

Chromothripsis-mediated small cell lung carcinoma

Natasha Rekhtman,*¹ Sam E Tischfield,*² Christopher A Febres-Aldana,*^{1,3} Jake June-Koo Lee,⁴ Jason C Chang,¹ Benjamin O Herzberg,^{4,5} Pier Selenica,¹ Hyung Jun Woo,² Chad M Vanderbilt,¹ Soo-Ryum Yang,¹ Fei Xu,¹ Anita S Bowman,¹ Edaise M da Silva,¹ Anne Marie Noronha,² Diana L Mandelker,¹ Miika Mehine,^{1,2} Semanti Mukherjee,⁴ Juan Blanco-Heredia,² John J Orgera,² Gouri J Nanjangud,⁶ Marina K Baine,¹ Rania G Aly,¹ Jennifer L Sauter,¹ William D Travis,¹ Omid Savari,^{1,7} Andre L Moreira,^{1,8} Christina J Falcon,^{4,9} Francis M Bodd,^{1,9} Christina E Wilson,^{1,9} Jacklynn V Sienty,^{4,10} Parvathy Manoj,⁴ Harsha Sridhar,⁴ Lu Wang,^{1,11} Noura J Choudhury,^{4,12} Michael Offin,⁴ Helena A Yu,^{4,12} Alvaro Quintanal-Villalonga,⁴ Michael F Berger,^{1,2,3} Marc Ladanyi,¹ Mark T A Donoghue,² Jorge S Reis-Filho,^{1,13} Charles M Rudin⁴

1. Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA
2. Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA
3. Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA
4. Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA
5. Current Affiliation: Department of Medicine, Division of Hematology/Oncology, Columbia University Irving Medical Center and the Herbert Irving Comprehensive Cancer Center, New York, NY, USA
6. Department of Molecular Cytogenetics Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA
7. Current Affiliation: Department of Pathology, University Hospitals Cleveland Medical Center- Case Western Reserve University, Cleveland, OH, USA
8. Current Affiliation: New York University Grossman School of Medicine, Department of Pathology, New York, NY, USA
9. Druckenmiller Center for Lung Cancer Research, Memorial Sloan Kettering Cancer Center, New York, NY, USA
10. Current Affiliation: Division of Biostatistics Research Scientists, New York University, New York, NY, USA
11. Current Affiliation: Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA
12. Department of Medicine, Weill Cornell Medical College, New York, NY, USA
13. Current Affiliation: AstraZeneca, Cancer Biomarker Development, Gaithersburg, MD, USA

*co-first authors

Corresponding authors:

Charles M Rudin, MD, PhD rudinc@mskcc.org, 1275 York Avenue New York, NY 10065, 646-888-4527

Natasha Rekhtman, MD, PhD rekhtman@mskcc.org, 1275 York Avenue New York, NY 10065, 212-639-6235

Key words: SCLC, carcinoid, chromothripsis, ecDNA, RB1, TP53

Running title: Chromothripsis in *RB1*⁺/*TP53*⁺ SCLC

Acknowledgments and Funding: This work was supported by NCI grants R35 CA263816, U24 CA213274, and P30 CA0087448, the Druckenmiller Center for Lung Cancer Research, and the Sharon and Jon Corzine Foundation.

Word count: 6,355

Total number of main figures and tables: 7

Total number of supplementary figures: 13

Total number of supplementary tables: 13

Authors' Disclosures: N. Rekhtman reports serving on the scientific advisory board of Merck. B.O. Herezberg reports consulting and receiving personal fees from Amgen, Astellas, Eli Lilly, Boxer Capital, MJH Life Sciences, Guardant Health, and IDEology health outside the current work; received grants from AZ. C.M. Vanderbilt reports Equity; Intellectual Property Rights; Professional Services and Activities (Uncompensated) from Paige.AI, Inc. SR. Yang reports uncompensated professional services and activities from Genentech. D. Mandelker reports consulting fees from AstraZeneca. W.D. Travis reports uncompensated professional services and activities from Genentech. A.L. Moreira reports a grant from BMS outside the scope of this work. N.J. Choudhury reports research funding from the following: AbbVie, Amgen, Harpoon Therapeutics, Merck, Monte Rosa Therapeutics, Roche/Genentech; serving on the scientific advisory board of Harpoon Therapeutics, AbbVie; Honoraria: G1 Therapeutics, Sanofi; receiving royalties from Wolter Kluwer, all of which are outside the scope of the submitted work. N.J. Choudhury reports research funding from the following: AbbVie, Amgen, Harpoon Therapeutics, Merck, Monte Rosa Therapeutics, Roche/Genentech; serving on the scientific advisory board of Harpoon Therapeutics, AbbVie; honoraria from G1 Therapeutics, Sanofi; royalties from Wolter Kluwer. M.D. Offin reports consulting roles/honorarium with Novartis, Jazz, Pfizer, Targeted Oncology, OncLive, American Society for Radiation Oncology; grant support from the Druckenmiller Foundation and LUNgevity Foundation; uncompensated scientific advisory board member for the Mesothelioma Applied Research Foundation. H.A. Yu reports grants from the NIH/NCI during the conduct of this study; personal fees from AstraZeneca, Daiichi, Janssen, Abvie, Amgen, Novocure, Black Diamond, Taiho, Takeda, Cullinan outside of the scope of this project. A. Quintanal-Villalonga reports receiving honoraria from Astra Zeneca. M.F. Berger reports personal fees from Eli Lilly, AstraZeneca, Paige.AI; research support from Boundless Bio; intellectual property rights from SOPHiA Genetics. M. Ladanyi reports receiving honoraria for advisory board participation from Merck, AstraZeneca, Bristol Myers Squibb, Blueprint Medicines, Janssen Pharmaceuticals, Takeda Pharmaceuticals, Lilly Oncology, LOXO Oncology, Bayer, ADC Therapeutics, Riken Genesis/Sysmex, and PaigeAI; research support from Rain Oncology, LOXO Oncology, Merus, Elevation Oncology, Helsinn Therapeutics, and ADC Therapeutics. J.S Reis-Filho reports employment at AstraZeneca. C.M. Rudin reports serving as consultant for Amgen, Astra Zeneca, Chugai, D2G, Daiichi Sankyo, Hoffman-La Roche, Jazz, and Legend; serves on the scientific advisory boards of Auron, Bridge Medicines, DISCO, Earli, and Harpoon Therapeutics. No disclosures were reported by the other authors.

ABSTRACT

Small cell lung carcinoma (SCLC) is a highly aggressive malignancy that is typically associated with tobacco exposure and inactivation of *RB1* and *TP53* genes. Here we performed detailed clinicopathologic, genomic and transcriptomic profiling of an atypical subset of SCLC that lacked *RB1* and *TP53* co-inactivation and arose in never/light smokers. We found that most cases were associated with chromothripsis – massive, localized chromosome shattering – recurrently involving chromosomes 11 or 12, and resulting in extrachromosomal (ecDNA) amplification of *CCND1* or co-amplification of *CCND2/CDK4/MDM2*, respectively. Uniquely, these clinically aggressive tumors exhibited genomic and pathologic links to pulmonary carcinoids, suggesting a previously uncharacterized mode of SCLC pathogenesis via transformation from lower-grade neuroendocrine tumors or their progenitors. Conversely, SCLC in never-smokers harboring inactivated *RB1* and *TP53* exhibited hallmarks of adenocarcinoma-to-SCLC derivation, supporting two distinct pathways of plasticity-mediated pathogenesis of SCLC in never-smokers.

STATEMENT OF SIGNIFICANCE

Here, we provide the first detailed description of a unique SCLC subset lacking *RB1/TP53* alterations, and identify extensive chromothripsis and pathogenetic link with pulmonary carcinoids as its hallmark features. This work defines atypical SCLC as a novel entity among lung cancers, highlighting its exceptional histogenesis, clinicopathologic characteristics and therapeutic vulnerabilities.

INTRODUCTION

Small cell lung carcinoma (SCLC) is an extremely aggressive malignancy, characterized by a nearly universal genomic inactivation of *RB1* and *TP53*, a high tumor mutation burden (TMB), and nearly invariable association with cigarette smoking (1). We recently described an uncommon subset of SCLC with *RB1*-proficiency, characterized by expression of wild-type *RB1* in tumors that otherwise exhibited classic properties of SCLC, including *TP53* mutations, high TMB, and history of smoking. Conversely, the characteristics of the exceptional cases of SCLC lacking both of its hallmark genomic alterations - *RB1* and *TP53* - remain almost entirely undefined.

In a prior whole-genome sequencing (WGS) study of 110 SCLC, only two cases were identified that harbored wild-type *RB1* and *TP53* (2). Although not the focus of that prior report, it is notable that both cases demonstrated evidence of chromothripsis – a process of localized massive chromosome shattering, involving chromosomes 11 and 3, and associated with overexpression of *CCND1* on chromosome 11. This suggested that the pathogenesis of SCLC with intact *RB1* and *TP53* may be mediated by an entirely distinct mechanism; however, further details on clinicopathologic characteristics of such tumors and analysis of this phenomenon at scale has been lacking. Furthermore, in a cohort of 3,600 SCLC samples submitted for broad targeted NGS (tNGS) at Foundation Medicine, 5.5% were reported to lack *RB1* and *TP53* alterations; this subset harbored instances of *CCND1* and *MDM2* amplification, but no WGS or detailed clinicopathologic characterization of these samples was available (3). Furthermore, as recently demonstrated, tNGS may miss deleterious *RB1* alterations in a substantial proportion of cases; therefore, combining tNGS with expression data is essential for establishing *RB1* proficiency (4).

Through enterprise-wide clinical application of a broad tNGS using MSK-IMPACT (5,6), integrated with expression-based analysis, we identified 20 SCLC patients whose tumors lacked *RB1* and *TP53* co-inactivation. Remarkably, all identified patients were never or light smokers, defined as the life-time smoking history of ≤ 10 pack-years, further highlighting the unique pathogenesis underlying these tumors. Based on prior studies, it is known that SCLC can rarely arise in never or light smokers via histologic transformation of lung adenocarcinoma (LUAD) either after treatment with targeted therapies or *de novo* (7,8), and that *RB1* and *TP53* co-deficiency is pre-requisite for this conversion (9,10). Conversely, SCLC in never/light smokers lacking *RB1* and *TP53* co-deficiency is a highly unusual and previously uncharacterized subset.

To better understand the pathogenetic mechanisms and clinicopathologic characteristics of this cohort, we performed a multifaceted analysis of clinical samples, which included WGS and RNAseq when sufficient tissue was available after tNGS. Here, we demonstrate that nearly all *RB1/TP53*-proficient SCLC are characterized by extensive chromothripsis, associated with extra-chromosomal (ecDNA) amplification of *CCND1*, *CCND2*, *CDK4*, and *MDM2*. Furthermore, we present evidence for a histogenetic relationship of these tumors with a separate class of neuroendocrine cancers – pulmonary carcinoids.

Pulmonary carcinoids, alternatively termed neuroendocrine tumors (NETs), are regarded as an entirely separate class of tumors from SCLC. These are generally indolent neuroendocrine neoplasms, that lack an association with tobacco exposure, and which arise in younger patients than SCLC (11,12). They are characterized genomically by a low TMB, recurrent alterations in *MEN1*, *EIF1AX*, and *ARID1A*, and notably, the lack of *RB1* and *TP53* alterations (12,13). Pathologically, pulmonary carcinoids exhibit minimal to intermediate proliferation rate, featuring Ki67 index of 1-30% (12). Although proliferative escalation has been documented metastasizing carcinoids (14), whether carcinoids can exhibit a full

phenotypic conversion to SCLC has remained a controversial concept, lacking robust clinicopathologic and molecular documentation. The data presented here suggest a new mode of SCLC pathogenesis through a histogenetic link to carcinoids or their progenitors through chromothripsis-mediated cancer gene deregulation in the context of *RB1/TP53* proficiency.

RESULTS

Analysis of *RB1*, *TP53* and smoking status to identify “atypical SCLC”

As depicted in **Figure 1A**, among the first 600 patients diagnosed with the *de novo* SCLC who underwent sequencing by MSK-IMPACT, 20 (3%) exhibited *RB1* and *TP53* dual proficiency, defined by an integrated genomic and immunohistochemical (IHC) approach (**Supplementary Table S1**). *RB1* proficient (*RB1*⁺) tumors lacked *RB1* genomic alterations and retained pRb protein expression by IHC, although, as discussed later, few cases exhibited subclonal or acquired *RB1* mutations in subsequent/metastatic samples. As additional support for *RB1* proficiency in initial samples, all cases had D-type cyclin^{high} and/or p16^{low} profile, which in prior studies was exclusive to *RB1*⁺ SCLC (4). *TP53* proficiency (*TP53*⁺) was defined by the lack of *TP53* genomic alterations and wild-type p53 pattern by IHC (see Methods). Remarkably, all patients with *RB1*⁺/*TP53*⁺ SCLC were never or light (<10 pack-year) smokers. The rest of the SCLC cohort was composed predominantly of smokers with *RB1*⁻/*TP53*⁻ SCLC (88%), smokers with the recently-described *RB1*-proficient SCLC (*RB1*⁺/*TP53*⁻; 6%) (4), and never-smokers with *RB1*⁻/*TP53*⁻ SCLC (3%; described later). Given the lack of *RB1* and *TP53* alterations, and the absent or low tobacco exposure, we designated this unique subset as atypical SCLC (aSCLC).

Clinicopathologic characteristics of atypical SCLC

The demographic and pathologic sample characteristics of patients with aSCLC are summarized in **Figures 1B, 1C** and **Supplementary Tables S2, S3**. Patients had 1-7 pathologic specimens obtained during the course of disease (total 49), all of which underwent detailed pathologic review, 31 (at least 1 per patient) were profiled by tNGS, 12 (from 11 patients) by WGS and 7 by RNAseq.

Patients presented with lung tumors measuring up to 7.2 cm (mean 3.3 cm). All patients had pathologically confirmed metastatic disease at presentation, including distant metastases in 16 patients and metastasis limited to regional lymph nodes in 4 patients. The most prevalent sites of distant metastases included brain (55%), bone (55%), liver (45%) and adrenal (40%) – the sites commonly involved by neuroendocrine lung cancers (**Figure 1D; Supplementary Table S2**).

Pathologically, aSCLC samples exhibited classic SCLC morphology, manifesting as primitive, crowded cells with a high nuclear-to-cytoplasmic ratio, extensive necrosis, and a high Ki67 proliferation index (in most cases 70% to >90%) (**Figures 1E-G**). All aSCLC expressed multiple neuroendocrine markers by IHC, including synaptophysin, chromogranin A, CD56/NCAM and INSM1 (full IHC results are summarized in **Supplementary Table S2**, and illustrated in **Supplementary Figures S1, S2, S3** and **S4A-4B**). Uniquely, in five patients (cases A02, A05, A16, A17 and A20), the SCLC histotype was present in metastatic sample(s), whereas primary/intrathoracic sample(s) exhibited a carcinoid histotype, characterized by well-differentiated bland, uniform cells with a lower nuclear-to-cytoplasmic ratio, and Ki67 index of mostly ≤20% (**Figure 1E, Supplementary Figures S3, S4A**).

Compared to other SCLC, patients with aSCLC exhibited several distinct clinicopathologic characteristics (**Figure 1H**). First, patients with aSCLC were significantly younger, as young as 19 years at presentation

(mean age 53 vs. 67 years, respectively; $P<0.0001$). Second, while Ki67 proliferation index in all aSCLC samples was in the range of conventional SCLC (50-100%), these tumors were enriched in Ki67 rates at the lower end of this spectrum compared to other SCLC (mean Ki67 rate 69% vs. 87%, respectively; $P<0.0001$). Lastly, the absent/low pack-year smoking history in aSCLC contrasted sharply with pack-year smoking in other SCLC ($P<0.0001$). Full comparison of the demographic and clinicopathologic characteristics of aSCLC versus other SCLC is shown in **Supplementary Table S4**.

Genomic profiling of atypical SCLC reveals low tumor mutation burden and recurrent oncogene amplifications

The overall mutational landscape in aSCLC was initially assessed using the MSK-IMPACT assay for all 20 patients, which was performed with a mean coverage depth of 613x (range 255-959x) (**Supplementary Table S5**). The results were compared with those in smoking-associated SCLC (sSCLC; $n=206$) and pulmonary carcinoids ($n=157$) analyzed on the same platform (**Figure 2A**).

Tumor mutation burden (TMB) in aSCLC was remarkably low, with most cases exhibiting TMB of <2 mutations per Mb; two outliers featuring high TMB were from a sample obtained after temozolomide (TMZ) treatment (A07) and a unique case in this study exhibiting microsatellite instability (MSI; case A20). Tumors with sufficient number of single nucleotide variants (SNV) to assess mutational signature consistently lacked a smoking signature, and primarily exhibited a dominant signature reflecting systemic therapy received (**Supplementary Figure S5A-C**). The low TMB and absence of a smoking signature in aSCLC contrasted sharply with the findings in sSCLC (**Figure 2A**).

Recurrent genomic alterations in aSCLC were dominated by mutations characteristic of lung carcinoids (*MEN1*, *EIF1AX*, *ARID1A*, *ATM*), with an enrichment in the rate of *ATM* mutations (30%) compared to both carcinoids (8%; $p=0.008$) and SCLC (3%; $p=0.0003$) (**Figure 2A**; see **Supplementary Tables S5, S6** for full SNV data).

A striking feature of aSCLC was the presence of recurrent and mutually exclusive amplifications of several key oncogenes controlling cell cycle and survival: *CCND1* (30%), *CCND2/CDK4/±MDM2* (15%) and *MYCL1* (10%) (**Figure 2A**; full CNA data is provided in **Supplementary Tables S5** and **S7**). These amplifications were mostly present at a high copy number (>10), and some exceeded 100 copies (such as case A02, harboring 125 copies of *CCND1*). Conversely, amplifications of these genes were seen only rarely in conventional SCLC and carcinoids; for example, *CDK4* amplifications were found in 0% of SCLC ($p=0.0006$) and 1% of carcinoids ($p=0.01$). Amplifications were consistently accompanied by overexpression of the corresponding proteins by IHC (**Figure 2B**) and mRNA (**Figure 2C**). Of note, two cases (A11 and A17) showed *CCND1* overexpression by mRNA and IHC in the absence of *CCND1* amplification at the genomic level, indicating that *CCND1* overexpression may be mediated by other mechanisms in some aSCLC, as discussed later.

The sole MSI-high case (A20) was characterized by somatic *MLH1* homozygous deletion and dual loss of *MLH1/PMS2* by IHC. Using the same methods, the MSI rate in conventional lung carcinomas was restricted to rare cases in a recent series from our institution (1.9% for SCLC and 0.4% for non-small cell lung carcinoma [NSCLC])(15); in the control set of lung carcinoids ($n=157$), none were found to be MSI-high. Pathologically, A20 was one of the cases with SCLC histotype in metastatic samples (Ki67 50-80%) and carcinoid histotype in the primary tumor (Ki67 5-20%) (**Supplementary Figure S4**).

Chromothripsis as the underlying mechanism for amplifications

To elucidate the underlying genomic processes leading to aSCLC, samples from 11 patients with sufficient residual DNA after MSK-IMPACT were further analyzed by WGS. This confirmed the lack of pathogenic *RB1* and *TP53* alterations in all tested samples, except for unusual subclonal *RB1* mutation in A12, discussed later (**Supplementary Table S8**). Remarkably, all but one case (MSI-high, A20) revealed the hallmark features of chromothripsis, as evidenced by clustered massive structural variants (SV) (**Figures 3A and 3B; Supplementary Figure S6**) associated with co-localized CNAs (summarized in **Figures 3C and 3D**). Within the regions affected by chromothripsis, CNAs exhibited the characteristic oscillating pattern featuring alternating retained and lost genomic material, reflecting the loss of DNA segments during re-ligation of the shattered chromosomes (16). The average number of SVs per case was 565, extending to >1900 (**Figure 3E**; see **Supplementary Table S9** for the full list of SVs).

Next, using cases with matched WGS and tNGS, we developed an approach for the manual detection of chromothripsis in tNGS based on the distinctive oscillating CNA pattern, which showed 100% specificity and 77% sensitivity of tNGS for the detection of chromothripsis using WGS as a gold standard (**Supplementary Figure S7A-C**). Using this approach, we identified chromothripsis in six of nine additional cases analyzed by tNGS only. Overall, for the combined WGS and tNGS analyses, evidence of chromothripsis was identified in 16 of 19 (84%) non-MSI aSCLC.

Across cancer types, chromothripsis typically localizes to one or a few chromosomes (17,18). Indeed, in aSCLC, on average, two chromosomes were affected per case (range 1 – 4) (**Figure 3C**). Chromosome 11 was the most commonly affected (35% of cases), which was invariably accompanied by chromothripsis on chromosome 3. This was followed by recurrent chromothripsis on chromosome 12 (20%) and chromosome 1 (15%).

Notably, chromothripsis of specific chromosomes correlated with the amplifications of genes localized to these chromosomes. Namely, chromothripsis on chromosome 11 was observed in 5 of 6 cases with *CCND1* (11q13) amplification, chromothripsis on chromosome 12 was seen in 3 of 3 cases with *CCND2* (12q13)/*CDK4* (12q14)/±*MDM2* (12q15) co-amplification, and chromothripsis on chromosome 1 was seen in 2 of 2 cases with *MYCL1* (1p34) amplification. Of note, the sole case with *CCND1* amplification (23 copies) in the absence of detectable chromothripsis (A06) was analyzed by tNGS only; therefore, a false-negative result cannot be excluded given the incomplete sensitivity of this assay for detecting chromothripsis.

Interestingly, in two cases, chromothripsis on chromosome 11 (A12, A13) and chromosome 12 (A12) lacked the amplifications of *CCND1* or *CCND2/CDK4/MDM2*, respectively. Thus, chromothripsis on individual chromosomes is strongly but not invariably associated with the amplification of specific oncogenes.

Beyond the recurrent amplification of *CCND1*, *CCND2*, *CDK4*, *MDM2* and *MYCL1*, chromothriptic chromosomes were associated with the amplification or loss of other well-established cancer genes. This included *HRAS* (11p15) amplification on chromothriptic chromosome 11, *KRAS* (12q12) and *ERBB3* (12q13) amplification on chromosome 12, as well as deletions of key carcinoid-associated tumor suppressor genes *MEN1* (11q13) and *ARID1A* (1p36) on chromosomes 11 and 1, respectively. These CNAs were consistently associated with the overexpression or loss of expression of the corresponding mRNA by RNAseq (**Supplementary Figure S8**). On average, three (up to eight) genes regarded as oncogenic or likely oncogenic by the OncoKB annotation (19) exhibited amplifications or (in the case of tumor suppressors) losses on chromothriptic chromosomes (full list included in **Supplementary Table**

S7). Overall, of the 16 cases with chromothripsis, 12 harbored oncogenic CNAs in the regions of chromothripsis.

Although each instance of chromothripsis on chromosome 11 was accompanied by chromothripsis on chromosome 3 (7/7 cases), the potential target gene(s) on chromosome 3 are unclear. The only recurrent event on chromothriptic chromosome 3 was the deletion of *TGFBR2* (3p24) in two cases; suppression of this gene has been implicated in SCLC progression (20) (**Figure 3C, Supplementary Figure S8**). Loss of the 3p chromosome arm is a well-known recurrent event in SCLC, occurring in >90% of SCLC cases (21). Although the relevant target gene(s) on 3p are not well established, *ROBO1* and *FHIT5* have been suggested as potential candidates (2); no deletions (**Supplementary Table S7**) or loss of expression by RNAseq was identified in these genes in our cases.

To directly compare the prevalence and extent of chromothripsis in aSCLC versus other major lung cancer types, we analyzed publicly available WGS datasets for LUAD (18) and *RB1⁻/TP53⁻* SCLC (2), using similar computational methods as applied to aSCLC. We found that, as previously reported for LUAD (18), chromothripsis was present in a subset of cases; however, for both LUAD and *RB1⁻/TP53⁻* SCLC, its prevalence and extent, as reflected by the number of SV per chromothriptic chromosomes, were significantly lower than in aSCLC (**Supplementary Figure S9A,B**).

These findings establish aSCLC as a tumor that is predominantly characterized by extensive chromothripsis, leading to amplification of oncogenes and loss of tumor suppressors, some of which have established critical roles in the biology of lung neuroendocrine cancers.

Analysis of oncogenic fusions

In addition to CNAs, the other major potential functional consequence of chromothripsis is the formation of oncogenic fusions (22-24). Therefore, we investigated whether a massive number of SVs in aSCLC was associated with the formation of in-frame fusions. For cases with available RNAseq data, we found that despite the massive number of rearrangements by WGS, cases with chromothripsis harbored on average only 16 fusions per case (range 0-29), of which only a small fraction (mean 3 per case, range 0-5) was in-frame (**Figure 3F**). All in-frame fusions involved genes with unknown oncogenic function by OncoKB, except for a single fusion involving *ALDH1L2* - a folate regulatory enzyme considered likely oncogenic, but with no established role in lung cancer or neuroendocrine cancers (25). Furthermore, none of the fusions was recurrent (**Supplementary Table S10**). We therefore conclude that while fusions could have a contributory role, they are unlikely to be the dominant oncogenic drivers in aSCLC.

CCND1 upregulation due to a genomic rearrangement resulting in enhancer hijacking

We further asked whether chromothripsis-associated SVs could deregulate cancer genes by disrupting their regulatory sequences. In case A17, harboring chromothripsis on chromosome 2 but no associated oncogene amplifications, there was a genomic rearrangement between *SH3RF3* gene on chromosome 2 and a region upstream of *CCND1*. *SH3RF3* is a highly expressed gene in lung tissue (ID: ENSG00000172985 in the GTEx RNA dataset (26)), and the region juxtaposed upstream of *CCND1* is rich in active chromatin marks based on the data extrapolated from multiple tissue types (Epilogos search tool (27)), suggesting a putative enhancer hijacking mechanism to upregulate *CCND1* in the absence of amplification (**Figure 3G**). Indeed, marked overexpression of cyclin D1 was detected in this case by both IHC and RNAseq (**Figure 2C**). This illustrates an alternative mechanism of oncogene upregulation in aSCLC, in line with prior cancer-wide evidence for SVs involving upstream regulatory sequences representing an efficient mechanism of oncogene activation (28,29).

Chromothripsis through micronucleation and formation of extrachromosomal DNA

Next, we explored the architecture of oncogene amplification and potential generative mechanisms of chromothripsis by integrative analysis of SV breakpoints and associated CNAs in WGS. This confirmed the cardinal features of chromothripsis, including characteristic copy-number oscillations and numerous SVs interleaved in random orientations, indicating ligations between DNA fragments in a random manner (**Figure 4A**). Previous studies reported that chromothripsis can be observed either in a whole chromosome or more locally (18); in seven of our cases, the chromothripsis events involved the entire length of the affected chromosome(s) (**Supplementary Figure S10**). Whole-chromosome involvement with evidence of heavy fragmentation and re-ligation suggests that chromosomal mis-segregation during mitosis, micronucleation, and subsequent massive DNA damage and rearrangements represent an underlying mechanism of chromothripsis in aSCLC (**Figure 4B**) (16,30).

In addition, a group of genomic segments exhibited far higher copy numbers than the other chromothriptic segments, while maintaining other features of chromothripsis. This suggested the direct formation of circular extrachromosomal DNA (ecDNA) from chromothripsis rather than amplicon formation subsequent to chromothripsis (31). Some oncogenes were further amplified up to twice the copy-number level of the other ecDNA fragments (e.g., *CCND2* in A08; **Figure 4A, lower panel**), indicating early duplication of the oncogenes within the ecDNA and subsequent high-level amplification.

To further corroborate the nature of amplification, we performed fluorescence in situ hybridization (FISH) for *CCND1* or *CCND2/CDK4/MDM2* in cases with corresponding amplifications and sufficient residual tissue for this analysis (n=5). In all cases, this confirmed ecDNA (“double-minute”) amplification of the tested genes, with one case exhibiting both extra- and intra-chromosomal amplifications (**Figure 4C, 4D**).

Analysis of potential predispositions to chromothripsis in atypical SCLC, including germline analysis

We next sought to examine whether patients with aSCLC harbored potential predispositions that were previously implicated as risk factors for chromothripsis.

Given the previously described association of chromothripsis with germline mutations in *ATM* and *TP53* (32-34), we reviewed the germline data for the aSCLC cohort. Germline variant calls were evaluable for 17 of the 20 patients and revealed no pathogenic germline variants in any of the 90 tested genes associated with hereditary cancer susceptibility. While no germline *ATM* mutations were identified, the enrichment in somatic *ATM* mutations may be of interest. *ATM* blocks cell cycle progression in the presence of DNA double-strand breaks (35), and disabling this checkpoint could provide a permissive environment for the development of aSCLC. In model systems, *ATM* inhibition can lead to increased formation of micronuclei (36).

The association of viral DNA with chromosome pulverization and chromothripsis has been previously suggested (37,38). We screened aSCLC sequences for a wide range of human viruses by tNGS (39), and found no viral DNA in any samples.

Finally, chromothripsis has been associated with telomere dysfunction, resulting from either up-regulation of telomerase reverse transcriptase (*TERT*)(34) or through a repair-based pathway called alternative lengthening of telomeres (*ALT*) (40). No *TERT* overexpression was identified in aSCLC cases with available RNAseq (**Supplementary Figure S8**), and no *ALT* footprints were identified by WGS, suggesting that these phenomena are unlikely to play a role in the development of these tumors.

Spatial and temporal conservation of chromothripsis and associated amplifications

In prior studies, chromothripsis has been postulated to represent an early causative event in tumorigenesis (17,34), but it can also occur as an acquired event later in disease progression or in association with systemic therapy resistance (31,34). In aSCLC cohort, serial pathologic samples from different locations and time-points were available for 14 patients; of those, 8 patients had multiple samples analyzed by tNGS (2-4 per patient), and one patient (A17) had WGS performed on both primary and metastatic samples (**Figures 1B-C, 5A**). First, all samples from individual patients were clonally related based on shared genomic alterations (**Supplementary Table S5**). Furthermore, they exhibited concordant presence or absence of chromothripsis involving identical chromosomes, as well as concordance in associated gene amplifications and overexpression by IHC (**Figure 5A**). Lastly, presence of chromothripsis was unrelated to the administration of systemic therapy. This indicates that in our cohort, chromothripsis represents an early and stable event that is conserved temporally and spatially through disease progression.

To further interrogate the timing of chromothripsis and associated amplifications in the clonal evolution of aSCLC, we assessed the ratio of duplicated to non-duplicated mutations within amplified regions using MutationTimeR (41). Across all samples, all or nearly all chromothripsis-associated amplifications were estimated to occur early in tumorigenesis (**Supplementary Table S11**), in line with our multi-sample data.

We next reviewed the matched NGS data for two patients with chromothripsis and histotype heterogeneity, in whom sequencing was performed on primary lung tumor samples with carcinoid histotype and metastatic samples with SCLC histotype (A05 and A17). This demonstrated matching chromothripsis patterns in primary and metastatic tumors (**Supplementary Figures S11A, B**), and no acquired unique mutations in established oncogenes or tumor suppressor genes in samples with the SCLC histotype (**Supplementary Figure S11C**). This suggests that other events superimposed on chromothripsis, possibly epigenetic, may facilitate the carcinoid-to-SCLC phenotypic transition, which, interestingly, parallels the findings in adenocarcinoma-to-SCLC plasticity where transition to SCLC phenotype is thought to be mediated by epigenetic reprogramming rather than additional genomic events (9,10,42).

Multimodal cell cycle deregulation in atypical SCLC

Given that chromothripsis-associated amplifications were centered on upstream regulators of pRb protein (CDK4, D-type cyclins), we further explored cell cycle deregulation using integrated DNA, RNA and IHC results. As summarized in **Supplementary Figures S12A,B**, pRb pathway deregulation at the DNA, mRNA and/or protein levels was evident in 75% of the samples in this cohort.

pRb pathway in aSCLC was further investigated via gene pathway analysis by RNAseq, which revealed major upregulation of the pro-proliferative G2M checkpoint and E2F target signatures in aSCLC, at a level markedly exceeding that found in a control set of carcinoid tumors and in some cases reaching the level seen in conventional smoking-associated SCLC (**Supplementary Figure S12C**).

Interestingly, in the multi-sample analysis, we identified acquired *RB1* alterations in two aSCLC patients (A16, A20) (**Figure 5B**); in both cases, *RB1* was wild-type and expressed in primary lung tumors, whereas metastatic/subsequent sample(s) exhibited acquired, private *RB1* mutations and concurrent loss of pRb expression (**Figure 5C**). In addition, one primary tumor (A12) exhibited a subclonal *RB1* mutation (cancer cell fraction 71%) with a corresponding subclonal loss of pRb expression by IHC in ~70% of the tumor

cross-sectional area. In contrast, in conventional SCLC, *RB1* mutations/loss are consistently clonal events (4). The functional significance of acquired and subclonal *RB1* alterations in aSCLC is unclear, given that there was no overt escalation in the proliferation rate associated with pRb loss (**Figure 5C**). It is possible that *RB1* inactivation in these samples plays a non-canonical role outside of E2F-mediated cell cycle control, as has been suggested for other cancer types that exhibit acquired *RB1* alterations later in disease progression (43).

In prior studies, expression pRB and p16 was found to be consistently reciprocal in conventional SCLC (12). Interestingly, some aSCLC exhibited a unique disjoining of this reciprocity by IHC (**Figure 5B**) and RNAseq (**Figure 5D**). First, all cases with chromothripsis on chromosome 12 were paradoxically p16^{high} despite expressing a wild-type pRB (A07, A08, A09) (**Supplementary Figure S2** illustrates case A08). We hypothesize that in such cases post-translational inhibition of pRB by ecDNA-amplified *CDK4* may be so extreme that it results in a pRb-null-like state, leading to reciprocal p16 overexpression. Intriguingly, p16 overexpression was similarly noted in liposarcomas lacking *RB1* alterations but harboring ecDNA-based *CDK4* amplification (44). Also, aSCLC samples with acquired or subclonal *RB1* mutations, remarkably, lacked p16 overexpression. This could reflect epigenetic silencing of *CDKN2A* – the gene encoding p16 – in the distinct precursors of aSCLC, in line with the reported *CDKN2A* promoter methylation in carcinoids/NETs (45), which, as we discuss next, may represent putative progenitors of aSCLC. Overall, these findings further highlight unique features of pRb pathway deregulation in aSCLC compared to conventional SCLC.

Lastly, unlike *RB1*, *TP53* remained consistently wild-type in multi-sample analysis. Notably, in cases with chromothripsis on chromosome 12, co-amplification of *MDM2* (a potent p53 antagonist), together with *CDK4* and *CCND2* (pRb antagonists), may represent a full phenocopy of the *TP53* and *RB1* genomic co-inactivation in conventional SCLC.

Histogenetic relationships in atypical SCLC versus SCLC in never-smokers with *RB1*⁻/*TP53*⁻

The observation that five aSCLC patients had both SCLC and low-proliferative carcinoid histotypes detected suggested the notion that these SCLC have a histogenetic relationship with lower-grade neuroendocrine tumors. We also found that additional aSCLC, including those without evidence of carcinoid histology, expressed Orthopedia Homeobox Protein (OTP) by IHC and RNAseq (**Figure 6A**) – a marker that is uniquely expressed in a subset of pulmonary carcinoids, but not in conventional SCLC (46). Overall, based on histopathology, OTP expression, and/or presence of genomic alterations that, while not entirely specific, are highly characteristic of carcinoid tumors (*MEN1*, *EIF1AX*), 55% of aSCLC exhibited features of a histogenetic relationship with pulmonary carcinoids.

Among those five patients with SCLC and carcinoid histotypes, two patients had completely resected primary lung tumors (A17, A20), and both were entirely composed of lower-proliferative carcinoid histology (**Supplementary Figures S3, S4**), suggesting that at least in some cases, aSCLC may arise through major clonal selection or dedifferentiation during metastatic progression. Also notable is that some aSCLC patients had only SCLC histotype identified in all samples. While this could be a result of limited sampling by small biopsies in some cases, three patients (A02, A07, A08) had completely resected primary tumors composed entirely of SCLC histology. This conversely suggests that some aSCLC may arise via divergence at an early progenitors stage, similar to the phenomenon recognized among molecularly-defined dedifferentiated sarcomas that may lack detectable differentiated components (47). Overall, we document a histogenetic relationship with carcinoids in the majority of aSCLC, with a spectrum of pathologic manifestations, that may reflect plasticity occurring at different time-points of disease evolution.

We next aimed to better characterize the differences in the genomic and clinicopathologic characteristics of aSCLC versus the *de novo* SCLC in never-smokers harboring *RB1* and *TP53* genomic mutations (nsSCLC-*RB1*⁻/*TP53*⁻; n=18). In contrast to aSCLC, none of those tumors exhibited features of chromothripsis based on the lack of oscillating CNAs by tNGS. Instead, 56% of nsSCLC-*RB1*⁻/*TP53*⁻ harbored either canonical *EGFR* (39%) or *KRAS* (6%) mutations, and/or displayed histologic components of NSCLC, most commonly in the form of adenocarcinoma (**Figure 6B**). Furthermore, nsSCLC-*RB1*⁻/*TP53*⁻ was enriched in the apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) mutagenesis signature, in contrast to the lack of this signature in aSCLC (**Figure 6C**; **Supplementary Figures S5B, S5C**). These data support and expand on prior observations on the association with *EGFR* mutations, adenocarcinoma histologic components and APOBEC mutagenesis in *RB1*⁻/*TP53*⁻ SCLC in never-smokers (10,42,48,49), and highlight the contrast between these cases and aSCLC.

Last, we compared aSCLC and nsSCLC-*RB1*⁻/*TP53*⁻ for the distribution of ASCL1, NEUROD1 and POU2F3 – recently identified markers of transcriptional subtypes in SCLC (50) (**Figure 6D**). Strikingly, all evaluated aSCLC were exclusively ASCL1-positive, suggesting a specific relationship between aSCLC and an ASCL1-expressing subset of carcinoids (51). In contrast, nsSCLC-*RB1*⁻/*TP53*⁻ had comparable distribution of transcriptional subtypes to that of smoking-associated SCLC. YAP1 – a marker associated with NSCLC-to-SCLC plasticity (52) – was enriched in nsSCLC-*RB1*⁻/*TP53*⁻, whereas it was entirely absent in aSCLC (**Figure 6E**). These data further support the distinct properties of aSCLC compared to those of other SCLC.

In aggregate, our data suggest a model of SCLC pathogenesis in never-smokers involving two distinct plasticity-mediated pathways (**Figure 6F**) – one, known from prior studies but expanded in our cohort, with a histogenetic link to NSCLC precursors, where conversion to SCLC occurs in *RB1*⁻/*TP53*⁻ background in association with APOBEC mutagenesis, and the other, a novel pathway, with a histogenetic link to carcinoid precursors, where the plasticity occurs in an *RB1*⁺/*TP53*⁺ background through chromothripsis-mediated amplification of oncogenes, some of which have key roles in pRb and p53 suppression.

Survival, treatment outcomes, and expression of therapeutic biomarkers in atypical SCLC

Having identified the highly distinct genomic and clinicopathologic features of aSCLC compared to other SCLC, we next sought to characterize the clinical outcomes and treatment responses associated with this newly defined entity. The survival of patients with aSCLC was compared to that of patients with other types of SCLC (sSCLC and nsSCLC-*RB1*⁻/*TP53*⁻) and atypical carcinoids – the subset of carcinoids defined by increased mitotic rate (see **Methods**); patients with typical carcinoids were excluded from this analysis because they are well established to have only rare tumor-associated mortality (53). Kaplan-Meier analysis revealed that aSCLC was associated with distinct survival characteristics, which were intermediate between those of SCLC in smokers and atypical carcinoids (**Figure 7A**). In contrast, nsSCLC-*RB1*⁻/*TP53*⁻ had an outcome similar to that of smoking-related SCLC.

We also investigated the profile of therapeutic sensitivity in aSCLC, focusing on platinum-based therapies – the mainstay first-line therapy for SCLC. Of the 15 patients with available response data for platinum-based systemic therapy, complete or partial responses were achieved in 13% and 20% of patients, respectively (**Figure 7B**), which is well below the historical rates expected for conventional SCLC (70%) (1). Using time-to-next treatment as a surrogate of clinical benefit, a subset of patients appeared to exhibit sensitivity to immune checkpoint inhibitors, as reflected by the length of treatment of 2 to >5 years for three of the five treated patients (**Figure 7C**). An additional four of the six patients treated with TMZ had >10 months on treatment, ranging up to 2 years. The long time on treatment for

these two therapies stands in contrast to the clinical experience with standard SCLC patients, and suggests that distinct pathways of treatment sensitivity may be present in aSCLC. The cohort size was too small to evaluate specific genomic or clinicopathologic features associated with survival and treatment outcomes, although we note that patients with longest time on TMZ (A01 and A07) had the lowest expression of O⁶-methylguanine-DNA methyltransferase (MGMT) by RNAseq – the marker that is inversely associated with TMZ sensitivity (54) (**Supplementary Figure S13A**).

Several targeted therapies are currently under clinical investigation for SCLC, including agents directed at DLL3 (55) and SEZ6 (56). To determine whether aSCLC patients might be candidates for these therapies, we analyzed the expression of these markers in samples with available tissue. All evaluated samples (n=9) exhibited high levels of DLL3 (H score 200-300; mean 278) and SEZ6 (H score 140-300; mean 240) (**Supplementary Table S2; Supplementary Figure S13B**). The high level of DLL3 expression is in line with consistently high ASCL1 expression in aSCLC, as the *DLL3* gene is a downstream target of ASCL1 (57).

DISCUSSION

Here, we provide the first detailed genomic and clinicopathologic description of a rare subtype of SCLC that lacks concomitant *RB1* and *TP53* inactivation and arises in the absence of smoking-induced carcinogenesis. We identify extensive chromothripsis with recurrent ecDNA amplification of several oncogenes involved in the regulation of pRb and p53 as a hallmark feature in these cases and microsatellite instability as an alternative genomic context. Furthermore, we identify a histogenetic link between these tumors and pulmonary carcinoids, suggesting a new pathway for the development of SCLC via progression from lower-grade neuroendocrine tumors or their progenitors. This study defines aSCLC as a novel entity among lung cancers, highlighting its exceptional etiology, distinct clinicopathologic properties, and unique therapeutic vulnerabilities.

Since its initial description in 2011, chromothripsis has emerged as a major driver of tumorigenesis that can exert a profound impact on tumor genomes by generating diverse genomic alterations (17). A common consequence of chromothripsis is oncogene amplification, which often occurs as ecDNA – the autonomously replicating unit that enables the accumulation of a remarkably high copy number of amplified genes (31,58). Similar to aSCLC, *CCND1*, *CDK4* and *MDM2* are among the most common amplifications associated with chromothripsis across tumor type (18). In addition to ecDNA, we also found evidence of *CCND1* upregulation through putative enhancer hijacking resulting from chromothripsis-mediated rearrangement of regulatory elements, highlighting the diversity of mechanisms by which chromothripsis may contribute to cancer development. In a minority of cases, established driver gene(s) targeted by chromothripsis could not be identified; it is possible that cumulative effect of alterations in multiple genes could underly carcinogenesis in some tumors.

Chromothripsis has been identified across a wide spectrum of cancer types, with a particularly high prevalence in sarcomas and gliomas (18,34). Although identified as a hallmark feature of aSCLC, chromothripsis is not unique to this subset of SCLC. In fact, the initial description of chromothripsis included an *RB1*⁻/*TP53*⁻ SCLC cell line (SCLC-21H), which harbored chromothripsis involving chromosome 8 with ecDNA amplification of *MYC* (17,59). Furthermore, recent studies have identified ecDNA amplification of *MYC* paralogs in a minority of *de novo* SCLC tumor samples (60) and post-treatment SCLC models (61), although the prevalence of underlying chromothripsis in such cases remains to be determined. In our analysis of a set of published whole genomes of *RB1*⁻/*TP53*⁻ SCLC, chromothripsis

was indeed identified in a minority of cases, but was substantially more limited in extent than in aSCLC. Overall, while chromothripsis and ecDNA oncogene amplification are not exclusive to aSCLC, this cohort is distinguished by its high prevalence and extent, recurrent involvement of chromosomes 3, 11, and 12, and *RB1*⁺/*TP53*⁺ context with low number of other mutations, suggesting that in these tumors chromothripsis represents a central driver of tumorigenesis.

Analysis of chromothripsis and its clinicopathologic significance at scale has been hampered by the lack of established methods for the detection of this phenomenon in panel NGS assays utilized for clinical sequencing. To date, large-scale studies on chromothripsis have been based primarily on WGS, utilizing integrated assessment of SVs and CNAs (18,34); although CNA only-based approaches have also been applied to WGS (62,63) or genomic array platforms (61). To our knowledge, this study is the first to demonstrate the feasibility of detecting chromothripsis in a panel NGS assay based on a distinctive oscillating CNA pattern. Although a larger set of samples is needed to fully validate the performance metrics, this advance should facilitate wider recognition of chromothripsis in NGS panel sequencing.

Currently, the understanding of lung neuroendocrine cancers centers on SCLC and carcinoids representing entirely unrelated tumor entities, occurring in distinct patient populations (younger, never-smokers for carcinoids vs. older smokers for SCLC), and characterized by highly distinct genomic profiles, most notably separated by *RB1* and *TP53* genomic status and TMB (12). Here we identified that most of *RB1*⁺/*TP53*⁺ tumors that have histomorphology of *bona fide* SCLC exhibited evidence of a histogenetic link with carcinoids, including harboring genomic alterations and marker expression characteristic of pulmonary carcinoids, and exhibiting co-occurring carcinoid histotype in some cases. Small cell transformation – that may occur in *de novo* tumors or as a form of acquired resistance to therapy – is a well-established phenomenon in several organs, including lung and prostate adenocarcinomas (64), in line with the concept of small cell carcinoma representing a convergent phenotype that may arise from a variety of precursors (65). Here we add carcinoid-SCLC transition as a rare form of plasticity associated with small cell phenotype.

Pulmonary carcinoids are generally indolent tumors, but a subset can metastasize and behave aggressively (12). Aggressive subsets of carcinoids have been suggested based on gene expression characteristics (51), genomic features (66) and high TERT expression (67). We postulate that chromothripsis may represent a novel tractable risk factor for carcinoid progression and dedifferentiation. Previously, isolated instances of chromothripsis in lung carcinoids were reported, including cases with chromothripsis on chromosomes 3, 12 and 13 (68), 2, 11 and 20 (68), 11 and 20 (51) and 11 (69). Only limited clinicopathologic information is available for most of these cases, but some exhibited aggressive clinical behavior (68).

In our series, all patients with *RB1*⁺/*TP53*⁺ SCLC were never or light smokers. However, in the George et al study, one of the two patients with chromothripsis-associated *RB1*⁺/*TP53*⁺ SCLC was reported to have a 30 pack-year smoking history (2). Thus, in aggregate, while apparently uncommon, tobacco exposure may not preclude chromothripsis-mediated pathogenesis in SCLC.

In this study, we contrasted aSCLC with *de novo* SCLC in never-smokers harboring *RB1* and *TP53* genomic alterations – another rare and incompletely characterized subset of SCLC. The genomic and pathologic data generated here support predominant adenocarcinoma-to-SCLC plasticity in this subset and its association with *EGFR* mutations and APOBEC mutagenesis, and contrast this pathway with carcinoid-to-SCLC plasticity associated with chromothripsis in aSCLC. A translational implication of this model is that *RB1*/*TP53* mutation status – which can be assessed by routine immunohistochemical methods – may

serve as surrogate markers for putative tumor progenitors in the never-smoker SCLC population. This is clinically relevant since aSCLC is associated with a distinct prognosis, and might ultimately define a category with distinct therapeutic approaches.

Our study suggests several potential therapeutic vulnerabilities in aSCLC based on the unique genomic underpinnings of these tumors. These may include agents under clinical development that target ecDNA-based oncogene amplification (70), as well as agents targeting amplification and overexpression of CDK4 (71) and MDM2 (72). Furthermore, consistently high expression of DLL3 and SEZ6 suggests that these patients could benefit from the emerging therapies targeting these cell surface determinants (55,56). Such targets may be of particular importance given the relative platinum insensitivity of aSCLC. The apparent efficacy of TMZ is also notable, as this agent is active in both SCLC (54) and lung carcinoids (73), given the dual histologic characteristics in this cohort.

In conclusion, here we describe a new pathway for the development of SCLC mediated predominantly by chromothripsis in tumors with a histogenetic link with lower-grade carcinoid tumors or their progenitors. This study establishes a novel concept in lung tumorigenesis with potential therapeutic implications.

METHODS

Study Design

This study was approved by the Memorial Sloan Kettering Cancer Center (MSKCC) Institutional Review Board (IRB). In accordance with the Declaration of Helsinki, all patients included in this study signed a written informed consent form following the IRB's recommendations. All patients included in the study had tumors analyzed prospectively by MSK-IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets) as part of routine clinical care at the MSKCC. A detailed review of the demographic, radiologic, pathologic and clinical information was performed retrospectively. Pathology slides were retrieved and re-reviewed. If sufficient residual DNA or formalin-fixed paraffin-embedded (FFPE) tissue was available, the samples were further analyzed by whole-genome sequencing (WGS), RNAseq, FISH and additional IHC. For patients with multiple samples, WGS was performed on the chronologically earliest sample or based on sufficiency for additional testing. For comparison with the study group, cohorts of SCLC (4) and pulmonary carcinoids that were prospectively sequenced using MSK-IMPACT were included in the analysis.

Clinicopathologic Assessment

Clinical patient characteristics were annotated by reviewing the electronic medical record. The baseline characteristics included age, sex, tobacco smoking history, date of diagnosis, pathology sample site, and pathology sample type. Radiology data were reviewed to collect information on the metastatic site distribution at presentation. Smoking history was collected from the patient-completed smoking questionnaire. Pack years of smoking was derived as: [(average number of cigarettes smoked per day/20) X years of smoking]. Never-smokers were defined as patients who had smoked <100 cigarettes, and light smokers were defined as those who had a ≤ 10 pack-year smoking history.

Tumor classification was performed according to the criteria in the 2021 World Health Organization (WHO) classification of thoracic tumors (74). All tumors underwent central pathologic review by pathologists with expertise in thoracic tumors (NR, CFA, and JC). The criteria for SCLC included

undifferentiated morphology with high a nuclear/cytoplasmic ratio and nearly imperceptible cytoplasm, cell molding, mitotic count of >10 per 2 mm^2 (if sufficient well-preserved tissue available to perform the counts) and Ki67 proliferation index of $\geq 50\%$, commonly associated with extensive necrosis. The criteria for carcinoids included a well-differentiated morphology composed of uniform cells with lower nuclear/cytoplasmic ratios and readily visible cytoplasm, mitotic count of ≤ 10 per 2 mm^2 , and absence of extensive necrosis. For survival analysis, carcinoids were classified as typical or atypical using the WHO criteria: <2 mitoses per 2 mm^2 and no necrosis for typical carcinoids, and ≥ 2 mitoses per 2 mm^2 and/or focal necrosis for atypical carcinoids. Samples that were too crushed or poorly preserved to evaluate the morphology, or that had equivocal morphological features, were regarded as unclassified. For all patients, pathologic specimens and clinical records were reviewed in detail to exclude the possibility of an alternative tumor type or non-pulmonary origin.

Immunohistochemistry

Immunohistochemistry was performed by previously established and validated protocols, as summarized in detail in **Supplementary Table S12**. Primary antibodies included synaptophysin (SNP88, Bio Genex), chromogranin A (LK2H10, Ventana), CD56/NCAM/neural cell adhesion molecule (MRQ42, Cell Marque), INSM/Insulinoma-associated protein 1 (A-8, Santa Cruz), Ki67 (MIB1, Dako), OTP/Orthopedia homeobox protein (EPR22178-17, Abcam), ASCL1 (24B72D11.1, BD Biosciences), NEUROD1 (EPR17084, Abcam), POU2F3 (6D1, Santa Cruz), YAP1 (63.7, Santa Cruz), DLL3 (SP347, Ventana), SEZ6 (SC17.14, Creative Biolabs), cyclin D1 (SP4, Lab vision), cyclin D2 (M20, Santa Cruz), Mdm2 (IF2, Milipore), Cdk4 (CDS-156, Biosciences), pRb (13A10, Leica), p53 (D07, Ventana) and p16 (E6H4, Ventana). Transcriptional subtype (SCLC-A, -N, -P, -TN/triple negative) was assigned based on the predominant expression of ASCL1, NEUROD1, POU2F3 or triple-negative by IHC, respectively, as described previously (52). For semi-quantitative scoring of DLL3 and SEZ6, the histologic (H) score was derived by multiplying the intensity of staining (1+ weak, 2+ moderate, 3+ strong) by the percentage of cells staining (1-100%), yielding H scores from 0 to 300. The Ki67 proliferation index was assessed as the percentage of positive cells in hot-spot areas – regions with the highest Ki67 rate counted in at least 500 tumor cells (75).

Survival and Treatment Outcome Analysis

Data on patient treatments and outcomes were collected by reviewing electronic medical record, including treatments administered and best response to platinum/etoposide, date of death, or last follow-up. Disease-specific overall survival was estimated from the date of diagnosis to the date of documented death from disease or the last follow-up using the Kaplan-Meier approach. For comparative survival analysis, control groups of SCLC in smokers and atypical carcinoids were generated, consisting of consecutively-encountered patients with available clinical follow-up. For the swimmer plots, treatment times were depicted from the first administration date of one therapy to a next therapy, counting maintenance therapies (such as platinum/etoposide/atezolizumab followed by maintenance atezolizumab) as one regimen. Imaging studies and reports were manually reviewed to generate a real-world response rate by comparing on-treatment scans to pre-treatment scans. Patients were considered to have partial response, complete response, stable disease or progressive disease on the basis of clinician interpretation of the change in disease burden on subsequent scans from the first pre-treatment scans.

Targeted NGS by MSK-IMPACT

Genomic sequencing was performed on tumor DNA extracted from FFPE tissue and normal DNA was sequenced in all patients using the FDA-authorized MSK-IMPACT targeted sequencing panel using methods and analysis as previously described (5,6). Briefly, the MSK-IMPACT assay is a clinically validated FDA-authorized custom hybridization capture-based platform that sequences the entire coding region and select noncoding regions of 341 (v3 – 1 sample), 468 (v5 – 2 samples), or 505 (v6 – 17 samples) genes for the detection of SNVs, indels, CNAs, and select SVs (gene list provided in **Supplementary Table S13**). Somatic alterations were classified as oncogenic, likely oncogenic, or unknown using OncoKB (19,76).

Tumor mutational burden (TMB) was calculated as the number of non-synonymous mutations in canonical exons per megabase. Tumor Purity was estimated by FACETS and the “hisense” solution is reported, cval parameter = 50. Microsatellite instability (MSI) was analyzed using MSIsensor (<https://github.com/ding-lab/msisensor>) from the MSK-IMPACT sequencing data. A minimum of 800 loci and tumor-normal matched sequencing were required for MSI evaluation. Tumors with MSIsensor scores of <3, ≥3 to <10, and ≥10 were classified as MS stable (MSS), MSI indeterminate (MSI-ind), and MSI-H, respectively using previously validated cutoffs (77). MSI-ind were adjudicated by MiMSI, an independent algorithm for MSI calling based on multiple instance learning (<https://www.biorxiv.org/content/10.1101/2020.09.16.299925v1.full.pdf>).

To assess subclonal mutations, cancer cell fraction (CCF) was estimated for selected single nucleotide variants as a function of variant allele frequency (VAF), tumor purity (p), and allelic CN state, as previously reported: $VAf(CCF) = p * CCF / [CN_{diploid} * (1-p) + CN_{mut} * p]$ and using a binomial distribution and maximum likelihood estimation normalized to produce posterior probabilities (78). CN_{mut} was calculated using the expected number of copies for each mutation generated based on the observed VAF and local copy number (via FACETS, see below)(79); a CCF of ≥ 0.8 was regarded as clonal or near-clonal.

Assessment of *RB1* and *TP53* by integrated genomic and immunohistochemical analysis

The screening of consecutive SCLC analyzed by MSK-IMPACT for *RB1* and *TP53* status was performed using an integrated approach utilizing genomic alterations and expression of pRb and p53 proteins by IHC, respectively. Based on prior work (4), for cases lacking detectable *RB1* (NM_000321) genomic alterations by MSK-IMPACT routine clinical pipeline (which covers all exons of *RB1* gene in all versions of MSK-IMPACT and 5'UTR and introns 6, 8 and 23 in latest version V6), manual review was performed to identify noncanonical splice-site mutations. Also, pRb IHC was performed, and cases with complete loss of pRb protein expression were classified as *RB1* deficient (*RB1*⁻), whereas only those lacking *RB1* genomic alterations and exhibiting retained pRb expression were designated as *RB1* proficient (*RB1*⁺). Also, as reported previously (4), we assessed the expression of D-type cyclins and p16^{INK4A} – the upstream pRb regulators – to further corroborate pRb proficiency, which was supported by D-type cyclin^{high} (H score >50) and/or p16^{low} (H score <100) profile.

p53 IHC was also performed for all cases lacking *TP53* (NM_000546) genomic alterations by MSK-IMPACT (which covers all exons of *TP53* in all versions of MSK-IMPACT and 5'UTR in latest version V6) to confirm the wild-type expression pattern. Using the standard criteria, p53 mutant/aberrant pattern included any one of the 3 patterns: strong nuclear intensity in 80-100% tumor cells (overexpression pattern, reflecting aberrant degradation of p53, usually resulting from missense *TP53* mutations), complete lack of immunoreactivity (null pattern, reflected degradation of p53 harboring truncating *TP53*

mutations), or strong cytoplasmic reactivity only (reflecting mutations disrupting nuclear localization domain), whereas p53 wild-type pattern was defined as any staining other than the 3 mutant patterns (80,81).

Cases lacking *RB1* or *TP53* genomic alteration but with insufficient tissue for IHC confirmation were excluded.

Analysis of Copy Number Alterations in MSK-IMPACT

The CNAs were evaluated by MSK-IMPACT using coverage-based method (82). The FACETS algorithm (83) was also applied to define copy number states, the total copy number in gene amplifications, chromosome-level alterations, and screening for chromothripsis. FACETS was run on matched normal mode using two critical segmentation values (CVAL): a “purity” output using a CVAL= 150 and “hisens” output with CVAL= 50. Other refitting parameters such as diplogR and minNhet were adjusted on a case-by-case basis (83).

To assess potential chromothripsis by MSK-IMPACT, FACETS hisens output plots were manually inspected to identify chromosomal segments oscillating between predominantly two copy number states. A minimum number of five consecutive oscillating uniformly sized segments with estimated cancer cell fraction $\geq 50\%$ and matched patterns between plots (log-ratio copy number, odds-ratio copy number, and integer copy number) were considered suggestive of chromothripsis. A threshold of five consecutive oscillations was chosen to minimize the risk of overcalling chromothripsis in cases where two gene-level amplifications or deletions occurred in close proximity. Chromothripsis calls and method performance were evaluated by comparison with CN segments and SVs using WGS in cases with available material (see Results).

Analysis of Mutational Signatures in MSK-IMPACT

The mutational spectra of the SNVs were calculated using a custom algorithm designed in house (<https://github.com/mskcc/DeepSig/>). To obtain maximal sensitivity for single base substitution (SBS) signatures in SCLC which has no established signature catalog, we first employed *de novo* signature detection using a combined panel of 11 WGS samples from this study, 10 WGS samples that were *RB1/TP53* deficient from George et al. (2) (processed through an in-house pipeline TEMPO, <https://github.com/mskcc/tempo>), and SNV calls from 101 additional SCLC samples available from the supplemental table in George et al. (2). Briefly, using the Bayesian marginal likelihood method, we determined the most likely number of signatures present, 10. Next, signature decomposition was performed on the optimal number of signatures, and *de novo* signatures were subsequently annotated to known reference signatures using cosine similarity. *De novo* signatures with cosine similarity >0.7 to known Cosmic V3 signatures were called as the Cosmic signatures. Signatures with a common etiology were merged, e.g. SBS4 and SBS92; SBS31 and subtypes. Significance thresholds (alpha) were defined for each signature as follows: SBS44: $1e-4$; SBS11: $1e-3$; SBS40: $1e-3$; all others: $5e-2$. Dominant signatures were assigned to each sample as the signature with the largest mutational attribution. Only samples with ≥ 5 SNVs and those that had a matched normal were included in the analysis.

DNA Virus Read Detection in MSK-IMPACT

The presence of DNA viruses in tumor samples was determined by the analysis of off-target reads, as previously described (39). Briefly, all tNGS reads were aligned to the human genome (hg19). Paired unmapped reads from the processed BAM files were extracted into the FASTA files. Unmapped reads

from each sample were queried for viral content using blastn 2.9.0+ (parameters: strand both, word_size 28, e-value 1 e-10, perc_identity 90) and mapped to the genomes of selected human DNA viruses (EBV type 1-2, HHV type 1-8, Merkel Cell Polyomavirus, and HPV types 3, 5, 8, 9, 20, 21, 29, 33, 36, 45, 62, 71, 72, 74, 77, 81, 82, 86, 92, 105, 107, 115, 117, 118, 147, 150, 152, 174, and 178) from the National Center for Biotechnology Information Virus database. Each paired read that aligned with >90% identity was quantified as a read for the specific virus. Samples with >2 paired reads for a specific virus were considered positive.

Assessment of Germline Variants

To interrogate germline variants, a modified sequencing pipeline for paired tumor/normal MSK-IMPACT was utilized, which has been validated for clinical use in the context of an IRB-approved protocol (84), which covers 90 well-established cancer predisposing genes for pathogenic and likely pathogenic germline variants (listed in **Supplementary Table S13**). This analysis was performed for 15 evaluable patients. Furthermore, two additional patients underwent clinical germline testing using peripheral blood, performed using a New York State Department of Health approved germline test covering up to 90 hereditary cancer predisposition genes (82,85).

Whole-Genome Sequencing

After PicoGreen quantification and quality control by Agilent TapeStation, 293-500 ng of genomic DNA was sheared using an LE220-plus Focused-ultrasonicator (Covaris catalog # 500569), and sequencing libraries were prepared using the KAPA Hyper Prep Kit (Kapa Biosystems KK8504) with modifications. Briefly, libraries were subjected to a 0.5X size selection using aMPure XP beads (Beckman Coulter catalog # A63882) after post-ligation cleanup. Libraries were either not amplified by PCR and were pooled equivolume and quantitated based on their initial sequencing performance or were amplified with 5 cycles of PCR and pooled equimolar. Samples were run on a NovaSeq 6000 in a PE150 run using the NovaSeq 6000 SBS Kit and an S4 flow cell (Illumina). The average number of read pairs per sample was 1.4/1.1 billion for tumors and normal, respectively, corresponding to 102X/83X coverage. Coverage range for tumor and normal WGS samples was: 96.4X-192.1X and 59.2X-123.6X, respectively.

The WGS data were processed and analyzed using the TEMPO pipeline (<https://github.com/mskcc/tempo>). Briefly, the FASTQ files were aligned to the b37 assembly of the human reference genome from the GATK (<https://software.broadinstitute.org/gatk/>) bundle using BWA mem (v0.7.17) (<http://bio-bwa.sourceforge.net/>). The aligned reads were converted and sorted into BAM files using samtools (v1.9) (<http://htslib.org/>) and marked for PCR duplicates using GATK MarkDuplicates (v3.8-1). CNAs and loss of heterozygosity were determined using the FACETS (83) (<https://github.com/mskcc/facets>) and FACETS-suite (<https://github.com/mskcc/facets-suite>). The FACETS parameters (CVAL, diplogR) were adjusted on a case-by-case basis according to visual inspection. Structural variants were called using Manta (86) (<https://github.com/Illumina/manta>), SvABA (87) (<https://github.com/walaj/svaba>) and BRASS (<https://github.com/cancerit/BRASS>). Variants were normalized to a common representation and merged using a fixed window size of 200 bps using mergesvvcf (<https://github.com/papaemmelab/mergeSVvcf>). Genes were annotated as oncogenic or likely oncogenic according to the OncoKB Cancer Gene List (<https://www.oncokb.org/cancer-genes>, update 7/2023) (19). The merged SV calls were annotated using iAnnotateSV (<https://github.com/rhshah/iAnnotateSV>). Telomere content and shortening was estimated using TelomereHunter (88) (<https://github.com/linasieverling/TelomereHunter>).

RB1 and TP53 Assessment by Whole-Genome Sequencing

Non-coding variants identified in *RB1* and *TP53* by WGS were annotated using three algorithms: namely Combined Annotation Dependent Depletion (CADD) (89), Functional Analysis through Hidden Markov Models with extended Features (FATHMM-XF) (90) and SpliceAI (91). Overall pathogenicity (likely pathogenic or benign) for a variant was determined based on the majority consensus from all three prediction algorithms.

Assessment of Chromothripsis in Whole-Genome Sequencing

Merged SV and CN calls from the FACETS were inputted in Shatterseek (18)m and regions were determined to be chromothripsis if one or more of the following criteria were met: (1) At least 6 interleaved intrachromosomal SVs, 7 contiguous segments oscillating between 2 CN states, the fragment joins test, and either the chromosomal enrichment or the exponential distribution of the breakpoints test. (2) At least 3 interleaved intrachromosomal SVs and 4 or more interchromosomal SVs, 7 contiguous segments oscillating between 2 CN states and the fragment joins test. (3) At least 40 interleaved intrachromosomal SVs and the fragment joins test. (4) At least 100 SVs (intrachromosomal+ interchromosomal) and at least 5 contiguous segments oscillating between 2 CN states (5) At least 6 interleaved intrachromosomal SVs, 4, 5 or 6 adjacent segments oscillating between 2 CN states, the fragment joins test, and either the chromosomal enrichment or the exponential distribution of breakpoints test. If a region only passed criterion 5, it was regarded as “low density” for comparison with tNGS chromothripsis calling. Circos plots to visualize SV and CN were created using signature.tools.lib in R (<https://github.com/Nik-Zainal-Group/signature.tools.lib>).

Integrative Analysis of Structural Variant Breakpoints and Associated Copy-Number Alterations by Whole-Genome Sequencing

To infer mechanistic patterns in the WGS, we applied the Hartwig Medical Foundation (HMF) bioinformatics pipeline for our analysis (<https://github.com/hartwigmedical/hmftools>) (92). This pipeline was chosen because, in their PURPLE algorithm (v2.54), the boundaries of copy-number segments were determined by jointly analyzing the regional depth of coverage (COBALT v1.11), B-allele frequency (AMBER v3.5), and, most importantly, SVs. This integration resulted in near-complete concordance between the rearrangement breakpoints and the copy-number boundaries, which was pivotal in analyzing the SVs at the amplification boundaries. SVs were called primarily by GRIDSS2 (v2.12.0) (93) (<https://github.com/PapenfussLab/20rids>), annotated with RepeatMasker (v4.1.2-p1) (<http://repeatmasker.org/>) and Kraken2 (v2.1.2) (94) (<https://github.com/DerrickWood/kraken2/>), filtered by GRIPSS (v1.9), and further annotated and analyzed with LINX (95) (v1.15). Complex genomic rearrangements were reconstructed, as previously described (96).

Assessment of the Timeline of Amplifications in Whole-genome Sequencing

MutationTimeR (<https://github.com/gerstung-lab/MutationTimeR>) was run with default settings to estimate the timing for copy number alterations in chromothriptic regions. Mutation MAF files from TEMPO were first converted to VCF format using maf2vcf (<https://github.com/mskcc/vcf2maf/blob/main/maf2vcf.pl>). Copy number alteration segments were generated from FACETS as previously described. Subclonal cluster information was estimated using CliPP (<https://github.com/wwylab/CliPP>), except in sample A07 where due to the large number of TMZ induced mutations, computational resources limited CliPP from finishing. In this case a placeholder for subclonal mutations at 50% purity was used per the authors suggestion. Timing of copy number gains

and amplifications for key genes (e.g. *CCND1*, *CDK4*, etc.) were computed on a scale from 0 (early) to 1 (late), corresponding to the proportion of mutations before the gain.

Assessment of Chromothripsis in other Whole-Genome Sequencing Cohorts

To compare chromothripsis events in aSCLC to other major lung cancer types, 21 WGS of *RB1/TP53* co-mutated SCLC were downloaded from a previous study (2). BAM files were first converted to FASTQ (GATK v4.1.9.0 SamToFastq), then processed using the same TEMPO pipeline as the aSCLC samples. Similar to aSCLC samples, Shatterseek was used to determine chromothripsis events across all samples with the same thresholds as previously described. Included SCLC samples IDs were: S00830, S00945, S02065, S02209, S02219, S02237, S02243, S02248, S02274, S02402, S00838, S01297, S01366, S01861, S01873, S02139, S02234, S02277, S02328, S02376, S02241.

For comparison with LUAD, WGS data were obtained from a previous study (18). While the pipeline for processing the WGS data differed from the one used for aSCLC samples, both used Shatterseek to obtain chromothripsis metrics, and are expected to yield comparable calls for SV and CAN used in the calling of chromothripsis. The same criteria for calling chromothripsis events in the aSCLC samples was used in calling chromothripsis in LUAD. Included LUAD sample IDs were: TCGA-55-6986, TCGA-50-6597, TCGA-67-3771, TCGA-64-1680, TCGA-49-6742, TCGA-55-8299, TCGA-05-4398, TCGA-55-6982, TCGA-50-5930, TCGA-97-8171, TCGA-05-4420, TCGA-78-7535, TCGA-73-4666, TCGA-78-7158, TCGA-49-4486, TCGA-05-5429, TCGA-55-7281, TCGA-05-4397, TCGA-05-4395, TCGA-91-6847, TCGA-50-6591, TCGA-05-4396, TCGA-49-4512, TCGA-05-4389, TCGA-75-6203, TCGA-44-2659, TCGA-55-6972, TCGA-38-4628, TCGA-73-4659, TCGA-75-5147, TCGA-91-6840, TCGA-67-6215, TCGA-75-7030, TCGA-50-5932, TCGA-55-6984, TCGA-64-1678.

RNAseq

After RiboGreen quantification and quality control by Agilent BioAnalyzer, 1 µg of total RNA with DV200 percentages varying from 30-69% underwent ribosomal depletion and library preparation using the TruSeq Stranded Total RNA LT Kit (Illumina catalog # RS-122-1202) according to instructions provided by the manufacturer with 8 cycles of PCR. Samples were barcoded and run on a NovaSeq 6000 in a PE100 run, using the NovaSeq 6000 S4 Reagent Kit (200 cycles) (Illumina). On average, 131 million paired reads were generated per sample and 26% of the data were mapped to the transcriptome.

Reads were aligned and processed using the MSKCC FORTE pipeline (<https://github.com/mskcc/forte>). Briefly, raw FASTQ files from Rnaseq paired-end sequencing were aligned to the ENSEMBL GRCh37 Homo sapiens release 99 transcriptomes using Kallisto (97) and filtered to remove transcripts with low counts. Gene expression levels were then calculated as transcripts per million (TPM) using Sleuth (98). Fusions were determined by Arriba (97) and Fusion Catcher (99), and calls that were only found in a single caller were culled. Fusion frame was determined by the individual callers. Single sample gene set enrichment analysis (ssGSEA) was performed using the GSVA package in R. (100) Limma (101) was used to apply a linear model to the data to determine pathways significantly different between the cohorts.

Fluorescence in situ hybridization

FISH was performed on 5 µm sections from FFPE tissue for all evaluated cases. Additionally, FISH was performed on cells from a patient-derived xenograft (PDX) corresponding to case A07 (P-0039208), established as previously described (102) under the approval of the MSKCC Animal Care and Use Committee (IACUC Protocol 04-03-009). FISH analysis was performed using a 2-color *CCND1*/Cen11

probe and a 3-color CCND2/CDK4/MDM2 probe to confirm gene amplifications detected by NGS analysis. The 2-color probe mix consisted of bacterial artificial chromosome (BAC) clones spanning CCND1 (RP11-300I6 chr11: 69,453,281- 69,614,785; RP11-804L21 chr11: 69,589,482- 69,628,306; labeled with Red dUTP) and the centromeric repeat plasmid for chromosome 11 served as the control (clone pLC11A; labeled with Green dUTP). The 3-color probe mix consisted of BAC clones spanning CDK4 (RP11-571M6 chr12: 57,999,870- 58,211,408; RP11-970A5 chr12: 58,136,289- 58,353,071; labeled with Green dUTP), MDM2 (RP11-611O2 chr12: 69,192,689- 69,343,255; RP11-630N19 chr12: 69,337,168- 69,510,888; labeled with Red dUTP) and CCND2 (purchased from Empire Genomics, sequenced per proprietary; labeled with Orange dUTP). Probe labeling, tissue processing, hybridization, post-hybridization washing, and fluorescence detection were performed according to the standard laboratory procedures. Slides were scanned using a Zeiss Axioplan 2i epifluorescence microscope equipped with a MetaSystems (Waltham, MA) imaging system. The Metafer and Vslide modules within the system were used to generate virtual images of H&E- and DAPI-stained sections. The H&E sections served as guides to identify corresponding tumor regions in the DAPI-stained slides. The entire hybridized area or section was scanned under a 63x objective to assess the amplification status and representative regions imaged through the depth of the tissue. Amplification was defined as ≥ 6 copies of the gene and further categorized as double minute (extrachromosomal DNA segments often observed as paired signals) or homogeneously staining region (intrachromosomal DNA segments observed as medium to large, clustered signal).

Quantification and Statistical analysis

Statistical analyses were conducted using R version 4.2.3 (R Project for Statistical Computing, Vienna, Austria) (<http://www.r-project.org/>) and GraphPad Prism V10 (GraphPad Software). Fisher's exact test and Mann-Whitney *U* test were used to assess the categorical and continuous variables, respectively. All parametric and nonparametric tests were two-tailed, with $P < 0.05$ considered statistically significant.

Data Availability

The cBioPortal repository was created for all aSCLC samples in this study and can be accessed at https://www.cbioportal.org/study/summary?id=asclc_msk_2024. Raw data for WGS and RNAseq are available in dbGAP: accession # phs003676.v1.p1 (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003676.v1.p1).

Code Availability

The code used to analyze the WGS from Tempo is available at <https://github.com/mskcc/tempo>. The code for the reconstruction of complex genomic events can be found at <https://github.com/parklab/focal-amplification>. The code for mutational signature analysis can be found at <https://github.com/mskcc/DeepSig/>. The pipeline for the transcriptome analysis can be found at <https://github.com/mskcc/forte>.

REFERENCES

1. Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. *Nat Rev Dis Primers* 2021;7:3.
2. George J, Lim JS, Jang SJ, Cun Y, Ozretić L, Kong G, *et al.* Comprehensive genomic profiles of small cell lung cancer. *Nature* 2015;524:47-53.
3. Sivakumar S, Moore JA, Montesion M, Sharaf R, Lin DI, Colón CI, *et al.* Integrative Analysis of a Large Real-World Cohort of Small Cell Lung Cancer Identifies Distinct Genetic Subtypes and Insights into Histologic Transformation. *Cancer Discov* 2023;13:1572-91.
4. Febres-Aldana CA, Chang JC, Ptashkin R, Wang Y, Gedvilaite E, Baine MK, *et al.* Rb Tumor Suppressor in Small Cell Lung Cancer: Combined Genomic and IHC Analysis with a Description of a Distinct Rb-Proficient Subset. *Clin Cancer Res* 2022;28:4702-13.
5. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* 2017;23:703-13.
6. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, *et al.* Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J Mol Diagn* 2015;17:251-64.
7. Varghese AM, Zakowski MF, Yu HA, Won HH, Riely GJ, Krug LM, *et al.* Small-cell lung cancers in patients who never smoked cigarettes. *J Thorac Oncol* 2014;9:892-6.
8. Marcoux N, Gettinger SN, O'Kane G, Arbour KC, Neal JW, Husain H, *et al.* EGFR-Mutant Adenocarcinomas That Transform to Small-Cell Lung Cancer and Other Neuroendocrine Carcinomas: Clinical Outcomes. *J Clin Oncol* 2019;37:278-85.
9. Quintanal-Villalonga A, Taniguchi H, Zhan YA, Hasan MM, Chavan SS, Meng F, *et al.* Multiomic Analysis of Lung Tumors Defines Pathways Activated in Neuroendocrine Transformation. *Cancer Discov* 2021;11:3028-47.
10. Lee JK, Lee J, Kim S, Kim S, Youk J, Park S, *et al.* Clonal History and Genetic Predictors of Transformation Into Small-Cell Carcinomas From Lung Adenocarcinomas. *J Clin Oncol* 2017;35:3065-74.
11. Caplin ME, Baudin E, Ferolla P, Filosso P, Garcia-Yuste M, Lim E, *et al.* Pulmonary neuroendocrine (carcinoid) tumors: European Neuroendocrine Tumor Society expert consensus and recommendations for best practice for typical and atypical pulmonary carcinoids. *Ann Oncol* 2015;26:1604-20.
12. Rekhtman N. Lung neuroendocrine neoplasms: recent progress and persistent challenges. *Mod Pathol* 2022;35:36-50.
13. Kawasaki K, Rekhtman N, Quintanal-Villalonga Á, Rudin CM. Neuroendocrine neoplasms of the lung and gastrointestinal system: convergent biology and a path to better therapies. *Nat Rev Clin Oncol* 2023;20:16-32.
14. Rekhtman N, Desmeules P, Litvak AM, Pietanza MC, Santos-Zabala ML, Ni A, *et al.* Stage IV lung carcinoids: spectrum and evolution of proliferation rate, focusing on variants with elevated proliferation indices. *Mod Pathol* 2019;32:1106-22.
15. Yang SR, Gedvilaite E, Ptashkin R, Chang J, Ziegler J, Mata DA, *et al.* Microsatellite Instability and Mismatch Repair Deficiency Define a Distinct Subset of Lung Cancers Characterized by Smoking Exposure, High Tumor Mutational Burden, and Recurrent Somatic MLH1 Inactivation. *J Thorac Oncol* 2023;19:409-24.

16. Zhang CZ, Spektor A, Cornils H, Francis JM, Jackson EK, Liu S, *et al.* Chromothripsis from DNA damage in micronuclei. *Nature* 2015;522:179-84.
17. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 2011;144:27-40.
18. Cortés-Ciriano I, Lee JJ, Xi R, Jain D, Jung YL, Yang L, *et al.* Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nat Genet* 2020;52:331-41.
19. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* 2017;1:1-16.
20. Jin Y, Xiao T, Feng Y, Yang J, Guo C, Hu L, Ji H. A mesenchymal-like subpopulation in non-neuroendocrine SCLC contributes to metastasis. *J Genet Genomics* 2021;48:571-81.
21. Wistuba, II, Gazdar AF, Minna JD. Molecular genetics of small cell lung carcinoma. *Semin Oncol* 2001;28:3-13.
22. Northcott PA, Shih DJ, Peacock J, Garzia L, Morrissy AS, Zichner T, *et al.* Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* 2012;488:49-56.
23. Parker M, Mohankumar KM, Punchihewa C, Weinlich R, Dalton JD, Li Y, *et al.* C11orf95-RELA fusions drive oncogenic NF- κ B signalling in ependymoma. *Nature* 2014;506:451-5.
24. Lee JJ, Park S, Park H, Kim S, Lee J, Lee J, *et al.* Tracing Oncogene Rearrangements in the Mutational History of Lung Adenocarcinoma. *Cell* 2019;177:1842-57.e21.
25. Xia J, Li S, Liu S, Zhang L. Aldehyde dehydrogenase in solid tumors and other diseases: Potential biomarkers and therapeutic targets. *MedComm (2020) 2023*;4:e195.
26. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580-5.
27. Boix CA, James BT, Park YP, Meuleman W, Kellis M. Regulatory genomic circuitry of human disease loci by integrative epigenomics. *Nature* 2021;590:300-7.
28. Northcott PA, Lee C, Zichner T, Stütz AM, Erkek S, Kawachi D, *et al.* Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* 2014;511:428-34.
29. Zhang Y, Chen F, Fonseca NA, He Y, Fujita M, Nakagawa H, *et al.* High-coverage whole-genome analysis of 1220 cancers reveals hundreds of genes deregulated by rearrangement-mediated cis-regulatory alterations. *Nat Commun* 2020;11:736.
30. Ly P, Brunner SF, Shoshani O, Kim DH, Lan W, Pyntikova T, *et al.* Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements. *Nat Genet* 2019;51:705-15.
31. Shoshani O, Brunner SF, Yaeger R, Ly P, Nechemia-Arbely Y, Kim DH, *et al.* Chromothripsis drives the evolution of gene amplification in cancer. *Nature* 2021;591:137-41.
32. Ratnaparkhe M, Hlevnjak M, Kolb T, Jauch A, Maass KK, Devens F, *et al.* Genomic profiling of Acute lymphoblastic leukemia in ataxia telangiectasia patients reveals tight link between ATM mutations and chromothripsis. *Leukemia* 2017;31:2048-56.
33. Rausch T, Jones DT, Zapatka M, Stütz AM, Zichner T, Weischenfeldt J, *et al.* Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. *Cell* 2012;148:59-71.
34. Voronina N, Wong JKL, Hübschmann D, Hlevnjak M, Uhrig S, Heilig CE, *et al.* The landscape of chromothripsis across adult cancer types. *Nat Commun* 2020;11:2320.
35. Simovic M, Ernst A. Chromothripsis, DNA repair and checkpoints defects. *Semin Cell Dev Biol* 2022;123:110-4.

36. Westbrook AM, Schiestl RH. Atm-deficient mice exhibit increased sensitivity to dextran sulfate sodium-induced colitis characterized by elevated DNA damage and persistent immune activation. *Cancer Res* 2010;70:1875-84.
37. Schütze DM, Krijgsman O, Snijders PJ, Ylstra B, Weischenfeldt J, Mardin BR, *et al.* Immortalization capacity of HPV types is inversely related to chromosomal instability. *Oncotarget* 2016;7:37608-21.
38. Krupina K, Goginashvili A, Cleveland DW. Scrambling the genome in cancer: causes and consequences of complex chromosome rearrangements. *Nat Rev Genet* 2023;25:196-210.
39. Vanderbilt CM, Bowman AS, Middha S, Petrova-Drus K, Tang YW, Chen X, *et al.* Defining Novel DNA Virus-Tumor Associations and Genomic Correlates Using Prospective Clinical Tumor/Normal Matched Sequencing Data. *J Mol Diagn* 2022;24:515-28.
40. Ernst A, Jones DT, Maass KK, Rode A, Deeg KI, Jebaraj BM, *et al.* Telomere dysfunction and chromothripsis. *Int J Cancer* 2016;138:2905-14.
41. Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, *et al.* The evolutionary history of 2,658 cancers. *Nature* 2020;578:122-8.
42. Offin M, Chan JM, Tenet M, Rizvi HA, Shen R, Riely GJ, *et al.* Concurrent RB1 and TP53 Alterations Define a Subset of EGFR-Mutant Lung Cancers at risk for Histologic Transformation and Inferior Clinical Outcomes. *J Thorac Oncol* 2019;14:1784-93.
43. Dick FA, Goodrich DW, Sage J, Dyson NJ. Non-canonical functions of the RB protein in cancer. *Nat Rev Cancer* 2018;18:442-51.
44. Kobayashi A, Sakuma T, Fujimoto M, Jimbo N, Hirose T. Diagnostic Utility and Limitations of Immunohistochemistry of p16, CDK4, and MDM2 and Automated Dual-color In Situ Hybridization of MDM2 for the Diagnosis of Challenging Cases of Dedifferentiated Liposarcoma. *Appl Immunohistochem Mol Morphol* 2019;27:758-63.
45. Karpathakis A, Dibra H, Thirlwell C. Neuroendocrine tumours: cracking the epigenetic code. *Endocr Relat Cancer* 2013;20:R65-82.
46. Moonen L, Mangiante L, Leunissen DJG, Lap LMV, Gabriel A, Hillen LM, *et al.* Differential Orthopedia Homeobox expression in pulmonary carcinoids is associated with changes in DNA methylation. *Int J Cancer* 2022;150:1987-97.
47. Kilpatrick SE. Dedifferentiated Liposarcoma: A Comprehensive Historical Review With Proposed Evidence-based Guidelines Regarding a Diagnosis in Need of Further Clarification. *Adv Anat Pathol* 2021;28:426-38.
48. Tatematsu A, Shimizu J, Murakami Y, Horio Y, Nakamura S, Hida T, *et al.* Epidermal growth factor receptor mutations in small cell lung cancer. *Clin Cancer Res* 2008;14:6092-6.
49. Thomas A, Mian I, Tlemsani C, Pongor L, Takahashi N, Maignan K, *et al.* Clinical and Genomic Characteristics of Small Cell Lung Cancer in Never Smokers: Results From a Retrospective Multicenter Cohort Study. *Chest* 2020;158:1723-33.
50. 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;19:289-97.
51. Alcalá N, Leblay N, Gabriel AAG, Mangiante L, Hervas D, Giffon T, *et al.* Integrative and comparative genomic analyses identify clinically relevant pulmonary carcinoid groups and unveil the supra-carcinoids. *Nat Commun* 2019;10:3407.
52. Baine MK, Hsieh MS, Lai WV, Egger JV, Jungbluth AA, Daneshbod Y, *et al.* SCLC Subtypes Defined by ASCL1, NEUROD1, POU2F3, and YAP1: A Comprehensive Immunohistochemical and Histopathologic Characterization. *J Thorac Oncol* 2020;15:1823-35.
53. Lou F, Sarkaria I, Pietanza C, Travis W, Roh MS, Sica G, *et al.* Recurrence of pulmonary carcinoid tumors after resection: implications for postoperative surveillance. *Ann Thorac Surg* 2013;96:1156-62.

54. Zauderer MG, Drilon A, Kadota K, Huberman K, Sima CS, Bergagnini I, *et al.* Trial of a 5-day dosing regimen of temozolomide in patients with relapsed small cell lung cancers with assessment of methylguanine-DNA methyltransferase. *Lung Cancer* 2014;86:237-40.
55. Rudin CM, Reck M, Johnson ML, Blackhall F, Hann CL, Yang JC, *et al.* Emerging therapies targeting the delta-like ligand 3 (DLL3) in small cell lung cancer. *J Hematol Oncol* 2023;16:66.
56. Wiedemeyer WR, Gavriilyuk J, Schammel A, Zhao X, Sarvaiya H, Pysz M, *et al.* ABBV-011, A Novel, Calicheamicin-Based Antibody-Drug Conjugate, Targets SEZ6 to Eradicate Small Cell Lung Cancer Tumors. *Mol Cancer Ther* 2022;21:986-98.
57. Kim JW, Ko JH, Sage J. DLL3 regulates Notch signaling in small cell lung cancer. *iScience* 2022;25:105603.
58. Rosswog C, Bartenhagen C, Welte A, Kahlert Y, Hemstedt N, Lorenz W, *et al.* Chromothripsis followed by circular recombination drives oncogene amplification in human cancer. *Nat Genet* 2021;53:1673-85.
59. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;483:603-7.
60. Pongor LS, Schultz CW, Rinaldi L, Wangsa D, Redon CE, Takahashi N, *et al.* Extrachromosomal DNA Amplification Contributes to Small Cell Lung Cancer Heterogeneity and Is Associated with Worse Outcomes. *Cancer Discov* 2023;13:928-49.
61. Choudhuri SP, Girard L, Lim JYS, Wise JF, Freitas B, Yang D, *et al.* Acquired Cross-resistance in Small Cell Lung Cancer due to Extrachromosomal DNA Amplification of MYC paralogs. *Cancer Discov* 2023;14:804-27.
62. Maclachlan KH, Rustad EH, Derkach A, Zheng-Lin B, Yellapantula V, Diamond B, *et al.* Copy number signatures predict chromothripsis and clinical outcomes in newly diagnosed multiple myeloma. *Nat Commun* 2021;12:5172.
63. Steele CD, Abbasi A, Islam SMA, Bowes AL, Khandekar A, Haase K, *et al.* Signatures of copy number alterations in human cancer. *Nature* 2022;606:984-91.
64. Quintanal-Villalonga Á, Chan JM, Yu HA, Pe'er D, Sawyers CL, Sen T, Rudin CM. Lineage plasticity in cancer: a shared pathway of therapeutic resistance. *Nat Rev Clin Oncol* 2020;17:360-71.
65. Balanis NG, Sheu KM, Esedebe FN, Patel SJ, Smith BA, Park JW, *et al.* Pan-cancer Convergence to a Small-Cell Neuroendocrine Phenotype that Shares Susceptibilities with Hematological Malignancies. *Cancer Cell* 2019;36:17-34.e7.
66. Pelosi G, Melocchi V, Dama E, Hofman P, De Luca M, Albin A, *et al.* An in-silico analysis reveals further evidence of an aggressive subset of lung carcinoids sharing molecular features of high-grade neuroendocrine neoplasms. *Exp Mol Pathol* 2024;135:104882.
67. Werr L, Bartenhagen C, Rosswog C, Cartolano M, Voegelé C, Sexton-Oates A, *et al.* TERT expression defines clinical outcome in pulmonary carcinoids. *J Clin Oncol In Press* 2024.
68. Fernandez-Cuesta L, Peifer M, Lu X, Sun R, Ozretić L, Seidal D, *et al.* Frequent mutations in chromatin-remodelling genes in pulmonary carcinoids. *Nat Commun* 2014;5:3518.
69. Robb TJ, Tsai P, Fitzgerald S, Shields P, Houseman PS, Patel R, *et al.* Complex Patterns of Genomic Heterogeneity Identified in 42 Tumor Samples and ctDNA of a Pulmonary Atypical Carcinoid Patient. *Cancer Res Commun* 2023;3:31-42.
70. Chowdhry S, Milutinovic S, Tse E, Garcia S, Perusse D, Ritland M, *et al.* Abstract 1626: Tumors driven by oncogene amplified extrachromosomal DNA (ecDNA) demonstrate enhanced sensitivity to cell cycle checkpoint kinase 1 (CHK1) inhibition. *Cancer Research* 2023;83:1626-.
71. Knudsen ES, Shapiro GI, Keyomarsi K. Selective CDK4/6 Inhibitors: Biologic Outcomes, Determinants of Sensitivity, Mechanisms of Resistance, Combinatorial Approaches, and Pharmacodynamic Biomarkers. *Am Soc Clin Oncol Educ Book* 2020;40:115-26.

72. Zhu H, Gao H, Ji Y, Zhou Q, Du Z, Tian L, *et al.* Targeting p53-MDM2 interaction by small-molecule inhibitors: learning from MDM2 inhibitors in clinical trials. *J Hematol Oncol* 2022;15:91.
73. Uprety D, Halfdanarson TR, Molina JR, Leventakos K. Pulmonary Neuroendocrine Tumors: Adjuvant and Systemic Treatments. *Curr Treat Options Oncol* 2020;21:86.
74. Tsao MS, Nicholson AG, Maleszewski JJ, Marx A, Travis WD. Introduction to 2021 WHO Classification of Thoracic Tumors. *J Thorac Oncol* 2022;17:e1-e4.
75. Naheed S, Holden C, Tanno L, Pattini L, Pearce NW, Green B, *et al.* Utility of KI-67 as a prognostic biomarker in pulmonary neuroendocrine neoplasms: a systematic review and meta-analysis. *BMJ Open* 2022;12:e041961.
76. Suehnholz SP, Nissan MH, Zhang H, Kundra R, Nandakumar S, Lu C, *et al.* Quantifying the Expanding Landscape of Clinical Actionability for Patients with Cancer. *Cancer Discov* 2024;14:49-65.
77. Middha S, Zhang L, Nafa K, Jayakumaran G, Wong D, Kim HR, *et al.* Reliable Pan-Cancer Microsatellite Instability Assessment by Using Targeted Next-Generation Sequencing Data. *JCO Precis Oncol* 2017;1:1-17.
78. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci Transl Med* 2015;7:283ra54.
79. Dentre SC, Wedge DC, Van Loo P. Principles of Reconstructing the Subclonal Architecture of Cancers. *Cold Spring Harb Perspect Med* 2017;7.
80. Jimbo N, Ohbayashi C, Fujii T, Takeda M, Mitsui S, Tsukamoto R, *et al.* Implication of cytoplasmic p53 expression in pulmonary neuroendocrine carcinoma using next-generation sequencing analysis. *Histopathology* 2024;84:336-42.
81. Köbel M, Ronnett BM, Singh N, Soslow RA, Gilks CB, McCluggage WG. Interpretation of P53 Immunohistochemistry in Endometrial Carcinomas: Toward Increased Reproducibility. *Int J Gynecol Pathol* 2019;38 Suppl 1:S123-s31.
82. Cheng DT, Prasad M, Chekaluk Y, Benayed R, Sadowska J, Zehir A, *et al.* Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. *BMC Med Genomics* 2017;10:33.
83. Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res* 2016;44:e131.
84. Srinivasan P, Bandlamudi C, Jonsson P, Kemel Y, Chavan SS, Richards AL, *et al.* The context-specific role of germline pathogenicity in tumorigenesis. *Nat Genet* 2021;53:1577-85.
85. Mandelker D, Zhang L, Kemel Y, Stadler ZK, Joseph V, Zehir A, *et al.* Mutation Detection in Patients With Advanced Cancer by Universal Sequencing of Cancer-Related Genes in Tumor and Normal DNA vs Guideline-Based Germline Testing. *Jama* 2017;318:825-35.
86. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, *et al.* Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 2016;32:1220-2.
87. Wala JA, Bandopadhyay P, Greenwald NF, O'Rourke R, Sharpe T, Stewart C, *et al.* SvABA: genome-wide detection of structural variants and indels by local assembly. *Genome Res* 2018;28:581-91.
88. Feuerbach L, Sieverling L, Deeg KI, Ginsbach P, Hutter B, Buchhalter I, *et al.* TelomereHunter - in silico estimation of telomere content and composition from cancer genomes. *BMC Bioinformatics* 2019;20:272.

89. Meyer-Lehnert H, Tsai P, Schrier RW. Atrial natriuretic factor (ANF) inhibits arginine vasopressin-stimulated Ca²⁺ fluxes and cell contraction in vascular smooth muscle cells. *Klin Wochenschr* 1987;65 Suppl 8:115-21.
90. Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR, Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics* 2018;34:511-3.
91. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell* 2019;176:535-48.e24.
92. Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, *et al.* Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 2019;575:210-6.
93. Cameron DL, Baber J, Shale C, Valle-Inclan JE, Besselink N, van Hoeck A, *et al.* GRIDSS2: comprehensive characterisation of somatic structural variation using single breakend variants and structural variant phasing. *Genome Biol* 2021;22:202.
94. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 2019;20:257.
95. Shale C, Cameron DL, Baber J, Wong M, Cowley MJ, Papenfuss AT, *et al.* Unscrambling cancer genomes via integrated analysis of structural variation and copy number. *Cell Genom* 2022;2:100112.
96. Lee JJ, Jung YL, Cheong TC, Espejo Valle-Inclan J, Chu C, Gulhan DC, *et al.* ER α -associated translocations underlie oncogene amplifications in breast cancer. *Nature* 2023;618:1024-32.
97. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525-7.
98. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods* 2017;14:687-90.
99. Edgren H, Murumagi A, Kangaspeska S, Nicorici D, Hongisto V, Kleivi K, *et al.* Identification of fusion genes in breast cancer by paired-end RNA-sequencing. *Genome Biol* 2011;12:R6.
100. Hännelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
101. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
102. Caesar R, Egger JV, Chavan S, Socci ND, Jones CB, Kombak FE, *et al.* Genomic and transcriptomic analysis of a library of small cell lung cancer patient-derived xenografts. *Nat Commun* 2022;13:2144.

MAIN FIGURES LEGENDS

Figure 1. Cohort selection and clinicopathologic characteristics. **A**, Schematic diagram of *RB1* and *TP53* inactivation (outer doughnut) and smoking status (inner doughnut) in a cohort of 600 consecutive SCLC that underwent targeted NGS by MSK-IMPACT. *RB1* and *TP53* status was determined based on integrated genomic and IHC data (see Results). **B**, Tabular and **C**, pictorial summary of pathologic sample characteristics and sequencing assays performed for 20 aSCLC patients. See Methods for histopathologic criteria. FNA fine needle aspiration, Core bx core biopsy, ? Ki67 not available. **D**, Summary of metastatic site distribution based on radiologic-pathologic findings (full data in **Supplementary Table S2**). **E, F, G**, Compilation of radiologic and pathologic findings for three representative patients. **E**, (A01): 19-year-old patient with brain metastasis showing classic undifferentiated morphology of SCLC: small, crowded/molding cells, numerous mitotic figures ($>50/2 \text{ mm}^2$), apoptotic bodies, and Ki67 of 90%, whereas thoracic sample shows well-differentiated, nested morphology of a carcinoid tumor lacking mitotic figures or apoptotic bodies, and exhibiting Ki67 of 5-20%. **F**, Case A08 and **G**, case A09 illustrate classic SCLC morphology in primary lung tumor samples, featuring Ki67 rate of 80% and 90%, respectively. n indicates areas of extensive necrosis – a hallmark feature of SCLC. For extended pathologic illustrations see **Supplementary Figures S1-S4**. Scale bars are 100 μm for Ki67, 50 μm for H&E except top panel in F which is 1 mm. **H**, Comparison of age, Ki67 proliferation index and smoking pack-years in aSCLC vs other SCLC ($n=224$ for age and Ki67; $n=200$ for pack-year smoking). For patients with multiple samples, Ki67 represents the maximal (max) hot-spot rate among all evaluated samples.

Figure 2. Mutation and copy number alteration (CNA) landscape of atypical SCLC. **A**, Mutations and CNAs in aSCLC in comparison to the control groups of smoking-associated SCLC (sSCLC; $n=206$) and pulmonary carcinoids ($n=157$) analyzed by MSK-IMPACT. OncoPrint summarizing assays performed and overall genomic features: microsatellite instability (MSI) vs microsatellite stability (MSS), tumor mutation burden (TMB), and genomic signature (full signature analysis shown in **Supplementary Figure S5**). Case A09 lacked a matched normal DNA sample, and was therefore excluded from TMB, MSI and signature analysis. Displayed genomic alterations include selected recurrently altered genes (see **Supplementary Tables S5 and S6** for a full list of detected alterations). For patients with multiple samples, the index sample is displayed (see Methods). $*p<0.05$; $**p<0.01$, $***p<0.001$, $****p<0.0001$. Abbreviations: tNGS targeted NGS (MSK-IMPACT), TMZ temozolomide, N/A not available, MMR mismatch repair. **B**, Expression of recurrently amplified genes by immunohistochemistry (IHC) in two representative cases (scale bars: 100 μm) and **C**, by RNAseq in cases with available data. Arrows in **C** indicate cases harboring corresponding amplifications. TPM: Transcripts Per Million.

Figure 3. Chromothripsis and corresponding cancer gene amplifications and losses. **A**, Representative Circos plots of SVs and CNAs across the genome by WGS. Outer band shows an ideogram of chromosome positions and cytogenetic bands. Second band depicts total copy number and third band shows minor allele copy number. The inner circle depicts SVs as arcs connecting the two relevant genomic points as identified by three algorithms (see Methods). CNAs in key cancer genes in the regions of chromothripsis (red, amplifications; blue, deletions) are displayed. Circos plots for all cases are shown in **Supplementary Figure S6**. **B**, Copy number log ratio plots from the FACETS algorithm displaying the distinctive oscillating copy number states on chromosomes with chromothripsis. Copy number segments are shown in red. Focal segments ($<2\text{MB}$ in size) are shown as enlarged points for visual purposes. Selected amplifications are indicated (yellow). **C**, Summary of chromosomal location of chromothriptic events in cases analyzed by WGS and targeted NGS (tNGS). Also shown are selected amplifications and losses in oncogenes and tumor suppressors, respectively, localized to the chromothriptic chromosomes. Full list is provided in **Supplementary Table S7**. Abbreviations: ChrT chromothripsis. **D**, Schematic

summary for the rate of major genomic mechanisms detected in the set analyzed by WGS (n=11) and in the full cohort (n=20). In the lower diagram, major chromosomes involved by chromothripsis are indicated in the inner doughnut, and corresponding recurrent gene amplifications are indicated in the outer doughnut. **E**, Total number of SVs identified in samples analyzed by WGS. Variants are color coded by type. **F**, Number of fusions predicted in samples with available RNAseq. **G**, Diagram illustrating putative enhancer hijacking in case A17 with chromothripsis on chromosome 2 resulting in translocation between *SH3RF3* on chromosome 2 and upstream regulatory region of *CCND1* on chromosome 11. Epigenetic landscape surrounding the breakpoint was extrapolated from data from multiple tissue types (Epilogs search tool).

Figure 4. Chromosomal architecture of chromothripsis and amplifications. **A**, Integrated SV and CNA analysis in two representative cases confirming the hallmark features of chromothripsis and suggesting formation of circular extra-chromosomal DNA (ecDNA) and micronucleation as a mechanism of chromothripsis (see Results). Allele-specific copy number for each genomic segment is shown on the left y-axis. Genomic segments where the copy number is greater than 5 times of chromosomal baseline are indicated as orange shades to highlight the most significant amplification events. Vertical lines and arcs indicate genomic breakpoints and connections of SVs, of which the number of supporting read fragments is shown on the right y-axis. Color of the SVs are based on the orientation of paired reads. All cases are shown in **Supplementary Figure S10**. **B**, A conceptual diagram explaining the mechanism of chromothripsis and oncogene amplification in aSCLC. **C**, FISH analysis of case A07 with 3-color probes: *MDM2* (red), *CDK4* (green), and *CCND2* (orange), revealing ecDNA location of amplified genes (so-called “double-minutes”). Sample is of a metaphase spread from a patient-derived xenograft (see Methods). **D**, Summary of FISH results for all tested samples.

Figure 5. Multi-sample analysis from all temporally or spatially distinct samples obtained for each patient. **A**, Conservation of chromothripsis and corresponding amplifications across all samples with available data from individual patients. ^ indicates lower purity samples compared to the other sample from same patient (**Supplementary Table S5**); copy number values in such samples may be falsely low. * indicates chromothripsis detected by WGS only, but not detected by targeted NGS. N/A not available: sample without NGS or IHC. Sample numbering is chronological (**Supplementary Table S3**). **B**, Multi-sample analysis showing *RB1* and *TP53* genomic alterations and expression of pRb, p53 and p16 by IHC, illustrating a subset of patients with acquired or subclonal *RB1* alterations/loss (details in **Supplementary Table 1**) and a subset with unusual patterns of p16 expression (see Results). **C**, Illustration of a case with acquired *RB1* mutations and loss of expression in metastatic samples. pRb-positive cells admixed with negative tumor cells in panels on the right are benign lymphocytes and stromal cells, serving as internal controls. Scale bars are 100 um, except left lower panel, which is 1 mm. **D**, RNAseq for *CDKN2A*, encoding p16. Arrows indicate cases with chromothripsis of chromosome 12 and ecDNA amplification of *CDK4*. TPM: Transcripts Per Million.

Figure 6. Comparison of atypical SCLC (aSCLC) versus *RB1*⁻/*TP53*⁻ SCLC in never smokers (nsSCLC): the dual model of SCLC pathogenesis in never smokers. **A**, Lines of evidence for a histogenetic relationship between aSCLC and carcinoids or their progenitors, including histologic evidence of carcinoid histotype in at least one sample (see also Figure 1C and Figure 5A), expression of a lung carcinoid-specific gene OTP (Orthopedia Homeobox Protein) by immunohistochemistry (IHC), or the presence of genomic alterations characteristic of carcinoids. Lower panels illustrate expression of OTP by RNAseq and IHC. N/A: OTP IHC not available. Scale bar in OTP image: 100 um. **B**, Lines of evidence for a histogenetic relationship of nsSCLC with non-small cell lung carcinoma (NSCLC). **C**, The distribution of APOBEC mutational signature in aSCLC, nsSCLC and sSCLC. See also **Supplementary Figure S5** and Methods for

details. **D**, Expression of transcriptional subtype markers and **E**, YAP1 by IHC. A ASCL1-dominant, N NEUROD1-dominant, P POU2F3, TN triple negative. Subtype marker data was available for 15 aSCLC, 13 nsSCLC and 142 sSCLC. **F**, A conceptual diagram depicting dual pathogenetic pathways underlying SCLC in never smokers, highlighting carcinoid-SCLC pathway associated with chromothripsis versus NSCLC-SCLC pathway associated with *EGFR* mutations and APOBEC signature. Abbreviations: aSCLC atypical SCLC, nsSCLC never-smoker SCLC with *RB1*⁻/*TP53*⁻, sSCLC smoking-associated SCLC; TSG tumor suppressor genes.

Figure 7. Outcome and treatment responses. **A**, Kaplan-Meier analysis of the disease-specific overall survival assessed from the time of diagnosis. **B**, Pie chart summarizing radiologic treatment response to platinum-based chemotherapy received in any line of therapy. CR complete response, PR partial response, SD stable disease, PD progressive disease. **C**, Swimmer plot summarizing treatment modalities used and time-on-treatment, with time of pathologic samples collected and analysis performed indicated.

Figure 1

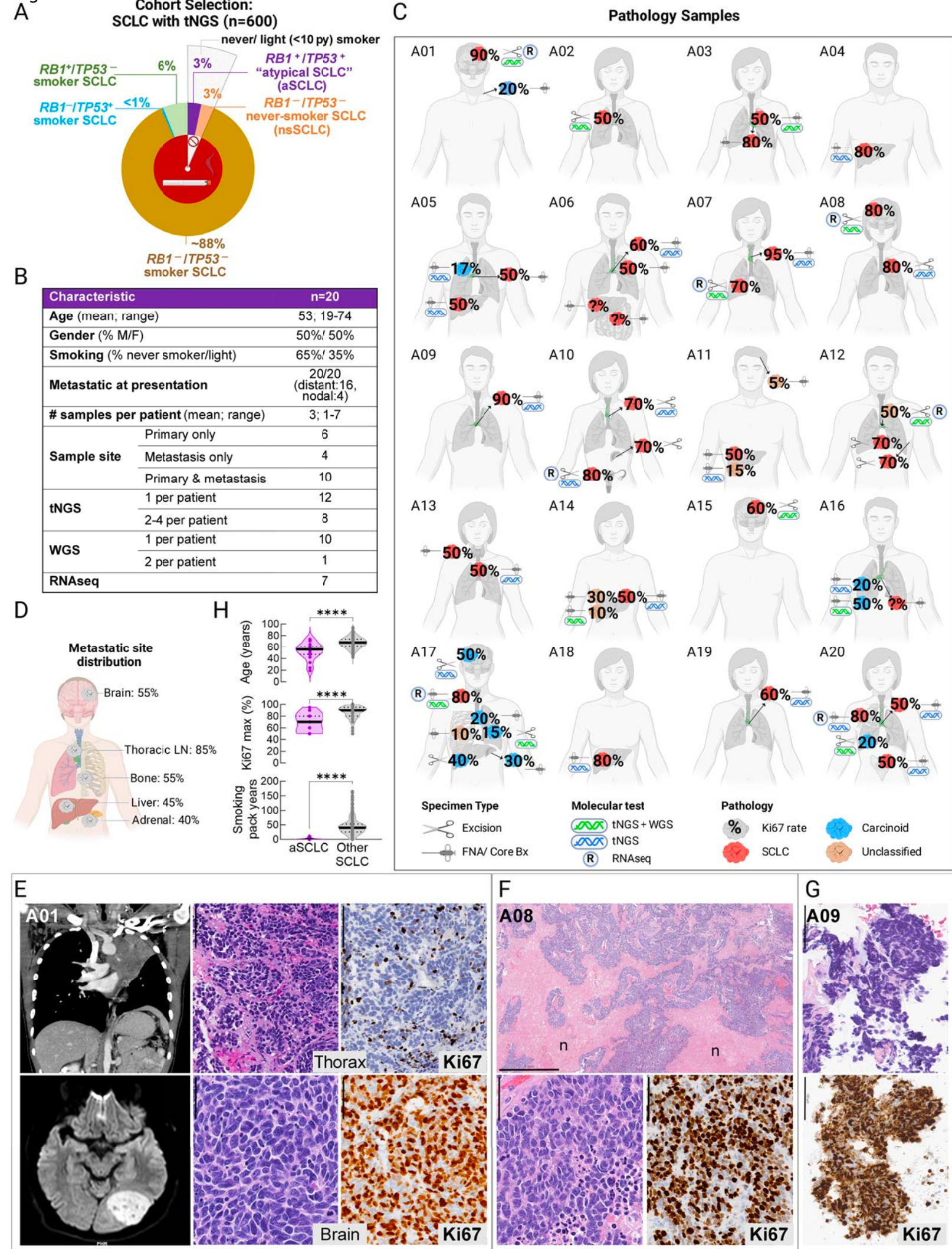


Figure 2

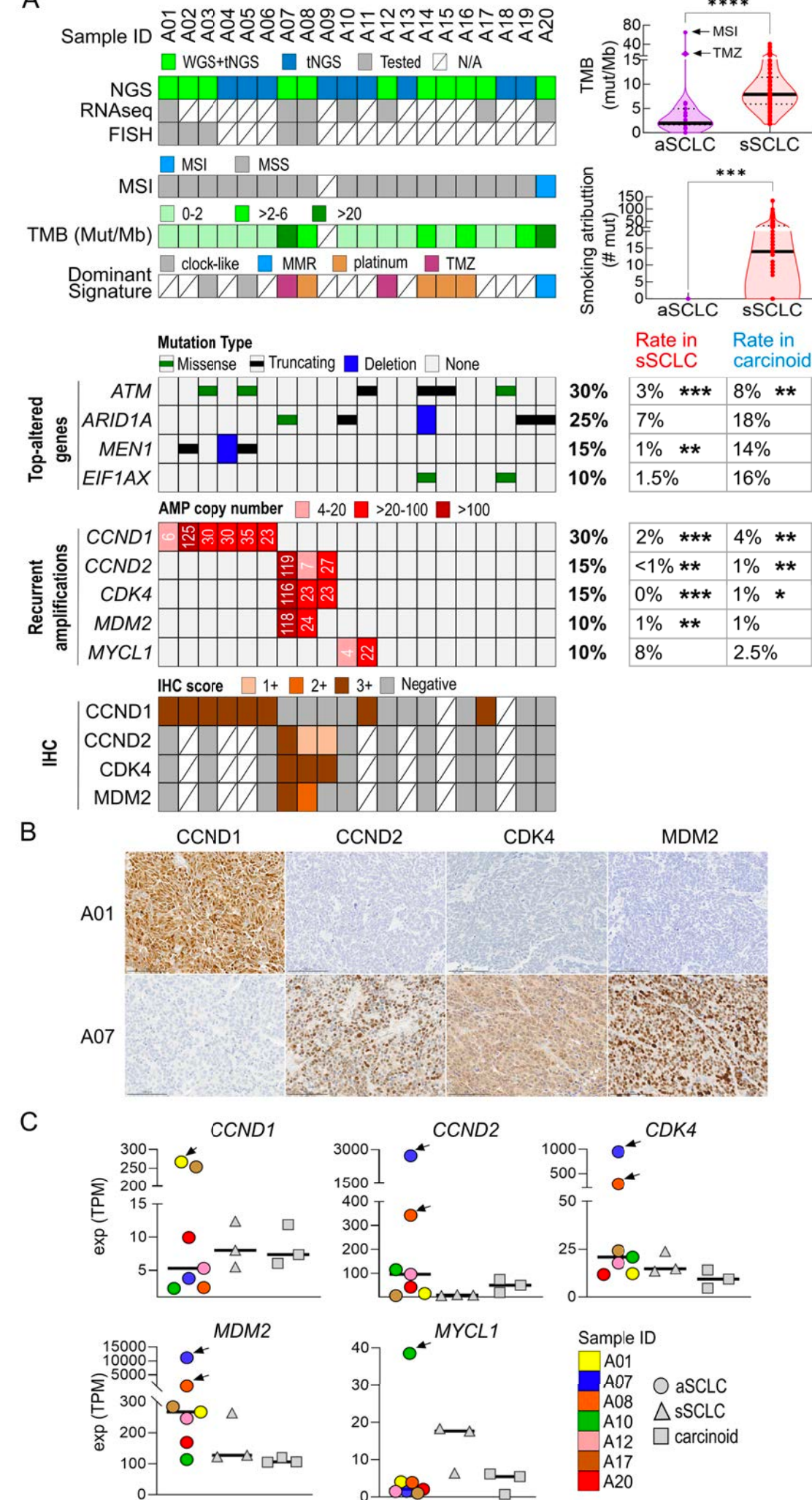


Figure 3

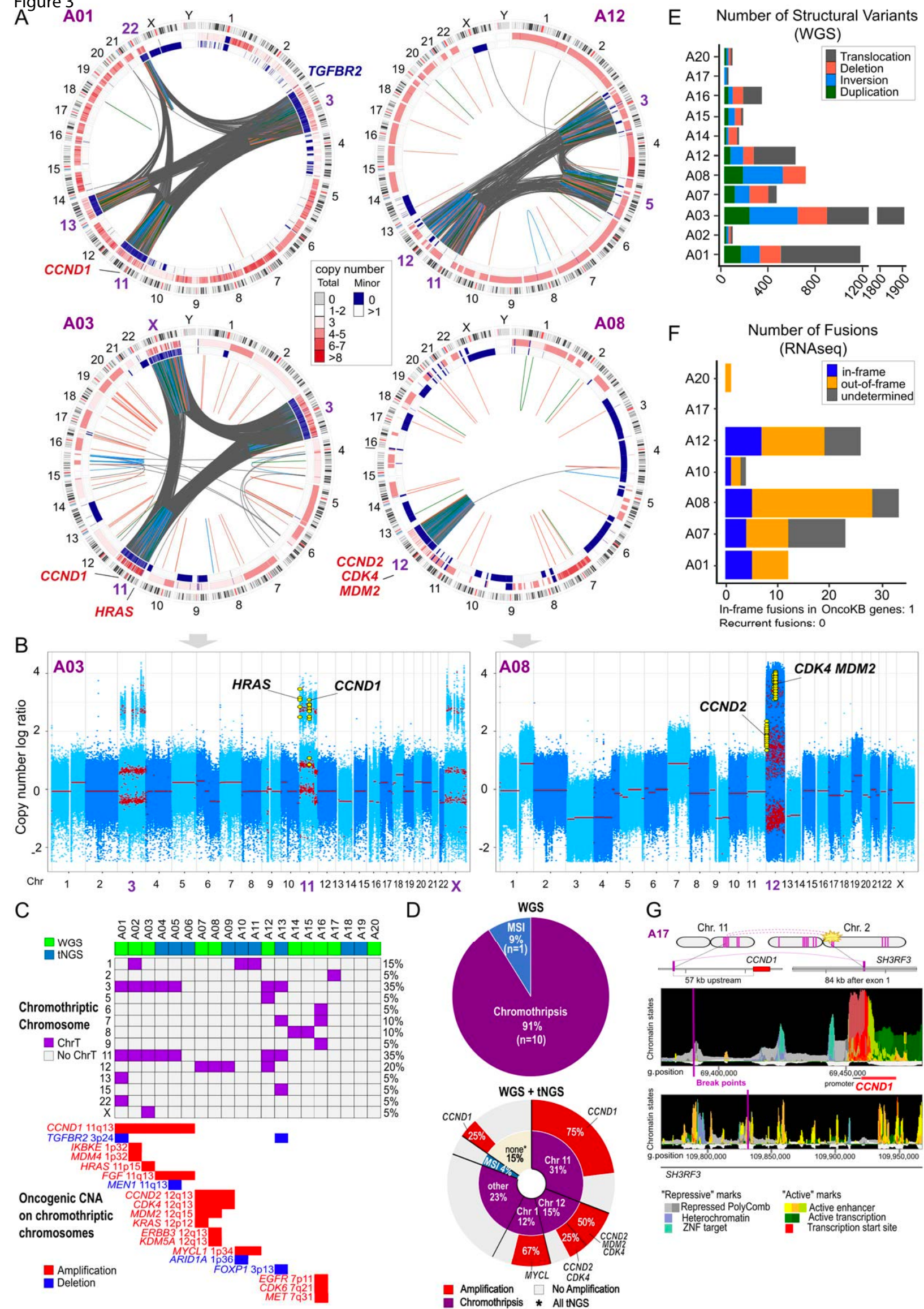


Figure 4

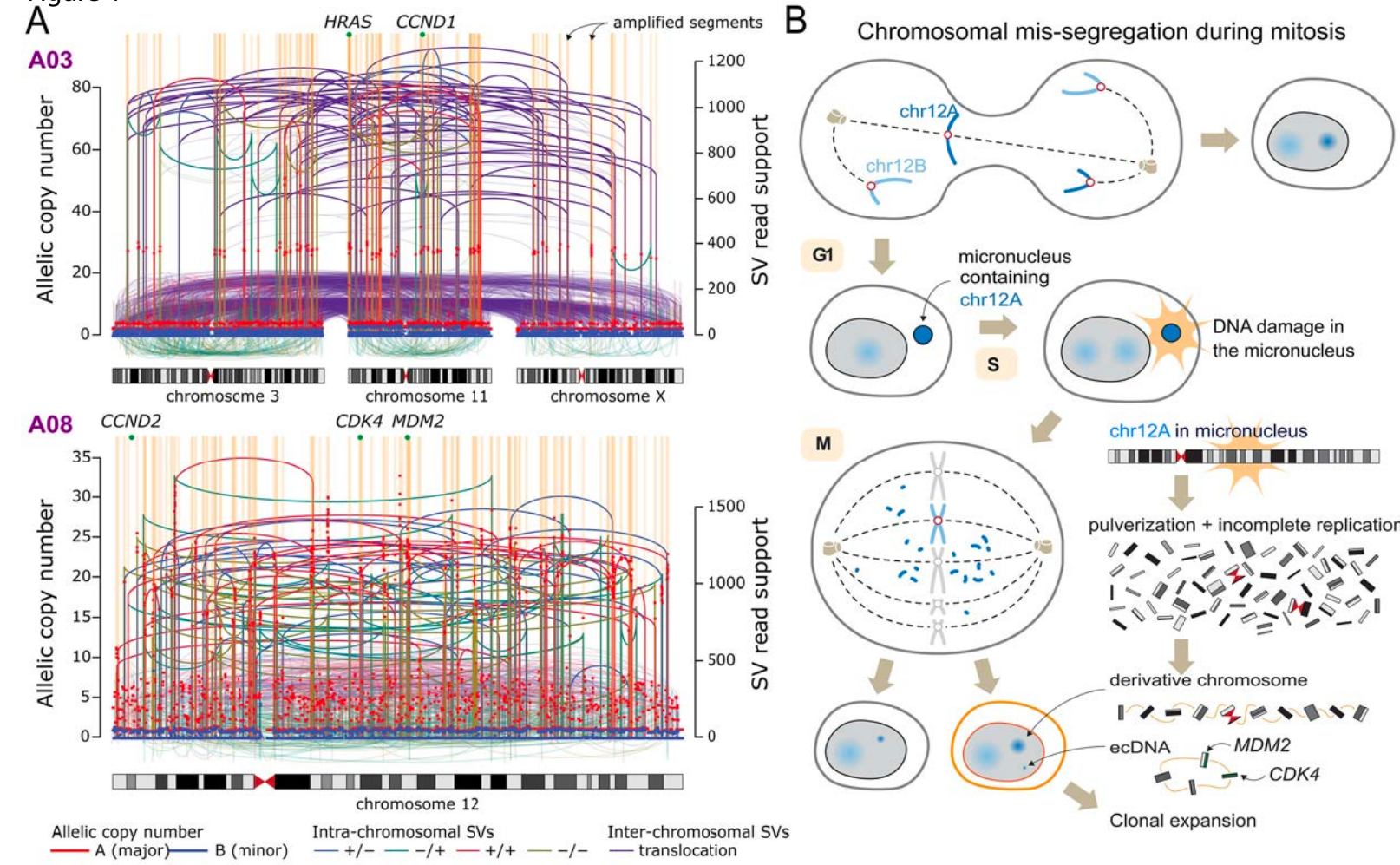
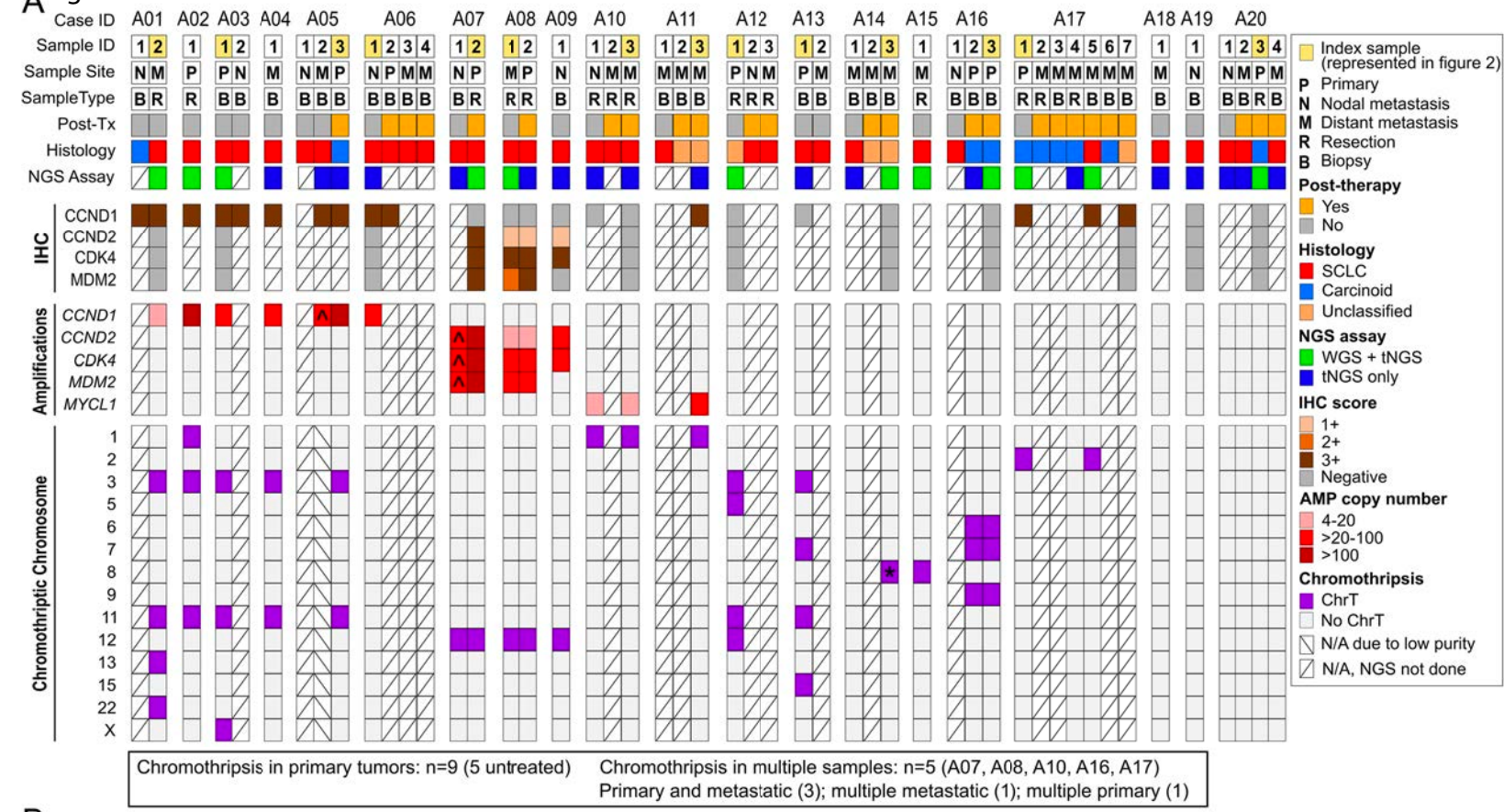
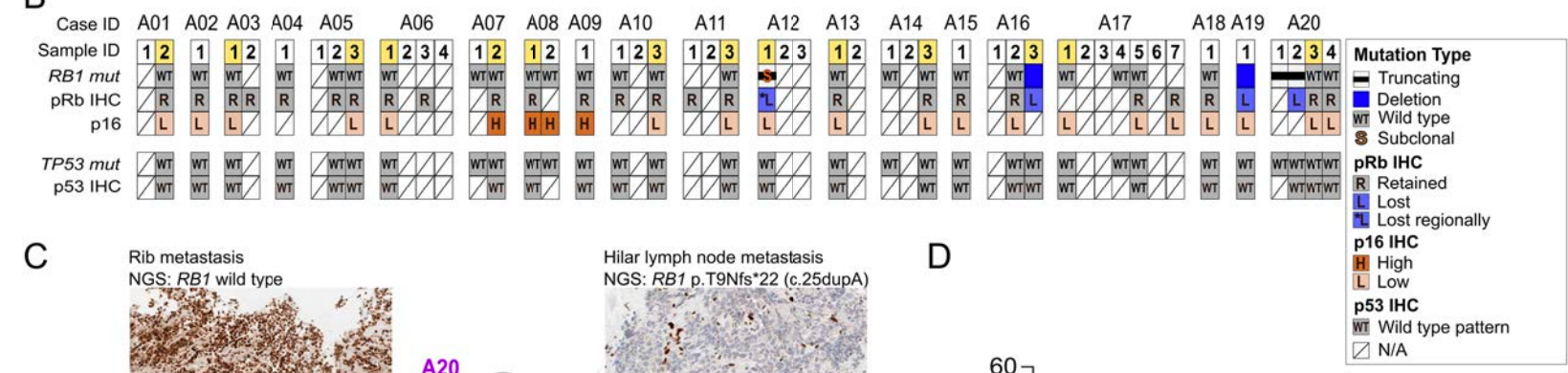


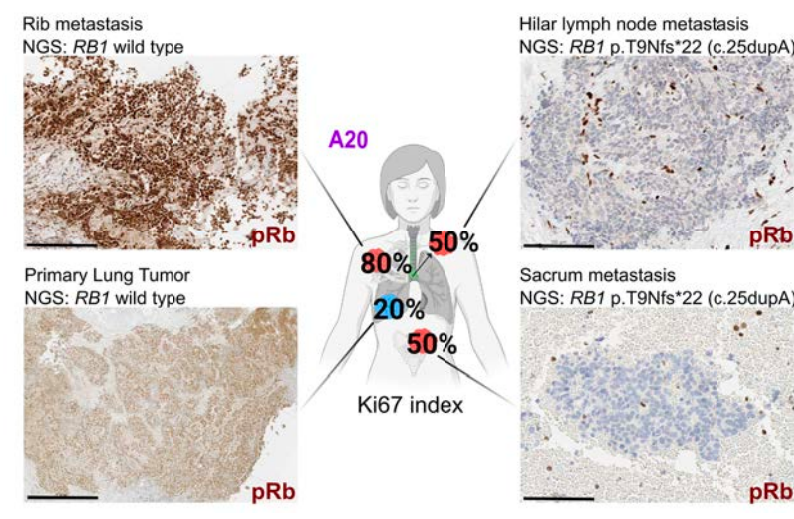
Figure 5



B



C



D

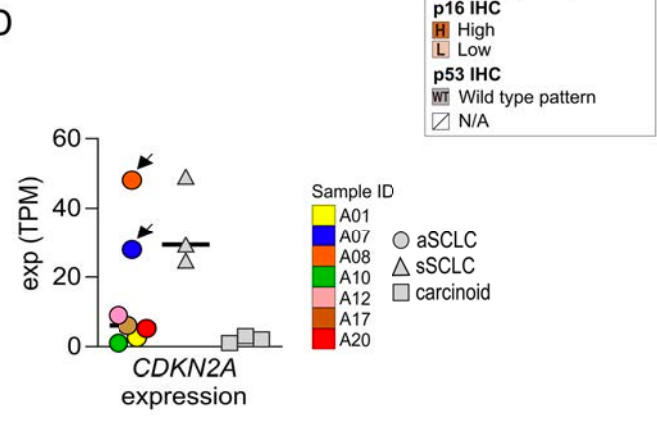
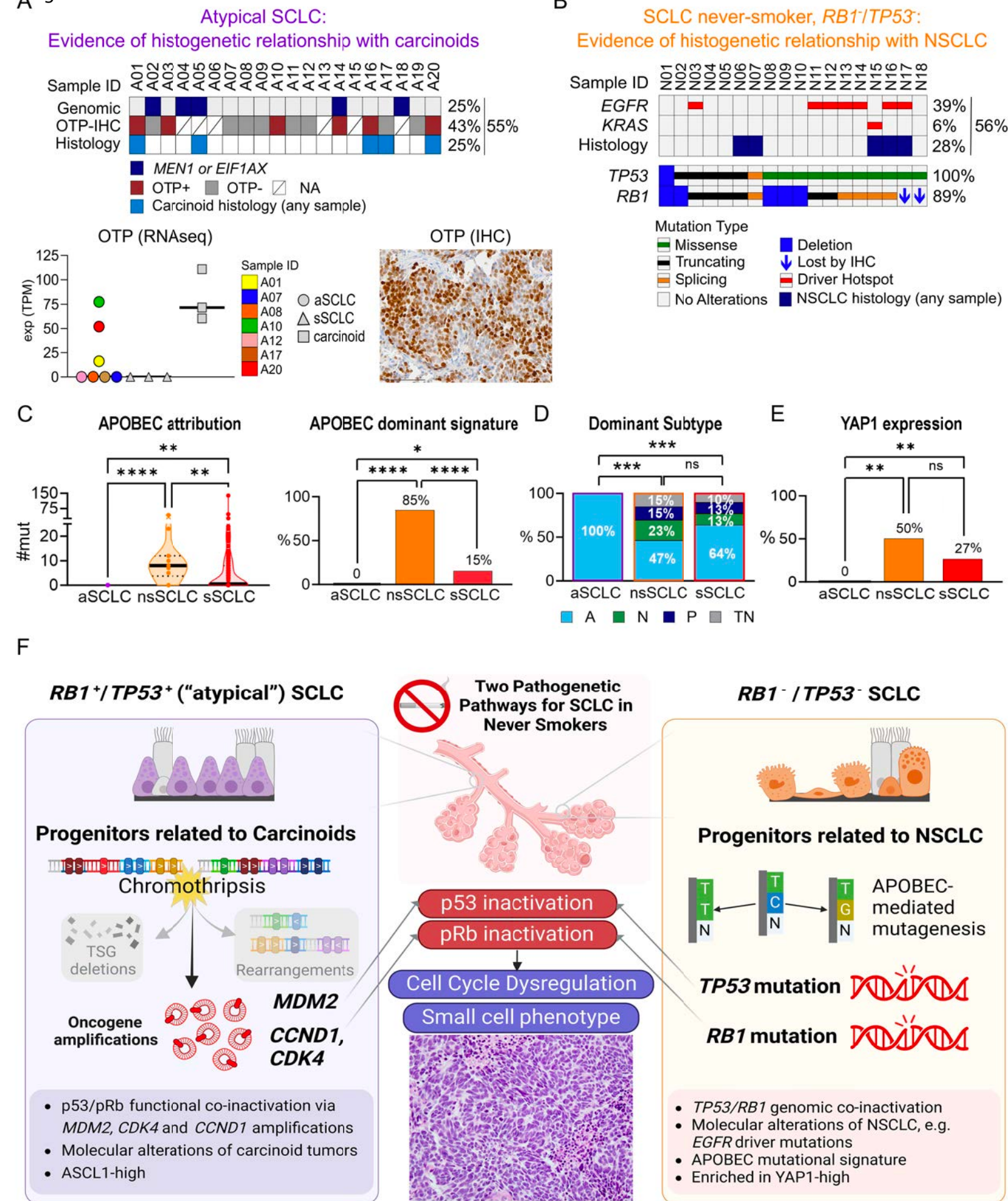
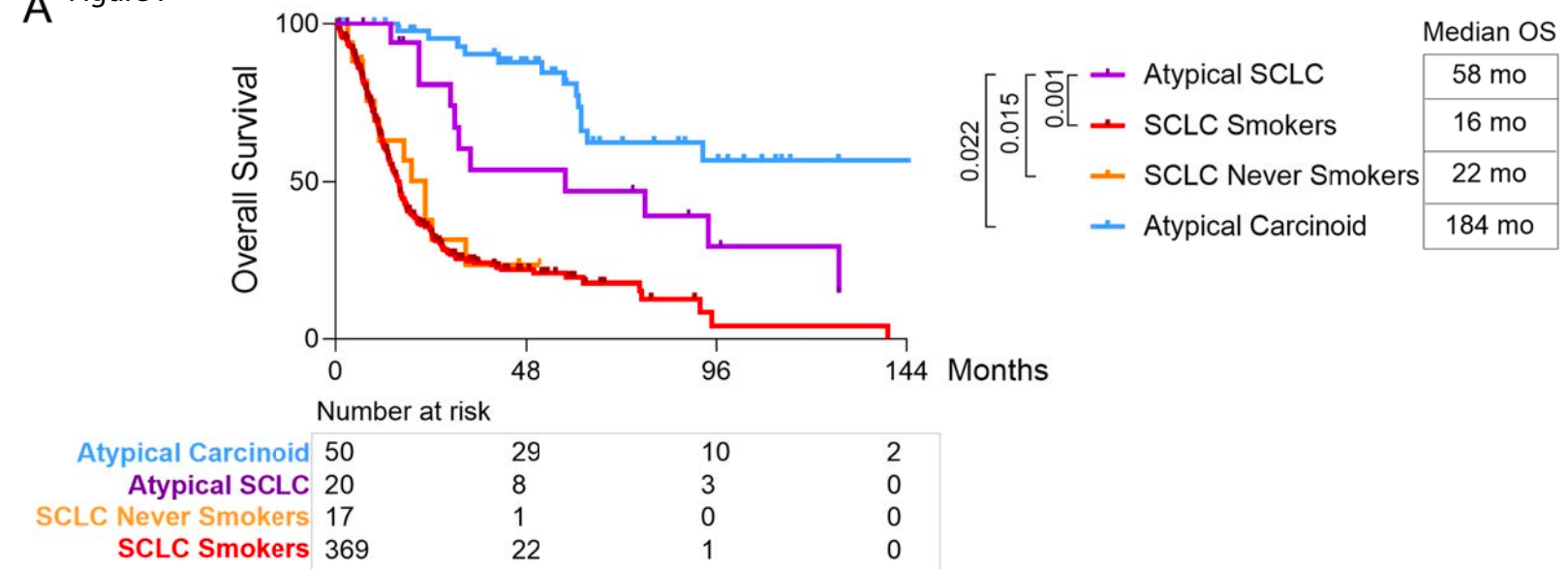


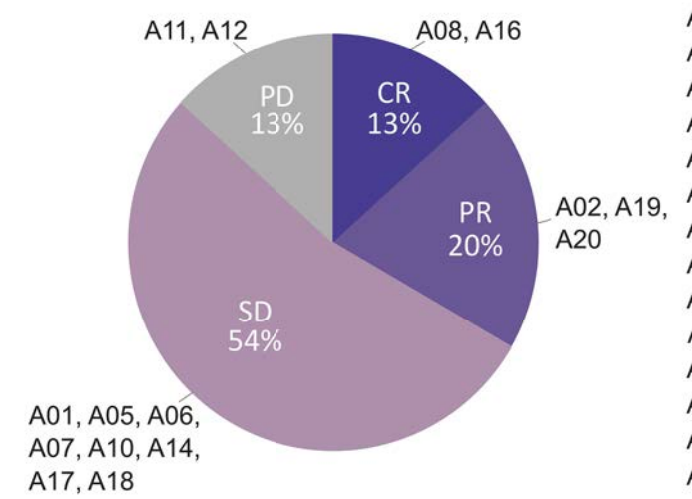
Figure 6



A Figure 7



B



C

