Fully accessible fitness landscape of oncogene-negative lung adenocarcinoma

Maryam Yousefi[1], Laura Andrejka[1], Monte M. Winslow[1,2,3*], Dmitri A. Petrov[2,4*], and Gábor Boross[4*]

[1] Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA
[2] Cancer Biology Program, Stanford University School of Medicine, Stanford, CA, USA
[3] Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA
[4] Department of Biology, Stanford University, Stanford, CA, USA

[*] Corresponding authors:
Monte M. Winslow, Stanford University School of Medicine | 279 Campus Drive, Beckman Center B256, Stanford, CA 94305. Phone: 650-725-8696 | Fax: 650-725-1534 E-mail: mwinslow@stanford.edu

Dmitri Petrov, Stanford University Department of Biology | 327 Campus Drive, Bass Biology 232, Stanford, CA 94305 | dpetrov@stanford.edu

Gábor Boross, Stanford University Department of Biology | 327 Campus Drive, Bass Biology 230A, Stanford, CA 94305. | E-mail: gaboross@stanford.edu

ABSTRACT

Cancer genomes are almost invariably complex with genomic alterations cooperating during each step of carcinogenesis. In cancers that lack a single dominant oncogene mutation, cooperation between the inactivation of multiple tumor suppressor genes can drive tumor initiation and growth. Here, we shed light on how the sequential acquisition of genomic alterations generates oncogene-negative lung tumors. We couple tumor barcoding with combinatorial and multiplexed somatic genome editing to characterize the fitness landscapes of three tumor suppressor genes NF1, RASA1, and PTEN, the inactivation of which jointly drives oncogene-negative lung adenocarcinoma initiation and growth. The fitness landscape was surprisingly accessible, with each additional mutation leading to growth advantage. Furthermore, the fitness landscapes remained fully accessible across backgrounds with additional tumor suppressor mutations. These results suggest that while predicting cancer evolution will be

challenging, acquiring the multiple alterations required for the growth of oncogene-negative tumors can be facilitated by the lack of constraints on mutational order.

INTRODUCTION

Fitness landscapes are a helpful conceptual aid in understanding evolutionary processes[1]. In a fitness landscape, genotypes are represented as nodes on a graph such that two genotypes are connected by an edge if they differ by a single allele and thus each genotype can be reached from the other by a single alteration[2]. Fitness values of each genotype are shown as an extra (height) dimension with adaptation represented as a hill-climbing adaptive walk. The key properties of evolutionary adaptation, such as its predictability and the ability of adaptation to reach the fittest genotype in a deterministic walk, are determined by the shape of the landscape, specifically by its ruggedness[3]. The ruggedness of the fitness landscape is defined by the frequency of sign epistasis, a phenomenon where a mutation is advantageous in one genetic context but disadvantageous in another. Sign epistasis determines whether an adaptive walk can end up stalled on a local fitness peak and whether only some combinations of mutations are sequentially adaptive. Intriguingly, many previous studies found that rugged fitness landscapes are common across diverse evolutionary systems[3–7]. This suggests that early adaptive mutations can strongly affect the direction of the adaptive process and limit the accessible adaptive paths.

Cancers are driven by the sequential acquisition of genetic driver mutations and thus are canonical examples of adaptive walks on a fitness landscape. Nonetheless, full fitness landscapes for cancer remain poorly understood due to the difficulty of inferring the fitness of growing tumors, the combinatorially large numbers of intermediate genotypes for even moderately genetically complex tumors, and the fact that tumors are generally analyzed once they have

57 accumulated a large number of mutations with the order of the early steps hidden from the study.
58 Despite the importance of defining the properties of cancer fitness landscapes for understanding
59 tumorigenesis, we currently lack experimental systems to quantify the fitness of nascent tumors
60 of diverse sequential genotypes growing *in vivo*. Defining cancer fitness landscapes could help to
61 determine cancer progression and its therapeutic vulnerabilities. Human cancer genomics has
62 uncovered genomic alterations that are mutually exclusive with one another, and we and others
63 have found sign epistasis between oncogene and tumor suppressor mutations that suggests that
64 cancer fitness landscapes might be rugged and not fully accessible[8–10].

65 Oncogene-negative tumors represent ~20% of lung adenocarcinoma cases and affect
66 ~150,000 patients each year worldwide[11–14]. These tumors lack oncogene mutations, and a large
67 fraction of them could be driven by the combinatorial inactivation of tumor-suppressive
68 genes[14,15]. We recently identified combinatorial inactivation of *Nf1, Rasa1*, and *Pten* and
69 activation of MAPK and PI3K pathway as a potent driver of a subset of oncogene-negative lung
70 tumors [14]. While oncogene-negative tumors represent a large fraction of lung cancer, they likely
71 require the coincident inactivation of multiple tumor suppressor genes to drive growth
72 comparably to a single oncogene alteration. This raises the question of the extent to which
73 genetic interaction among individual tumor suppressor mutations affect fitness during cancer
74 evolution. Does inactivation of each gene confer a growth benefit regardless of genetic context,
75 and hence form a fully accessible fitness landscape, or do they form a rugged fitness landscape
76 where only certain orders of mutations are favored by selection?

77 Here, we characterized the full fitness landscape of autochthonous oncogene-negative
78 lung tumors with combinatorial inactivation of the NF1, RASA1 and PTEN tumor suppressor
79 genes. We used combinatorial and multiplexed somatic CRISPR/Cas9 genome editing and tumor

80 barcoding to quantify the accessibility of the Nf1, Rasa1, Pten triple-mutant genotype via all
81 possible single steps. We further investigated how the shape of the fitness landscape is affected
82 by the inactivation of other tumor suppressor genes. We uncover an unexpectedly smooth
83 landscape that remains accessible even in the presence of additional tumor suppressor alterations.
84

85 **RESULTS**

86 **Quantitative analysis of the size and number of tumors of each stepwise genotype towards**
87 ***Nf1;Rasa1;Pten* triple-mutant lung cancer**

88 To quantify the impact of diverse genotypes on tumor initiation and growth, we
89 developed a method based on tumor barcoding coupled with high-throughput barcode
90 sequencing (Tuba-seq)[16–18]. Tuba-seq uses lentiviral-based DNA barcoding coupled with
91 CRISPR-Cas9 genome editing and genetically engineered mouse alleles to generate and quantify
92 the number and size of tumors of many different genotypes in parallel (**Supplementary Fig. 1a)**.
93 Each lentiviral vector encodes Cre recombinase, a sgRNA, as well as a two-component barcode
94 region with a fixed sgRNA identifier (sgID) that labels the vector and a random barcode (BC)
95 that uniquely tags each clonal tumor that arises from the initial transduced cell.

96 We previously used Tuba-seq to identify combinations of tumor suppressor genes whose
97 inactivation is capable of driving lung adenocarcinoma in the absence of oncogene mutations.
98 Specifically, we discovered that coincident inactivation of *Nf1*, *Rasa1*, and *Pten* is a potent driver
99 of oncogene-negative lung adenocarcinoma[14]. To investigate the contribution of single, double,
100 and triple mutations of *Nf1*, *Rasa1*, and *Pten* to the genesis of oncogene-negative lung
101 adenocarcinomas, we previously initiated tumors in $R26^{LSL\text{-}Tomato}$;$H11^{LSL\text{-}Cas9}$ (*TC*) and
102 $Trp53^{f/f}$;*TC* mice with a pool of eight uniquely dual barcoded lentiviral vectors with sgRNAs

103 targeting *Nf1, Rasa1,* and *Pten* (**Fig. 1a)**. This pool contains lentiviral triple sgRNA vectors that
104 target each gene alone, in pairwise combinations, and all three together (Lenti-sg$TS^{Triple\text{-}pool}$/*Cre*,
105 **Fig. 1a**)[14]. These mice developed thousands of clonal tumors of different genotypes
106 (**Supplementary Fig. 1b-c**). Our previous Tuba-seq analysis indicated that most of the tumor
107 burden in *TC* and *Trp53*$^{f/f}$*;TC* mice was from tumors with concomitant inactivation of all three
108 tumor suppressor genes[14]. However, we had not previously used our high-resolution tumor
109 barcoding data to build a fitness landscape to understand the potential evolutionary paths to the
110 triple mutant state, nor had we compared the tumorigenic potential between the p53-proficient
111 and -deficient backgrounds.
112

113 **The fitness landscape of *Nf1;Rasa1;Pten* triple-deficient lung cancer is completely**
114 **accessible**

115 Generating Tuba-seq data on tumors initiated with lentiviral vectors that create all
116 possible combinatorial mutations of *Nf1*, *Rasa1*, and *Pten* should enable the generation of a
117 complete fitness landscape across all genotypes. We approximated the relative fitness of each
118 genotype by calculating the growth rate using data on tumor number and size (**Fig. 1b** and
119 **Methods**). Across these three tumor suppressor genes there are six possible paths on the
120 landscape, *i.e* six possible routes from the genotype with the lowest fitness to the one with the
121 highest fitness (**Supplementary Fig. 2a)**. As anticipated, we found that *Nf1;Rasa1;Pten* triple
122 mutant tumors had the highest relative fitness in *TC* mice (**Fig. 1c**). Interestingly, all mutational
123 routes leading to the *Nf1;Rasa1;Pten* triple mutant genotype were accessible, with each
124 additional mutation increasing fitness beyond the fitness of the less complex genotypes of origin
125 (**Fig. 1c**). These results are consistent with a model in which inactivation of each tumor

126 suppressor gene is advantageous on all backgrounds regardless of whether the other two tumor
127 suppressor genes are wild-type or mutant.
128 To assess the robustness of our relative fitness estimates, we performed a series of
129 additional experiments. As an initial negative control for our method, we analyzed the tumors
130 initiated with Lenti-sg*TS*$^{Triple\text{-}pool}$*/Cre* in control *Kras*$^{LSL\text{-}G12D}$*;T* (*KT*) mice, which lack Cas9 (**Fig.**
131 **1a**). In these mice, all sgRNAs are inert, and tumors with different sgIDs have the same
132 genotype. As anticipated, our method uncovered a fitness landscape that was not accessible and
133 that lacked the gains in fitness for tumors initiated with vectors with sgRNA targeting increasing
134 numbers of tumor suppressor genes (**Fig. 1d**). Next, we performed a different method in which
135 we used bootstrap resamples of mice as well as tumors, which had minimal impact on the overall
136 significance of the fitness differences across the landscape in *TC* mice (**Supplementary Fig. 2b**).
137 In lung cancer models using lentiviral vectors, a small percent of tumors arise from cells
138 that were transduced by more than one vector[14,16]. In the context of our experiments to map
139 fitness, these multiple transduction events could impact tumor growth (except for tumors with
140 the *Nf1;Rasa1;Pten* triple vector as they have the maximum numbers of possible mutations). We
141 know that multiple infections must occur as we observe a small number of tumors containing the
142 triple sgInert vector. These tumors are certainly the result of this vector co-transducing an initial
143 cell with a vector that encodes one or more tumor suppressor gene targeting sgRNAs. We
144 performed a series of analyses taking into account the expected rates of multiple transduction
145 events and corrected our data to account for this effect. The number of sgInert tumors in *TC* mice
146 was used to calculate the rate of coinfection in our experiments. We then used the rate of
147 coinfection to remove a fraction of tumors of each genotype, picking the tumors to remove based
148 on the size distribution of *Nf1*;*Rasa1*;*Pten* triple mutant size distribution (**Methods**). Importantly,

149 the landscape remained accessible in the absence of correction for multiple transduction events
150 (**Supplementary Fig. 2c**) as well as when we used an alternate method to correct for the impact
151 of multiple transduction events (**Supplementary Fig. 2d**, see **Methods**). Our results show that
152 the observed accessibility of the fitness landscape is robust to different methods to correct for
153 multiple transduction events. These results are also consistent with experimental evidence from
154 our previous study in which single and double mutant genotypes lead to some neoplastic growths
155 *in vivo*[14].
156
157 ***Trp53*-deficiency reduced tumor initiation without impacting the fully accessible landscape**
158 In oncogenic KRAS, EGFR, and BRAF-driven lung cancer models, *Trp53* inactivation
159 increases tumor initiation/early tumor growth and tumor size[9,17,19]. Our data allowed us to test
160 whether Trp53 inactivation has a similar effect on the growth of oncogene-negative tumors,
161 especially in case of the potent *Nf1;Rasa1;Pten* triple-deficient tumors. Similarly to the effects in
162 the oncogene-driven tumors, multiple metrics of tumor size indicated that all tumors, as well as
163 the *Nf1;Rasa1;Pten* triple-deficient tumors were larger in *Trp53*$^{f/f}$*;TC* mice relative to *TC* mice
164 (**Fig. 2a-b,f** and **Supplementary Fig. 3a-d,f**). In contrast, *Trp53* inactivation reduced overall
165 tumor number (*i.e* number of tumors estimated to have >50 neoplastic cells), as well as
166 *Nf1;Rasa1;Pten* triple-deficient tumor number (**Fig. 2c-d,f** and **Supplementary Fig. 3f**). These
167 results agree with our previous visual analysis of surface tumor number in these mice[14] (**Fig. 2e**
168 and **Supplementary Fig. 3e**). However, the reduction in tumor initiation is the opposite of the
169 effect of *Trp53*-deficiency in oncogenic KRAS, BRAF and EGFR-driven lung cancer
170 models[8,9,17]. Thus, *Trp53* inactivation has a uniquely divergent effect on tumor initiation and
171 tumor size in oncogene-negative tumors.

172 Next, we mapped the fitness landscape of *Nf1, Rasa1,* and *Pten* mutations on the
173 background of *Trp53* deficiency in *Trp53*$^{f/f}$*;TC* mice. Similar to our observations in *Trp53*
174 proficient *TC* mice, the *Nf1;Rasa1;Pten* triple mutant genotype had the highest relative fitness.
175 Importantly, despite the overall effects of *Trp53* deficiency on tumor number and size, the
176 landscape remained fully accessible (**Fig. 2g** and **Supplementary Fig. 4a-c**).
177
178 **Quantification of the overall tumorigenicity of the *Nf1*;*Rasa1*;*Pten* triple-mutant genotype**
179 **across *Lkb1* and *Keap1* deficient backgrounds**
180 While the fitness landscape of *Nf1*;*Rasa1*;*Pten* lung tumors was fully accessible in *Trp53*-
181 proficient and -deficient backgrounds, changes in the genetic background often influence
182 epistatic interactions, and thus change landscape accessibility. Therefore, we next quantified
183 whether the effects of combinatorial *Nf1*, *Rasa1*, and *Pten* inactivation change on other genetic
184 backgrounds. We chose loss-of-function mutations in *Keap1* and *Lkb1*(*Stk11*) which are known
185 tumor suppressor genes that are frequently mutated in human lung cancer and have been shown
186 to suppress growth in models of oncogenic KRAS-driven lung adenocarcinoma[16,17,20–22]. We
187 initiated tumors with Lenti-sg*TS*$^{Triple\text{-}pool}$*/Cre* in *TC*, *Keap1*$^{f/f}$*;TC, Lkb1*$^{f/f}$*;TC* and control *Kras*$^{LSL\text{-}G12D}$*;T* mice (**Fig. 3a**). After 3 months of tumor growth, we first analyzed tumor burden by direct
189 fluorescent imaging and histology. Unexpectedly, *Keap1*$^{f/f}$*;TC* mice developed many fewer
190 tumors than *TC* mice as assessed by direct fluorescence and histology (**Fig, 3b-c**). Conversely,
191 *Lkb1*$^{f/f}$*;TC* mice visually had many tumors that appeared somewhat larger than those in *TC* mice
192 (**Fig. 3b-c**).
193 Next, we performed Tuba-seq on tumor-bearing lungs and quantified the effect of *Keap1*
194 and *Lkb1* deficiency on overall tumor initiation and size. *Keap1* inactivation decreased overall

195 tumor number by ~80%, while modestly reducing tumor size (**Fig. 3d-g** and **Supplementary**
196 **Fig. 5b-c**). The impact of *Keap1* inactivation on tumor number and size was consistent across all
197 tumors, as well as when we assessed only *Nf1;Rasa1;Pten* triple-deficient tumors (**Fig. 3d-g**).
198 Thus, *Keap1* inactivation has an overall negative effect on tumorigenesis driven by inactivation
199 of *Nf1, Rasa1*, and/or *Pten* with the greatest impact on tumor initiation and/or very early tumor
200 growth.

201 The impact of *Lkb1* inactivation was even more interesting. *Lkb1* inactivation decreased
202 overall tumor number by >95% (**Fig. 3f-g**), which seemed at odds with the gross examination of
203 the tumor-bearing lung. However, the small number of tumors that did form were much larger
204 than those in *TC* mice, with *Nf1;Rasa1;Pten* triple-mutant tumors in *Lkb1;TC* having >15-fold
205 greater median tumor size than in *TC* mice (**Fig. 3d-e**). This dramatic increase in the size
206 reconciles the visual observation of similar numbers of tumors in *Lkb1;TC* and *TC* mice, while
207 quantitative analysis by Tuba-seq data indicated that *Lkb1;TC* mice had many fewer tumors (**Fig.**
208 **3h-i** and **Supplementary Fig. 5d-e**).
209

210 **Landscape accessibility is largely robust to changes in the genetic background**

211 As the *Lkb1* and *Keap1* deficient backgrounds had substantial effects on overall
212 oncogene-negative lung tumor growth in our experimental model, we next characterized their
213 impacts on the fitness landscapes. Due to the very low number of tumors for less complex (inert,
214 single and double mutant) genotypes on the *Lkb1*$^{f/f}$*;TC* background, we could not construct a
215 fitness landscape (**Supplementary Fig. 6a-b**). However, our method allowed us to generate
216 fitness landscapes for the tumors in *TC*, *Keap1;TC,* and control *KT* mice. The landscape in *TC*
217 mice was similar to that generated from our initial cohort of *TC* mice and the landscape in the

218 control *KT* Cas9-negative mice was again very flat (**Fig. 4a** and **Supplementary Fig. 7d**).
219 Interestingly, in the *Keap1*-deficient background, the landscape remained fully accessible (**Fig.**
220 **4b**). These effects were consistent across our different analysis methods (**Supplementary Fig.**
221 **7a-c**). Thus, despite greatly reducing tumor initiation, the ability of each additional mutation to
222 increase fitness on the path to the *Nf1;Rasa1;Pten* triple deficient state was maintained.
223 Despite the overall accessibility of the fitness landscape in both *Keap1*-proficient and -
224 deficient backgrounds, the impact of additional mutation was not identical. In particular, the
225 relative fitness of tumors of the different double mutant genotypes was different, with *Nf1;Rasa1*
226 being most fit in the *TC* and *Trp53;TC* background and *Nf1;Pten* being the most fit in the
227 *Keap1;TC* background (**Fig. 4c**). Thus, changes in the genetic background did impact the
228 topology of the landscape, but it remained fully accessible.
229
230 **DISCUSSION**
231 Our results suggest that the combinatorial inactivation of *Nf1, Rasa1,* and *Pten* forms a
232 fully accessible fitness landscape. Thus, any of the six evolutionary routes is possible, with the
233 inactivation of each gene increasing fitness beyond the preceding genotype. The overall lack of
234 sign epistasis makes predicting cancer evolution challenging. On the other hand, while
235 oncogene-negative tumors require the inactivation of multiple tumor suppressor genes to
236 substitute for the lack of a strong oncogene alteration, gaining fitness advantage from each
237 additional alteration increases the probability of acquiring all of the required mutations. Human
238 genomics data suggest that oncogene-negative lung adenocarcinoma with high MAPK and PI3K
239 pathway activation are likely generated from combinatorial mutations in diverse components of
240 these pathways further expanding the available paths to the tumorigenic state[14].

241 Contrary to the fitness landscapes we mapped in this study, the landscapes of many other
242 evolutionary systems are ruggedand exhibit sign epistasis between mutations[3–7]. Mutations often
243 have pleiotropic effects, with some effects being advantageous and other effects being
244 detrimental. In one genetic context the beneficial effects can outweigh the detrimental effects,
245 while in a different genetic context, this can be reversed. This leads to sign epistasis and a rugged
246 fitness landscape. Why is the oncogene-negative fitness landscape smooth and thus fully
247 accessible? We speculate that this could be because somatic cells and neoplastic cells *in vivo* are
248 far from the highest possible fitness and individual tumor suppressor gene mutations have
249 relatively weak effects. The combined effect could create sufficiently linear behavior of all the
250 individual tumor suppressor mutations without encountering any significant curvature of the
251 fitness function. This is often not the case for the large effect mutations observed in experimental
252 evolution of microorganisms or in the tumor suppressor effects on the oncogene-driven tumors.
253 While the fitness landscape of combinatorial *Nf1, Rasa1*, and *Pten* inactivation remained
254 fully accessible and robust to inactivation of *Trp53, Keap1* or *Lkb1*, those alterations
255 dramatically changed overall tumor initiation and growth. Comparing the effects of *Trp53,*
256 *Keap1* and *Lkb1* inactivation on *Nf1;Rasa1;Pten* mutant tumors with their impact on various
257 oncogene-driven tumors underscore the context-dependency of these effects[8,9,17]. *Trp53*
258 inactivation increased tumor size but had a negative effect on the initiation of oncogene-negative
259 tumors. This observation was in contrast to the universally positive effect of *Trp53* deficiency on
260 multiple models of oncogene-driven lung cancer[8,9,17,19]. Interestingly, while LKB1 and KEAP1
261 are generally regarded as tumor suppressors, several lines of evidence suggest that their role in
262 cancer is more complex and context dependent[24]. *Lkb1* and *Keap1* inactivation has been shown
263 to have detrimental effects on *in vivo* lung tumor initiation and growth across different genetic

264 contexts[8,9,25,26]. Thus, while the detrimental effect of *Lkb1* and *Keap1* inactivation on
265 initiation/early growth and the positive effect of *Lkb1* inactivation on the growth of oncogene-
266 negative tumors could not have been predicted, it is consistent with these "tumor suppressors"
267 having highly context-dependent effects.

268 Human tumors likely evolve through states in which tumor suppressor genes are
269 heterozygously inactivated. However, as CRISPR/Cas9-genome editing generally generates
270 homozygous gene inactivation, we could only study homozygous mutations in tumor suppressor
271 genes. In the future, deciphering where the multiple middle steps of heterozygous mutants fall
272 within cancer fitness landscapes would improve the resolution. The effects of heterozygous
273 genotypes are likely to fall within the range of their homozygous counterpart and therefore are
274 unlikely to render the fitness landscape inaccessible.

275 Most human tumors have complex genotypes, often including multiple inactivating
276 alterations. Modeling complex cancer genotypes in genetically engineered mouse models has
277 been greatly facilitated by somatic CRISPR/Cas9 genome editing, however broad-scale
278 quantitative analysis of complex cancer genotypes remains limited[14,19]. Generating defined
279 combinatorial alterations *in vivo* allows not only the analysis of tumorigenesis driven by these
280 complex genotypes, but also the assessment of contribution of each stepwise genomic alteration
281 to tumor fitness and evolution. Quantitative and comprehensive fitness maps of other genotypes
282 (within and beyond lung cancer) could uncover general rules of how driver alterations interact
283 and affect cancer evolution by creating rugged and smooth cancer fitness landscapes. While
284 libraries of vectors with pairs of sgRNAs have been broadly employed[27–31] generating large
285 pools of vectors to inactivate greater numbers of genes while maintaining the ability to perform
286 highly-quantitative clonal-level analysis will require additional methods. In the future, the

287 analysis of tumors with a comprehensive combination of many other genotypes coupled with
288 molecular profiling and external interventions like therapies should illuminate the underlying
289 biology of adaptation and uncover exploitable vulnerabilities.
290

291 **METHODS**

292 **Animal Studies**

293 The use of mice for the current study has been approved by the Institutional Animal Care
294 and Use Committee at Stanford University, protocol number 26696. *Kras*$^{LSL\text{-}G12D/+}$ (Jax # 008179
295 (*K*)), *R26*$^{LSL\text{-}tdTomato}$(*ai9*) (Jax # 007909 (*T*)), and *H11*$^{LSL\text{-}Cas9}$ (Jax # 026816), *Keap1*^{flox}, *Lkb1*^{flox}
296 (Jax # 014143), and *Trp53*^{flox} (Jax # 008462) mice have been previously described[32–38]. All mice
297 were on a C57BL/6:129 mixed background. Tumors were initiated by intratracheal delivery of
298 Lenti-Triple-sgRNA/*Cre* vectors. Mice were allowed to develop tumors for 3 months after viral
299 delivery.
300

301 **Tumor barcode sequencing and analysis**

302 Tuba-seq libraries were generated as described previously[14]. Briefly, genomic DNA was
303 isolated from bulk tumor-bearing lung tissue followed by PCR amplification of the sgID-BC
304 region from 32 μg of bulk lung genomic DNA using Q5 Ultra II High-Fidelity 2x Master Mix
305 (New England Biolabs, M0494X). Unique dual-indexed primers were used to amplify each
306 sample followed by purification using Agencourt AMPure XP beads (Beckman Coulter,
307 A63881). The libraries were pooled based on lung weights to ensure even reading depth and
308 sequenced (read length 2x150bp) on the Illumina HiSeq 2500 or NextSeq 500 platform (Admera

309 Health Biopharma Services). Tuba-seq analysis of tumor barcode reads was performed as
310 previously described (*37, 39*).
311
312 **Histology and immunohistochemistry**
313 Lung lobes were inflated with 4% formalin and fixed for 24 hours, stored in 70% ethanol,
314 paraffin-embedded, and 4 μm thick sections were used for Hematoxylin and Eosin (H&E)
315 staining and immunohistochemistry (IHC). Anti-RFP (Rockland, 600-401-379), anti-
316 TTF1(Abcam, ab76013), anti-UCHL1(Sigma, HPA005993), anti-TP63 (Cell Signaling
317 Technology, 13109) IHC was performed using Avidin/Biotin Blocking Kit (Vector Laboratories,
318 SP-2001), Avidin-Biotin Complex kit (Vector Laboratories, PK-4001), and DAB Peroxidase
319 Substrate Kit (Vector Laboratories, SK-4100) following standard protocols.
320
321 **Data availability**
322 Tuba-seq barcode sequencing data from the experiment in Figure 1 and 2 has been
323 previously deposited in NCBI's Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/)
324 (GSE174393). Tuba-seq barcode sequencing data from the experiment in Figure 3 and 4 has
325 been deposited in NCBI's Gene Expression Omnibus (GSE223678).
326
327 **Calculation of relative fitness of each genotype of tumors**
328 Fitness of each tumor genotype was approximated by the growth rate of tumors. Fitness
329 was calculated for each tumor genotype in a given mouse strain by considering the expected
330 number of cells with each lentiviral vector at the start of the experiment ($\#start_{genotype_mouse\text{-}strain}$)

331 and number of neoplastic cells with each lentiviral vector after tumor growth at the end of the

332 experiment (#end$_{genotype_mouse\text{-}strain}$).

333 $$Fitness_{genotype_{mouse\ strain}} = log_2(\frac{\#end_{genotype_{mouse-strain}}}{\#start_{genotype_{mouse-strain}}})$$

334

335

336

337 #end$_{genotype_mouse\text{-}strain}$ is the sum of all neoplastic cells in all tumors of that genotype at the

338 end of the experiment from all mice of a given strain (for example *TC* mice in Figure 1).

339 $$\#end_{genotype_{mouse-strain}} = sum(neoplastic\ cell\#_{genotype_{mouse-strain}})$$

340

341 #start$_{genotype_mouse\text{-}strain}$ is the expected number of cells of that genotype at the start of the

342 experiment. #start$_{genotype_mouse\text{-}strain}$ was determined from the number of tumors with each viral

343 vector (titer$_{genotype}$) from *KT* Cas9-negative control mice. In *KT* mice, the number of tumors with

344 each vector (number of unique BCs associated with each sgID) represents the exact titer of each

345 vector (titer$_{genotype}$) in the Lenti-sg*TS*$^{Triple\text{-}pool}$/*Cre* pool. To put #start$_{genotype}$ in the context of the

346 most potent genotype in the given mouse strain, we normalized the number of tumors with each

347 lentiviral vector in *KT* mice (titer$_{genotype}$) to the number of *Nf1;Rasa1;Pten* triple mutant tumors

348 in *KT* mice (titer$_{Nf1;Rasa1;Pten}$) and then multiplied this by the total number of *Nf1;Rasa1;Pten*

349 triple mutant tumors across all mice of a given strain (TumorNumber$_{Nf1;Rasa1;Pten_mouse\text{-}strain}$). Thus,

350 #start$_{genotype_mouse\text{-}strain}$ represents the titer-corrected expected number of tumors of that genotype if

351 that genotype was as potent as triple *Nf1:Rasa1;Pten* mutation in that strain. As an example,

352 #start for single mutant genotype *Nf1* in *TC* mice (#start $_{Nf1_TC}$) was calculated as;

353

354 $$\#start_{Nf1_{TC}} = \frac{titer_{Nf1}}{titer_{Nf1;Rasa1;Pten}} \times TumorNumber_{Nf1;Rasa1;Pten_{TC}}$$

355

356 Finally, we calculated "Relative Fitness" within each mouse strain by normalizing the
357 fitness of each genotype to the fitness of tumors with the lentiviral vector with three Inert
358 sgRNAs (Lenti-sgNeo1-sgNeo2-sgNT/Cre; inert tumors).

359

360 $$relative\ fitness_{genotype_{mouse-strain}} =$$

361 $$\frac{log_2(\#end_{genotype\ mouse-strain} / \#start_{genotype\ mouse-strain})}{log_2(\#end_{inert_{mouse-strain}} / \#start_{inert_{mouse-strain}})}$$

362 **Methods to control for multiple transductions**

363 The most potent genotype in our experiments was the *Nf1;Rasa1;Pten* triple mutant
364 generated by transduction with the Lenti-sg*Nf1*-sg*Rasa1*-sg*Pten*/Cre vector. We found that in our
365 *in vivo* experiments that a small percent of tumors arise from lung epithelial cells that were
366 initially transduced by more than one lentiviral vector[14]. We refer to these as multiple
367 transductions. Multiple transductions influence our ability to estimate fitness as a subset of
368 tumors identified as a given genotype (e.g. *Nf1* single mutant tumors which arise from a cell
369 transduced with the Lenti-sg*Nf1*/Cre vector) are in fact double mutants or even *Nf1;Rasa1;Pten*
370 triple mutant due to multiple transduction of the initial cells with a vector with complementary
371 sgRNAs (e.g. with the Lenti-sgNf1-sg*Rasa1-sgPten*/Cre vector). To quantify the fraction of
372 tumors that arise from multiple transductions (multiple transduction rate), we calculated the ratio
373 of the number of tumors with greater than a given number of neoplastic cells (minimum cell
374 number cutoff, X) for a given genotype versus the number of tumors with greater than the same
375 minimum cell number cutoff for triple mutants:

376 $$\frac{number\ of\ tumors > X\ neoplastic\ cells_{genotype}}{number\ of\ tumors > X\ neoplastic\ cells_{Nf1;Rasa1;Pten}}$$

377 As we increase the minimum cell number cutoff (X), the ratio will decrease and
378 asymptotically approach the true multiple transduction rate. The reason behind this is that as we
379 increase the minimum cell number cutoff, the proportion of tumors co-transduced with the most
380 potent triple mutant genotype among all our tumors increases, as the larger the tumor, the less
381 likely it is to be from a genotype other than the *Nf1;Rasa1;Pten* triple mutant. Therefore, the
382 minimum ratio across all minimum cell number cutoff (X) values was used as the multiple
383 transduction rate. The multiple transduction rate provides us the expected number of tumors (n)
384 for any given genotype that are the result of multiple transductions. Next, to control for the effect
385 of multiples transduction, we used the distribution of *Nf1;Rasa1;Pten* triple mutant tumors to
386 remove the n tumors from the data for each other genotype. For the n tumors to remove, we took
387 the n-quantiles of the *Nf1;Rasa1;Pten* triple mutant tumor size distribution. For each of these
388 quantile values, we chose the tumor of the given genotype that was closest to the quantile in size
389 and remove that tumor from the dataset for the given genotype. We used the sgInert multiple
390 transduction rate in *TC* mice for all genotypes and all mouse strains in all analysis and figures
391 except **Figure 2e-f, 3h-i** and **Supplementary Figure 1b-c, 2c-d, 3e-f, 4a-b, 5d-e, 6a-b** and **7a-**
392 **b**. This is logical, as multiple transductions should not be influenced by the sgIDs, sgRNA
393 sequences, or mouse genotypes. Furthermore, we do not expect growth of tumors with Lenti-
394 sgInert/vectors in *TC* mice, unless those initial cells have multiple transductions that occur with
395 vectors with tumor suppressor gene targeting lentiviral vectors.

396 As an alternative method, we also quantified the multiple transduction rate for all
397 genotypes of tumors separately and used that for multiple transduction correction (multiple
398 transduction correction Method #2, see **Supplementary Fig. 2c, Supplementary Fig. 4b,**

399 **Supplementary Fig. 7b**). The results and conclusions were unchanged when we used this
400 alternate multiple transduction correction method.
401
402 **Statistical analysis**
403 All statistical analyses were performed using the R software environment. For all bar
404 plots showing relative tumor size and tumor number and for fitness landscape plots, p-values and
405 95% confidence intervals (represented by whiskers) were calculated using bootstrap resampling
406 (10,000 repetitions). Bootstrapping was done by random resampling with replacement all the
407 tumors in all mice of a given strain. When indicated in the figure legend and in cases of barplots
408 showing tumor number, a nested bootstrap approach was used where first all the mice from a
409 given strain were resampled with replacement and then all tumors in this virtual mouse cohort
410 were resampled again with replacement.
411
412 **ACKNOWLEDGEMENTS**
413 We thank the Stanford Veterinary Animal Care Staff for expert animal care; Leo Chen
414 and Min K. Tsai for technical assistant; Chuan Li, Paloma Ruiz and other members of the
415 Winslow and Petrov laboratories for helpful discussions and reviewing the manuscript. M.
416 Yousefi was supported by a Stanford University School of Medicine Dean's fellowship, an
417 American Lung Association grant, and an NIH Ruth L. Kirschstein National Research Service
418 Award (F32-CA236311). G. Boross was supported by Tobacco-Related Disease Research
419 Program Postdoctoral Fellowships (T31FT-1619). This work was supported by NIH R01-
420 CA231253 (to M.M. Winslow and D.A. Petrov), NIH R01-CA230919 (to M.M. Winslow) and

NIH R01-CA234349 (to M.M. Winslow and D.A. Petrov), as well as by the Stanford Cancer Institute, an NCI-designated Comprehensive Cancer Center.

REFERENCES

1. Fragata, I., Blanckaert, A., Dias Louro, M. A., Liberles, D. A. & Bank, C. Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82 (2019).
2. Wright, S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proc. Sixth Intl Cong. Genet.* **1**, 356–366 (1932).
3. de Visser, J. A. G. M. & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490 (2014).
4. Bank, C., Hietpas, R. T., Jensen, J. D. & Bolon, D. N. A. A Systematic Survey of an Intragenic Epistatic Landscape. *Mol. Biol. Evol.* **32**, 229–238 (2015).
5. Karageorgi, M. *et al.* Genome editing retraces the evolution of toxin resistance in the monarch butterfly. *Nature* **574**, 409–412 (2019).
6. Weinreich, D. M., Delaney, N. F., DePristo, M. A. & Hartl, D. L. Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. *Science* **312**, 111–114 (2006).
7. Szendro, I. G., Schenk, M. F., Franke, J., Krug, J. & Visser, J. A. G. M. de. Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech. Theory Exp.* **2013**, P01005 (2013).
8. Foggetti, G. *et al.* Genetic Determinants of EGFR-Driven Lung Cancer Growth and Therapeutic Response In Vivo. *Cancer Discov.* **11**, 1736–1753 (2021).
9. Blair, L. M. *et al.* Oncogenic context shapes the fitness landscape of tumor suppression. 2022.10.24.511787 Preprint at https://doi.org/10.1101/2022.10.24.511787 (2022).

444 10. van de Haar, J. *et al.* Identifying Epistasis in Cancer Genomes: A Delicate Affair. *Cell* **177**,
445 1375–1383 (2019).
446 11. Jorge, S. E. D. C., Kobayashi, S. S. & Costa, D. B. Epidermal growth factor receptor (EGFR)
447 mutations in lung cancer: preclinical and clinical data. *Braz. J. Med. Biol. Res. Rev. Bras.*
448 *Pesqui. Medicas E Biol.* **47**, 929–939 (2014).
449 12. Saito, M. *et al.* Gene aberrations for precision medicine against lung adenocarcinoma.
450 *Cancer Sci.* **107**, 713–720 (2016).
451 13. Skoulidis, F. & Heymach, J. V. Co-occurring genomic alterations in non-small-cell lung cancer
452 biology and therapy. *Nat. Rev. Cancer* **19**, 495–509 (2019).
453 14. Yousefi, M. *et al.* Combinatorial Inactivation of Tumor Suppressors Efficiently Initiates Lung
454 Adenocarcinoma with Therapeutic Vulnerabilities. *Cancer Res.* **82**, 1589–1602 (2022).
455 15. Carrot-Zhang, J. *et al.* Whole-genome characterization of lung adenocarcinomas lacking
456 alterations in the RTK/RAS/RAF pathway. *Cell Rep.* **34**, (2021).
457 16. Rogers, Z. N. *et al.* A quantitative and multiplexed approach to uncover the fitness
458 landscape of tumor suppression *in vivo*. *Nat. Methods* **14**, 737–742 (2017).
459 17. Cai, H. *et al.* A functional taxonomy of tumor suppression in oncogenic KRAS-driven lung
460 cancer. *Cancer Discov.* (2021) doi:10.1158/2159-8290.CD-20-1325.
461 18. Li, C. *et al.* Quantitative In Vivo Analyses Reveal a Complex Pharmacogenomic Landscape in
462 Lung Adenocarcinoma. *Cancer Res.* **81**, 4570–4580 (2021).
463 19. Rogers, Z. N. *et al.* Mapping the in vivo fitness landscape of lung adenocarcinoma tumor
464 suppression in mice. *Nat. Genet.* **50**, 483 (2018).

465 20. Collisson, E. A. *et al.* Comprehensive molecular profiling of lung adenocarcinoma. *Nature*
466 **511**, 543–550 (2014).
467 21. Romero, R. *et al.* Keap1 loss promotes Kras-driven lung cancer and results in dependence
468 on glutaminolysis. *Nat. Med.* **23**, 1362–1368 (2017).
469 22. Ji, H. *et al.* LKB1 modulates lung cancer differentiation and metastasis. *Nature* **448**, 807–810
470 (2007).
471 23. Fisher, R. A. *The genetical theory of natural selection*. (Oxford University Press, 1930).
472 24. Bourouh, M. & Marignani, P. A. The Tumor Suppressor Kinase LKB1: Metabolic Nexus. *Front.*
473 *Cell Dev. Biol.* **10**, 881297 (2022).
474 25. Schabath, M. B. *et al.* Differential association of STK11 and TP53 with KRAS mutation-
475 associated gene expression, proliferation and immune surveillance in lung adenocarcinoma.
476 *Oncogene* **35**, 3209–3216 (2016).
477 26. Kang, Y. P. *et al.* Cysteine dioxygenase 1 is a metabolic liability for non-small cell lung
478 cancer. *eLife* **8**, e45572 (2019).
479 27. Najm, F. J. *et al.* Orthologous CRISPR–Cas9 enzymes for combinatorial genetic screens. *Nat.*
480 *Biotechnol.* **36**, 179–189 (2018).
481 28. Murray, C. W. *et al.* An Lkb1-Sik axis suppresses lung tumor growth and controls
482 differentiation. *Cancer Discov.* (2019) doi:10.1158/2159-8290.CD-18-1237.
483 29. Adamson, B. *et al.* A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic
484 Dissection of the Unfolded Protein Response. *Cell* **167**, 1867-1882.e21 (2016).
485 30. Zetsche, B. *et al.* Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat.*
486 *Biotechnol.* **35**, 31–34 (2017).

487 31. Han, K. *et al.* Synergistic drug combinations for cancer identified in a CRISPR screen for
488 pairwise genetic interactions. *Nat. Biotechnol.* **35**, 463–474 (2017).
489 32. Liu, C. *et al.* Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell*
490 **146**, 209–221 (2011).
491 33. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional
492 expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
493 34. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system
494 for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
495 35. Chiou, S.-H. *et al.* Pancreatic cancer modeling using retrograde viral vector delivery and in
496 vivo CRISPR/Cas9-mediated somatic genome editing. *Genes Dev.* **29**, 1576–1585 (2015).
497 36. Okawa, H. *et al.* Hepatocyte-specific deletion of the keap1 gene activates Nrf2 and confers
498 potent resistance against acute drug toxicity. *Biochem. Biophys. Res. Commun.* **339**, 79–88
499 (2006).
500 37. Bardeesy, N. *et al.* Loss of the Lkb1 tumour suppressor provokes intestinal polyposis but
501 resistance to transformation. *Nature* **419**, 162–167 (2002).
502 38. Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional
503 mouse model for breast cancer. *Nat. Genet.* **29**, 418–425 (2001).
504
505

506 **FIGURE LEGENDS**

507 **Figure 1. The fitness landscape of *Nf1;Rasa1;Pten* oncogene-negative lung adenocarcinoma**
508 **is fully accessible**
509 **a**, Schematic of barcoded triple sgRNA vectors in the Lenti-sg*TS*$^{Triple\text{-}pool}$*/Cre* pool for
510 CRISPR/Cas9-mediated inactivation of all combinations of *Nf1*, *Rasa1*, and *Pten* in *R26*$^{LSL\text{-}Tomato}$*;H11*$^{LSL\text{-}Cas9}$ (*TC*), *Trp53*$^{flox/flox}$*;TC* mice, and Cas9-negative *Kras*$^{LSL\text{-}G12D}$*;T* mice. sgNeo1,
512 sgNeo2, and sgNT are inert sgRNAs. The number and size of tumors initiated with each vector
513 can be quantified by tumor barcoding coupled with high-throughput sgID-BC sequencing (Tuba-
514 seq). Mouse genotype, mouse number, and viral titer (ifu; infectious units) are indicated. Tuba-
515 seq was performed on bulk tumor-bearing lungs 3 months after tumor initiation. Data was
516 generated in Yousefi, Boross *et al.*, Cancer Research 2022.
517 **b**, Overview of our method to calculate Relative Fitness of each cancer genotype (Methods).
518 **c-d**, Fitness landscape for *TC* (**c**) and *Kras*$^{LSL\text{-}G12D}$*;T* (Cas9-negative control)(**d**) mice. Fitness for
519 tumors of each single and double mutant genotype, as well as those with all three tumor
520 suppressor gene targeting sgRNAs are shown relative to the triple inert vector. Green arrows
521 indicate increased fitness, red arrows indicate reduced fitness, solid line indicates significance (p-
522 value < 0.05). Whiskers show 95% confidence intervals from bootstrap resampling of tumors.
523

524 **Figure 2. The fitness landscape of *Nf1;Rasa1;Pten* oncogene-negative lung adenocarcinoma**
525 **remains fully accessible despite otherwise strong impacts of Trp53-deficiency**
526 **a-b**, Effect of *Trp53* inactivation on the median size of all tumors (**a**) and of *Nf1;Rasa1;Pten*
527 tumors (**b**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals
528 from bootstrap resampling. ***, P < 0.001 based on a bootstrap resampling.
529 **c-d**, Effect of *Trp53* inactivation on the number of all tumors (**c**) and of *Nf1;Rasa1;Pten* tumors
530 (**d**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals from
531 bootstrap resampling. ***, P < 0.001 based on a bootstrap resampling.
532 **e**, Comparison of Surface tumor number (from direct counting of Tomato^{positive} tumors >0.5 mm
533 in diameter) with the Relative tumor number from Tuba-seq (the number of clonal exapansions
534 with >50 neoplastic cells relative to the median of *TC* values). Each dot represents a mouse. Data
535 from all tumors is shown.
536 **f**, Comparison of the Relative tumor size determined from Tuba-seq (the median number of
537 neoplastic cells relative to the median of *TC* values) with the Relative tumor number from Tuba-
538 seq (the number of clonal expansions with >50 neoplastic cells relative to the median of *TC*
539 values). Each dot represents a mouse. Data from all tumors is shown.
540 **g**, Fitness landscape for *Trp53*$^{f/f}$*;TC* mice. Fitness for tumors of each single and double mutant
541 genotype, as well as those with all three tumor suppressor targeting sgRNAs are shown relative
542 to the triple inert vector. Green arrows indicate increased fitness, red arrows indicate reduced
543 fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence
544 intervals from bootstrap resampling of tumors.
545

546 **Figure 3. Inactivaton of the Keap1 or Lkb1 "tumor suppressors" reduces *Nf1;Rasa1;Pten***
547 **oncogene-negative tumor initiation with divergent effects on tumor size**
548 **a**, Schematic of experiments to assess the effect of *Keap1* and *Lkb1* inactivating background
549 mutations on combinations of *Nf1;Rasa1;Pten* inactivation. Mouse genotype, mouse number,

550 and viral titer (ifu; infectious units) are indicated. Tuba-seq was performed on bulk tumor-
551 bearing lungs 3 months after tumor initiation.
552 **b**, Representative light and fluorescence images of lung lobes from the indicated genotypes of
553 mice. Lung lobes are outlined with white dotted lines. Scale bars = 4 mm
554 **c**, The number of tumors (Tomato^{positive} tumors >0.5mm in diameter) quantified by direct
555 counting. Each dot represents a mouse, and the bar is the median. The genotypes of the mice are
556 indicated. **, P < 0.01; and n/s, P > 0.05 (not significant) based on a bootstrap resampling.
557 **d-e**, Effect of *Keap1* or *Lkb1* inactivation on the median size of all tumors (**d**) and of
558 *Nf1;Rasa1;Pten* tumors (**e**). Effects are shown relative to *TC* mice. Whiskers show 95%
559 confidence intervals from bootstrap resampling. ***, P < 0.001 based on a bootstrap resampling.
560 **f-g**, Effect of *Keap1* or *Lkb1* inactivation on the number of all tumors (**f**) and of *Nf1;Rasa1;Pten*
561 tumors (**g**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals
562 from bootstrap resampling. ***, P < 0.001 based on a bootstrap resampling.
563 **h**, Comparison of Surface tumor number (from direct counting of Tomato^{positive} tumors >0.5mm
564 in diameter) with the Relative tumor number from Tuba-seq (the number of clonal expansions
565 with >50 neoplastic cells relative to the median of *TC* values). Each dot represents a mouse. Data
566 from all tumors is shown.
567 **i**, Comparison of the Relative tumor size determined from Tuba-seq (the median number of
568 neoplastic cells relative to the median of *TC* values) with the Relative tumor number from Tuba-
569 seq (the number of clonal expansions with >50 neoplastic cells relative to the median of *TC*
570 values). Each dot represents a mouse. Data from all tumors is shown.
571
572 **Figure 4. Keap1 inactivation changes the topology of the fitness landscape of**
573 ***Nf1;Rasa1;Pten* oncogene-negative lung adenocarcinoma**
574 **a-b**, Fitness landscape for *TC* (**a**) and *Keap1$^{flox/flox}$;TC* (**b**) mice. Fitness for tumors of each single
575 and double mutant genotype, as well as those with all three tumor suppressor gene targeting
576 sgRNAs are shown relative to the triple inert vector. Green arrows indicate increased fitness, red
577 arrows indicate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers show
578 95% confidence intervals from bootstrap resampling of tumors.
579 **c**, Relative fitness of the double mutant genotypes from the *TC* and *Trp53$^{f/f}$;TC* mice from Figure
580 1/2 and from *TC* and *Keap1$^{f/f}$;TC* mice from Figure 4. *, P < 0.05; ***, P < 0.001; and n/s, P >
581 0.05 (not significant), based on a bootstrap resampling.

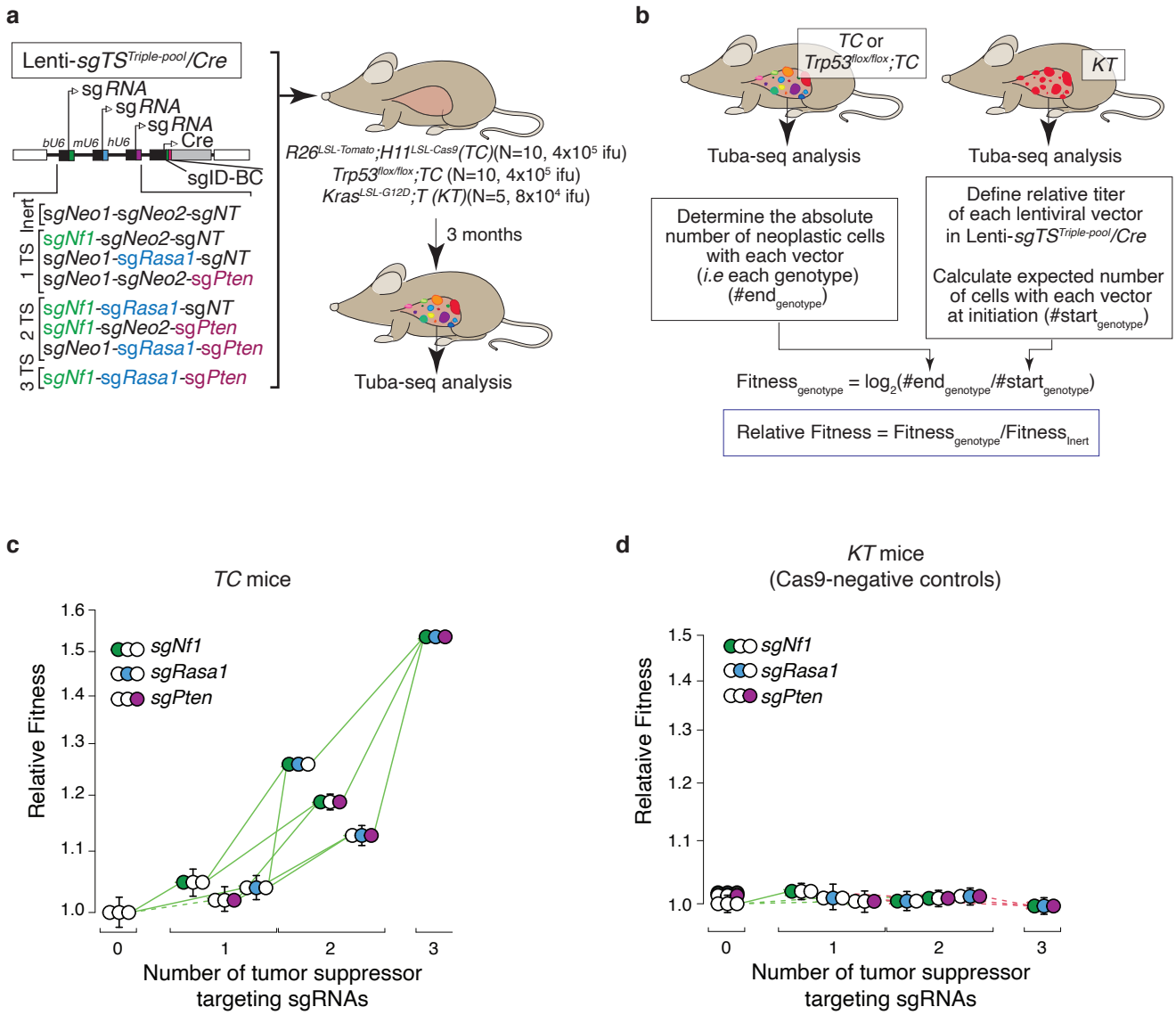

Figure 1. The fitness landscape of *Nf1;Rasa1;Pten* oncogene-negative lung adenocarcinoma is fully accessible
a, Schematic of barcoded triple sgRNA vectors in the Lenti-sg*TS*$^{Triple\text{-}pool}$*/Cre* pool for CRISPR/Cas9-mediated inactivation of all combinations of *Nf1, Rasa1*, and *Pten* in *R26*$^{LSL\text{-}Tomato}$*;H11*$^{LSL\text{-}Cas9}$ (*TC*), *Trp53*$^{flox/flox}$*;TC* mice, and Cas9-negative *Kras*$^{LSL\text{-}G12D}$*;T* mice. sgNeo1, sgNeo2, and sgNT are inert sgRNAs. The number and size of tumors initiated with each vector can be quantified by tumor barcoding coupled with high-throughput sgID-BC sequencing (Tuba-seq). Mouse genotype, mouse number, and viral titer (ifu; infectious units) are indicated. Tuba-seq was performed on bulk tumor-bearing lungs 3 months after tumor initiation. Data was generated in Yousefi, Boross *et al.*, Cancer Research 2022.
b, Overview of our method to calculate Relative Fitness of each cancer genotype (Methods).
c-d, Fitness landscape for *TC* (**c**) and *Kras*$^{LSL\text{-}G12D}$*;T* (Cas9-negative control)(**d**) mice. Fitness for tumors of each single and double mutant genotype, as well as those with all three tumor suppressor gene targeting sgRNAs are shown relative to the triple inert vector. Green arrows indicate increased fitness, red arrows indicate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals from bootstrap resampling of tumors.

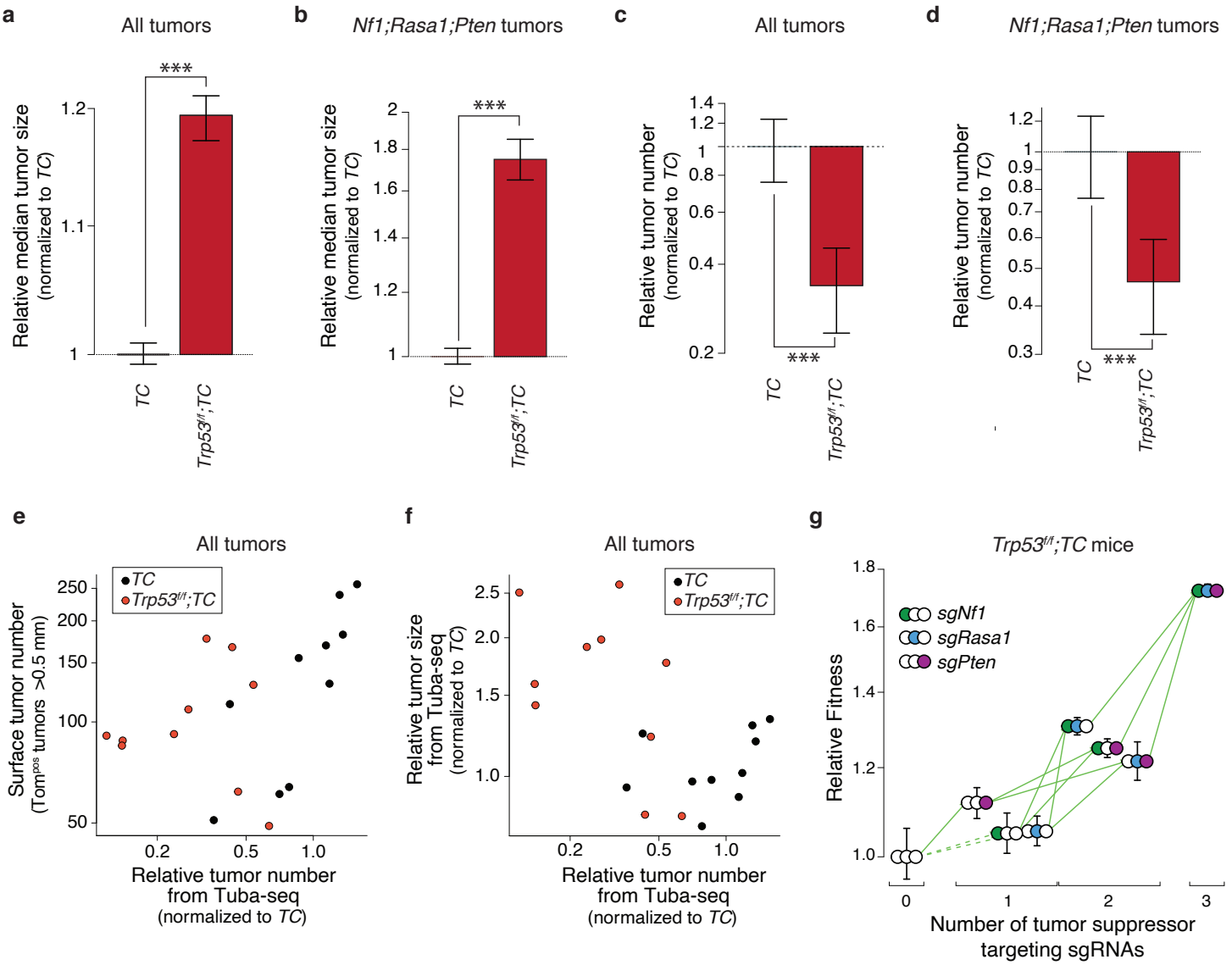

Figure 2. The fitness landscape of *Nf1;Rasa1;Pten* oncogene-negative lung adenocarcinoma remains fully accessible despite otherwise strong impacts of Trp53-deficiency

a-b, Effect of *Trp53* inactivation on the median size of all tumors (**a**) and of *Nf1;Rasa1;Pten* tumors (**b**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals from bootstrap resampling. ***, $P < 0.001$ based on a bootstrap resampling.

c-d, Effect of *Trp53* inactivation on the number of all tumors (**c**) and of *Nf1;Rasa1;Pten* tumors (**d**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals from bootstrap resampling. ***, $P < 0.001$ based on a bootstrap resampling.

e, Comparison of Surface tumor number (from direct counting of Tomato^{positive} tumors >0.5 mm 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 all 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 all tumors is shown.

g, Fitness landscape for *Trp53$^{f/f}$;TC* 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 increased fitness, red arrows indicate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals from bootstrap resampling of tumors.

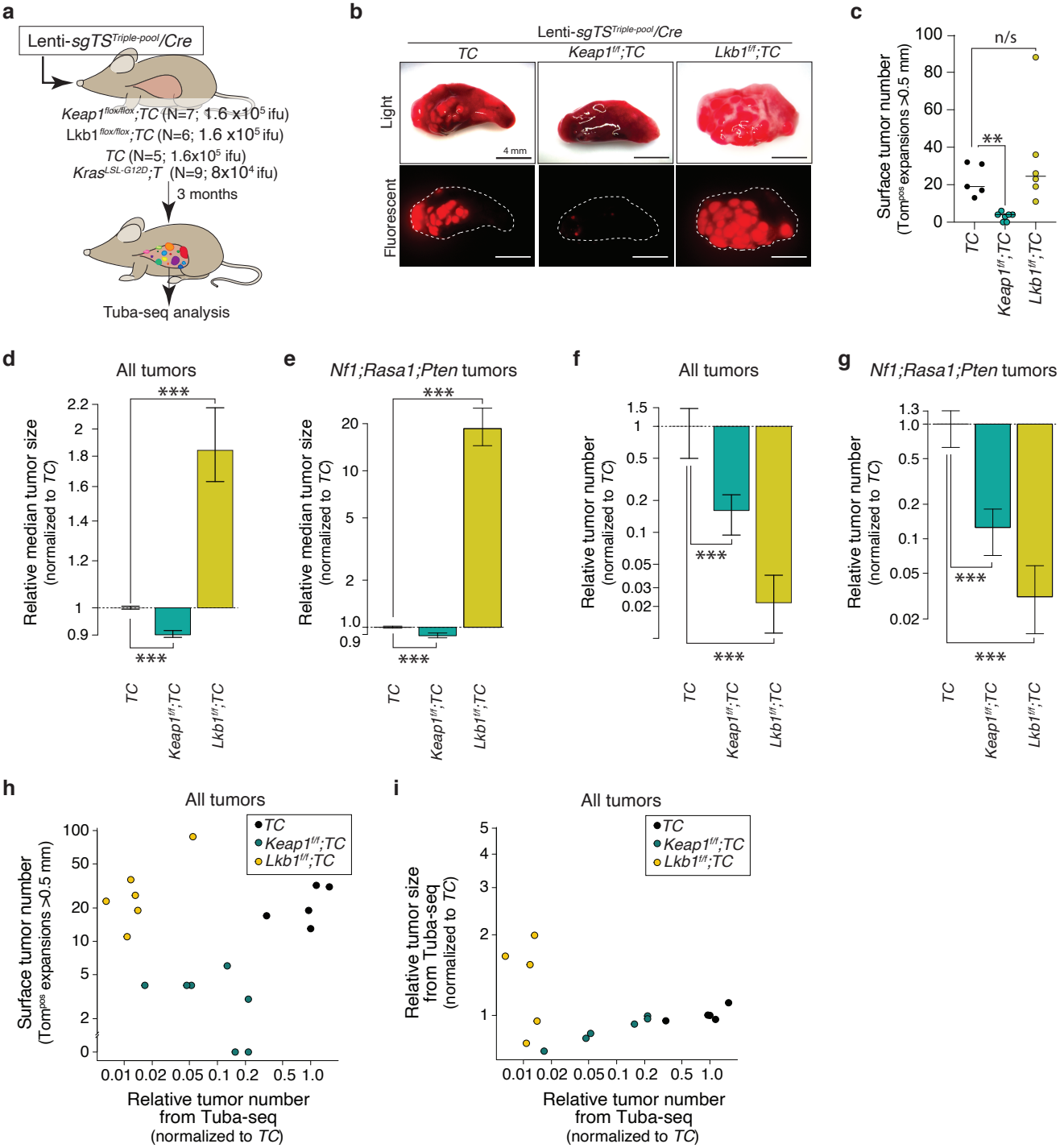

Figure 3. Inactivaton of the Keap1 or Lkb1 "tumor suppressors" reduces *Nf1;Rasa1;Pten* oncogene-negative tumor initiation with divergent effects on tumor size

a, Schematic of experiments to assess the effect of *Keap1* and *Lkb1* inactivating background mutations on combinations of *Nf1;Rasa1;Pten* inactivation. Mouse genotype, mouse number, and viral titer (ifu; infectious units) are indicated. Tuba-seq was performed on bulk tumor-bearing lungs 3 months after tumor initiation.

b, Representative light and fluorescence images of lung lobes from the indicated genotypes of mice. Lung lobes are outlined with white dotted lines. Scale bars = 4 mm

c, The number of tumors (Tomato^{positive} tumors >0.5mm in diameter) quantified by direct counting. Each dot represents a mouse, and the bar is the median. The genotypes of the mice are indicated. **, $P < 0.01$; and n/s, $P > 0.05$ (not significant) based on a bootstrap resampling.

d-e, Effect of *Keap1* or *Lkb1* inactivation on the median size of all tumors (**d**) and of *Nf1;Rasa1;Pten* tumors (**e**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals from bootstrap resampling. ***, $P < 0.001$ based on a bootstrap resampling.

f-g, Effect of *Keap1* or *Lkb1* inactivation on the number of all tumors (**f**) and of *Nf1;Rasa1;Pten* tumors (**g**). Effects are shown relative to *TC* mice. Whiskers show 95% confidence intervals from bootstrap resampling. ***, $P < 0.001$ based on a bootstrap resampling.

h, 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 all tumors is shown.

i, 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 all tumors is shown.

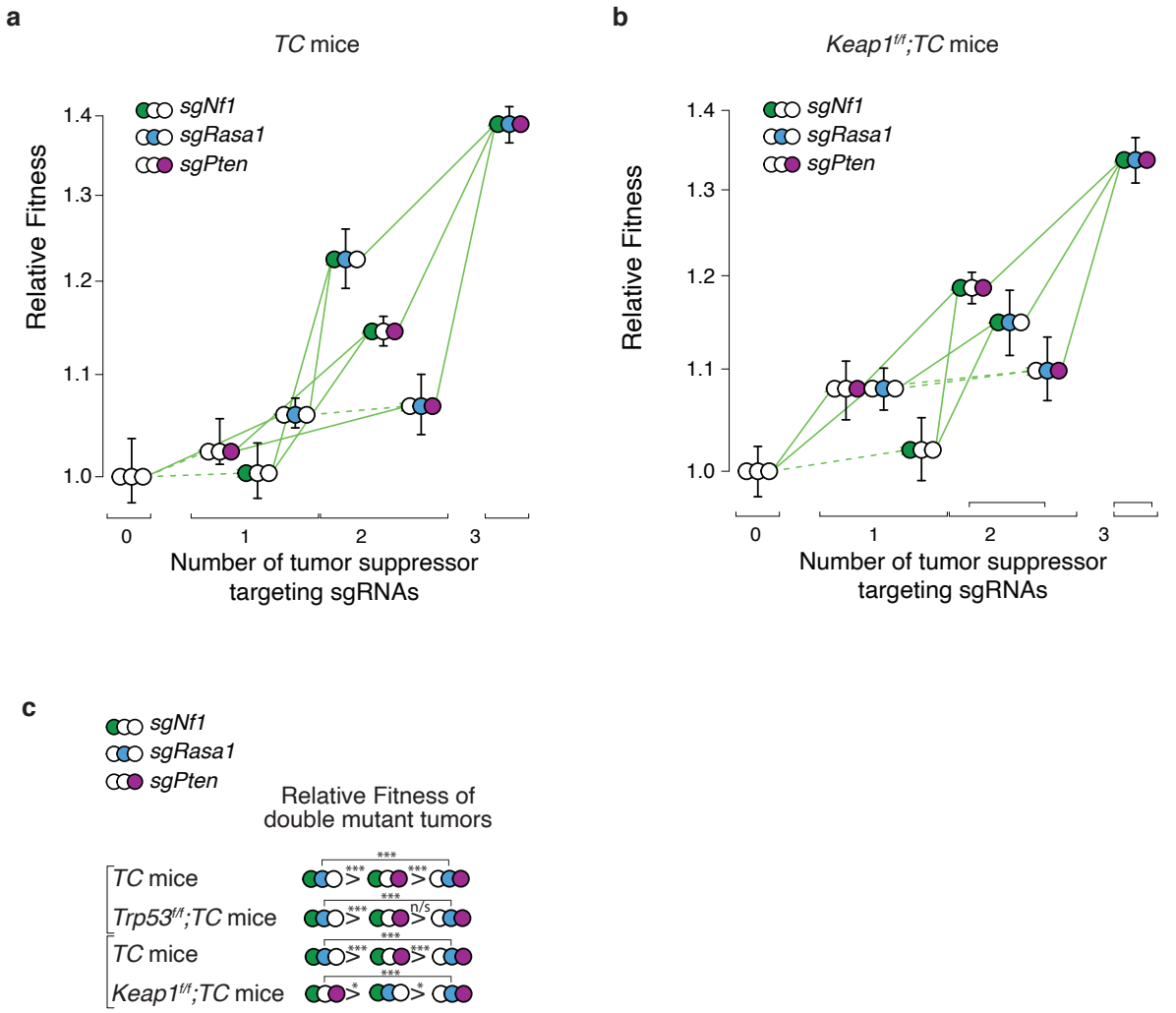

Figure 4. *Keap1* inactivation changes the topology of the fitness landscape of *Nf1;Rasa1;Pten* oncogene-negative lung adenocarcinoma.

a-b, Fitness landscape for *TC* (**a**) and *Keap1*$^{flox/flox}$*;TC* (**b**) mice. Fitness for tumors of each single and double mutant genotype, as well as those with all three tumor suppressor gene targeting sgRNAs are shown relative to the triple inert vector. Green arrows indicate increased fitness, red arrows indicate reduced fitness, solid line indicates significance (p-value < 0.05). Whiskers show 95% confidence intervals from bootstrap resampling of tumors.

c, Relative fitness of the double mutant genotypes from the *TC* and *Trp53*$^{f/f}$*;TC* mice from Figure 1/2 and from *TC* and *Keap1*$^{f/f}$*;TC* mice from Figure 4. *, P < 0.05; ***, P < 0.001; and n/s, P > 0.05 (not significant), based on a bootstrap resampling.