitrile
MeOH, Methanol
NMR, Nuclear magnetic resonance
$Pd(dppf)Cl_2$, [1,1'-Bis(diphenylphosphino)ferrocene]dichloropalladium(II)
PE, Petroleum ether
RT, Room temperature or ambient temperature
TFA, Trifluoroacetic acid
THF, Tetrahydrofuran

Synthetic route to I-138: (2-(2-isopropylphenyl)-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one)



Reagents and conditions: (**a**) i. NaOAc, water, 100 °C, 45 min; ii. NH_4OH, MeOH, RT, 40 min then 100 °C, 1 hr; (**b**) i. NaH, THF, 0 °C, 1 hr; ii. MeI, 0 °C, 2 hr; (**c**) H_2 (15 atm), Raney Nickel, 7 M NH_3 in MeOH. 5 hr; (**d**) DIEA, THF -70 °C to RT, 1 hr; (**e**) Fe, NH_4Cl, THF/EtOH/H_2O, 80 °C, 1 hr; (**f**) CDI, DCM, 40 °C, 1 hr; (**g**) $Pd(dppf)Cl_2$, K_2CO_3, dioxane/water (5:1), 120 °C, 1 hr.

General Procedure

Step (a). *4-(4-(trifluoromethyl)-1H-imidazol-2-yl)benzonitrile* ***(3).*** A mixture of 3,3-dibromo-1,1,1-trifluoropropan-2-one **1** (122 g, 912 mmol), sodium acetate (37.6 g, 458 mmol) and water (60 mL) was stirred for 45 min at 100 °C. After cooling to ambient temperature, the mixture was added to a solution of 4-formylbenzonitrile **2** (30.0 g, 458 mmol) and ammonium hydroxide (400 mL, 25% solution in water) in MeOH (1000 mL) and the resulting mixture was stirred for 40 min at ambient temperature, then 1 hr at 100 °C. After cooling to ambient temperature, the reaction mixture was concentrated under reduced pressure, filtered. The solid obtained by filtration was dried in vacuo to give the crude product (88 g). Further purification by trituration with EtOAc (50 mL) afforded the 4-(4-(trifluoromethyl)-1H-imidazol-2-yl)benzonitrile (50 g, 46%) as a yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ 13.54 (s, 1H), 8.17-8.12 (m, 2H), 8.06 (d, *J* = 1.0 Hz, 1H), 8.00-7.95 (m, 2H); LC-MS $(M+H)^+$: m/z = 238.2.

Step (b). *4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzonitrile* ***(4).*** A solution of 4-(4-(trifluoromethyl)-1H-imidazol-2-yl)benzonitrile **3** (70.0 g, 295 mmol) in THF (800 mL) at 0 °C was treated portion wise with sodium hydride (60% dispersion in mineral oil, 11.8 g, 443 mmol). After stirring for 1 hr at 0 °C, iodomethane (28.0 mL, 443 mmol) was added dropwise and resulting mixture was stirred for 2 h at 0 °C. The reaction mixture was poured into water (500 mL) and concentrated in vacuo until a solid was formed. The solid was filtered and dried in vacuo to give the desired product 4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzonitrile as a yellow solid (65 g, 88%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.09-7.93 (m, 5H), 3.85 (s, 3H); LC-MS $(M+H)^+$: m/z = 252.2.

Step (c). *(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)phenyl)methanamine* ***(5)***. A solution of 4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzonitrile **4** (57.5 g, 229 mmol) and Raney Nickel (58 g) in a solution of 7 M NH_3 in MeOH (600 mL) was evacuated and backfilled with hydrogen several times and was then charged with hydrogen. The resulting mixture was stirred for 5 hr at ambient temperature under 15 atm, then was filtered and concentrated under vacuum to afford the crude (4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)phenyl)methanamine as a yellow oil (58 g) which was taken forward without further purification. LC-MS $(M+H)^+$: m/z = 256.2.

Step (d). *2-chloro-N-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-5-nitropyrimidin-4-amine* ***(7)***. A mixture of 2,4-dichloro-5-nitropyrimidine **6** (43.8 g, 227 mmol) and DIEA (57.3 g, 568 mmol) in THF (1.0 L) was treated with dropwise addition of a solution of (4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)phenyl)methanamine **5** (58.0 g, crude) in THF (300 mL) at -70 °C. The solution was stirred for 1 hr at -70 °C and allowed to warm to ambient temperature gradually over 1 hr before being concentrated under vacuum. The residue was purified by silica gel chromatography (eluting with 0-50% EtOAc/PE) and the resulting residue was further purified by trituration with PE/EtOAc (1:1, 100 mL) to afford 2-chloro-*N*-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-5-nitropyrimidin-4-amine (52 g 56% for two steps) as a yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.70 (t, *J* = 6.0 Hz, 1H), 9.07 (s, 1H), 7.93 (d, *J* = 1.2 Hz, 1H), 7.69 (d, *J* = 8.0 Hz, 2H), 7.52 (d, *J* = 8.0 Hz, 2H), 4.82 (d, *J* = 6.0 Hz, 1H), 3.78 (s, 3H);

LC-MS $(M+H)^+$: m/z = 413.2.

Step (e). *2-chloro-N-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)pyrimidine-4,5-diamine* ***(8)***. A mixture of 2-chloro-*N*-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-5-nitropyrimidin-4-amine **7** (52 g, 126 mmol), iron powder (35.2 g, 630 mmol) and ammonium chloride (13.3 g, 252 mmol) in THF/EtOH/H_2O (3:3:1, 640 mL) was stirred for 1 hr at 80 °C. After cooling to ambient temperature the reaction mixture was filtered and concentrated under vacuum. The resulting residue

was purified by flash chromatography (SiO_2, gradient elution: 0-100% EtOAc/PE) to afford 2-chloro-*N*-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)pyrimidine-4,5-diamine (40 g, 83%) as a yellow oil. ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.93 (d, *J* = 1.1 Hz, 1H), 7.73-7.67 (m, 2H), 7.52-7.43 (m, 4H), 4.99 (br, 2H), 4.66 (d, *J* = 5.7 Hz, 2H), 3.79 (s, 3H); LC-MS $(M+H)^+$: m/z = 383.3.

Step (f). *2-chloro-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one **(9)***. A mixture of 2-chloro-*N*-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)pyrimidine-4,5-diamine **8** (40 g, 105 mmol) and CDI (68 g, 420 mmol) in DCM (500 mL) was stirred for 1 hr at 40 °C. The reaction mixture was concentrated under vacuum and the resulting residue was added portion-wise to ice-water (1.0 L) with vigorous stirring. The solid was filtered, washed with water, and dried in vacuum to give the 2-chloro-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one (37 g, 94%) as a white solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.20 (s, 1H), 7.93 (d, *J* = 1.0 Hz, 1H), 7.71 (d, *J* = 8.3 Hz, 2H), 7.46 (d, *J* = 8.3 Hz, 2H), 5.07 (s, 2H), 3.77 (s, 3H); LC-MS $(M+H)^+$: m/z = 409.1.

Step (g). *2-(2-isopropylphenyl)-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one **(I-138)***. A 1 L autoclave reactor was charged with 2-chloro-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one **9** (33 g, 80.9 mmol), (2-isopropylphenyl)boronic acid **10** (26.5 g, 161.8 mmol), potassium carbonate (33.5 g, 242.7 mmol), $Pd(dppf)Cl_2$ (8.8 g, 12.1 mmol), dioxane (500 mL) and water (125 mL). The reaction mixture was stirred at 120 °C for 10 hr under nitrogen atmosphere. After cooling to ambient temperature, the reaction mixture was filtered through a pad of diatomaceous earth. The filtrate was diluted with EtOAc, the phases separated, and the organic layer dried over Na_2SO_4, filtered, and concentrated in vacuo under reduce pressure to dryness. The resulting residue was subject to flash chromatography (SiO_2, gradient elution: 0-10% MeOH/DCM) and pure fractions were combined and concentrated. The resulting solid was suspended in MeCN (400 mL), concentrated in vacuo to remove most of the MeCN. Water (100 mL) was added and the mixture was concentrated under reduce pressure at 50 °C. The resulting solid was triturated with water (200 mL), collected by filtration, and washed with water (100 mL), then dried in vacuo to give the title compound 2-(2-isopropylphenyl)-9-(4-(1-methyl-4-(trifluoromethyl)-1H-imidazol-2-yl)benzyl)-7,9-dihydro-8H-purin-8-one (22.08 g, 50%, HPLC purity: 98.6%) as an off-white solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.60 (s, 1H), 8.42 (s, 1H), 7.92 (d, *J* = 1.1 Hz, 1H), 7.70 (d, *J* = 8.3 Hz, 2H), 7.57-7.36 (m, 5H), 7.25 (td, *J* = 7.7, 1.6 Hz, 1H), 5.12 (s, 2H), 3.76 (s, 3H), 3.46 (dq, *J* = 13.7, 6.7 Hz, 1H), 1.11 (d, *J* = 6.9 Hz, 6H); ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.42 (s, 1H), 7.79-7.64 (m, 5H), 7.57-7.47 (m, 3H), 7.34 (td, *J* = 7.6, 1.7 Hz, 1H), 5.31 (s, 2H), 3.84 (s, 3H), 3.45-3.41 (m, 1H), 1.25 (d, *J* = 6.9 Hz, 6H); LC-MS $(M+H)^+$: m/z = 493.2.

Mutual exclusivity test

Mutual exclusivity experiments were performed to understand the mode of inhibition of I-138 on USP1-UAF1. The results were fit to modified Yontenani and Theorell equation (equation 1) as shown below:

$$\frac{v_{ij}}{v_0} = \frac{1}{(1+\frac{I}{K_i}+\frac{J}{K_j}+\frac{I\times J}{\alpha\times K_i\times K_j})}$$

Where Vij represents the enzyme initial velocity in the presence of both compounds at concentrations [I] and [J], and V0 represents the enzyme initial velocity in the absence of compounds. Ki and Kj define the binding affinities of two compounds I and J, respectively. The value of α defines the degree to which

binding of one compound perturbs the binding of the other one. If α=1, the two compounds bind completely independently of one another. If α >> 1, the two compounds bind with a mutual exclusive fashion. The binding of one compound prevents the binding of the other one. While if α < 1, the two compounds bind in synergetic way where the binding of one compound augments the binding of the other one.

CRISPR-Cas9 screen analysis

Comparison of screens performed in BRCA1/2 WT and mutant cell lines

FASTQ files are mapped to the reference file of guide sequences using custom python scripts. After mismatches and low-quality reads are removed, only perfect aligned reads are included in further analysis. The file of read count of sgRNAs and a design matrix file were created as inputs for MAGeCK MLE analysis[1]. Z/beta scores and FDR adjusted p-values are estimated by 100 rounds of permutation of the dataset.

UMI Bayesian Beta-binomial (UMIBB) analysis pipeline

1) Read count normalization and sgRNA level tests of depletion or enrichment events.

FASTQ files are mapped to the reference file consisting of both guide sequences and barcode sequences using custom scripts. After mismatches and low-quality reads are removed, only aligned reads are used to measure the abundance of UMIs. Suppose we are to compare UMI counts between an experimental condition (E) and a control condition (C) in which cells were infected by the same CRISPR-UMI library. Raw counts of UMIs are normalized by a sequencing depth factor estimated by the ratio of total UMI counts between the two conditions. The combined normalized counts of the two conditions for each UMI(*i*) are required to be greater than a threshold (>=3 reads at default) to remove background noise UMIs with few reads. A pseudo-count of 0.5 was added if there were no reads for a UMI in either experimental or control condition. For each UMI(*i*) of sgRNA (*s*), the clonal fold changes (FC_{si}) is calculated by equation (1),

$$FC_{si} = \frac{n_{Esi}}{n_{Csi}} \quad (1)$$

Where n_{Esi} is the normalized counts of UMI(*i*) in E and n_{Csi} is the normalized counts of UMI(*i*) in C.

For a given sgRNA (*s*), there are m_s distinct barcodes (clones) in E and C combined, d_s number of clones show count depletion (FC_{si}<1) and e_s number clones show count enrichment FC_{si}>1 in the comparison between condition E vs C.

Within each individual sgRNA (*s*), the likelihood of encountering d_s number of depletion events out of m_s independent observations can be characterized by a binomial distribution:

$$p(X_s = d_s|\theta) = \binom{m_s}{d_s} \theta^{d_s}(1-\theta)^{m_s-d_s} \qquad (\theta \in [0,1]) \quad (2)$$

where the event probability θ follow a conjugate beta prior distribution:

$$p(\theta) = Beta(\theta|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} \quad (3)$$

where Γ is the gamma function. α and β are two positive parameters, which can be estimated using the observed number of depletion events in the non-targeting sgRNA guides. The mean depletion probability of prior (θ_{0d}) are usually equal or very close to 0.5, because the clones of non-targeting sgRNAs are expected to have same chance of enrichment or depletion if there is no systematic bias.

A posterior model of θ can then be constructed by combining the information from the prior with the information from the clonal depletion events observed in the given sgRNA(*s*), which is:

$$p(\theta|X_s = d_s) \propto p(X_s = d_s|\theta)p(\theta) \quad (4).$$

Markov chain Monte Carlo (MCMC) approximation is used to approximate the posterior distribution of θ. For each sgRNA(s), the MCMC estimates the mean posterior θ of depletion events (θ_{sd}) and the probability (p_{sd}) that θ_{sd} is higher than the mean prior (θ_{0d}) estimated from the depletion events of non-targeting sgRNAs. $\theta_{sd} > \theta_{0d}$ indicates that the clones in sgRNA(s) have p_{sd} probability of showing higher count depletion events frequency than expected from the non-targeting sgRNA clones.

Similarly, we can model the number of enrichment events of each sgRNA using a similar approach. The mean posterior θ of enrichment events (θ_{se}) and the probability (p_{se}) that θ_{se} is higher than the median prior θ_{0e} are derived accordingly.

2) Gene level tests by combining sgRNA level statistics

Suppose sgRNA (*s*) is one of the *k* sgRNAs (*s*= 1, 2, …, *k*) targeting a gene (g), the gene level test statistics of the depletion test (θ_{gd}, p_{gd}) and enrichment test (θ_{ge}, p_{ge})are calculated using the Stouffer's Z-score method weighted by the number of clones in each sgRNA (m_s).

$$\theta_{gd} = \frac{\sum_{s=1}^{k} m_s \theta_{sd}}{\sum_{s=1}^{k} m_s} \quad , \quad p_{gd} = 1 - \Phi\left(\frac{\sum_{s=1}^{k} m_s \Phi^{-1} p_{sd}}{\sqrt{\sum_{s=1}^{k} m_s^2}}\right)$$

$$\theta_{ge} = \frac{\sum_{s=1}^{k} m_s \theta_{se}}{\sum_{s=1}^{k} m_s} \quad , \quad p_{ge} = 1 - \Phi\left(\frac{\sum_{s=1}^{k} m_s \Phi^{-1} p_{se}}{\sqrt{\sum_{s=1}^{k} m_s^2}}\right) \quad (5)$$

Notice that $\theta_{gd} + \theta_{ge} \sim 1$, since there are usually very few clones with FC_{si}=1. We further consolidate the statistics by using the lower p-value from the depletion and enrichment tests as the gene-level p-value of UMIBB test. The relative frequency of clonal enrichment frequency (θ_g) is defined by equation (6), which is plotted as the effect size of genes in the volcano-plot.

$$p_g = \min(p_{gd}, p_{ge})$$

$$\theta_g = \begin{cases} 1 - \theta_{gd}, p_{gd} < p_{ge} \\ \theta_{ge}, p_{gd} \geq p_{ge} \end{cases} \quad (6)$$

crUMI analysis

The raw counts of UMIs after mapping and QC were also analyzed using an analysis pipeline described previously[2]. In summary, the median of the clonal fold changes (FC_{si}) is used as the fold change estimate of each sgRNA guide. The p-values are estimated by the MAGeCK RRA test[3] using UMIs counts as inputs[1]. Negative test p-values are used as the guide-level p-value if FC<1, otherwise positive test p-value are used as the guide-level p-value. A summary guide-level crUMI score is the product of $\log_{10}$FC and minus $\log_{10}$pvalue of the guide. The gene-level crUMI score is the average of crUMI scores of guides targeting a particular gene.

CRISPR analysis using guide-level data without UMI

sgRNA counts are derived from aggregated UMIs counts of each sgRNA. Beta scores and FDR adjusted p-values are estimated by a conventional CRISPR analysis pipeline (MAGeCK MLE) using the sgRNA count as inputs. Similarly, we performed the MAGeCK RRA test[3]. Similarly, the p-values of the MAGeCK RRA test are derived from the Negative test p-values or positive test p-values depends on gene level fold change. To compare with crUMI score, , we defined a conventional MAGeCK test score by the product of $\frac{log_{10}FC}{quantile(abs(log_{10}FC),95\%)}$ and $\frac{log_{10}(-pvalue)}{quantile(-log_{10}pvalue,95\%)}$.

High content immunofluorescence analysis

For γH2AX immunofluorescence, cells were rinsed with PBS 2X and fixed with 4% paraformaldehyde/0.25% Triton X-100 in PBS for 15 min, washed with PBS twice and permeabilized with 0.5% triton X-100/PBS for 10 min at room temp. Cells were washed with PBS and incubated with blocked with 10% goat serum in PBS for 1h at room temp followed by incubation with primary antibody against Anti-phospho Histone H2AX Ser 139 (1:500 dilution) in antibody buffer (3%BSA/0.05% triton/PBS) overnight at 4°C. Cells were washed with PBS-T and incubated with secondary antibody, alexa-488 conjugated goat anti rabbit IgG (Invitrogen, A-11008, 1:1000) in antibody buffer along with Hoechst 33342 (Thermofisher, H3570, 1:1000 dilution) for 1h at room temp protected from light. Cells were washed with PBS-T and plates were sealed. All cells were imaged on Phenix high content imaging system, and image analysis was performed by using Harmony software v4.9.

REFERENCES

1. Li, W. *et al.* Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol* 16, 281 (2015).

2. Michlits, G. *et al.* CRISPR-UMI: single-cell lineage tracing of pooled CRISPR–Cas9 screens. *Nat Methods* 14, 1191–1197 (2017).

3. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15, 554 (2014).