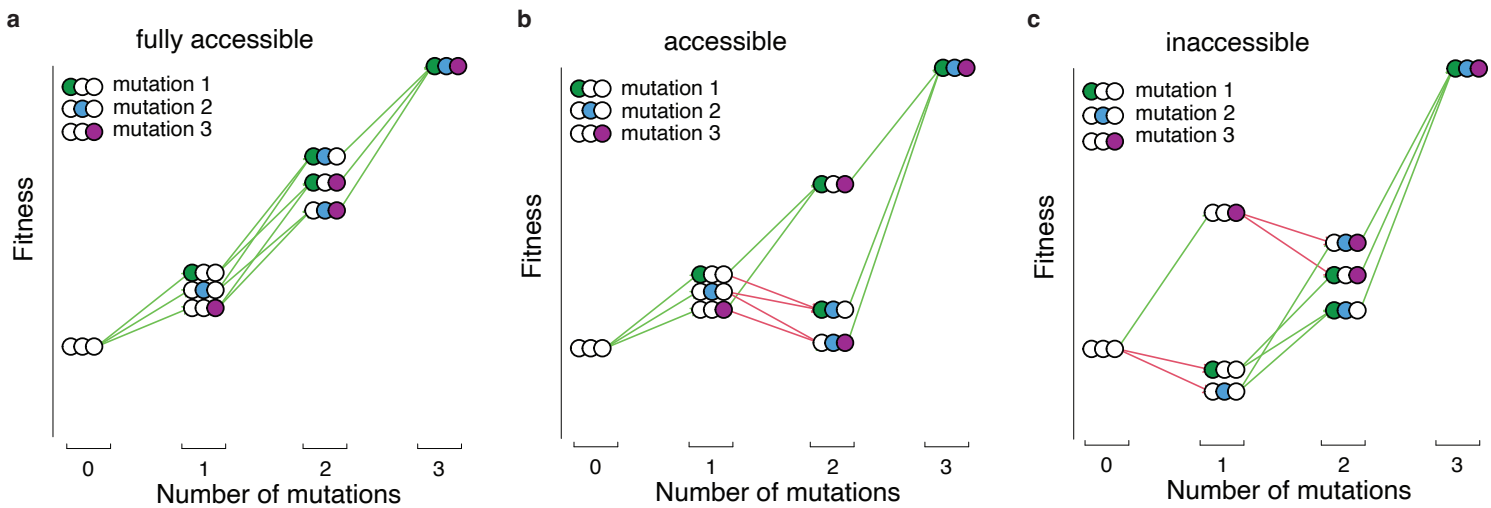
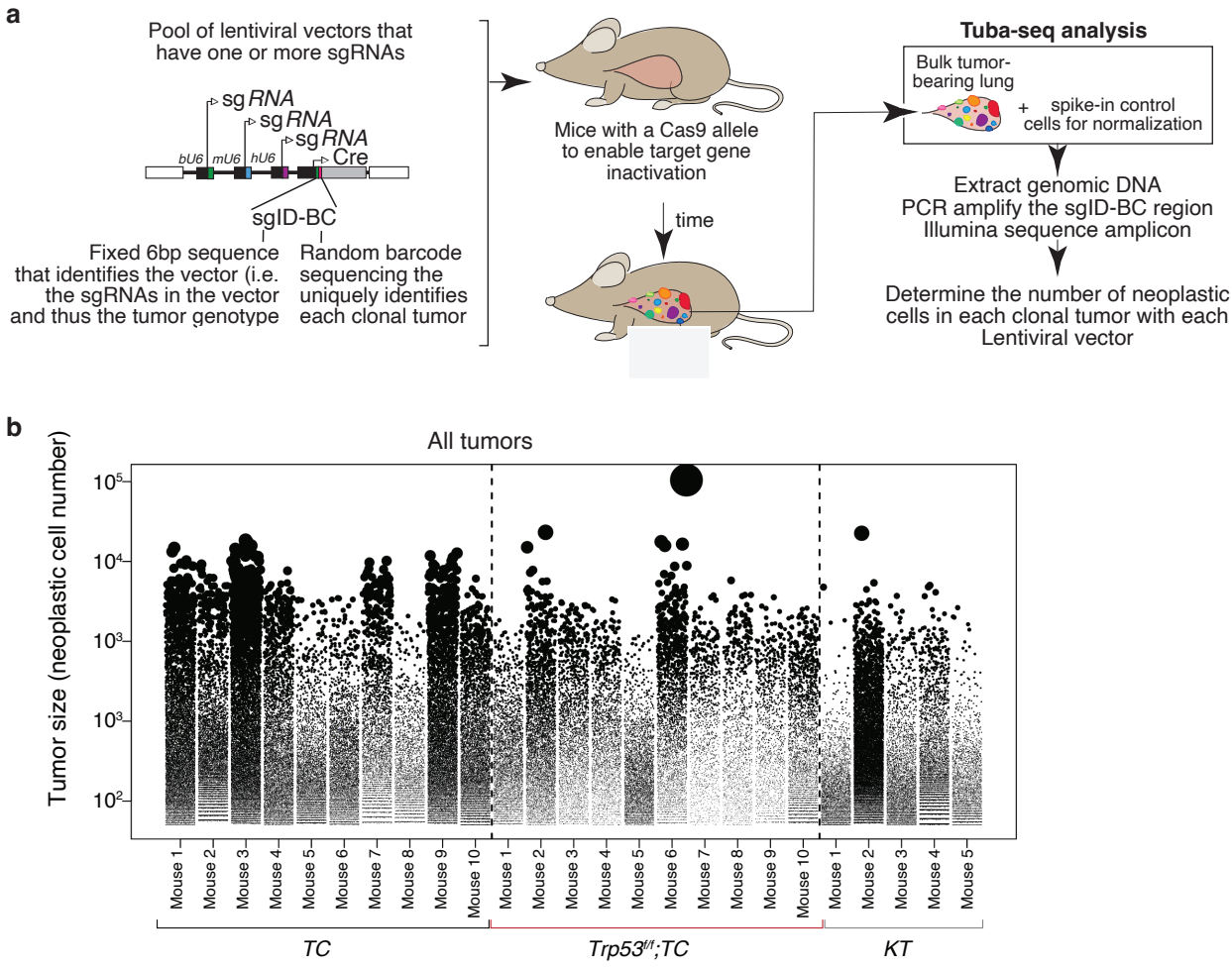
Supplementary Figures

Fully accessible fitness landscape of oncogene-negative lung adenocarcinoma

Maryam Yousefi, Laura Andrejka, Márton Szamecz, Monte M. Winslow, Dmitri A. Petrov, and Gábor Boross



Supplementary Figure 1. Accessibility on fitness landscapes
Hypothetical examples of a fully accessible (**a**), accessible (**b**) and inaccessible (**c**) fitness landscape. In a fully accessible landscape each mutation increases fitness across all other backgrounds, hence all 6 paths to the triple mutant genotype are possible. In an accessible landscape, not all mutaitons increase fitness across all other backgrounds, however, at least one order of mutations increases fitness with each new mutation and hence the triple mutant genotype can be achieved. In this example there are only 2 paths that can lead to the triple mutant genotype. In an inaccessible landscape despite the triple genotype being the most fit, there is no path where each sequential mutation increased fitness.



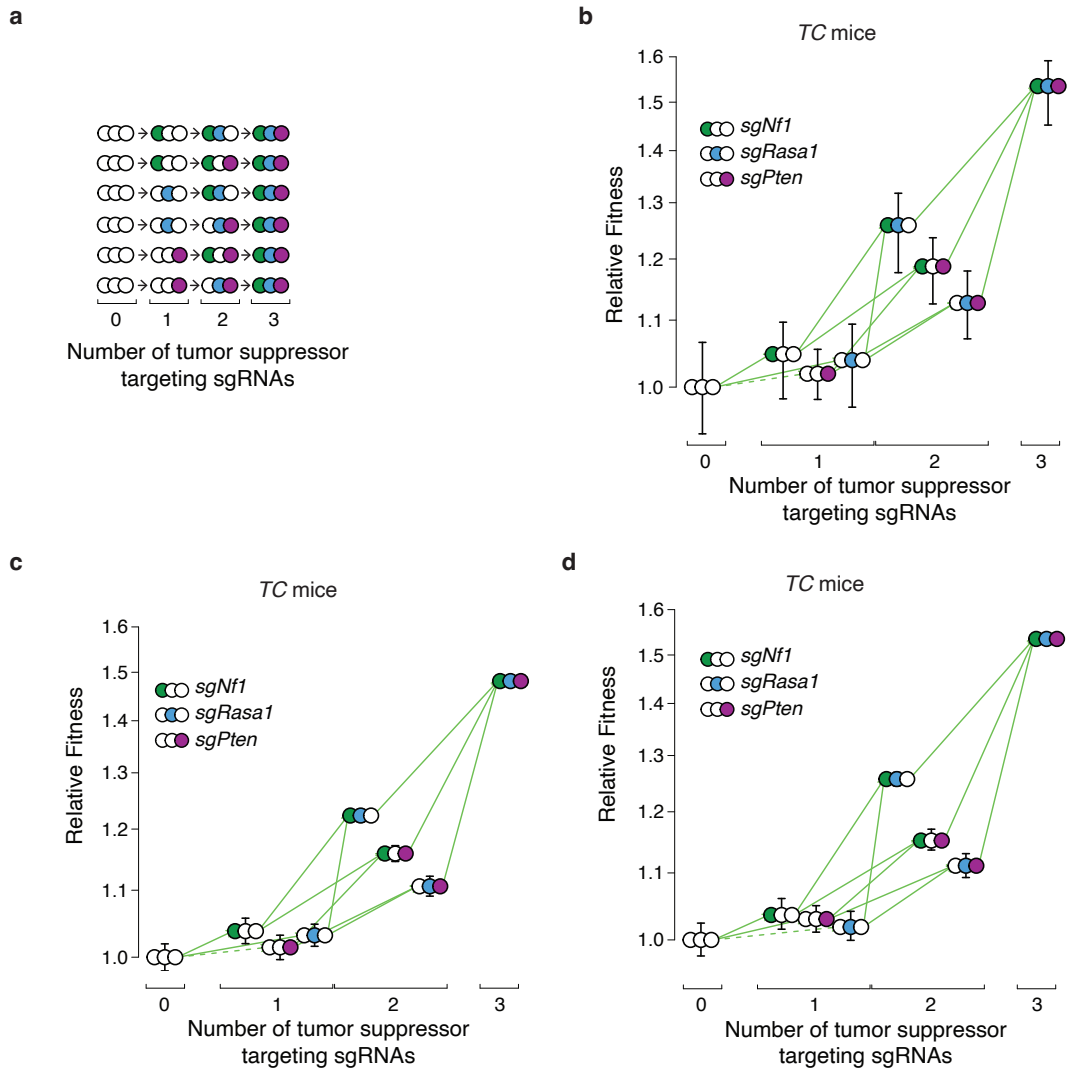
c

	TC	*Trp53$^{f/f}$;TC*	*KT*
Nf1;Rasa1;Pten	2438	1126	1836
Nf1;Rasa1	1317	525	1486
Nf1;Pten	1926	562	1887
Rasa1;Pten	1244	392	1520
Nf1	959	254	1706
Rasa1	1458	408	2740
Pten	699	223	1024
Inert	868	237	1634

d

	TC	*Trp53$^{f/f}$;TC*	*KT*
Nf1;Rasa1;Pten	191	335	111
Nf1;Rasa1	182	193	114
Nf1;Pten	137	142	112
Rasa1;Pten	131	142	112
Nf1	141	137	113
Rasa1	133	136	108
Pten	126	135	111
Inert	131	121	113

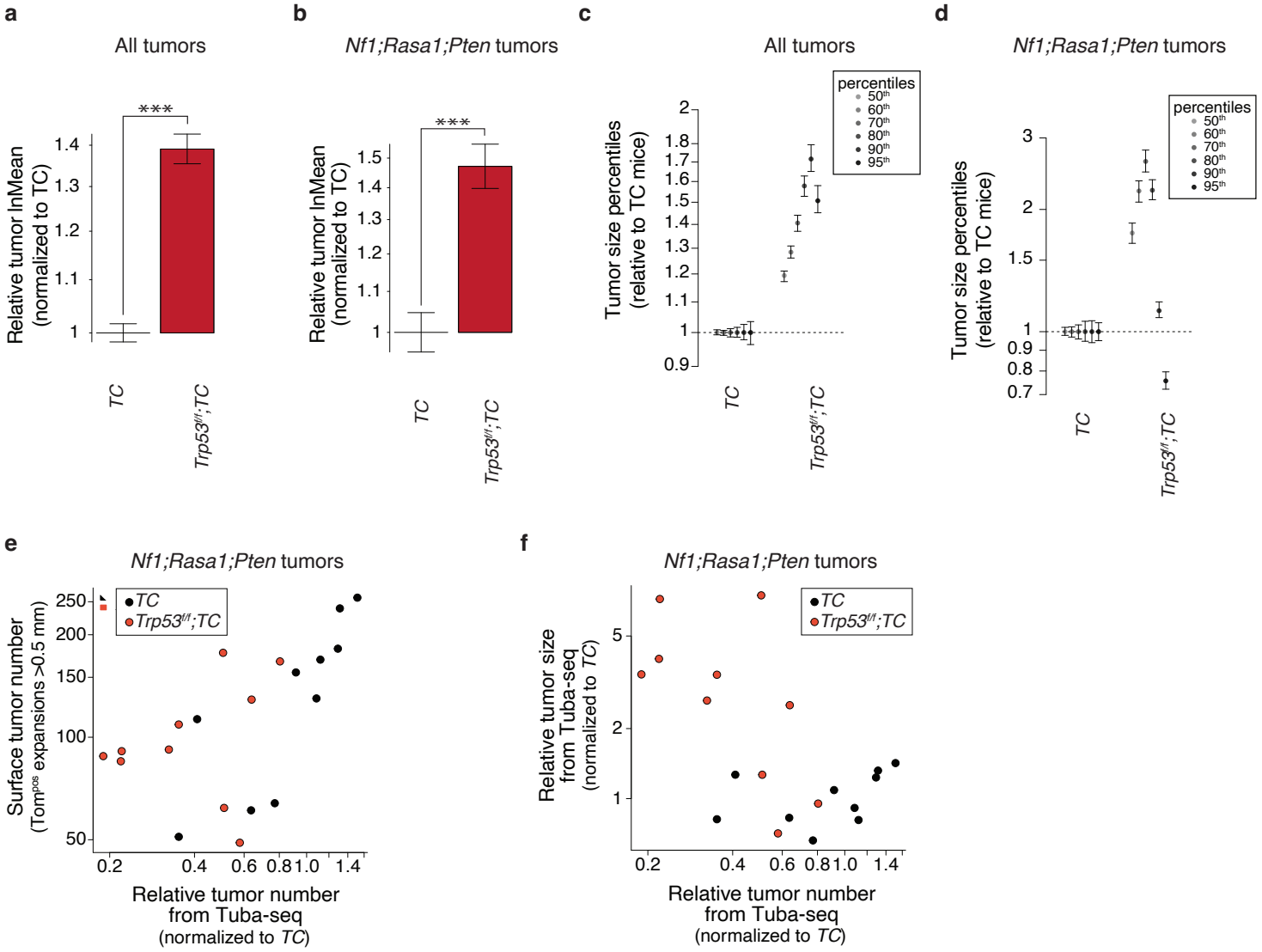
Supplementary Figure 2. *TC* and *Trp53$^{f/f}$;TC* mice with Lenti-sgTS$^{triple\text{-}pool}$/Cre-induced tumors develop many clonal tumors
a, Overview of tumor barcoding coupled with high-throughput barcode sequencing (Tuba-seq).
b, Gitter plot of tumor sizes in each mouse of each genotype. Each dot represents a tumor (>50 neoplastic cells). Sizes of dots scale with tumor size. Each column represents a mouse. All tumors generated with any of the lentiviral vectors are shown.
c-d, Average total number (**c**) and median size (**d**) of clonal tumors with each lentiviral vector in each mouse of each strain. Tumor numbers are the sum of all tumor with the given lentiviral vector across all mice of each genotype, divided by the number of mice of that genotype. Tumors are defined as clonal expansions estimated to have >50 neoplastic cells. These numbers reflect the total number of barcoded tumors prior to removal of tumors to reduce the impact of multiple transduction events (see **Methods**).



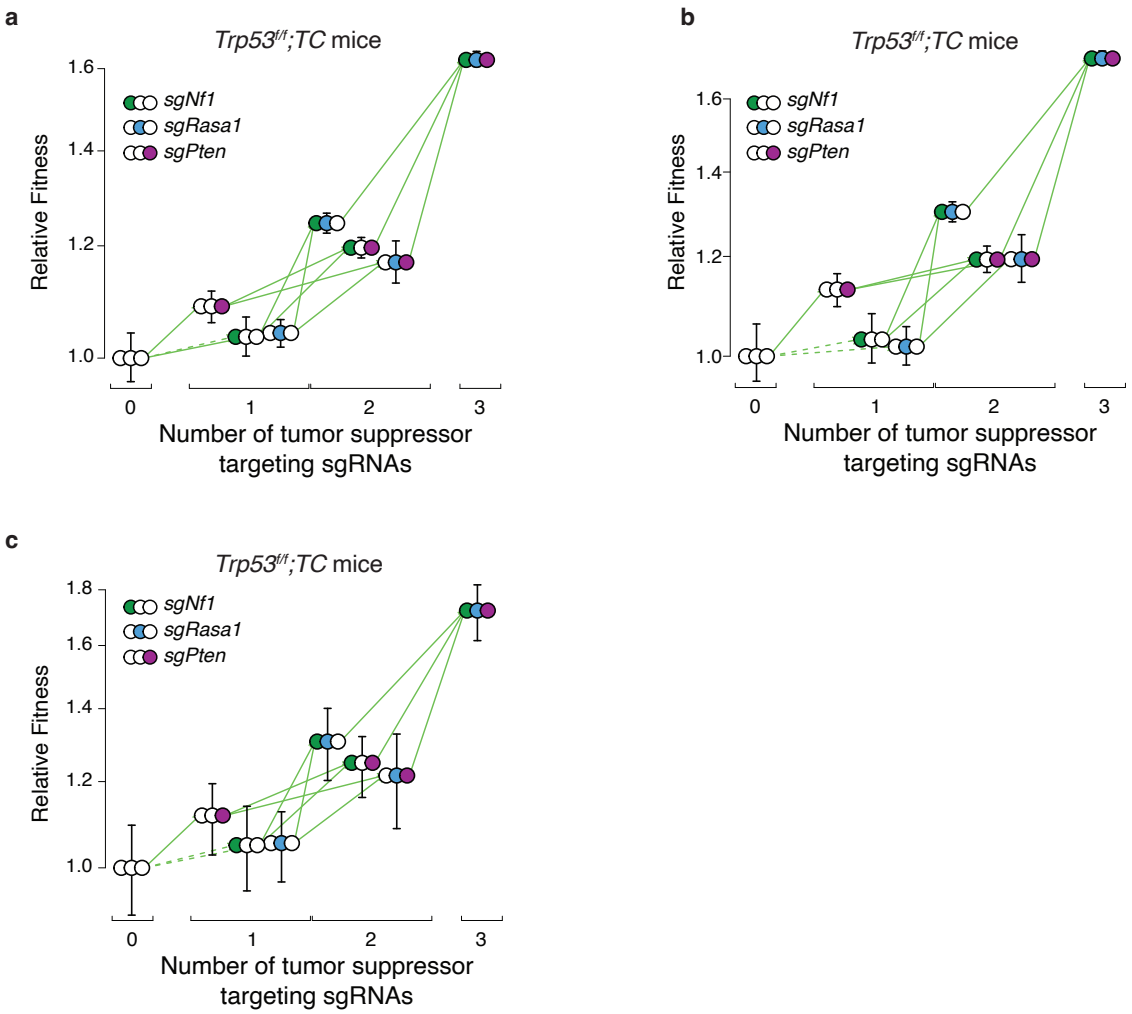
Supplementary Figure 3. Estimates of fitness for tumors in *TC* mice are robust to different methods for multiple transduction correction

a, There are six possible paths from wild type cells to the *Nf1;Rasa1;Pten* triple mutant genotype.

b-d, Fitness landscape of tumors in *TC* mice with multiple transduction correction as in Figure 1c but with 95% confidence intervals calculated through bootstrap resampling of both mice and tumors (**b**), without any multiple transduction corrrection (**c**), and with multiple transduction correction using method #2 (Methods)(**d**). Fitness for tumors of each single and double mutant genotype, as well as those with all three tumor suppressor targeting sgRNAs are shown relative to the triple inert vector. Green arrows indicate increased fitness, solid line indicates significance (p-value < 0.05). Whiskers in c and d show 95% confidence intervals from bootstrap resampling of tumors.

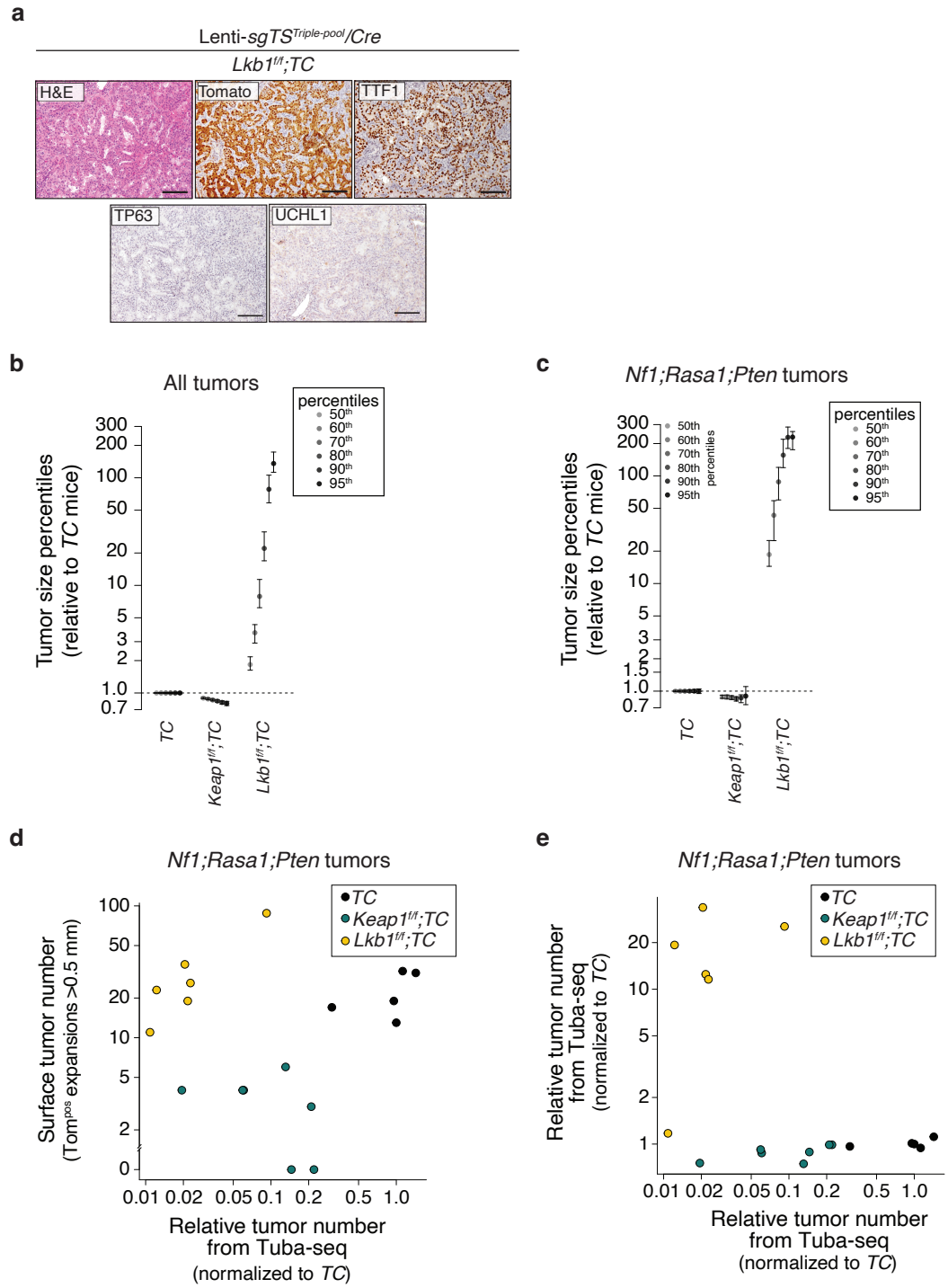


Supplementary Figure 4. *Trp53* inactivation increases tumor size but reduce tumor initiation/early expansion
a-b, Estimates of mean tumor size fitting a log-normal distribution (lnMean) are shown relative to the lnMean values in *TC* mice. Whiskers show 95% confidence intervals from bootstrap resampling. ***, $P < 0.001$ based on a bootstrap resampling.
c-d, Tumor sizes at indicated percentiles relative to the same percentile of tumor sizes in *TC* mice. Whiskers show 95% confidence intervals from bootstrap resampling.
e, Comparison of Surface tumor number (from direct counting of Tomato^{positive} tumors >0.5mm in diameter) with the Relative tumor number from Tuba-seq (the number of clonal expansions with >50 neoplastic cells relative to the median of *TC* values). Each dot represents a mouse. Data from *Nf1;Rasa1;Pten* triple mutant tumors is shown.
f, Comparison of the Relative tumor size determined from Tuba-seq (the median number of neoplastic cells relative to the median of *TC* values) with the Relative tumor number from Tuba-seq (the number of clonal expansions with >50 neoplastic cells relative to the median of *TC* values). Each dot represents a mouse. Data from *Nf1;Rasa1;Pten* triple mutant tumors is shown.



Supplementary Figure 5. Estimates of fitness for tumors in *Trp53$^{f/f}$;TC* mice are robust to different methods for multiple transduction correction.

a-c, Fitness landscape of tumors in *Trp53$^{f/f}$;TC* mice without any multiple transduction corrrection (**a**), with multiple transduction correction using method #2 (Methods)(**b**), and with multiple transduction correction as in Figure 2g but with 95% confidence intervals calculated through bootstrap resampling of both mice and tumors (**c**). Fitness for tumors of each single and double mutant genotype as well as those with all three genes inactivated are shown relative to inert. Green arrows indicate increased fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals.



Supplementary Figure 6. Inactivaton of the *Keap1* or *Lkb1* differentially impacts *Nf1;Rasa1;Pten* oncogene-negative tumor initiation/early expansion and tumor size

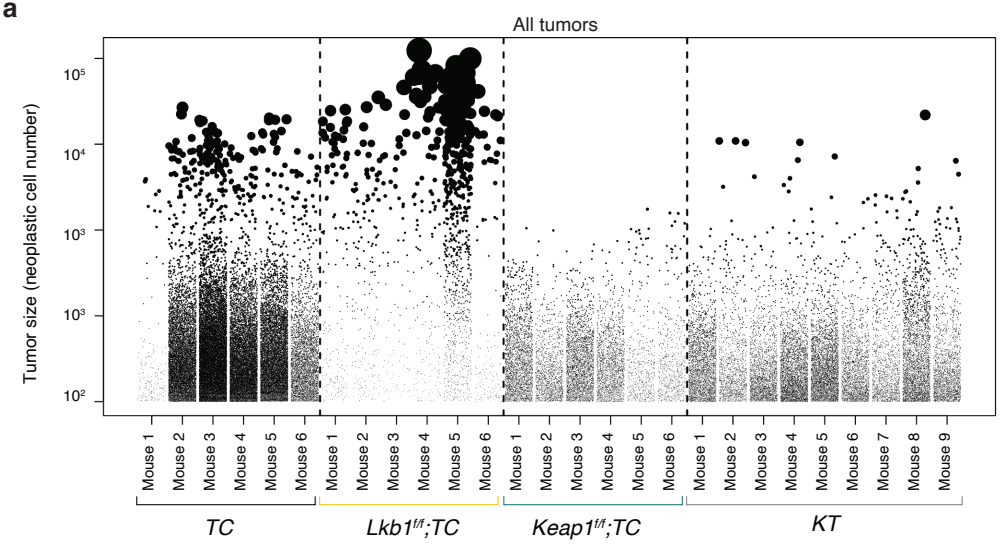
a, Representative H&E and Tomato, TTF-1 (adenocarcioma marker), P63 (squamouse cell carcinoma marker), and UCHL1(small cell lung cancer marker) stained sections of a lung tumor from a *Lkb1*$^{flox/flox}$*;TC* mouse. Scale bar = 100 µm.

b-c, Tumor sizes at indicated percentiles relative to the same percentile of tumor sizes in TC mice. Whiskers show 95% confidence intervals from bootstrap resampling.

d, Comparison of Surface tumor number (from direct counting of Tomato^{positive} tumors >0.5mm in diameter) with the Relative tumor number from Tuba-seq (the number of clonal exapansions with >50 neoplastic cells relative to the median of *TC* values). Each dot represents a mouse. Data from *Nf1;Rasa1;Pten* triple mutant tumors is shown.

e, Comparison of the Relative tumor size determined from Tuba-seq (the median number of neoplastic cells relative to the median of *TC* values) with the Relative tumor number from Tuba-seq (the number of clonal expansions with >50 neoplastic cells relative to the median of *TC* values). Each dot represents a mouse. Data from *Nf1;Rasa1;Pten* triple mutant tumors is shown.

a



b

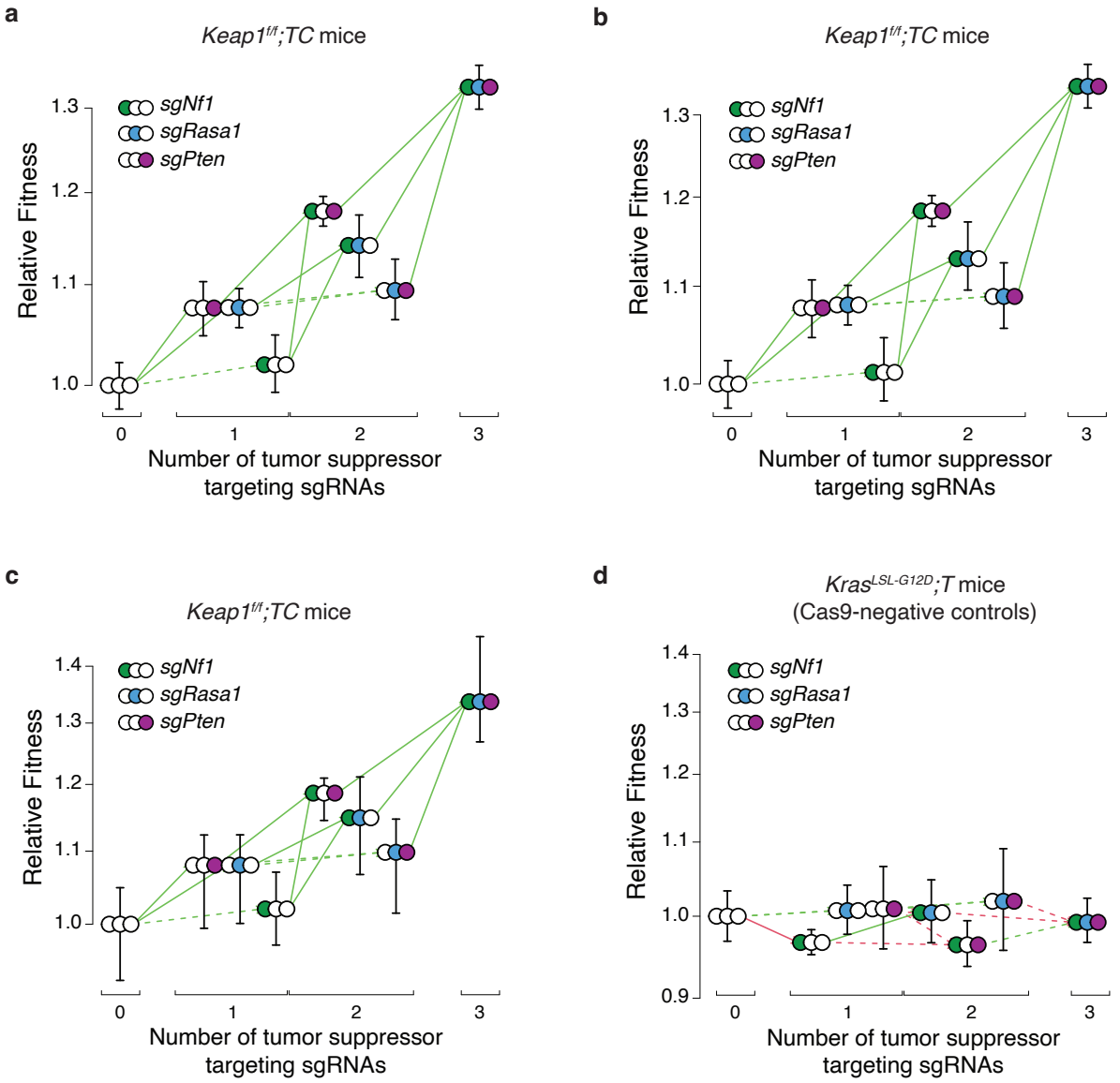
	TC	*Lkb1$^{f/f}$;TC*	*Keap1$^{f/f}$;TC*	*KT*
Nf1;Rasa1;Pten	2901	108	496	537
Nf1;Rasa1	669	39	90	131
Nf1;Pten	3198	66	526	594
Nf1	1974	21	315	697
Rasa1;Pten	1168	22	199	302
Rasa1	1854	24	284	552
Pten	825	12	137	223
Inert	1323	13	194	423

c

	TC	*Lkb1$^{f/f}$;TC*	*Keap1$^{f/f}$;TC*	*KT*
Nf1;Rasa1;Pten	109	2034	97	80
Nf1;Rasa1	91	660	88	82
Nf1;Pten	104	136	92	80
Rasa1;Pten	101	97	89	82
Nf1	103	94	90	84
Rasa1	98	76	88	82
Pten	100	82	93	82
Inert	99	68	93	83

Supplementary Figure 7. *Keap1* and *Lkb1* inactivation dramatically impacts the number and size of tumors initiated with Lenti-sgTS$^{triple\text{-}pool}$/Cre.

a, Gitter plot of tumor sizes in each mouse of each genotype. Each dot represents a tumor (>50 neoplastic cells). Sizes of dots scale with tumor size. Each column represents a mouse. All tumors generated with any of the lentiviral vectors are shown.
b-c, Average total number (**b**) and median size (**c**) of clonal tumors with each lentiviral vector in each mouse strain per mouse. Tumor numbers are the sum of all tumor with the given lentiviral vector avvcross all mice of each genotype, divided by the number of mice of that genotype. Tumors are defined as clonal expansions estimated to have >50 neoplastic cells. These numbers reflect the number of barcoded tumors prior to removal of tumors to reduce the impact of multiple transduction events (see **Methods**).

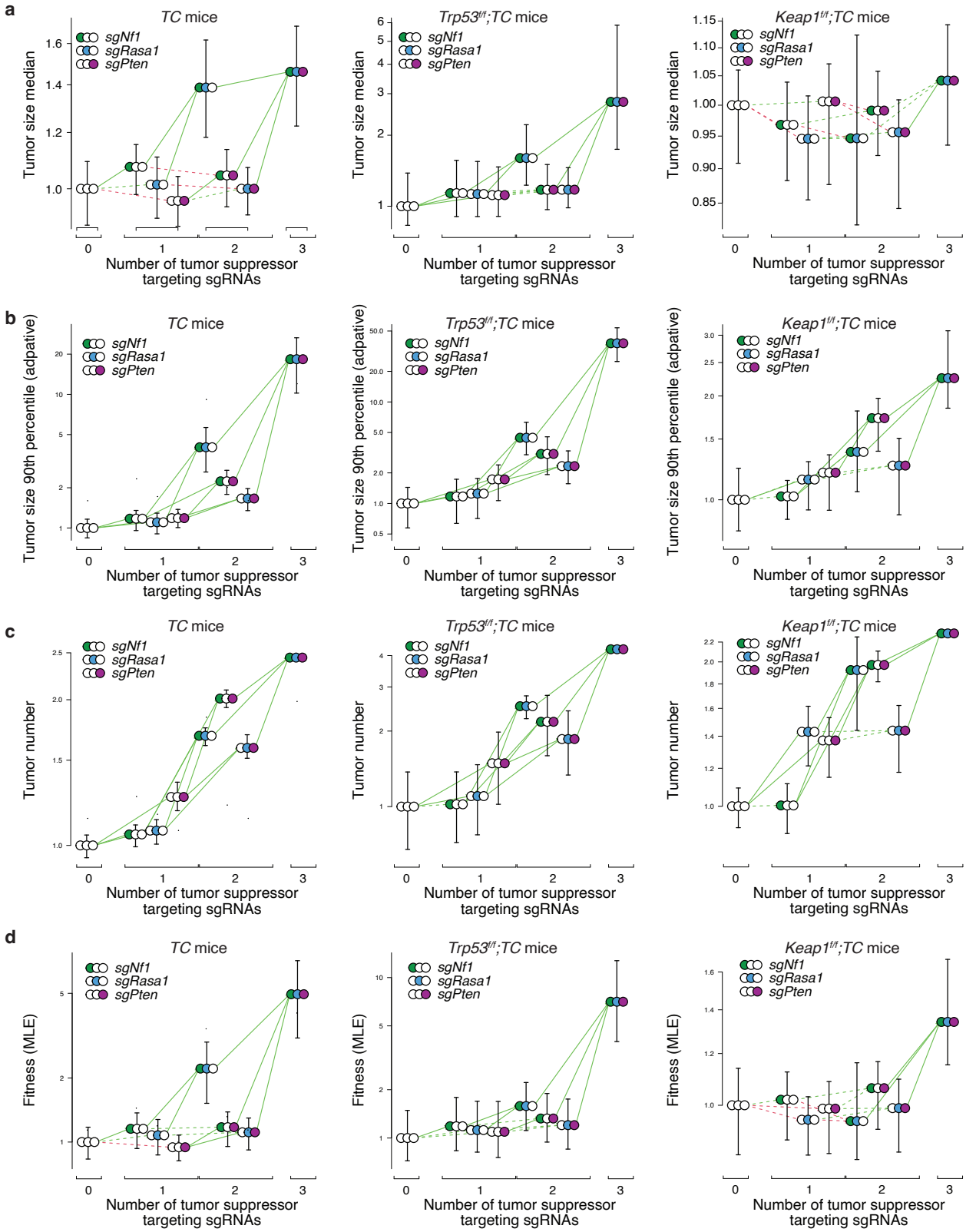


Supplementary Figure 8. Estimates of fitness are robust to methods that incorporate different methods for multiple transduction correction as well as different methods of

a-c, Fitness landscape of tumors in *Keap1$^{f/f}$;TC* mice without any multiple transduction corrrection (**a**), with multiple transduction using method #2 (Methods)(**b**) and with multiple transduction correction as in Figure 4b but with 95% confidence intervals calculated through bootstrap resampling of both mice and tumors (**c**). Green arrows indicate increased fitness, red arrows indicate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers in **b and c** show 95% confidence intervals from bootstrap resampling of tumors.

d, Fitness landscape for *Kras$^{LSL-G12D}$;T* (Cas9-negative control) mice. Fitness for tumors of each single and double mutant genotype, as well as those with all three tumor suppressor targeting sgRNAs are shown relative to the triple inert vector. Green arrows indicate increase in fitness, red arrows indicate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals from bootstrap resampling of tumors.

Fitness for tumors of each single and double mutant genotype as well as those with all three genes inactivated are shown relative to inert. Green arrows indicate increased fitness, red arrows inducate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals.



Supplementary Figure 9. Other estimates of fitness

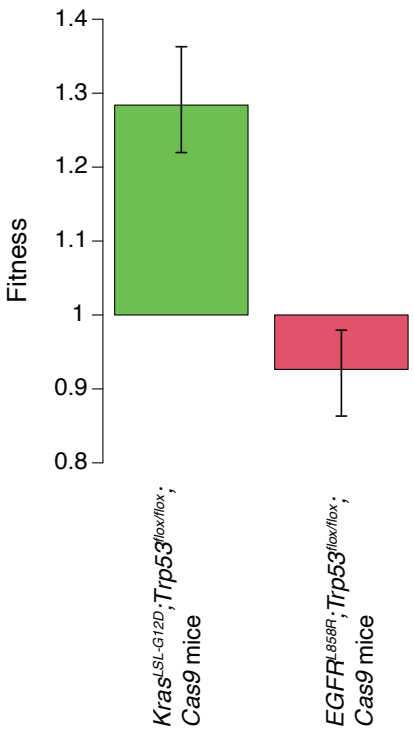
Adaptive landscape of tumors in mice of the indicated genotype with multiple transduction correction as in Figure 1c but with 95% confidence intervals calculated through bootstrap resampling of both mice and tumors. Instead of fitness (see Methods) other approximations of tumor growth are shown. Green arrows indicate increase, red arrows indicate decrease, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals from bootstrap resampling of tumors.

a, Median tumor size does not differentiate the genotypes well. Median size does not take into account tumor initation (tumor number). As an example, if a given genotype will generate half as many tumors as another, but those tumors show a similar size distribution, median tumor size will show no difference.

b, In contrast to simple median tumor size, here we take into account the expected number of tumors (adaptive cutoff) for a given genotype (expected_tumor_number$_{genotype}$ = titer$_{genotype}$ / titer$_{Nf1;Rasa1;Pten}$ x tumor_number$_{Nf1;Rasa1;Pten}$) and calculate 90th percentile tumor size as the size of the (expected_tumor_number$_{genotype}$ x 0.1)th tumor. This measure, similarly to our fitness measure, integrates both tumor number and tumor size, and reproduces fully accessible landscapes.

c, Landscapes based on entirely on tumor number (normalized to relative tumor titer based on *KT* mice, see Methods) are fully accessible.

d, We created an alternative fitness measure where using a maximum likelihood estimation (MLE) of mean tumor sizes assuming exponential growth of tumor genotypes with normally distributed growth rates (see Rogers *et al.,* 2017). This more complex fitness measure takes variable growth rates across tumors and within genotypes into account, but does not differentiate as well across genotypes (fewer significant changes in fitness across genotypes).



Supplementary Figure 10. Our fitness measure is able to detect sign epistasis
To test whether our experimental and computational pipeline is capable of detecting sign epistasis, we reanalyzed data from a past study (Foggetti *et al.* 2021), which also use the Tuba-seq methodology and employed our fitness measure to test for a known case of sign epistasis; Lkb1 inactivation has a positive effect on oncogenic KRAS-driven tumor growth (in *Kras*$^{LSL\text{-}G12D}$*;Trp53*$^{flox/flox}$*; Cas9* mice) and a negative effect on oncogenic EGFR-driven tumor growth (in *EGFR*^{L858R}*;Trp53*$^{flox/flox}$*;Cas9* mice). Our fitness measure find this sign epistasis (p values from left to right: < 0.001 and 0.004).