Rlf-Mycl gene fusion drives tumorigenesis and metastasis in a mouse model of small cell lung cancer

Metamia Ciampricotti[1,2], **Triantafyllia Karakousi**[3,13], **Allison L Richards**[4,13], **Àlvaro Quintanal-Villalonga**[1,2], **Angeliki Karatza**[5], **Rebecca Caeser**[1,2], **Emily A Costa**[6], **Viola Allaj**[1,2], **Parvathy Manoj**[1,2], **Kyle B Spainhower**[7], **Faruk E Kombak**[8], **Francisco J Sanchez-Rivera**[9], **Janneke E Jaspers**[10], **Anastasia-Maria Zavitsanou**[3], **Danilo Maddalo**[9,11], **Andrea Ventura**[9], **William M Rideout III**[12], **Elliot H Akama-Garren**[12], **Tyler Jacks**[12], **Mark T A Donoghue**[4], **Triparna Sen**[1,2], **Trudy G Oliver**[7], **John T Poirier**[5], **Thales Papagiannakopoulos**[3,5,*], **Charles M Rudin**[1,2,14,*]

[1]Department of Medicine, Thoracic Oncology Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[2]Molecular Pharmacology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[3]Department of Pathology, New York University School of Medicine, New York, NY, USA.

[4]Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[5]Perlmutter Cancer Center, New York University Langone Health, New York, NY, USA.

[6]Weill Cornell Graduate School of Medical Sciences, Weill Cornell Medicine, New York, NY, USA.

[7]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA.

[8]Precision Pathology Biobanking Center, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[9]Department of Cancer Biology and Genetics, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[10]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[*]Correspondence: Thales Papagiannakopoulos, NYU Grossman School of Medicine, 550 First Ave., Smilow 505, New York, NY 10016, thales.papgiannakopoulos@nyulangone.org, 646-501-0042, Charles M Rudin, Memorial Sloan Kettering Cancer Center, 530 E 74th St., Rm 22-222, New York, NY 10021, rudinc@mskcc.org, 646-608-2870.

[11]Current address: Department of Translational Oncology, Genentech, South San Francisco, CA, USA.

[12]David H. Koch Institute for Integrative Cancer Research, Department of Biology, Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA, USA.

[13]These authors contributed equally

[14]Lead contact

Abstract

Small cell lung cancer (SCLC) has limited therapeutic options and an exceptionally poor prognosis. Understanding the oncogenic drivers of SCLC may help define novel therapeutic targets. Recurrent genomic rearrangements have been identified in SCLC, most notably an in-frame gene fusion between *RLF* and *MYCL* found in up to 7% of the predominant ASCL1-expressing subtype (SCLC-A). To explore the role of this fusion in oncogenesis and tumor progression, we used CRISPR/Cas9 somatic editing to generate a *Rlf-Mycl*-driven mouse model of SCLC. RLF-MYCL fusion accelerated transformation and proliferation of murine SCLC and increased metastatic dissemination and the diversity of metastatic sites. Tumors from the RLF-MYCL genetically-engineered mouse model displayed gene expression similarities with human RLF-MYCL SCLC. Together our studies support RLF-MYCL as the first demonstrated fusion oncogenic driver in SCLC and provide a new preclinical mouse model for the study of this subtype of SCLC.

Keywords

RLF-MYCL; gene fusion; SCLC; CRISPR-Cas9; ASCL1; genetically engineered mouse model; metastasis

INTRODUCTION

Small cell lung cancer (SCLC) is an exceptionally aggressive malignancy that comprises 13% of all lung cancer cases and causes an estimated 250,000 deaths globally per year (1,2). Most patients have metastatic disease at the time of diagnosis. While SCLC tumors are initially sensitive to standard cytotoxics, nearly all patients develop recurrent and chemoresistant disease, leading to a median survival of slightly over one year (1). The recent addition of immune checkpoint inhibitors to standard chemotherapy has led to durable responses in a small minority of patients, but has extended median survival by only 2 months (3), underscoring the need to deepen our understanding of this disease in pursuit of more effective treatment strategies.

SCLC is recognized histologically by a characteristic cellular morphology, and frequently expresses markers of neuroendocrine differentiation (4). Data from both human tumors and mouse models of SCLC have defined distinct subtypes of SCLC based on differential expression of key transcriptional regulators including achaete-scute homologue 1 (ASCL1), neurogenic differentiation factor 1 (NEUROD1), POU class 2 homeobox 3 (POU2F3) and yes-associated protein 1 (YAP1) (5). These subtypes have been referred to respectively

as SCLC-A, SCLC-N, SCLC-P, and SCLC-Y. Canonical "classic" and "variant" subtypes of SCLC are associated with upregulation of ASCL1 and NEUROD1, respectively (6,7). Common genetic hallmarks of SCLC include near universal bi-allelic loss of tumor suppressors *TP53* and *RB1*, amplification of *MYC* family genes, and inactivating mutations of *NOTCH* genes (8,9). Recent work has advanced our understanding of how these molecular features may define subtype- or genotype-specific therapeutic vulnerabilities (2,7,10,11). Due in part to a paucity of primary tumor samples, such studies have relied heavily on established cell lines and genetically-engineered mouse models (GEMMs), the latter of which recapitulate many of the key mutations, markers, and metastatic patterns observed in human SCLC (7,12–17).

A number of GEMMs have been developed to model SCLC biology. Most include *Rb1* and *Trp53* alleles flanked by loxP sites, with lung tumor growth initiated through intranasal or intratracheal delivery of adenoviruses expressing Cre recombinase (17). While concomitant overexpression of MYCL has been shown to accelerate tumor growth in this context (18), long latency of these models has spurred the adoption of faster growing models generated through the additional deletion of *Rbl2* or *Pten*, or stabilization of MYC (7,13– 16). These and other double- and triple-knockout models have been used to interrogate SCLC tumorigenesis, providing insight to putative cell types of origin and defining a requirement for *Ascl1* in SCLC tumorigenesis (6,19,20). A stabilized and overexpressed MYC allele elicits a rapidly growing, metastatic phenotype and can drive tumor evolution from *Ascl1*^{high} to *Neurod1*^{high} and then to *Yap1*^{high} SCLC (7,21), suggesting distinct roles for these transcriptional regulators in tumor initiation and progression.

GEMMs have provided insight into the molecular underpinnings of SCLC metastasis. Multiple studies identified the neural transcription factor NFIB as a putative metastatic driver across different mouse models (20,22,23). *Nfib* overexpression and/or amplification accelerates oncogenesis in the context of *Mycl* amplification and enhances metastasis in GEMMs; expression of the human ortholog NFIB correlates with poorer prognosis in patients (12,22,23). Induction of Cre under different lung epithelial lineage-specific promotors has suggested that NFIB-driven metastasis is lineage-restricted, such that tumors that arise from distinct cells of origin metastasize either dependently or independently of NFIB (24). Clonal heterogeneity is likely to influence the modes of metastasis available to SCLC tumors (25). Continued delineation of the drivers of metastasis in SCLC may reveal distinct therapeutic opportunities.

Chromosomal translocations, insertions, and deletions resulting in gene fusions represent a common pathway of oncogenesis in solid tumors (26). While SCLC demonstrates extensive genomic structural aberrancies, the role of gene fusions in driving SCLC development and metastasis has not been defined. Several intrachromosomal rearrangements have been identified in SCLC cell lines and tumors, most notably a recurrent in-frame gene fusion between *RLF* and *MYCL* (8,27,28). *MYCL* expression is associated with SCLC-A (6) and has been shown to support tumorigenesis and proliferation (18,29). However, the pathological significance of *RLF-MYCL* fusions in SCLC remains unknown. The functional interrogation of fusion genes as oncogenic drivers in SCLC GEMMs has been challenging due to the lack of tractable somatic engineering strategies. However,

with the development of CRISPR-Cas9 methods, it is now possible to engineer complex chromosomal rearrangements *in vivo* (30,31). Here, we report the first gene fusion GEMM of SCLC and demonstrate that RLF-MYCL fusion promotes tumorigenesis and enhances metastasis in SCLC.

RESULTS

RLF-MYCL is a recurrent gene fusion in SCLC with high *MYCL* expression

To assess the frequency of the *RLF-MYCL* gene fusion in human SCLC, we first analyzed RNA-seq data from 105 SCLC samples (8,9) using the Arriba and FusionCatcher fusion transcript discovery algorithms (https://githubcom/suhrig/arriba) (32). We found that in-frame *RLF-MYCL* gene fusions had the highest frequency of fusion reads, observing recurrent fusion events in 4.8% of all SCLC cases (5/105) and 6.9% of those classified as SCLC-A (5/72). *RLF* and *MYCL* are encoded on opposing DNA strands in close proximity on chromosome 1p, approximately 300 kb apart. RLF-MYCL fusion events in SCLC primary tumors and cell lines demonstrate multiple intronic breakpoints, all of which result in an inversion event that at the transcript level leads to splicing of the first exon 1 of *RLF* to the last 2 exons of *MYCL* (Fig. 1A). The resulting mature transcript encodes the first 79 amino acids of RLF followed by all but the first 27 amino acids of the MYCL protein, generating a 446-residue fusion protein (Fig. 1B).

We confirmed the presence of the in-frame *RLF-MYCL* fusion transcript and protein by PCR and immunoblot analysis in 5 human SCLC cell lines: NCI-H889, CORL47, NCI-H1092, NCI-H1963 and NCI-H1836 (Fig. 1C; Supplementary Table S1). In three of the cell lines harboring the *RLF-MYCL* fusion, immunoblotting suggests that the fusion protein is present at a substantially higher level than endogenous MYCL in any other line examined. Analysis of RNA-seq data from published datasets (8,9) confirmed that the identified SCLC tumors and cell lines with RLF-MYCL fusion events all belong to the SCLC-A (ASCL1[high]) subtype (Supplementary Fig. S1). The RLF-MYCL fusion was associated with higher levels of detectable *MYCL* transcripts relative to other SCLC-A lines (Fig. 1D). While the fusion protein could have neomophic function(s), these data suggest that the *RLF-MYCL* gene fusion might primarily contribute to the cancers in which it is present in part by increasing functional MYCL expression, thereby increasing MYCL-driven oncogenic signaling in SCLC. To determine whether this fusion has a functional role in oncogenesis and tumor progression, we investigated the effects of this fusion event in mouse models of SCLC.

Rlf-Mycl fusion promotes tumorigenesis of pre-neoplastic neuroendocrine cells

In the mouse genome, *Rlf* and *Mycl* are located on chromosome 4 (qD2.2), in a region that is syntenic to human chromosome 1 (p34.2) (Fig. 2A). The proximity and opposing orientation of the two genes in both mouse and human enables modeling the human translocation event in the mouse genome by Cas9-mediated editing: in both species this fusion requires an inversion event without loss of intervening genetic material. We attempted to induce *Rlf-Mycl* fusion in mouse cells through the use of sgRNAs promoting Cas9-mediated double-strand DNA breaks in the first introns of *Rlf* and *Mycl* (Fig. 2B). Plasmids expressing individual sgRNAs and Cas9 (33) (Fig. 2B) were co-transfected into NIH/3T3

cells and editing was confirmed by surveyor assays (Supplementary Fig. S2A), followed by PCR and immunoblot analyses demonstrating the presence of the RLF-MYCL fusion (Fig. 2C). The presence of the desired *Rlf-Mycl* inversion was confirmed by sequencing the corresponding *Rlf-Mycl* fusion genomic DNA and transcript (Fig. 2C).

We initially sought to assess the role of the RLF-MYCL fusion in early stages of tumorigenesis using derivatives from the *Rb1$^{fl/fl}$;Trp53$^{fl/fl}$;Rbl2$^{fl/fl}$* (RPR2) transgenic model of SCLC (16). It has been previously demonstrated that isolated cells from early neoplastic lesions in the lungs of these animals that have not fully transformed to malignancy contain precursors of SCLC ("preSC cells"; genotype *Rb1$^{-/-}$;Trp53$^{-/-}$;Rbl2$^{-/+}$*) in which the effects of putative oncogenic drivers can be assessed (29). We found that preSC cells in which we induced the *Rlf-Mycl* fusion by introduction of Cas9 with sgRNAs targeting intronic sequences of both *Rlf* and *Mycl* formed larger and more colonies in soft agar than controls transfected with either single sgRNA construct (Fig. 2D). To determine the effect of the fusion on tumor initiation *in vivo*, we injected preSC cells transfected with each one or both sgRNA constructs into the flanks of immunodeficient mice (n=5/group). Animals injected with preSC cells transfected with both sgRNAs to induce *Rlf-Mycl* rearrangement developed tumors earlier and faster than tumors with either single sgRNA control (Supplementary Fig. S2B–D).

To more readily support subsequent *in vivo* transfection experiments, we modified a lentiviral vector (USEC) to drive expression of both *Rlf* and *Mycl* sgRNAs, or control sgRNAs, from tandem U6 promoters along with Cre recombinase (34,35) (Fig. 2E). To test this system, we co-transduced preSC cells with lentiviruses containing Cas9 and either USEC with sgRNAs targeting *Rlf* and *Mycl* (herein sgRlfsgMycl) or USEC with sgRNAs against neomycin (hereafter sgNeosgNeo), and injected these transduced preSC cells into the flanks of immunodeficient mice (n=5 and 4, with sgNeosgNeo and sgRlfsgMycl, respectively). Mice injected with sgRlfsgMycl preSC cells developed larger tumors (Fig. 2F) and had shorter survival (Fig. 2G) compared to sgNeosgNeo controls (*P* = 0.007, log rank test). Droplet PCR of cDNA confirmed presence of the *Rlf-Mycl* fusion transcript in tumors, albeit with lower expression in tumors of viral transduced preSC cells than achieved with vector transfected preSC cells (Supplementary Fig. S2E). Hematoxylin/eosin (H&E) staining and synaptophysin (SYP) immunohistochemistry (IHC) showed typical SCLC characteristics in tumors (Supplementary Fig. S2F&G). Together these data confirm that we can successfully engineer the *Rlf-Mycl* fusion in mouse cells and that the RLF-MYCL fusion accelerates oncogenic transformation and tumor growth in preSCs.

RLF-MYCL endogenous induction accelerates SCLC tumor formation *in vivo*

To further investigate the contribution of the *Rlf-Mycl* gene fusion in SCLC development *in vivo*, we employed an autochthonous SCLC model i.e. *Rb1$^{fl/fl}$;Trp53$^{fl/fl}$;Rbl2$^{fl/fl}$;Rosa26$^{LSL-Cas9-GFP}$* (herein RPR2C) (31). Cre-mediated inactivation of *Rb1*, *Trp53*, and *Rbl2* in this model has been previously reported to recapitulate morphologic characteristics and therapeutic vulnerabilities of human SCLC (16). Introduction of the *LSL-Cas9-GFP* cassette into the *Rosa16* locus allows Cre-mediated induction of Cas9 expression together with deletion of these key tumor suppressors. First,

we transduced a cohort of adult chimeric RPR2C mice via intratracheal instillation with USEC lentivirus expressing either sgRlfsgMycl or control sgNeosgNeo. Six months post-infection we collected lungs for histological analysis to assess tumor incidence and burden (Fig. 3A). Mice transduced with sgRlfsgMycl (n=17) had nearly three times higher tumor burden (p=0.012) (Fig. 3B) and larger tumor areas (p=0.041) (Fig. 3C) than sgNeosgNeo controls (n=20) estimated by quantitative histology (Fig. 3D). To analyze tumor progression over time with a higher potential penetrance of *Rlf-Mycl* rearrangement, we transduced a cohort of adult RPR2C mice by intratracheal delivery with 10 times higher viral titers of the USEC lentiviruses using both chimeric (n=14 sgNeosgNeo/ 13 sgRlfsgMycl) and fully transgenic (n=19 sgNeosgNeo/ 13 sgRlfsgMycl) mice, and monitored these animals for development of respiratory distress or other adverse symptoms requiring euthanasia (Fig. 3E&F). We assessed tumor burden in living mice at 6 months following USEC induction by MRI. We detected significantly greater tumor volume in sgRlfsgMycl-transduced mice in both chimeric (p=0.047) and fully transgenic cohorts (p=0.016) (Fig. 3G–J). A higher tumor burden was also evident upon histologic analysis of lungs of the sgRlfsgMycl as compared to sgNeosgNeo chimeric (p=0.01) and fully transgenic cohorts (p=0.03) (Supplementary Fig. S3A–D). Overall survival of chimeric sgRlfsgMycl mice was significantly shorter than that of sgNeosgNeo mice (p=0.001) (Supplementary Fig. S3E). No significant difference in overall survival was observed in the fully transgenic cohort (p=0.87) (Supplementary Fig. S3F). Relative to sgRlfsgMycl mice, in sgNeosgNeo mice we noted lower overall parenchymal tumor burden but also frequent occurrence of centrally located thoracic tumors leading to airway compromise, potentially explaining the similar duration of overall survival between these cohorts.

To evaluate whether a change in tumor spectrum might affect the acceleration of tumor progression in sgRlfsgMycl mice, histological analyses were performed by a pathologist. In both sgRlfsgMycl cohorts, the majority of tumors identified displayed typical histologic features of SCLC and stained positive for the neuroendocrine marker SYP (Supplementary Fig. S3G). Both 6-month and end-stage tumors showed a predominance of SCLC; in the chimeric cohort we observed some cases of admixed SCLC with large cell neuroendocrine carcinoma (LCNEC). The tumor architecture was classical and/or trabecular with lymphovascular invasion observed in most end-stage cases. We confirmed the presence of *Rlf-Mycl* fusion transcript by droplet PCR on cDNA from micro-dissected tumors from both chimeric (3 out of 3) and fully transgenic (5 out of 8) animals infected with sgRlfsgMycl (Supplementary Fig. S3H). *Mycl1* transcripts by droplet PCR on cDNA generally matched *Rlf-Mycl* fusion transcript levels of *Rlf-Mycl* fusion mice (Supplementary Fig. S3I). Taken together these data illustrate successful development of a novel SCLC GEMM harboring the *Rlf-Mycl* gene fusion and further support the oncogenic function of the RLF-MYCL fusion in SCLC.

RLF-MYCL fusion promotes metastasis in SCLC

Human SCLC has a remarkable predilection for metastatic spread: approximately two thirds of patients have distant metastases evident at the time of diagnosis. In addition to analyzing the role of RLF-MYCL in promoting SCLC initiation and growth, we sought to clarify its potential contributions to metastasis. This was prompted in part by the observation

that essentially all sgNeosgNeo-transduced mice ultimately required euthanasia due to respiratory distress from proximal airway compromise, whereas sgRlfsgMycl-transduced mice displayed a broader array of distress symptoms including abdominal distension and lethargy. At necropsy, 93% of sgRlfsgMycl chimeric mice and 15% of sgNeosgNeo chimeric mice displayed overt metastasis (p<0.0001) (Fig. 4A and Supplementary S4A). A similar pattern was observed in the fully transgenic animals, in which 92% of sgRlfsgMycl mice displayed overt metastasis compared to 39% of sgNeosgNeo mice (p=0.004) (Fig. 4A). The organ distribution of observed metastases also differed between cohorts. Most notably, sgRlfsgMycl-transduced mice demonstrated widespread metastatic disease. Distant metastases to multiple organs in individual mice was observed in approximately 30% of chimeric and over 40% of fully transgenic sgRlfsgMycl mice (Fig. 4B), while only liver metastases were observed in sgNeosgNeo mice. In sgRlfsgMycl-transduced mice, other sites of metastasis included thoracic, paraspinal and cervical lymph nodes, spleen, kidney and mesentery (Fig. 4C and Supplementary S4A). Liver metastases of mice transduced with sgRlfsgMycl were substantially more extensive than those of control sgNeosgNeo mice, as assessed by both histologic sections and evaluation of total liver weight (p=0.04) (Fig. 4D&E). Histopathological examination including staining for the neuroendocrine marker SYP confirmed SCLC phenotype in these metastases (Supplementary Fig. S4B). We also confirmed *Rlf-Mycl* fusion transcript by droplet PCR using cDNA from several metastatic nodules from both chimeric (3/3) and fully transgenic (5/6) sgRlfsgMycl-transduced animals (Supplementary Fig. S4C). To evaluate whether metastatic disease is an early event during SCLC tumor development in this *Rlf-Mycl* fusion model, we analyzed livers of chimeric RPR2C transduced mice at an early timepoint 6 months post-tumor initiation. We found that transduction with sgRlfsgMycl resulted in evident micro-metastases to the liver in 3/5 animals versus 0/5 control sgNeosgNeo transduced mice (p=0.038) (Supplementary Fig. S4D & S4E). Collectively, these findings suggest that RLF-MYCL both accelerates tumorigenesis and promotes early metastatic dissemination in SCLC.

Rlf-Mycl SCLC has an ASCL1^{high} phenotype consistent with SCLC-A subtype

Human RLF-MYCL tumors and cell lines are all SCLC-A subtype (Supplementary Fig. S1). To determine the SCLC subtype of mouse tumors with *Rlf-Mycl* fusion, we compared lung tumors generated by sgNeosgNeo and sgRlfsgMycl to tumors from previously developed SCLC mouse models including RPR2 mice and RPM *(Rb1$^{fl/fl}$;Trp53$^{fl/fl}$;Myc$^{LSL/LSL}$)* mice (7,13,16). Tumors in RPR2 mice demonstrate an exclusively ASCL1^{high} phenotype, while RPM mice develop a broader spectrum of phenotypes including a predominance of NEUROD1^{high}-tumors resembling SCLC-N and aggregates of POU2F3 and YAP1 expression. Tumors from both sgNeosgNeo- and sgRlfsgMycl-transduced animals demonstrated high expression of ASCL1 and a second neuroendocrine transcription factor INSM1; these findings were similar to and consistent with RLF-MYCL-positive human tumors (n=5/group) and RPR2 mice (Fig. 5A&B and Supplementary Fig. S5A). In contrast to tumors from RPM mice, tumors from the sgNeosgNeo and sgRlfsgMycl models did not express detectable NEUROD1, POU2F3 or YAP1 by IHC (n=5/group; p<0.0001) (Fig. 5C and Supplementary Fig. S5B, C&D). Human SCLC tumors have also been characterized into neuroendocrine (NE) and non-NE classes based on a 50-gene signature, identifying SCLC-A as NE^{high} tumors (36). Exploring

expression across a panel of orthologous murine genes did not reveal consistent differences in NE state between sgRlfsgMycl and sgNeosgNeo lung tumors.

NFIB expression has been reported in multiple mouse models and demonstrates inter- and intra-tumoral heterogeneity in NE[high] ASCL1[high] models (7,12,22). We confirmed a similarly heterogeneous distribution of NFIB expression in lung tumors at 6 months and in end stage lung tumors of both fully transgenic and chimeric mice (n=5/group) (Supplementary Fig. S5E). This pattern is reminiscent of what others have reported with amplification of chromosome 4, containing both *Mycl* and *Nfib* (12,22). NFIB expression increased significantly at later stages of tumor development, potentially associated with progressive metastasis, though no significant differences were observed between sgRlfsgMycl and sgNeosgNeo mice (n=5/group) (Supplementary Fig. S5E). Taken together, these data support that RLF-MYCL fusion, in contrast to MYC, does not shift SCLC subtype but rather leads to more rapid progression of ASCL1[high] SCLC, consistent with the observed phenotype of the corresponding fusion in human SCLC.

Expression profiling of murine RLF-MYCL and human RLF-MYCL SCLC

Given that RLF-MYCL fusion promotes both pro-tumorigenic and pro-metastatic phenotypes in mice, we sought to determine the underlying transcriptional programs driving *Rlf-Mycl* tumorigenesis, with reference to both sgNeosgNeo tumors and human RLF-MYCL-positive SCLC. We performed RNA-Seq from transgenic sgRlfsgMycl (n=4) and sgNeosgNeo (n=7) tumors. We observed higher *Mycl* expression levels in sgRlfsgMycl tumors relative to sgNeosgNeo lung tumors (Fig. 6A), consistent with our earlier observation in human *RLF-MYCL* fusion-positive SCLC versus other SCLC-A samples (Fig. 1D). We reanalyzed RNA-Seq data from our previously published dataset of human SCLC (8) to identify genes differentially expressed between human SCLC carrying the *RLF-MYCL* fusion (n=4) relative to other SCLC-A subtype SCLC (n=29) (Fig. 6B). The 10 most differentially expressed genes included *PPT1*, *PPIE*, *RLF*, *SNX9*, *OTUD1*, *MYO6*, *NBPF*, *C1orf220*, *PLEKHG1* and *HSD17B13* (Fig. 6B). Of these, neither *NBPF* nor *C1orf220* has a known mouse ortholog. We sought to determine whether the mouse *Rlf-Mycl* primary lung tumors differentially expressed any orthologs of the remaining 8 differentially expressed genes. We were able to detect nominally significant differences in the expression of 3 out of the 8 orthologous genes: similar to human RLF-MYCL samples, *Pp1t*, *Ppie* and *Rlf* were overexpressed in RLF-MYCL-containing mouse lung tumors relative to sgNeosgNeo lung tumors (Fig. 6C). Differential expression was not confirmed in the other 5 candidates. Palmitoyl-protein thioesterase (PPT1) has been previously associated with tumor growth and metastasis (37) and elevated expression of PPT1 in tumors correlates with poor patient survival in a variety of cancers (38). Peptidyl-prolyl cis-trans isomerase E (PPIE, or CYP33) has been implicated in processes including metabolism, apoptosis, inflammation and cancer (39).

Finally, to investigate pathways that discriminate fusion-associated from non-fusion-associated ASCL1-subtype SCLC, we performed gene-set enrichment analysis (GSEA) on both the human and mouse RNA-seq datasets. GSEA of human SCLC revealed significant enrichment for 6 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, all

downregulated in RLF-MYCL fusion tumors relative to non-fusion SCLC-A tumors (Fig. 6D; Supplementary Table S2), notably including pathways downregulated in contexts of invasion and metastasis. Negatively enriched KEGG pathways included cell adhesion molecules, ECM-receptor interaction and cytokine-cytokine receptor interaction, allograft rejection, complement and coagulation cascades and focal adhesion. We then explored whether any of these 6 KEGG pathways could similarly distinguish murine sgRlfsgMycl versus sgNeosgNeo lung tumors. Strikingly, GSEA revealed that 5 of the 6 KEGG pathways identified as significantly downregulated in human RLF-MYCL SCLC were also significantly suppressed in mouse RLF-MYCL tumors (Fig. 6D&E). These included 3 of the top 5 differentially regulated pathways in the mouse dataset: cell adhesion molecules, cytokine-cytokine receptor interaction, and ECM receptor interaction (Supplementary Table S3). Additional KEGG pathways significantly associated with sgRlfsgMycl vs. sgNeosgNeo lung tumors included upregulation of glycosaminoglycan biosynthesis heparan sulfate (40), consistent with the tumor growth phenotype we observed in *Rlf-Mycl* fusion mice (Fig. 6E). Both the individual gene data and in particular the KEGG pathway analyses point to hallmarks shared between human RLF-MYCL SCLC and the corresponding mouse model with CRISPR/Cas9-engineered RLF-MYCL fusion.

DISCUSSION

Recurrent structural gene rearrangements in SCLC have been only minimally cataloged (8,9,28), and the potential roles of fusion oncogenes as initiators and promoters of cancer growth and spread in SCLC have not been previously defined. In the work reported here, we sought to explore the role of the most commonly reported gene fusion in human SCLC, RLF-MYCL, primarily through analysis of GEMMs harboring this fusion event using CRISPR-mediated genome editing *in vivo*. The ability to effectively model this fusion was aided by the proximity of the relevant genes and by the fact that the local chromosome environments represent syntenic blocks in the mouse and human genomes. The opposite orientations of the *Rlf* and *Mycl* genes on mouse chromosome 4 and of the *RLF* and *MYCL* genes on human chromosome 1 both require a small inversion event to generate the relevant fusion gene. Using CRISPR/Cas9-based editing in a SCLC precursor model and GEMMs, we were able to successfully generate the *Rlf-Mycl* fusion and demonstrate its potential to accelerate oncogenesis and promote widespread metastasis of SCLC.

Chromosomal translocations resulting in gene fusions were first recognized as a mechanism of oncogenesis in hematologic malignancies and have been subsequently identified as drivers of a wide range of solid tumors (26). The identification and characterization of oncogenic fusions in non-small cell lung cancers (NSCLCs) has been critical not only in understanding the biology of these tumors, but in defining tumor-specific therapeutic targets for intervention. ALK, RET, ROS1, and the NTRK family members are all tyrosine kinases, which has facilitated rapid development of highly selective inhibitors for tumors driven by fusions activating or stabilizing these kinases (41). MYC family members have been known as oncogenic drivers for a much longer period of time than any of these non-small cell lung cancer targets, yet the development of selective targeted inhibitors for MYC family members remains an unmet challenge (42). Our data supporting that the RLF-MYCL fusion represents a driver of oncogenesis and metastasis provides rationale for exploring this fusion

as a therapeutic target. While inhibiting MYCL activity remains a challenge, the fusion would represent a uniquely tumor-specific target for strategies such as a proteolysis-targeting chimera (PROTAC) (43). The *Rlf-Mycl* fusion GEMMs described in this study can be used in future preclinical studies to test novel therapeutic strategies, and to assess the relative sensitivity to standard chemotherapy of these tumors.

Recent profiling of both human and mouse SCLC suggests that most tumors can be categorized into predominant subtypes based on differential expression of key transcription factors (5). MYCL expression is primarily associated with tumors of the SCLC-A subtype, expressing high levels of ASCL1; in contrast, MYC expression is associated with the other major subtypes. All reported examples of human SCLC harboring the *RLF-MYCL* fusion are of the SCLC-A subtype, which is recapitulated in both primary and metastatic lesions from the *Rlf-Mycl* fusion GEMMs reported here. That RLF-MYCL fusion is associated with the same subtype of SCLC as tumors overexpressing wildtype MYCL raises the question of whether the fusion protein is oncogenic only by increasing MYCL signaling (a hypermorphic function), or also by altering the nature of MYCL signaling (a neomorphic function). These are not mutually exclusive, but we do believe our data supports a neomorphic function – not only increasing the pace of oncogenesis but also changing the nature of progression, enhancing the frequency and distribution of metastases.

One distinction between human SCLC and most of the mouse models generated to date is that whereas the human disease is notorious for early and widespread metastases, most SCLC GEMMs generated to date develop primary tumors in the lung, with some demonstrating predominantly liver metastasis as recapitulated in our analysis of RPR2 and RPM models (14). The *Rlf-Mycl* GEMMs described here may more closely resemble the exceptionally broad metastatic tropism of human SCLC including distant nodal and multiorgan disease, representing a valuable model to facilitate research into the metastatic drivers of SCLC. However the distribution of metastatic sites in the *Rlf-Mycl* fusion GEMMs does not fully recapitulate that of the human disease, lacking evident predilection for some common metastatic sites in patients, most notably the central nervous system.

In summary, we present here the first mouse model of a recurrent gene fusion in SCLC, demonstrating that the RLF-MYCL fusion promotes oncogenic transformation in precursor cells, accelerates tumorigenesis *in vivo*, and facilitates frequent metastases to multiple organs. The resultant murine tumors consistently demonstrate an ASCL1[high] state, consistent with human SCLC harboring the homologous RLF-MYCL fusion. Further analysis of selective vulnerabilities of this mouse model may yield targeted strategies for better treatment of the human disease.

METHODS

Human genomics

RNA-Seq fastq files from primary tumors and cell lines were obtained from (8)(N=51) and (9)(N=54). Sequencing files were aligned to GRCh37 reference genome using the STAR aligner algorithm v2.7 (44). The resulting BAM files were interrogated to identify gene fusions by the Arriba fusion detection algorithm v1.1.0 (https://github.com/suhrig/arriba/)

and FusionCatcher v1.10 (https://github.com/ndaniel/fusioncatcher). Gene expression was quantified using Kallisto v0.45.0 (45) to Ensembl v75. We determined the subtype of each sample based on the highest expression of four genes: *ASCL1, NEUROD1, YAP1*, and *POU2F3*. All samples with *RLF-MYCL* gene fusion were classified as SCLC-A (ASCL1-high), and further comparative analyses were performed within the 33 samples of subtype SCLC-A from (8). We used Sleuth (46) and the following model to identify genes that were differentially expressed in the samples with *RLF-MYCL* fusions: Expression ~ SampleType + Fusion, where Fusion is 1 or 0 depending on detection by either Arriba or FusionCatcher and SampleType is cell line or primary tumor. We identified 10 genes that were differentially expressed (q-value < 0.05). Heat maps of differentially expressed genes represent z-scores calculated across samples. Clustering was performed using Manhattan Distance with Ward D method.

Cell Lines and in vitro assays

PreSC cells were generously provided by Dr. Kwon Park and were derived as previously described (29). Human SCLC cell lines were obtained from ATCC or Sigma and cultured in either RPMI (Corning #10-041-CV) supplemented with 10% fetal bovine serum (FBS) (Gemini Bio #900-108), 1% L-glutamine (Fisher #25030149), and 1% penicillin/ streptomycin antibiotic cocktail (Fisher #15070063) (H187, CORL88, H889, H209, H1963, H82, CORL47, H1092, preSC); modified HITES medium DMEM/F12 (ATCC #30-2006) supplemented with 0.005 mg/mL insulin (Fisher #12585-014), 0.01 mg/mL transferrin (Sigma #T2036), 30 nM sodium selenite (Sigma #S5261), 10 nM hydrocortisone (Sigma #H0888), 10 nM beta-estradiol (Sigma #E2758), 1% L-glutamine (Fisher #25030149), and 1% penicillin/streptomycin antibiotic cocktail (Fisher #15070063) (H1882 and H1836); or DMEM (Corning #10–014-CV) with 10% FBS, 1% L-glutamine, and 0.1% penicillin/ streptomycin antibiotic cocktail (HEK293T, NIH/3T3 and Green-Go cells). Soft agar assay was performed following standard protocols, seeding 1×10^5 preSC's per well in 6-well plates, resuspended in 1.5mL of the RPMI growth media containing 0.35% agar (Invitrogen #16500500) on bottom layer that contained 1.5mL of the growth media containing 0.5% agar (Invitrogen #16500500). The media was regularly changed every 3 days for 6 weeks. After incubation, colonies were counted by using Image J software.

RT-PCR and ddPCR analysis for gene fusion detection

Total RNA was extracted from fresh frozen cell pellets using RNeasy Plus Universal mini kit (QIAGEN # 15596026) following manufacturer's instructions). Extracted RNA samples were quantified using Qubit 2.0 Fluorometer (Life Technologies). cDNAs were prepared using the Superscript III Reverse Transcriptase kit (ThermoFisher #18080093) following manufacturer's instructions. For droplet digital PCR (ddPCR), total RNAs from cell pellets or tissue were extracted with Trizol (ThermoFisher #15596026) following manufacturer's instructions. Quantity was assessed by PicoGreen (ThermoFisher) and quality by BioAnalyzer (Agilent). Droplet generation was performed on a QX200 ddPCR system (Bio-Rad) using cDNA generated from 7 ng total RNA with the One-Step RT-ddPCR Advanced Kit for Probes (Bio-Rad #1864021) according to the manufacturer's protocol with reverse transcription at 42 °C and annealing/extension at 55 °C. Each sample was evaluated in technical duplicates. Reactions were partitioned into a median of approximately 30,000

droplets per well. Plates were read and analyzed with the QuantaSoft software to assess the number of droplets positive for the gene of interest and reference gene. PrimePCR ddPCR Expression Probe Assays were ordered through Bio-Rad for the following genes of interest: *Rlf-Mycl*, *Mycl* and *B2M*. Primers, primer pairs and probes used in the various PCR reactions (Supplementary Table S1).

Immunoblotting

Protein extraction and western blot were performed as previously described (47). After quantification of protein extracts using the Bradford method (Bio-Rad #5000205), we loaded 20–30 ug of total protein in the gels. Primary antibodies for MYCL (R&D #AF4050) and Vinculin (CST #13901S) were used at dilutions recommended by manufacturer. Donkey anti-Goat (Li-Cor Biosciences #925-32214) and Donkey anti-Rabbit (Li-Cor Biosciences #926-32213) secondary antibodies were used.

Genomic DNA isolation and Surveyor assay

Genomic DNA from cell lines or snap-frozen flank tumors of preSC's was isolated using the DNeasy Blood & Tissue Kit (Qiagen # 69504) following the manufacturer's guidelines. PCR products for surveyor assay were amplified using Phusion High-Fidelity DNA Polymerase (NEB #M0530S), gel purified and subsequently assayed with the Surveyor Mutation Detection Kit Integrated DNA Technologies, Inc (IDT #706020). Primers for fusion detection and surveyor assay are as indicated in Supplementary Table S1.

sgRNA cloning and lentiviral vector cloning

Oligos containing the sgRNAs sequences targeting *Rlf*, *Mycl or Neo* were designed and purchased from Integrated DNA Technologies, Inc (IDT). The pX330 vector expressing Cas9 (Addgene #42230) was digested with BbsI (ThermoFisher #ER1011) and ligated to annealed and phosphorylated sgRNA oligos targeting *Rlf* and *Mycl*. To deliver both sgRNAs in the same lentiviral vector we followed the cloning strategy described in (34). Briefly, sU6 promoter is inserted to 2sgRNA ultramer through PCR and Gibson assembly with pDonor_sU6 (Addgene plasmid #69351), then digested with BbsI restriction enzyme and ligated to pUSEC backbone that was digested with BsmBI restriction enzyme (ThermoFisher #ER0451) (35). Subsequently, colonies of transformed Stbl3 bacteria (ThermoFisher #C737303) were picked and sequenced by Sanger sequencing. Oligos sequences (Supplementary Table S1).

Lentiviral production

Lentiviruses were produced by co-transfection of HEK293T cells with lentiviral backbone of LentiCas9-blast (Addgene #52962) or USEC constructs, and packaging vectors (delta8.2 and VSV-G) using TransIT-LT1 Transfection Reagent (Mirus Bio #MIR2304). Supernatants were collected 48- and 72-hours post-transfection, concentrated by ultracentrifugation at 25,000 RPM for 120 minutes and resuspended in 50ul PBS (Gibco). Green-Go cells (48) were used to determine lentiviruses titer.

Transfections and transductions

NIH/3T3 and preSC cells were transfected in T75-flasks with 10 μg of total plasmid DNA using lipofectamine 3000 reagent (ThermoFisher #L3000001) following manufacturer's instructions. To enrich for transfected cells, transfections included 3 μg of a plasmid expressing the Puro-resistance gene (pSico) (Addgene #11578) and cells were incubated with 2 μg/ml Puromycin (ThermoFisher #A1113803) for 2 days. Transductions were performed by adding lentiviruses directly to each T75-flask.

Animal Studies

Nude mice were purchased from Charles River (stock #194). $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;$Rbl2^{fl/fl}$;$Rosa26^{LSL-Cas9-GFP}$ (RPR2C) embryonic stem cells (ESC) containing Rosa26-CAGGS-LSL-Cas9-GFP-Csy4 were generated by the Jacks laboratory at Massachusetts Institute of Technology as previously described (31). RPR2C chimeric animals were generated by the NYUMC Rodent Genetic Engineering Laboratory (RGEL). Briefly, RPR2C ES cells were injected into C57BL/6N blastocysts (stock #005304) (JAX) and cultured 2–3 hrs in KSOM+AA medium (Sigma # MR-101-D) to blastocyst stage and implanted into pseudopregnant CD-1 (stock #022) (Charles River laboratory) fosters. To generate approximately 10–20 chimeras, we implanted 30–60 embryos per ESC injection session. Chimeric pups that were over 6 weeks old displayed high degree of chimerism. To generate fully transgenic RPR2C animals, high degree RPR2C chimeric animals (described above) were crossed to $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;$Rbl2^{fl/fl}$ (RPR2) (16) (MMRRC: 043692-UCD) animals until RPR2 homozygous genes were acquired. Tumors in RPR2C mice were initiated by intratracheal instillation of 0.5×10^6 (chimeric) or 5×10^6 (chimeric and full transgenic) transduction units (TU) lentivirus expressing Cre recombinase, as previously described (49). Sections from previously described $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;$MycT58A^{LSL/LSL}$ (RPM) (JAX #029971) (7) and RPR2 mice (n=5/genotype) were placed on slides of immunohistochemistry. For all animal studies, both male and female mice were equally divided between treatment groups. Investigators were not blinded with respect to which lentivirus was injected. All studies and procedures were approved by the Memorial Sloan-Kettering Cancer Center Institutional Animal Care and Use Committee or the NYU Langone Medical Center Institutional Animal Care and Use Committee.

MRI analysis

Mouse scans were performed using a 9.4T 20-cm bore Bruker Biospec scanners (Bruker Biospin MRI GmbH) equipped with an ID 114 mm maximum strength of 530 mT/m Bruker gradient. An ID 40 mm Bruker volume resonator was used for RF excitation and MRI acquisition. The mice were anesthetized with 2% isoflurane (Baxter Healthcare Corp) gas in oxygen during MRI scanning. A small animal physiological monitoring system (SA Instruments) was used to monitor animal respiration during MRI scanning. Scout images along three orthogonal orientations were first acquired for animal positioning. For mouse lung imaging, respiratory gated T1-weighted axial images using the FLASH gradient echo sequence were acquired with TR 170 ms, TE 1.6 ms, slice thickness of 0.8 mm, FOV 35 × 25 mm, in-plane resolution of 182 × 130 um, and 10 averages. The tumors were measured from their Digital Imaging and Communications in Medicine (DICOM) files using Image J

software. Tumor burden was calculated by outlining the region of interest (ROI) of tumor structures, taking the output of the ROIs in mm^2 and multiplying each slice's ROI by its slice thickness.

Immunohistochemistry analysis

Lung, liver and other metastatic samples were fixed in 4% formalin and paraffin embedded. Hematoxylin and eosin staining was performed using standard methods. Formalin-fixed paraffin embedded (FFPE) sections at 4–5 μm were dewaxed, rehydrated and subjected to high-temperature antigen retrieval by boiling 15 min in a pressure cooker in 0.01 M citrate buffer at pH 6.0. Slides were quenched of endogenous peroxide in 3% H_2O_2 for 10 min, then blocked in 5% goat serum in PBS/0.1% Tween-20 (PBS-T) for 30 min, and then stained overnight with primary antibodies in blocking buffer (5% goat serum) (CST #8112). For non-CST primary antibodies, an HRP-conjugated secondary antibody (Vector Laboratories) was used at 1:200 dilution in PBS-T, incubated for 30 min at RT followed by DAB staining (Vector Laboratories). Alternatively, CST primary antibodies were detected using 200 μL of SignalStain Boost IHC Detection Reagent (CST #8114). All stainings were performed with Sequenza cover plate technology. For IHC, we used antibodies to NFIB (Sigma #HPA003956) 1:250; ASCL1 (BD #BD556604) 1:200; NEUROD1 (Abcam #109224) 1:150; INSM1 (Santa Cruz Biotechnology #sc-271408) 1:100; POU2F3 (Sigma #HPA019652) 1:300; YAP1 (CST #14074S) 1:400; and Synaptophysin (Neuromics #MO20000) 1:250. For manual H-score quantification, images were acquired on a Nikon Ci-L LED Microscope with DS-Fi3 Camera. H-score was quantified on a scale of 0–300 taking into consideration percent positive cells and staining intensity as described (50), where H Score = % of positive cells multiplied by intensity score of 0–3. For example, a tumor with 80% positive cells with high intensity of 3 = 240 H-Score. H&E and IHC-stained slides were digitally scanned with the Zeiss Axio Scope.A1 microscope using AxioVision SE64 software. Whole slide images containing 4–5 lung lobes per animal were analyzed using CaseViewer software (3DHISTECH). Tumor regions were manually annotated.

Mouse RNA Sequencing

RNA extraction of snap-frozen micro dissected lung tumors and sequencing was done in collaboration with Genewiz and Integrated Genomics Operation (IGO) at MSKCC. RNA of full transgenic sgNeosgNeo mice (n=7) and sgRlfgMycl mice (n=7) was subject to library construction prepared using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB) or TruSeq Stranded mRNA library Prep Kit (Illumina) following manufacturer's instructions. The sequencing libraries were validated on the Agilent TapeStation (Agilent Technologies), and quantified by using Qubit 2.0 Fluorometer (Invitrogen) as well as by quantitative PCR (KAPA Biosystems). The sequencing libraries were clustered on 2 lanes of a flowcell. After clustering, the flowcell was loaded on the Illumina HiSeq instrument (4000) according to manufacturer's instructions. The samples were sequenced using a 2×150bp Paired End (PE) configuration. Image analysis and base calling were conducted by the HiSeq Control Software (HCS). Raw sequence data (.bcl files) generated from Illumina HiSeq were converted into fastq files and de-multiplexed using Illumina's bcl2fastq 2.17 software.

All samples were processed with Kallisto (v0.45.0) (45) to GRCm38 cDNA (Ensembl release 100). We identified *Rlf-Mycl* gene fusion in 4 out of 7 samples via either FusionCatcher fusion detection algorithm (v1.10, http://code.google.com/p/fusioncatcher) (n=1) and/or ddPCR (n=4). We used Sleuth (46) and the following model to study differential expression in the samples with *Rlf-Mycl* fusions: Expression ~ Batch + Fusion, where Fusion is 1 or 0 depending on detection by either Arriba or FusionCatcher or ddPCR and Batch denotes sequencing at IGO or Genewiz.

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) was performed using KEGG pathways pulled from MSigDB using the R package msigdbr (51). When more than one Ensembl ID mapped to a given Hugo Symbol, the Ensembl ID with the lowest p-value from differential expression analysis was used to represent the gene. Ranking of genes was determined by effect size (beta) from Kallisto. GSEA was called from clusterProfiler R package (52) sing default parameters. Term was considered significant if BH-adjusted p-value < 0.05. GSEA plots were created by modifying gseaplot2 from enrichplot R package (https://github.com/YuLab-SMU/biomedical-knowledge-mining-book). Human and mouse KEGG pathways (Supplementary Table S2,3).

Statistical analysis

Remaining statistical analyses were performed using GraphPad Prism 7 software. Error bars show mean ± SD. Significance was determined by Student's two-tailed unpaired t tests with 95% confidence intervals and p values <0.05 considered statistically significant. Survival analysis was performed by comparing two survival curves using the Log-rank (Mantel-Cox) test. Statistical details are further described in respective Figure legends.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

1. Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. Nat Rev Dis Primers 2021;7(1):3 doi 10.1038/s41572-020-00235-0. [PubMed: 33446664]

2. Poirier JT, George J, Owonikoko TK, Berns A, Brambilla E, Byers LA, et al. New Approaches to SCLC Therapy: From the Laboratory to the Clinic. J Thorac Oncol 2020;15(4):520–40 doi 10.1016/j.jtho.2020.01.016. [PubMed: 32018053]

3. Rudin CM, Awad MM, Navarro A, Gottfried M, Peters S, Csoszi T, et al. Pembrolizumab or Placebo Plus Etoposide and Platinum as First-Line Therapy for Extensive-Stage Small-Cell Lung Cancer: Randomized, Double-Blind, Phase III KEYNOTE-604 Study. J Clin Oncol 2020;38(21):2369–79 doi 10.1200/JCO.20.00793. [PubMed: 32468956]

4. Gazdar AF, Bunn PA, Minna JD. Small-cell lung cancer: what we know, what we need to know and the path forward. Nat Rev Cancer 2017;17(12):725–37 doi 10.1038/nrc.2017.87. [PubMed: 29077690]

5. Rudin CM, Poirier JT, Byers LA, Dive C, Dowlati A, George J, et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. Nat Rev Cancer 2019;19(5):289–97 doi 10.1038/s41568-019-0133-9. [PubMed: 30926931]

6. Borromeo MD, Savage TK, Kollipara RK, He M, Augustyn A, Osborne JK, et al. ASCL1 and NEUROD1 Reveal Heterogeneity in Pulmonary Neuroendocrine Tumors and Regulate Distinct Genetic Programs. Cell Rep 2016;16(5):1259–72 doi 10.1016/j.celrep.2016.06.081. [PubMed: 27452466]

7. Mollaoglu G, Guthrie MR, Bohm S, Bragelmann J, Can I, Ballieu PM, et al. MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. Cancer Cell 2017;31(2):270–85 doi 10.1016/j.ccell.2016.12.005. [PubMed: 28089889]

8. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 2012;44(10):1111–6 doi 10.1038/ng.2405. [PubMed: 22941189]

9. George J, Lim JS, Jang SJ, Cun Y, Ozretic L, Kong G, et al. Comprehensive genomic profiles of small cell lung cancer. Nature 2015;524(7563):47–53 doi 10.1038/nature14664. [PubMed: 26168399]

10. Chalishazar MD, Wait SJ, Huang F, Ireland AS, Mukhopadhyay A, Lee Y, et al. MYC-Driven Small-Cell Lung Cancer is Metabolically Distinct and Vulnerable to Arginine Depletion. Clin Cancer Res 2019;25(16):5107–21 doi 10.1158/1078-0432.CCR-18-4140. [PubMed: 31164374]

11. Gay CM, Stewart CA, Park EM, Diao L, Groves SM, Heeke S, et al. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. Cancer Cell 2021;39(3):346–60 e7 doi 10.1016/j.ccell.2020.12.014. [PubMed: 33482121]

12. Dooley AL, Winslow MM, Chiang DY, Banerji S, Stransky N, Dayton TL, et al. Nuclear factor I/B is an oncogene in small cell lung cancer. Genes Dev 2011;25(14):1470–5 doi 10.1101/gad.2046711. [PubMed: 21764851]

13. Cui M, Augert A, Rongione M, Conkrite K, Parazzoli S, Nikitin AY, et al. PTEN is a potent suppressor of small cell lung cancer. Mol Cancer Res 2014;12(5):654–9 doi 10.1158/1541-7786.MCR-13-0554. [PubMed: 24482365]

14. Ko J, Winslow MM, Sage J. Mechanisms of small cell lung cancer metastasis. EMBO Mol Med 2021;13(1):e13122 doi 10.15252/emmm.202013122. [PubMed: 33296145]

15. McFadden DG, Papagiannakopoulos T, Taylor-Weiner A, Stewart C, Carter SL, Cibulskis K, et al. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. Cell 2014;156(6):1298–311 doi 10.1016/j.cell.2014.02.031. [PubMed: 24630729]

16. Schaffer BE, Park KS, Yiu G, Conklin JF, Lin C, Burkhart DL, et al. Loss of p130 accelerates tumor development in a mouse model for human small-cell lung carcinoma. Cancer Res 2010;70(10):3877–83 doi 10.1158/0008-5472.CAN-09-4228. [PubMed: 20406986]

17. Meuwissen R, Linn SC, Linnoila RI, Zevenhoven J, Mooi WJ, Berns A. Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell 2003;4(3):181–9 doi 10.1016/s1535-6108(03)00220-4. [PubMed: 14522252]

18. Huijbers IJ, Bin Ali R, Pritchard C, Cozijnsen M, Kwon MC, Proost N, et al. Rapid target gene validation in complex cancer mouse models using re-derived embryonic stem cells. EMBO Mol Med 2014;6(2):212–25 doi 10.1002/emmm.201303297. [PubMed: 24401838]

19. Sutherland KD, Proost N, Brouns I, Adriaensen D, Song JY, Berns A. Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. Cancer Cell 2011;19(6):754–64 doi 10.1016/j.ccr.2011.04.019. [PubMed: 21665149]

20. Yang D, Denny SK, Greenside PG, Chaikovsky AC, Brady JJ, Ouadah Y, et al. Intertumoral Heterogeneity in SCLC Is Influenced by the Cell Type of Origin. Cancer Discov 2018;8(10):1316–31 doi 10.1158/2159-8290.CD-17-0987. [PubMed: 30228179]

21. Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, et al. MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate. Cancer Cell 2020;38(1):60–78 e12 doi 10.1016/j.ccell.2020.05.001. [PubMed: 32473656]

22. Semenova EA, Kwon MC, Monkhorst K, Song JY, Bhaskaran R, Krijgsman O, et al. Transcription Factor NFIB Is a Driver of Small Cell Lung Cancer Progression in Mice and Marks Metastatic Disease in Patients. Cell Rep 2016;16(3):631–43 doi 10.1016/j.celrep.2016.06.020. [PubMed: 27373156]

23. Denny SK, Yang D, Chuang CH, Brady JJ, Lim JS, Gruner BM, et al. Nfib Promotes Metastasis through a Widespread Increase in Chromatin Accessibility. Cell 2016;166(2):328–42 doi 10.1016/j.cell.2016.05.052. [PubMed: 27374332]

24. Yang B, Zhou ZH, Chen L, Cui X, Hou JY, Fan KJ, et al. Prognostic significance of NFIA and NFIB in esophageal squamous carcinoma and esophagogastric junction adenocarcinoma. Cancer Med 2018;7(5):1756–65 doi 10.1002/cam4.1434. [PubMed: 29577671]

25. Shue YT, Lim JS, Sage J. Tumor heterogeneity in small cell lung cancer defined and investigated in pre-clinical mouse models. Transl Lung Cancer Res 2018;7(1):21–31 doi 10.21037/tlcr.2018.01.15. [PubMed: 29535910]

26. Schram AM, Chang MT, Jonsson P, Drilon A. Fusions in solid tumours: diagnostic strategies, targeted therapy, and acquired resistance. Nat Rev Clin Oncol 2017;14(12):735–48 doi 10.1038/nrclinonc.2017.127. [PubMed: 28857077]

27. Makela TP, Saksela K, Evan G, Alitalo K. A fusion protein formed by L-myc and a novel gene in SCLC. EMBO J 1991;10(6):1331–5. [PubMed: 1851085]

28. Iwakawa R, Takenaka M, Kohno T, Shimada Y, Totoki Y, Shibata T, et al. Genome-wide identification of genes with amplification and/or fusion in small cell lung cancer. Genes Chromosomes Cancer 2013;52(9):802–16 doi 10.1002/gcc.22076. [PubMed: 23716474]

29. Kim DW, Wu N, Kim YC, Cheng PF, Basom R, Kim D, et al. Genetic requirement for Mycl and efficacy of RNA Pol I inhibition in mouse models of small cell lung cancer. Genes Dev 2016;30(11):1289–99 doi 10.1101/gad.279307.116. [PubMed: 27298335]

30. Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature 2014;516(7531):423–7 doi 10.1038/nature13902. [PubMed: 25337876]

31. Ng SR, Rideout WM 3rd, Akama-Garren EH, Bhutkar A, Mercer KL, Schenkel JM, et al. CRISPR-mediated modeling and functional validation of candidate tumor suppressor genes in small cell lung cancer. Proc Natl Acad Sci U S A 2020;117(1):513–21 doi 10.1073/pnas.1821893117. [PubMed: 31871154]

32. Nicorici D, Satalan M, Edgren H, Kangaspeska S, Murumagi A, Kallioiemi O, et al. FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. bioRxiv 2014:https://doi.org/10/1101/011650.

33. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science 2013;339(6121):819–23 doi 10.1126/science.1231143. [PubMed: 23287718]

34. Vidigal JA, Ventura A. Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. Nat Commun 2015;6:8083 doi 10.1038/ncomms9083. [PubMed: 26278926]

35. Lignitto L, LeBoeuf SE, Homer H, Jiang S, Askenazi M, Karakousi TR, et al. Nrf2 Activation Promotes Lung Cancer Metastasis by Inhibiting the Degradation of Bach1. Cell 2019;178(2):316–29 e18 doi 10.1016/j.cell.2019.06.003. [PubMed: 31257023]

36. Zhang W, Girard L, Zhang YA, Haruki T, Papari-Zareei M, Stastny V, et al. Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. Transl Lung Cancer Res 2018;7(1):32–49 doi 10.21037/tlcr.2018.02.02. [PubMed: 29535911]

37. Anderson AM, Ragan MA. Palmitoylation: a protein S-acylation with implications for breast cancer. NPJ Breast Cancer 2016;2:16028 doi 10.1038/npjbcancer.2016.28. [PubMed: 28721385]

38. Rebecca VW, Nicastri MC, Fennelly C, Chude CI, Barber-Rotenberg JS, Ronghe A, et al. PPT1 Promotes Tumor Growth and Is the Molecular Target of Chloroquine Derivatives in Cancer. Cancer Discov 2019;9(2):220–9 doi 10.1158/2159-8290.CD-18-0706. [PubMed: 30442709]

39. Chen J, Santillan DA, Koonce M, Wei W, Luo R, Thirman MJ, et al. Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. Cancer Res 2008;68(15):6199–207 doi 10.1158/0008-5472.CAN-07-6514. [PubMed: 18676843]

40. Faria-Ramos I, Pocas J, Marques C, Santos-Antunes J, Macedo G, Reis CA, et al. Heparan Sulfate Glycosaminoglycans: (Un)Expected Allies in Cancer Clinical Management. Biomolecules 2021;11(2) doi 10.3390/biom11020136.

41. Ferrara MG, Di Noia V, D'Argento E, Vita E, Damiano P, Cannella A, et al. Oncogene-Addicted Non-Small-Cell Lung Cancer: Treatment Opportunities and Future Perspectives. Cancers (Basel) 2020;12(5) doi 10.3390/cancers12051196.

42. Duffy MJ, Crown J. Drugging "undruggable" genes for cancer treatment: Are we making progress? Int J Cancer 2021;148(1):8–17 doi 10.1002/ijc.33197. [PubMed: 32638380]

43. Nalawansha DA, Crews CM. PROTACs: An Emerging Therapeutic Modality in Precision Medicine. Cell Chem Biol 2020;27(8):998–1014 doi 10.1016/j.chembiol.2020.07.020. [PubMed: 32795419]

44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15–21 doi 10.1093/bioinformatics/bts635. [PubMed: 23104886]

45. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 2016;34(5):525–7 doi 10.1038/nbt.3519. [PubMed: 27043002]

46. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods 2017;14(7):687–90 doi 10.1038/nmeth.4324. [PubMed: 28581496]

47. Gardner EE, Lok BH, Schneeberger VE, Desmeules P, Miles LA, Arnold PK, et al. Chemosensitive Relapse in Small Cell Lung Cancer Proceeds through an EZH2-SLFN11 Axis. Cancer Cell 2017;31(2):286–99 doi 10.1016/j.ccell.2017.01.006. [PubMed: 28196596]

48. Sanchez-Rivera FJ, Papagiannakopoulos T, Romero R, Tammela T, Bauer MR, Bhutkar A, et al. Rapid modelling of cooperating genetic events in cancer through somatic genome editing. Nature 2014;516(7531):428–31 doi 10.1038/nature13906. [PubMed: 25337879]

49. DuPage M, Dooley AL, Jacks T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat Protoc 2009;4(7):1064–72 doi 10.1038/nprot.2009.95. [PubMed: 19561589]

50. Flowers JL, Burton GV, Cox EB, McCarty KS Sr., Dent GA, Geisinger KR, et al. Use of monoclonal antiestrogen receptor antibody to evaluate estrogen receptor content in fine needle aspiration breast biopsies. Ann Surg 1986;203(3):250–4 doi 10.1097/00000658-198603000-00005. [PubMed: 3954477]

51. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27(12):1739–40 doi 10.1093/bioinformatics/btr260. [PubMed: 21546393]

52. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284–7 doi 10.1089/omi.2011.0118. [PubMed: 22455463]

SATEMENT OF SIGNIFICANCE

The biological and therapeutic implications of gene fusions in small cell lung cancer, an aggressive metastatic lung cancer, are unknown. Our study investigates the functional significance of the in frame *RLF-MYCL1* gene fusion by developing a *Rlf-Mycl1*-driven genetically engineered mouse model and defining the impact on tumor growth and metastasis.
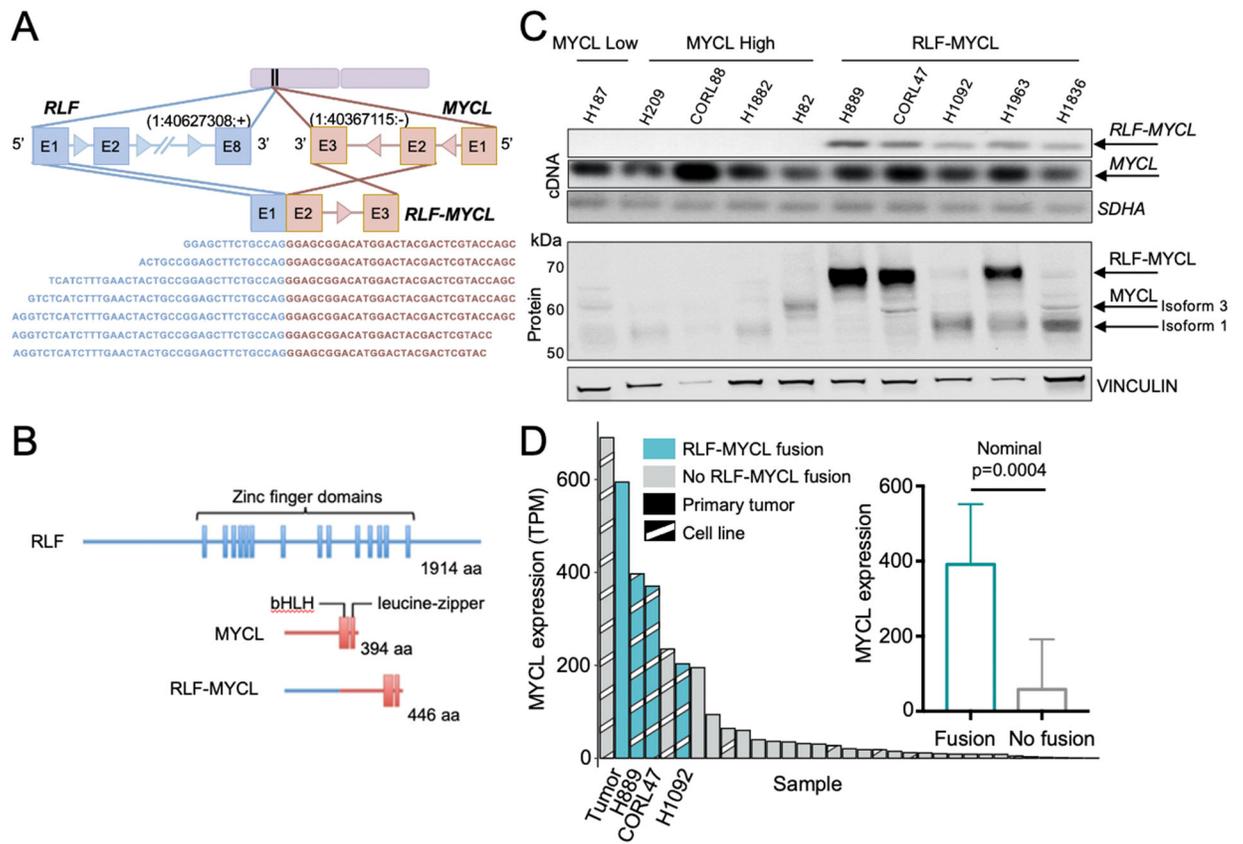
Figure 1. *RLF-MYCL* fusion gene samples in human SCLC are defined by high MYCL expression.

(A) Schematic of the in frame *RLF-MYCL* fusion identified using RNA-Seq. Multiple confirmatory junction reads from the cell line NCI-H889 are displayed.

(B) Schematic representation of the human RLF-MYCL fusion protein.

(C) *RLF-MYCL* gene and corresponding protein expression was assessed by cDNA PCR and Western blot of SCLC cell lines.

(D) *MYCL* expression by transcripts per million (TPM) in *RLF-MYCL* fusion-positive primary tumors and cell lines (N=4) relative to all other SCLC-A subtype (N=29) SCLC samples from (8) (Wald test, nominal p=0.0004).
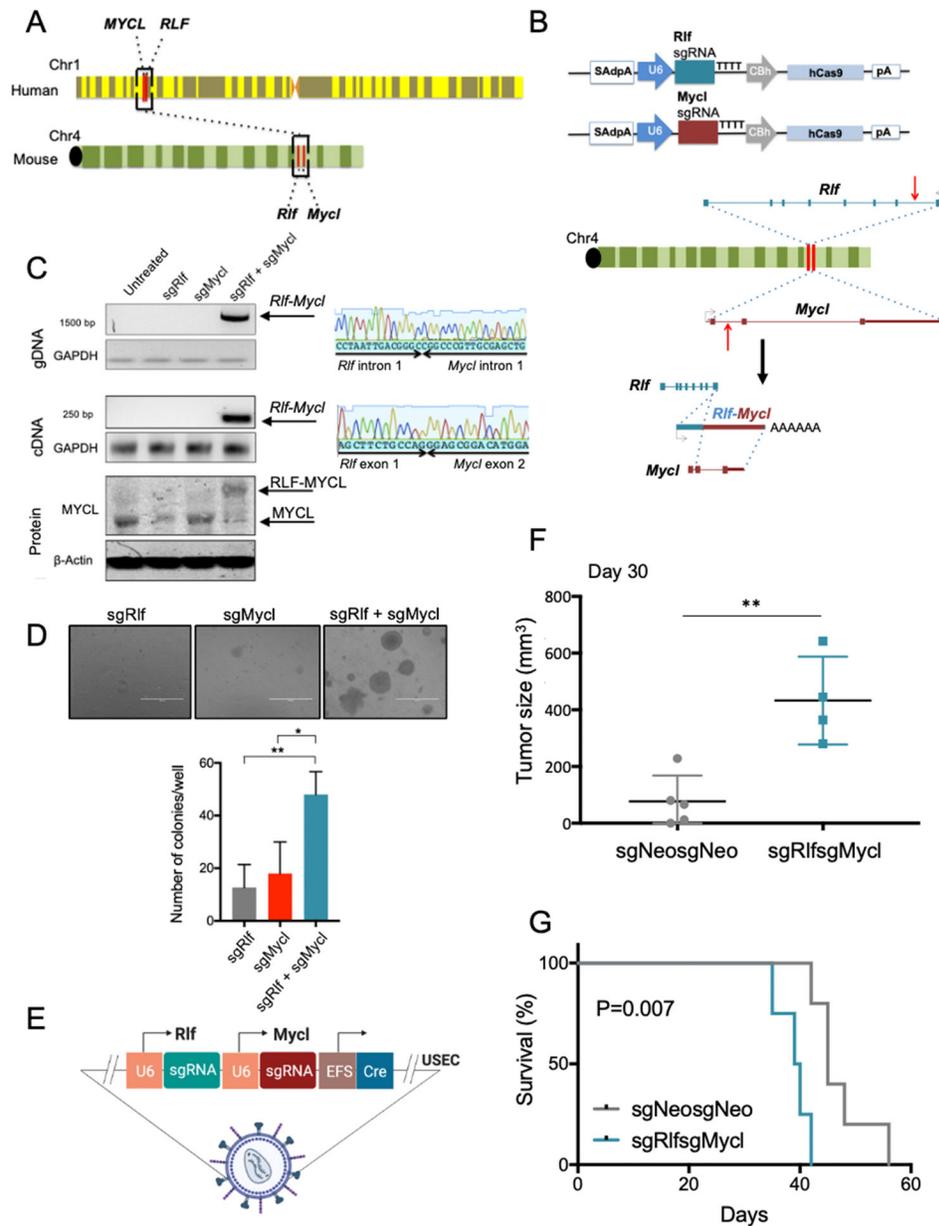
Figure 2. Induction of *Rlf-Mycl* fusion using CRISPR-Cas9 accelerates SCLC transformation.
(A) Schematic of the gene location of the *MYCL* and *RLF* loci on human chromosome 1 and of *Mycl* and *Rlf* on mouse chromosome 4.

(B) Diagrams of pX330 expression vectors used and schematic for generation of the *Rlf-Mycl* fusion gene and transcript. Red arrows indicate the sites recognized by the sgRNAs.

(C) PCR on genomic DNA and RNA, and Western blot on protein from NIH/3T3 cells transfected with the indicated pX330 constructs. Sanger sequence of the gDNA and cDNA PCR products confirming the expected *Rlf-Mycl* junction (right panels).

(D) Representative images of targeted preSC cells with the indicated PX330 constructs in soft agar 5 weeks after seeding of 1×10^5 cells/well (n=3). Scale bar = 2000 μm. Bottom:

quantification of colonies > 0.1 mm in diameter. Unpaired Student's t-test; *p=0.025, **p=0.008.

(E) Schematic of the USEC lentivirus sgRlfsgMycl containing the sgRNAs to induce the *Rlf-Mycl* rearrangement.

(F) Tumor sizes on day 30 in nude mice subcutaneous injected with LentiCas9-Blast + sgNeosgNeo-preSC cells (n=5) and with LentiCas9-Blast + sgRlfsgMycl-preSC cells (n=4) on day 30; 5×10^5 cells per mouse. Unpaired Student's t-test; **p= 0.004.

(G) Kaplan-Meier survival curves of nude mice subcutaneously injected with LentiCas9-Blast + sgNeosgNeo-preSC cells (n=5) and LentiCas9-Blast + sgRlfsgMycl-preSC cells (n=4); 5×10^5 cells per mouse. Mantel-Cox log-rank test; p=0.007.
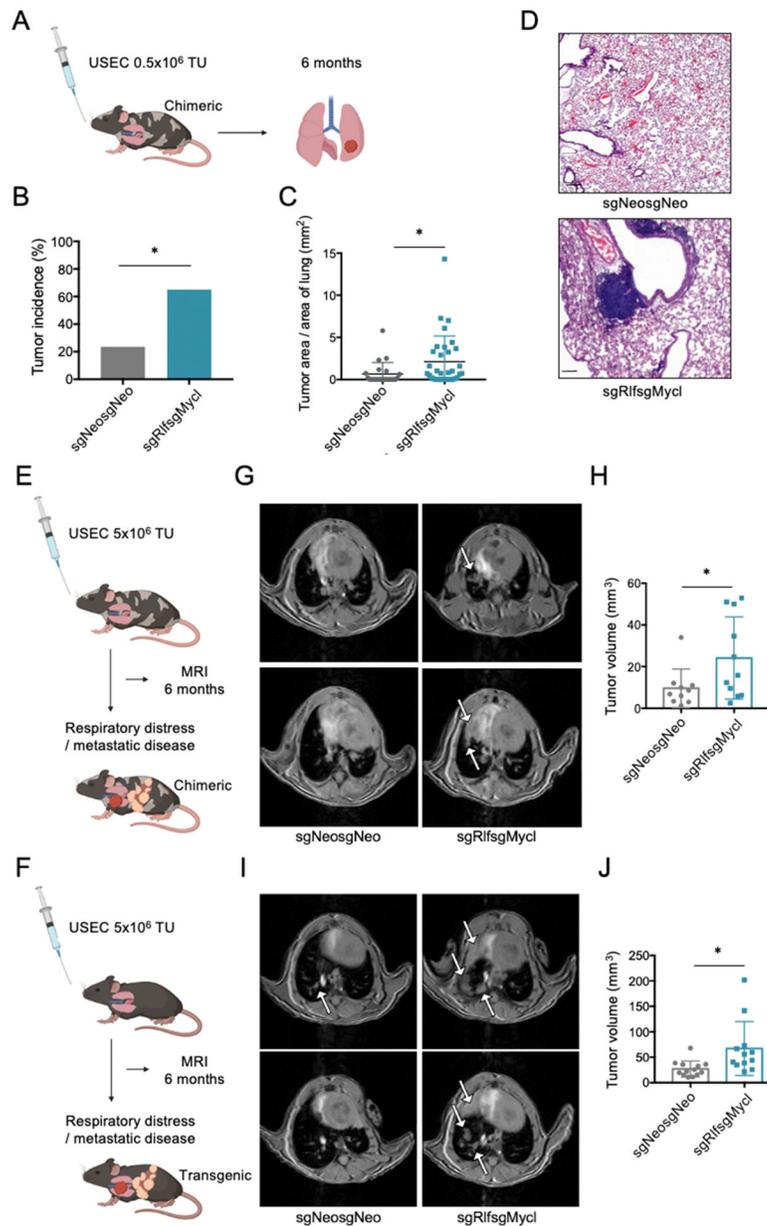
Figure 3. *Rlf-Mycl* induction in *Rb1/Trp53/Rbl2/Cas9* (RPR2C) mice accelerates primary SCLC tumor formation.

(A) Schematic of the experiment in chimeric RPR2C mice. Mice were euthanized 6 months after intratracheal administration of 0.5×10^6 transduction units (TU) per mouse USEC lentivirus and analyzed by H&E staining microscopy; sgNeosgNeo (n=17), sgRlfsgMycl (n=20).

(B) Tumor burden. Chi-square; *p=0.012.

(C) Quantification of tumor area (mm^2). Unpaired Student's t-test; *p=0.041.

(D) Representative H&E-stained lung section from a mouse of each group, scale bar = 100μm.

(E) Schematic of the experiment in chimeric RPR2C mice; sgNeosgNeo (n=13) and sgRlfsgMycl (n=14) mice transduced with USEC lentiviruses (5×10^6 TU per mouse).

(F) Schematic of the experiment in fully transgenic RPR2C mice; sgNeosgNeo (n=19) and sgRlfsgMycl (n=13).

(G) Representative MRI of lungs of chimeric mice at 6 months after intratracheal administration of USEC lentiviruses in each cohort. Lung tumors are indicated by arrows.

(H) Quantification of MRI tumor volume (mm^3) of chimeric RPR2C injected sgNeosgNeo (n=10) and sgRlfsgMycl (n=11) mice. Unpaired Student's t-test; *p=0.047.

(I) Representative MRI of GEMM mice at 6 months after intratracheal administration of USEC lentiviruses in each cohort. Lung tumors are indicated by arrows.

(J) Quantitative tumor volume (mm^3) determined by MRI of GEMM sgNeosgNeo (n=13) and sgRlfsgMycl (n=12) transduced mice. Unpaired Student's t-test; *p=0.016.
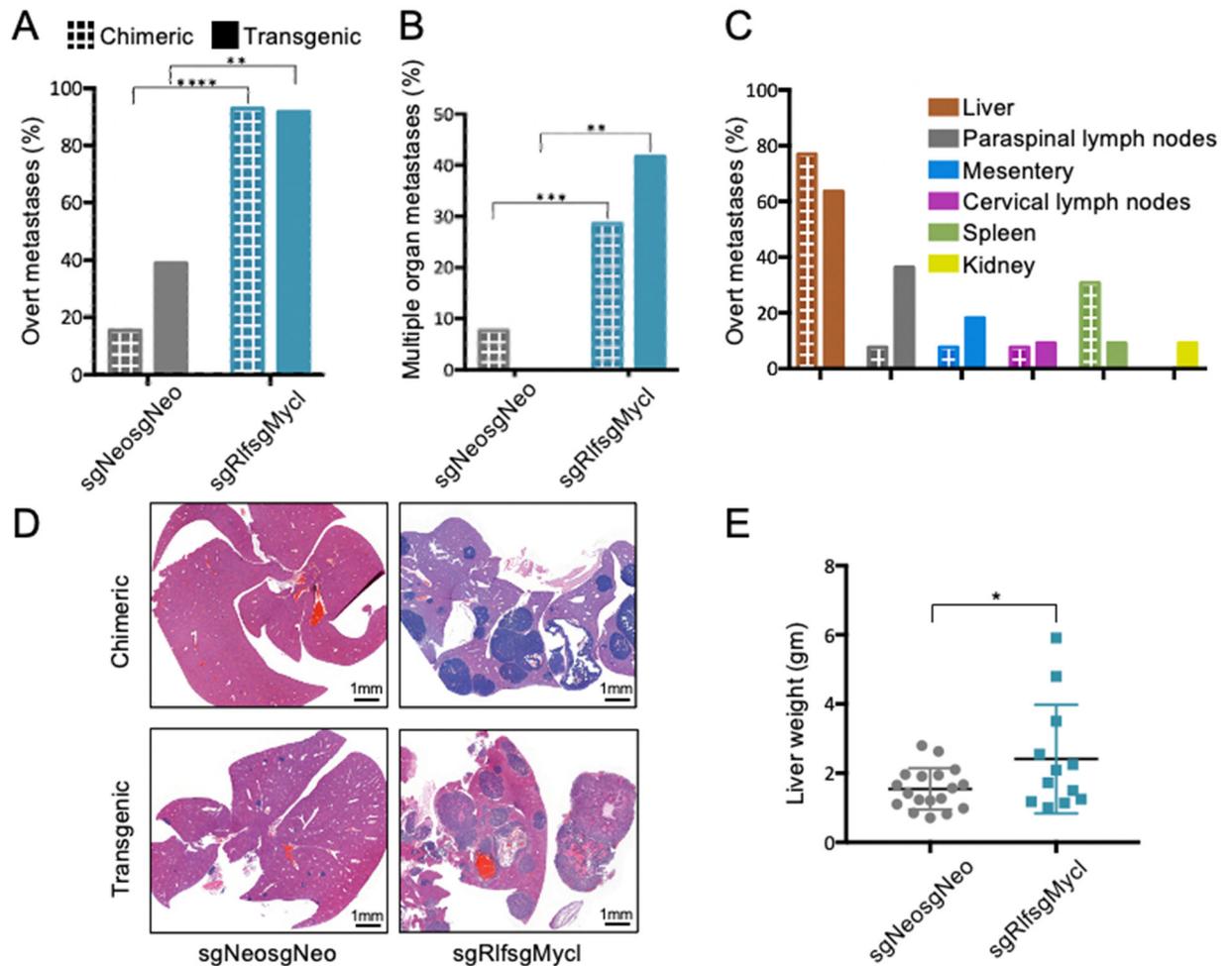
Figure 4. Induction of *Rlf-Mycl* drives metastasis formation in chimeric and fully transgenic RPR2C mice.

(A) Chimeric and fully transgenic mice with overt metastasis at the survival endpoint. Chimeric sgNeosgNeo (n=13) and sgRlfsgMycl (n=14), chi-square; ****p<0.0001; transgenic sgNeosgNeo (n=18) and sgRlfsgMycl (n=12); chi-square; **p=0.004.

(B) Mice with metastases in multiple organs at the survival end point. Chimeric chi-square; ***p=0.001; fully transgenic chi-square; **p=0.004.

(C) Distribution of overt metastasis in sgRlfsgMycl mice in chimeric and fully transgenic cohorts.

(D) Representative H&E-stained sections of end stage livers in each cohort of chimeric and fully transgenic mice. Scale bar = 1000μm.

(E) Weight in grams (gm) of endpoint livers of fully transgenic sgNeosgNeo (n=18) and sgRlfsgMycl (n=12) mice. Unpaired Student's t-test; *p=0.043.
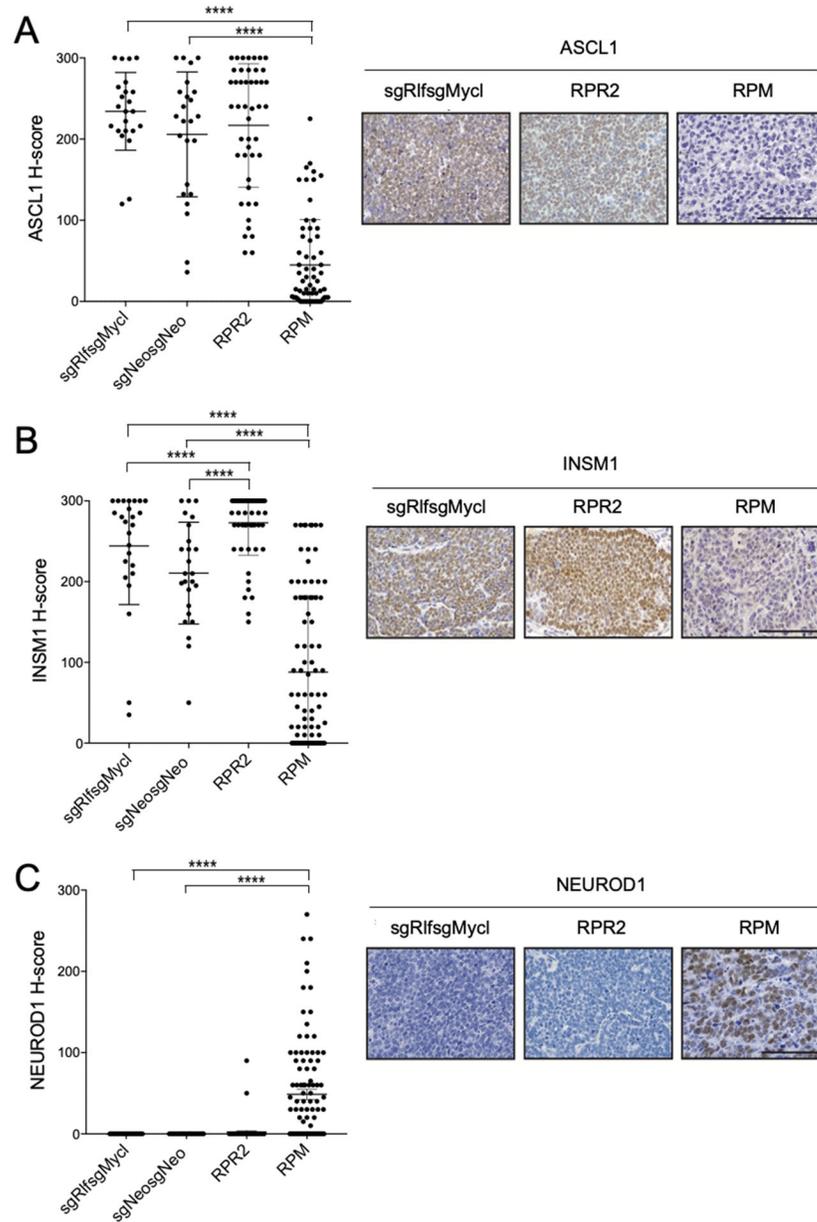
Figure 5. Transcription factor expression in *Rlf-Mycl* primary tumors resembles SCLC-A.
Quantitative measurement of neuroendocrine transcription factors (A) ASCL1, (B) INSM1 and (C) NEUROD1 in RPR2C transgenic sgRlfsgMycl, sgNeosgNeo, RPR2 and RPM mice (5 mice/group) using H-Score method. Unpaired Student's t-test; ****p<0.0001. Representative stained sections of end stage lung tumors. Scale bar = 100 μm.
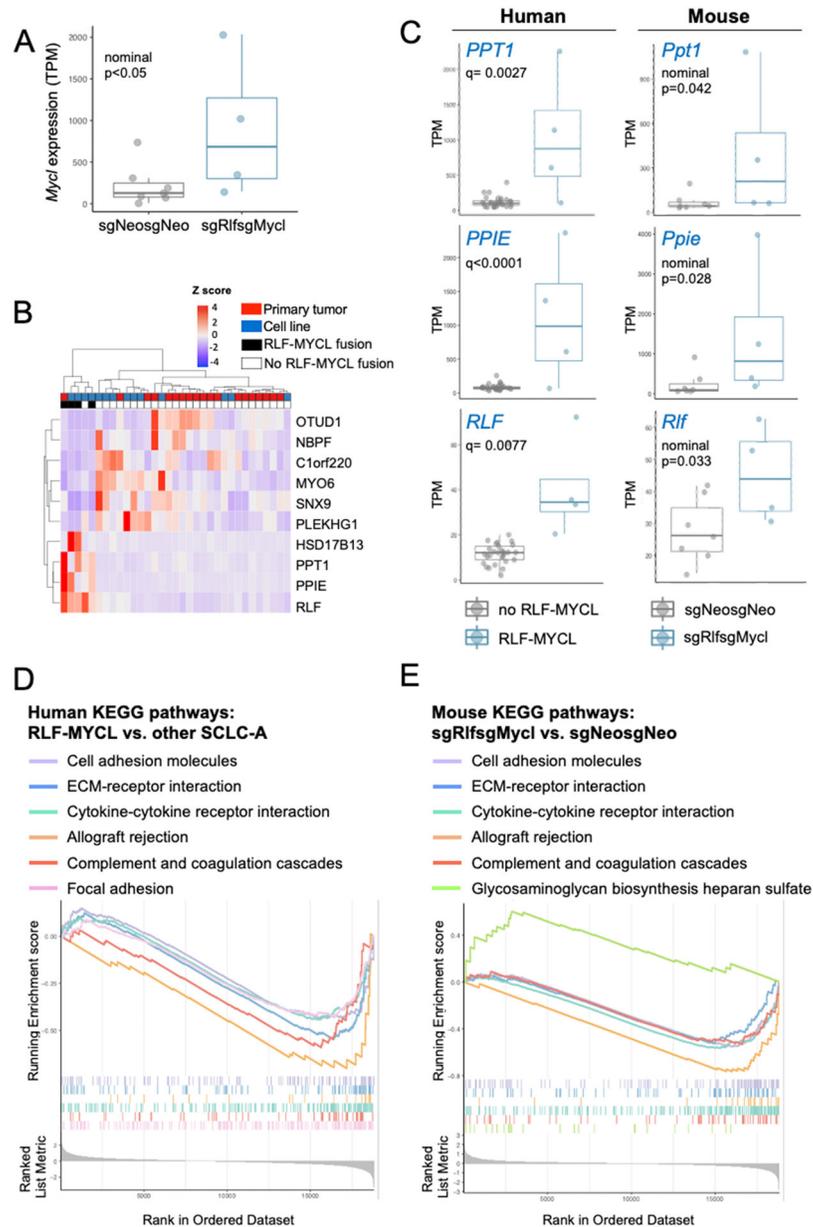
Figure 6. *Rlf-Mycl* tumor gene expression pattern resembles that of human *RLF-MYCL* SCLC.
(A) *Mycl* gene expression in primary tumors of sgNeosgNeo mice (n=7) and of fusion-detected sgRlfsgMycl mice (n=4). (Wald test, nominal p<0.05).

(B) Heat map of top differentially expressed genes between *RLF-MYCL* fusion-positive and fusion negative SCLC-A samples from (8). Z-score for expression of each gene was calculated and plotted.

(C) Expression by transcripts per million (TPM) corrected for sample type of *PPT1*, *PPIE* and *RLF* genes in human *RLF-MYCL* fusion samples (N=4) vs. non-fusion SCLC-A samples (N=29) from (8)(**q<0.05). Expression by TPM of *Ppt1*, *Ppie* and *Rlf* genes in mouse sgRlfsgMycl (n=4) vs. sgNeosgNeo (n=7) primary lung samples. (Wald test, nominal p<0.05).

(D) GSEA enrichment plot of KEGG pathways differentially enriched in RLF-MYCL fusion-positive (N=4) versus fusion-negative (N=29) human SCLC-A samples.

(E) GSEA enrichment plot of KEGG pathways differentially enriched between sgRlfsgMycl (n=4) versus sgNeosgNeo (n=7) primary tumor samples in mouse. Shared KEGG pathways identified in both sgRlfsgMycl mice and human RLF-MYCL samples include negative enrichment for cell adhesion molecules, ECM-receptor interaction, cytokine-cytokine receptor interaction, allograft rejection, and complement and coagulation cascades. In both (D) and (E) the top portion plots the running enrichment scores for each pathway and bottom portion shows value of the ranking metric in the ordered dataset.