Research Article

Genomic and TCR Repertoire Intratumor Heterogeneity of Small-cell Lung Cancer and its Impact on Survival

Authors: Ming Chen[1,2,3,12,13,*], Runzhe Chen[4,5,12], Ying Jin[1,2,3,12], Jun Li[5,12], Jiexin Zhang[6,12], Junya Fujimoto[7], Won-Chul Lee[5], Xin Hu[5], Shawna Maria Hubert[4,5], Julie George[8], Xiao Hu[1,2,3], Yamei Chen[1,2,3], Carmen Behrens[4], Chi-Wan Chow[7], Hoa H.N. Pham[9], Junya Fukuoka[9], Edwin Roger Parra[7], Carl M. Gay[4], Latasha D. Little[5], Curtis Gumbs[5], Xingzhi Song[5], Lixia Diao[6], Qi Wang[6], Robert Cardnell[4], Jianhua Zhang[5], Jing Wang[6], Don L. Gibbons[4], John V. Heymach[4], J. Jack Lee[6], William N. William Jr.[4], Bonnie Glisson[4], Ignacio Wistuba[7], P. Andrew Futreal[5], Roman K. Thomas[8,10,11,13], Alexandre Reuben[4,13], Lauren A. Byers[4,13], and Jianjun Zhang[4,5,13,*]

Affiliations:

[1] Cancer Hospital of University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, Zhejiang, China

[2] Institute of Cancer and Basic Medicine (IBMC), Chinese Academy of Sciences, Hangzhou, Zhejiang, China

[3] Zhejiang Key Laboratory of Radiation Oncology, Hangzhou, Zhejiang, China

[4] Department of Thoracic/Head and Neck Medical Oncology, the University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA

[5] Department of Genomic Medicine, the University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA

[6] Department of Biostatistics, the University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA

[7] Department of Translational Molecular Pathology, the University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA

[8] Department of Translational Genomics, Center of Integrated Oncology Cologne-Bonn, Medical Faculty, University of Cologne, 50931 Cologne, Germany

[9] Department of Pathology, Nagasaki University Graduate school of Biomedical Sciences

[10] Department of Pathology, Medical Faculty, University Hospital Cologne, 50937 Cologne, Germany

[11] DKFZ, German Cancer Research Center and German Cancer Consortium (DKTK), Heidelberg, Germany

[12] These authors contributed equally

[13] Lead Contact

* Correspondence: chenming@zjcc.org.cn (M.C.), jzhang20@mdanderson.org (J.Z.)

Corresponding authors

Dr. Ming Chen,

Cancer Hospital of University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, Zhejiang; Institute of Cancer and Basic Medicine (IBMC), Chinese Academy of Sciences; Zhejiang Key Laboratory of Radiation Oncology, Zhejiang, China

Phone: 86-571-88122391

Fax: 86-571-88122391

E-mail: chenming@zjcc.org.cn

Dr. Jianjun Zhang,

Departments of Thoracic and Head and Neck Medical Oncology; Genomic Medicine, the University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd Unit 428, Houston, TX 77030, USA

Phone: 01-713-563-6096

E-mail: jzhang20@mdanderson.org

Declaration of interests

L.A.B. serves on advisory committees for AstraZeneca, AbbVie, GenMab, BergenBio, Pharma Mar SA, Sierra Oncology, Merck, Bristol Myers Squibb, Genentech, and Pfizer and has research support from AbbVie, AstraZeneca, GenMab, Sierra Oncology and Tolero Pharmaceuticals. I.W reports grants and personal fees from Genentech/Roche, grants and personal fees from Bayer, grants and personal fees from Bristol-Myers Squibb, grants and personal fees from AstraZeneca/Medimmune, grants and personal fees from Pfizer, grants and personal fees from HTG Molecular, grants and personal fees from Merck, personal fees from GlaxoSmithKline, grants and personal fees from Guardant Health, personal fees from MSD, grants from Oncoplex, grants from DepArray, grants from Adaptive, grants from Adaptimmune, grants from EMD Serono, grants from Takeda, grants from Amgen, grants from Karus, grants from Johnson & Johnson, grants from Iovance, grants from 4D, grants from Novartis, grants from Oncocyte, grants from Akoya. J.Z reports research funding and personal fees from BMS, Johnson and Johnson, AstraZeneca, Geneplus, OrigMed, Innovent and Merck, outside the submitted work. The remaining authors declare no competing interests.

Running title: Immunogenomic intratumor heterogeneity of SCLC

Abstract

Small-cell lung cancer (SCLC) is speculated to harbor complex genomic intratumor heterogeneity (ITH) associated with high recurrence rate and suboptimal response to immunotherapy. Here, we revealed a rather homogeneous mutational landscape but extremely suppressed and heterogeneous T cell receptor (TCR) repertoire in SCLCs. Higher mutational burden, lower chromosomal copy number aberration (CNA) burden, less CNA ITH and less TCR ITH were associated with longer overall survival of SCLC patients. Compared to non-small cell lung cancers (NSCLCs), SCLCs had similar predicted neoantigen burden and mutational ITH, but significantly more suppressed and heterogeneous TCR repertoire that may be associated with higher CNA burden and

CNA ITH in SCLC. Novel therapeutic strategies targeting CNA could potentially improve the tumor immune microenvironment and response to immunotherapy in SCLC.

Keywords: small-cell lung cancer, intratumor heterogeneity, genomic, T cell receptor, survival

Introduction

Small-cell lung cancer (SCLC) accounts for ~15% of all newly diagnosed lung cancers leading to ~30,000 deaths in the United States annually[1]. SCLC is a highly aggressive cancer characterized by rapid growth and high rates of early local and distant metastases[2,3]. At initial diagnosis, around one third of SCLC patients present with cancer confined to one hemithorax, defined as limited-stage disease (LD) that can be treated with chemotherapy combined with radiotherapy or surgical resection, while the remaining patients present with extensive-stage disease (ED) exhibiting extensive lymph node involvement and/or distant metastases usually treated with palliative chemotherapy with or without immune checkpoint blockade (ICB)[4-6]. Although most SCLC patients experience an initial response, nearly all patients recur with rapidly progressing disease resistant to late-line treatments. Despite extensive research, only modest advances have been achieved in the treatment of SCLC over the past 30 years with median survival less than a year and 5-year overall survival (OS) is below 7% for ED SCLC[1,7-9]. Recently, the addition of ICB to chemotherapy has become a new standard of care for advanced SCLC, although it confers only an improvement of 2-3 months in survival[8]. National Cancer Institute (NCI) has identified SCLC as a recalcitrant malignancy[1]. Translational studies to understand the mechanisms underlying recurrence and therapeutic resistance remain an unmet need to design novel therapeutic strategies[8,10,11].

Tumors are composed of cancer cells and stromal cells of distinct molecular and phenotypic features, a phenomenon termed intratumor heterogeneity (ITH). ITH has been shown to impact response to therapy and patient survival[12-15]. We and others have previously delineated the ITH architecture of non-small cell lung cancers (NSCLCs) at genomic, epigenetic and gene expression levels utilizing multiregional sequencing and demonstrated that complex ITH was associated with inferior survival[16-22]. It has been speculated that SCLC has extremely complex ITH architecture that lead to poor prognosis[23]. Another plausible explanation for the poor outcome is that SCLC is associated with an immunosuppressive tumor microenvironment, particularly T cell responses[24,25]. In localized NSCLC, our recent work has revealed that a suppressed

and heterogeneous T cell receptor (TCR) repertoire is associated with inferior survival[18,26].

The genomic and TCR ITH architecture of SCLC and their potential clinical impact have not been well studied, largely due to lack of available tumor specimens[27-30]. Through international collaboration, we conducted multiregional whole exome sequencing (WES) and TCR sequencing of 50 tumor samples from 19 resected LD SCLCs **(Figure S1)** to depict the immunogenomic ITH architecture of SCLC**.** We further compared these SCLCs to a cohort of 216 localized NSCLCs (PROSPECT cohort) [26] and assessed the impact of genomic and TCR attributes on patient survival.

Results

Mutational landscape of LD SCLC tumors is overall homogeneous

A total of 50 tumor regions (hereafter referred as region) from 19 resected LD SCLC tumors (hereafter referred as tumor) were subjected to WES (**Figure S1**, **Supplementary Table 1**). In total, 3,773 nonsilent (nonsynonymous, stop-gain and stop-loss) mutations were identified from these 50 tumor regions (**Supplementary Data 1**) for a median nonsilent tumor mutational burden (TMB) of 4.69/Mb. TMB was similar between different tumor regions within the same tumors, but varied substantially between patients (**Figure S2**).

Next we constructed phylogenetic trees of 18 SCLCs for which multiregional WES data available (P13 only had one tumor region and was excluded from this analysis) to depict the genomic ITH and the evolutionary trajectory of these SCLCs as described previously[17]. A median of 80.4% (28%-93%) of mutations were mapped to the trunks of these 18 SCLCs (**Figure 1** and **Figure 2A**) representing ubiquitous mutations present in all tumor regions within the same tumor, comparing to 73% (8%-99.6%) trunk mutations in 100 early-stage NSCLC in TRACERx cohort (p=0.218)[19]. Furthermore, using PyClone[31], we classified mutations as clonal (defined as estimated cancer cell fraction=1, indicating mutations present in all cancer cells) or subclonal

(defined as estimated cancer cell fraction<1, indicating mutations only present in a subset of cancer cells) in each tumor specimen. A median of 92.8% (37%-99.9%) of mutations were clonal in these 50 SCLC specimens (**Figure 2B**). Taken together, these results suggested the mutational landscape is overall homogeneous in these SCLCs.

The most frequently mutated cancer genes in this cohort included *TP53*, *RB1*, and *LRP1B* identified in 14, 10 and 7 patients respectively (**Figure S3**). Of note, these mutations were trunk mutations detected in all regions from the same SCLC tumors (**Figure 1** and **Figure S3**) and these canonical mutations were clonal in each tumor specimen where they were identified. These results suggest that these canonical cancer gene mutations were all early genomic events during evolution of SCLC.

Mutational processes in this cohort of SCLC

Understanding how mutational processes shape cancer evolution may inform mechanisms underlying tumor adaptation. We next analyzed the mutational spectrum and signatures in these SCLCs[32]. C>A transversions were the most common nucleotide substitutions (**Figure S4**) and Cosmic Signature 4 (associated with cigarette smoking) was the predominant mutational signature (**Figure 3A**) as expected, given 15 of the 19 patients were smokers.

To further dissect the mutational processes associated with early clonal expansion versus subsequent subclonal diversification, we delineated the mutational signatures of trunk mutations representing early genomic events and non-trunk mutations representing later subclonal events, respectively. As shown in **Figure 3B-C,** Cosmic Signature 4 remained as the predominant signature in trunk mutations consistent with previous reports that smoking associated mutational processes play critical roles during early mutagenesis of lung cancers[17,19,33]. On the other hand, the contribution of Signature 4 was significantly reduced (p=0.002) while Cosmic Signature 3 (associated with defect of DNA double-strand break-repair) emerged as the predominant signature for non-trunk mutations (p< 0.0001). Taken together, these results highlight the dynamic nature of mutagenesis at different times during evolution

of SCLC and suggest that smoking-associated mutational processes play essential roles during early clonal expansion while subclonal diversification of this cohort of SCLC may be associated with other mutational processes such as DNA repair defect.

Suppressed TCR repertoire in SCLC

We next performed TCR sequencing in 36 tumor specimens (1 to 3 regions per tumor) and 16 tumor-adjacent lung tissues from patients with adequate DNA remaining. T-cell density, an estimate of the proportion of T cells in a specimen, ranged from 0.11% to 33% with a median of 1.7% (**Figure S5A**). T-cell richness, a measure of T-cell diversity, ranged from 38 to 8,286 unique T-cells (median: 510) per specimen (**Figure S5B)** and T-cell clonality, a metric indicating T-cell expansion and reactivity, ranged from 0.002 to 0.139 (median=0.009) (**Figure S5C**). Density, richness and clonality were positively correlated with each other (Density *vs.* Richness: r=0.87, p<0.0001; Density *vs.* Clonality: r=0.88, p<0.0001; Clonality *vs.* Richness: r=0.97, p<0.0001) (**Figure S5D**). Compared to tumor-adjacent lung tissues (≥2 cm from tumor margin), SCLC tumors demonstrated lower T-cell density, richness and clonality (p=0.0580, p=0.0067, p=0.0166, respectively; **Figure S6A-C**), indicating a suppressed T-cell repertoire in tumor tissues consistent with NSCLC[26]. Of particular interest, all three TCR metrics were significantly lower than those of the 216 localized NSCLCs from PROSPECT cohort (T-cell density 0.014 *vs.* 0.21, p<0.0001 (**Figure S7A**); diversity 510 *vs.* 3,246, p<0.0001 (**Figure S7B**); and T-cell clonality 0.009 *vs.* 0.14, p<0.0001 (**Figure S7C**))[26].

Additionally, we derived immune scores quantifying the density of immune-cells within tumors by deconvoluting RNA seq data of 81 SCLCs[28] and compared those to 1,027 NSCLCs from TCGA[33,34]. In line with TCR repertoire findings, the immune scores in SCLCs were significantly lower than NSCLCs (**Figure S8**, p<0.0001). Taken together, these results suggested that SCLC may have more suppressed immune microenvironment than NSCLC.

Substantial TCR repertoire heterogeneity in SCLC

To gain insights into TCR heterogeneity, we calculated Jaccard index (JI), a metric measuring the proportion of shared T-cell clonotypes between two samples. Substantial TCR heterogeneity was evident across all tumors, with a median JI of 0.05 (0.02 to 0.15) in the 10 SCLCs with multiregional TCR data available (**Figure 4A**), significantly lower than the 11 localized NSCLCs[18] (median 0.05 in SCLC *vs.* 0.16 in NSCLC, p<0.0001) (**Figure S7D**). Furthermore, 79.9%-97.7% of T-cell clones were restricted to individual tumor regions while only 0.2%-14.6% were identified in all regions within the same tumors **(Figure 4B)**, significantly lower than NSCLC (1.6% to 14.5%, p=0.0048)[18] demonstrating profound TCR ITH in SCLC even beyond NSCLC.

High-level and heterogeneous copy number alterations may be the underlying genomic basis for suppressed and heterogeneous TCR repertoire in SCLC

To identify genomic aberrations that could contribute to the suppressive TCR repertoire and ITH in SCLC, we first looked at somatic mutations that play central roles in anti-tumor T cell response by producing non-self proteins that can be recognized by T cells – so called neoantigens[35,36]. We performed *in silico* prediction of HLA-A-, -B-, and -C-presented neoantigens. A median of 78 (26-463) predicted neoantigens (IC_{50}< 500 nmol/L) per tumor were detected (**Figure S9A**), which was similar to NSCLCs from the PROSPECT cohort (median: 72/tumor, 2-801, **Figure S9B**, p=0.31). Similar to somatic mutations, 81% (43%-93%) of predicted neoantigens were present across different regions within the same tumors (**Figure S9C**) suggesting the suppressed and heterogeneous TCR repertoire in SCLC was unlikely due to low clonal neoantigen burden.

Next, we explored whether SCLC had higher incidence of loss of heterozygosity (LOH) of human leukocyte antigen (HLA), a potential immune evasion mechanism in cancer [37,38]. Evidence of HLA LOH was revealed in 9 of 19 SCLCs, higher than NSCLCs from the PROSPECT cohort, but the difference did not reach statistical difference (9/19 vs. 60/216, p=0.11) suggesting that HLA LOH is a common mechanism underlying immune evasion in both SCLC and NSCLC, but may not be the main determinant of more suppressed TCR repertoire in SCLC *versus* NSCLC.

As a higher chromosomal copy number aberration (CNA) burden has been reported to associate with immunosuppression in multiple cancer types[39,40], we next assessed the CNA burden in this cohort of SCLC. A median of 2,180 CNA events per tumor (26 to 7622) were identified from these SCLCs (**Figure S10A**), significantly higher than 622 per tumor (range: 0-7741) in NSCLCs from PROSPECT cohort (p<0.0001) (**Figure S10B**)[26]. Additionally, in this cohort of SCLC tumors, CNA burden was negatively associated with T-cell density, richness and clonality (r=-0.4, p=0.0157; r=-0.36, p=0.0317; r=-0.33, p=0.0484; respectively) (**Figure S11A-C**). Furthermore, CNA JI, a surrogate for CNA ITH was positively associated with TCR JI (r=0.74, p=0.0141) (**Figure 4C)**. Taken together, these data suggest that higher CNA burden and higher level of CNA ITH could be important genomic basis for profoundly suppressed and heterogeneous TCR ITH in these SCLCs.

Genomic and TCR ITH were associated with survival of SCLC

With small sample size fully acknowledged, we attempted to assess whether the genomic and T cell features impact clinical outcome. We focused on overall survival since recurrence status was unavailable for some patients. With a median of 45 months of postsurgical follow up, 9 patients have expired with a median of OS of 45 months, comparable to previous reports[41-44]. Interestingly, higher TMB was associated with significantly longer OS (**Figure 5A**, HR=0.13, p=0.0281), consistent with previous reports in NSCLC[45]. Conversely, higher CNA burden was associated with significantly shorter OS (**Figure 5B**, HR=13.8, p=0.0033), while significantly longer OS was observed in patients with low level of copy number ITH (high CNA JI, **Figure 5C**, HR=4.21E-10, p=0.0019). No TCR parameter (T-cell density, richness and clonality) was associated with OS (**Figure S12A-C**), however, patients with a more homogenous TCR repertoire (higher TCR JI) exhibited significantly longer OS (**Figure 5D**, HR=0.16, p=0.0496).

Discussion

Evolutionary theory suggests populations of high genetic variation have survival advantages[46]. Similarly, tumors of complex ITH may be difficult to eradicate. Higher level of molecular ITH has been demonstrated to associate with inferior outcome of cancer patients[12,16,17]. In SCLC, however, although pioneering studies have revealed some pivotal molecular features[27-30,47], the genomic ITH architecture has not been defined, primarily due to the lack of adequate tumor specimens for multiregional profiling. Because SCLC is sensitive to initial treatment but nearly all patients experience relapse with refractory disease, it has been speculated that SCLC may have profound mutational ITH, where cancer cells highly resistant to chemotherapy/radiotherapy hide in the treatment-naïve SCLC tumors as minor subclones that give rise to relapse[7,11]. Surprisingly, all SCLCs in the current study demonstrated homogeneous mutational ITH with the majority of mutations present in all regions within the same tumors (**Figure 1 and 2A**) and a median of 92.8% of mutations being clonal in each tumor specimen (**Figure 2B**). Additionally, previous work from Wagner and colleagues has demonstrated striking similarity of the mutational landscape between primary and relapsed SCLC[47]. Taken together, these data indicate that complex mutational ITH and selection of chemo-/radio-resistant minor subclones may not be the main mechanisms underlying therapeutic resistance in SCLC.

Cancer evolution with or without treatment may be shaped by the dynamic interaction between cancer cells and host factors, particularly through immune surveillance[48]. Our study delineates for the first time, the TCR repertoire of SCLC and demonstrates a suppressed T-cell repertoire in SCLC. All TCR attributes were extremely low quantitatively (density) and qualitatively (richness and clonality), compared to not only matched normal lung tissues (**Figure S6**) but also compared to NSCLC tumors (**Figure S7A-C**). Similarly, comparing a previously published large SCLC cohort (n=81)[28] to TCGA NSCLC cohorts (n=1,027) also revealed more suppressed immune contexture in SCLC than NSCLC (**Figure S8**).

In addition to the suppressed TCR repertoire, SCLC also demonstrated extremely heterogeneous TCR repertoire with only 0.2%-14.6% of all T cells identified across all

tumor regions within the same tumors. TCR ITH was even more pronounced than that in NSCLC (**Figure S7D**)[18], which may further impair the efficacy of anti-tumor immune response. Interestingly, even with such a small sample size, higher TCR JI indicating less TCR ITH was associated with better survival in these SCLC patients (**Figure 5B**) indicating the potential clinical impact of TCR ITH. SCLC is among the cancers with high TMB[49] and our study also demonstrated homogenous mutational landscape, both of which have been reported to associate with benefit from ICB[50]. However, compared to NSCLC and other tumor types, fewer SCLC patients benefit from ICB[51]. The suppressed and heterogeneous TCR repertoire may be one potential reason underlying suboptimal response to immunotherapy.

As the TCR repertoire attributes in this cohort of SCLC were significantly suppressed compared to NSCLC from PROSPECT cohort (**Figure S7A-C**), we compared the genomic landscape of tumors of these two cohorts to understand the potential genomic bases for the more suppressed TCR repertoire in SCLC. These analyses revealed significantly higher CNA burden in SCLC (**Figure S10B**). Moreover, the CNA burden was negatively associated with both T cell quantity (density) and quality (richness and clonality) (**Figure S11**) and CNA ITH was positively associated with TCR ITH (**Figure 4C**) in this cohort of SCLC. These results suggest that high CNA burden and high level of CNA ITH may be one of the underlying genomic bases for the suppressed and heterogeneous TCR repertoire in this cohort of SCLC.

High CNA burden has been reported to correlate with immunosuppressive microenvironment and inferior benefit from ICB across different cancer types[39,52,53]. The mechanisms underlying the association between high CNA burden and immunosuppression are not well understood. Several hypotheses have been proposed such as relatively low neoantigen concentration due to protein imbalance leading to impaired cancer cell signals in tumors[39,52]. From a therapeutic standpoint, the significantly higher CNA burden suggests targeting CNA could be a potential effective strategy for treating SCLC. Although CNA can potentially lead to gene dosage effects that could promote tumor growth and provide the immune evasive advantage for cancer

proliferation[39,52,53], excessive CNA beyond a certain level could be lethal to cancer cells[54,55]. Genes and pathways involved in CNA (e.g. spindle assembly checkpoint, supernumerary centrosome clustering, Aurora kinase family members, etc.) have been exploited as candidate therapeutic targets for different cancer types including SCLC[56]. Unfortunately, none of these agents has shown substantial efficacy to make the way to clinical practice in treating SCLC although anti-tumor activities have been observed from several agents of this class[54,56,57]. One plausible explanation is the profound CNA ITH in SCLC as observed in the current study where different cancer cells may have vastly different CNA profiles. As such, these CNA promoting agents could kill cancer cells with excessive CNA while spare cancer cells with less CNA leading to therapeutic failure. Moreover, increasing CNA potentially turns formerly CNA-low cancer cells into relatively CNA-high cells starting the cycle again that further suppressing host anti-tumor immune response. This is a similar quandary with inhibiting DNA damage response (DDR) pathway where deficient DDR pathways could increase DNA-damaging chemo-/radio-therapy sensitivity but conversely promote tumorigenesis[58,59]. Therefore, in order to effectively eliminate heterogeneous cancer cells with different CNA profiles, CNA targeting agents may be combined with ICB, which has already been tested in treating SCLC in both preclinical murine models [60] and clinical trials (NCT03041311)[61].

To the best of our knowledge, the current study is the first study on genomic and TCR ITH of SCLC. Our study was limited by the small sample size due to the scarcity of resected SCLC specimens. However, WES and TCR data from multiregional specimens made it valuable to the field. In summary, we demonstrate that despite a homogeneous mutational landscape, SCLC exhibits a suppressed and heterogeneous TCR repertoire that could lead to ineffective anti-tumor immune surveillance, which could be one potential molecular mechanism underlying high recurrence rate and suboptimal response to immunotherapy in SCLC. Our results also suggest that high CNA burden may be one of the underlying reasons for the suppressed T-cell repertoire, therefore a potential therapeutic target to improve the efficacy of immune checkpoint blockade in patients with SCLC.

Methods

Patients

A total of 19 patients with lymph node negative LD SCLC, who underwent surgical resection at Zhejiang Cancer Hospital, Hangzhou, China from 2010 to 2015 were enrolled. With a median of 45 months of postsurgical follow up, 7 patients have relapsed and deceased, 10 patients were still alive with no evidence of recurrence and 2 patients deceased with unknown recurrence status. The median survival of this cohort was 45 months. The study was approved by the Institutional Review Boards (IRB) at MD Anderson Cancer Center and Zhejiang Cancer Hospital.

Sample processing and DNA extraction

Hematoxylin and eosin slides from each tumor were reviewed by experienced lung cancer pathologists to confirm the diagnosis, assess necrosis, tumor purity and cell viability. Manual macro-dissection was conducted to enrich malignant cells. DNA was extracted using the AllPrep® DNA/RNA FFPE Kit (Qiagen, Hilden, Germany) from 50 spatially separated tumor regions (3 regions per tumor from 13 patients, 2 regions per tumor from 5 patients and 1 tumor piece from one patient) and paired matched adjacent normal lung (≥2 cm from tumor margin, morphologically negative for malignant cells assessed by two lung cancer pathologists independently) as previously described[62].

Whole exome sequencing

WES was performed using the Illumina protocol in MD Anderson. Exome capture was performed on 500ng of genomic DNA per sample based on KAPA library prep (Kapa Biosystems) using the Agilent SureSelect Human All Exon V4 kit according to the manufacturer's instructions and paired-end multiplex sequencing of samples was performed on the Illumina HiSeq 2000 sequencing platform. The average sequencing depth was 180x for tumor DNA (ranging from 64x to 224x), 161x for germline DNA (ranging from 96x to 194x).

Mutation calling

The BWA aligner (bwa-0.7.5a) was applied to map the raw reads to the human hg19 reference genome (UCSC genome browser: genome.ucsc.edu). The Picard (v1.112, http://broadinstitute.github.io/picard/) ''MarkDuplicates'' module was applied to mark the duplicate reads. Then the ''IndelRealigner'' and ''BaseRecalibrator'' modules of the Genome Analysis Toolkit were applied to perform indel realignment and base quality recalibration. Mutect (v1.1.4) [63] was applied identify somatic single nucleotide variants (SNVs) and small insertions/deletions. To ensure high-quality mutation calls, the following filtering criteria were applied: 1) sequencing depth ≥20× in tumor DNA and ≥10× in germline DNA; and 2) variant alleleve frequency (VAF) ≥ 0.02 in tumor DNA and < 0.01 in germline DNA; and 3) the total number of reads supporting the variant calls is ≥4; and 4) variant frequency is < 0.01 in ESP6500, 1000 genome and EXAC databases; and 5) LOD score >18 (MuTect default is 6.3). We kept the mutations that passed all filtering criteria except LOD score < 18 only if the identical mutations were present with LOD score >=18 in other regions within the same tumors.

Cancer gene mutations were defined as identical oncogene mutations previously reported; stop gains and frameshift of tumor suppressor genes; other non-synonymous mutations with Combined Annotation Dependent Depletion (ACDD) score>20[64].

Clonal and subclonal analysis

Tumor contents and major/minor copy number changes were estimated by Sequenza (v2.1.2).[65] The cancer cell fraction (CCF) and mutant allele copy number for each SNV was inferred using Pyclone 12.3[31]. In brief, PyClone implements a Dirichlet process clustering model that simultaneously estimates the distribution of the cellular prevalence for each mutation. Copy numbers of somatic mutations were inferred by integrating integer copy numbers determined by Sequenza on single sample basis. The outputs were cellular prevalence value distributions per SNV estimated from Markov-chain Monte Carlo (MCMC) sampling. The median value of the MCMC sampling-derived distribution was used as a representative cellular prevalence for each mutation. A given mutation was classified as “clonal” if the 95% confidence interval of CCF overlapped 1 and “subclonal” otherwise.

Phylogenetic analysis

Mutation profiles were converted into binary format with 1 being mutated and 0 otherwise. Ancestors were germ line DNA assuming with no mutations. Multistate discrete-characters Wagner parsimony method in PHYLIP (Phylogeny Inference Package) was used to generate phylogenic tree[66].

Mutational signature analysis

The R package "DeconstructSigs" package [67] was applied to estimate the proportions of 30 COSMIC mutational signatures (http://cancer.sanger.ac.uk/cosmic/signatures).

Somatic copy number analysis

Somatic copy number analysis were performed applying CNVkit (v0.9.6)[68], through which both the targeted reads and the nonspecifically captured off-target reads were used to infer copy number evenly across the genome, and DNA segmentation of log2 ratios in the tumor samples were calculated, then segment data were processed using the "CNTools" package to generate segmented DNA copy number profile at gene level by assigning segment means to the genes within the chromosome segments for each sample. Genes with mean segment more than 0.6 was defined as copy number gain and less than -0.6 was defined as copy number loss. Copy number gain and loss burden were defined as the number of genes located in the segments with copy number gains and losses.

Neoantigen prediction

WES data were reviewed for non-synonymous exonic mutations. The binding affinity with patient-restricted MHC Class I molecules of all possible 9- and 10-mer peptides was evaluated with the NetMHC3.4 algorithm based on patient HLA-A, HLA-B, and HLA-C alleles[69-71]. Candidate peptides were considered HLA binders when IC50<500 nM.

TCRβ sequencing and comparison parameters

Immunosequencing of the CDR3 regions of human TCRβ chains was performed using ImmunoSeq (Adaptive Biotechnologies, hsTCRβ Kit)[18,26]. T-cell density was calculated by normalizing TCR-β template counts to the total amount of DNA for TCR sequencing, where the amount of total DNA was determined by PCR-amplification and sequencing of housekeeping genes expected to be present in all nucleated cells. T-cell richness is calculated using the unique rearrangements. T-cell clonality is defined as 1-Peilou's evenness and is calculated on productive rearrangements as previously described[18,26]. Jaccard index (JI) was calculated by the number of rearrangements shared/sum of total number of rearrangements between any two specimens.

Human leukocyte antigen loss of heterozygosity analysis

For Human Leukocyte Antigen Loss Of Heterozygosity (HLA LOH) analysis, we first performed HLA typing using PHLAT[72]. For each patient, we merged tumor and normal BAM files and inferred 4-digit HLA types for the major class I HLA genes (HLA-A, HLA-B and HLA-C). To evaluate HLA loss, we used a computational tool, LOHHLA [73] using purity and ploidy information estimated by Sequenza[74]. Sample as being subject to HLA loss was defined when any of the two alleles of HLA-A, HLA-B or HLA-C showed a copy number < 0.5 with a paired Student's *t* test $p < 0.01$.

Analysis of published data

RNA sequencing data from 81 SCLCs[28] and 1,027 NSCLCs from TCGA[33,34] were downloaded. Immune scores were calculated by taking the average of normalized expression levels of genes including cytolytic markers, HLA molecules, IFN-γ pathway, chemokines and adhesion molecules as previously described[75].

Statistical Analysis

Graphs were generated with GraphPad Prism 8.0 (La Jolla, CA). Pearson's correlations were calculated to assess association between 2 continuous variables. Wilcoxon signed-rank test was applied to compare paired TCR metrics. Mann-Whitney test was used to compare differences between two independent groups. Log-rank test was used for survival analysis.

Acknowledgements

Barbanti Small Cell Lung Cancer Award, Conquer Cancer Foundation ASCO Young Investigator Award, MD Anderson Physician Scientist Award, University Cancer Foundation Sister Institution Network Fund, Cancer Prevention & Research Institute of Texas (CPRIT) Multiple Investigator Award, TJ Martell Foundation, NIH/NCI R01-CA207295, NIH/NCI U01-CA213273, Department of Defense (LC170171), National Natural Science Foundation of China (grant No. 81672972); Major Program of Provincial Ministry of Health Science Foundation (grant No. WKJ-ZJ-1701).

Author contributions

M.C. and J.Z. conceived the study. R.C., J.L. and Jiexin Z. led the data analysis. J.Y., J.F, H.P., CW.C. and J.F. led the pathological assessment, multi-region sample preparation and DNA extraction. J.Y, Y.C. and X.H. collected resected specimens and clinical data. L.L and C.G. performed DNA preparation and whole-exome sequencing. X.S. and Jianhua Z. performed sequencing raw data processing. J.L., W.L. and X.H. performed downstream bioinformatics analyses. R.C., S.M.H, J.G., C.B., E.R.P., C.G., Robert C., D.G., J.H., W.W., B.G., I.W., P.A.F., R.K.T., A.R., L.A.B. and J.Z. interpreted the data for clinical and pathological correlation. L.D., Q.W., J.W., and JJ.L. performed statistical analyses. R.C., A.R., C.G. and J.Z. wrote the paper. All authors edited the manuscript.

References

1. Howlader, N.*, et al.* SEER Cancer Statistics Review, 1975-2016. *National Cancer Institute* (2019).
2. Alvarado-Luna, G. & Morales-Espinosa, D. Treatment for small cell lung cancer, where are we now?—a review. *Translational lung cancer research* **5**, 26 (2016).
3. Govindan, R.*, et al.* Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic, and end results database. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **24**, 4539-4544 (2006).
4. Kalemkerian, G.P.*, et al.* Small cell lung cancer. *Journal of the National Comprehensive Cancer Network* **11**, 78-98 (2013).

5. Micke, P.*, et al.* Staging small cell lung cancer: Veterans Administration Lung Study Group versus International Association for the Study of Lung Cancer—what limits limited disease? *Lung cancer* **37**, 271-276 (2002).
6. Horn, L.*, et al.* First-Line Atezolizumab plus Chemotherapy in Extensive-Stage Small-Cell Lung Cancer. *The New England journal of medicine* **379**, 2220-2229 (2018).
7. Pietanza, M.C. & Ladanyi, M. Bringing the genomic landscape of small-cell lung cancer into focus. *Nature genetics* **44**, 1074 (2012).
8. Horn, L.*, et al.* First-line atezolizumab plus chemotherapy in extensive-stage small-cell lung cancer. *New England Journal of Medicine* **379**, 2220-2229 (2018).
9. Paz-Ares, L.*, et al.* Durvalumab plus platinum–etoposide versus platinum–etoposide in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): a randomised, controlled, open-label, phase 3 trial. *The Lancet* **394**, 1929-1939 (2019).
10. Byers, L.A. & Rudin, C.M. Small cell lung cancer: where do we go from here? *Cancer* **121**, 664-672 (2015).
11. Gazdar, A.F., Bunn, P.A. & Minna, J.D. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nature Reviews Cancer* **17**, 725 (2017).
12. Yap, T.A., Gerlinger, M., Futreal, P.A., Pusztai, L. & Swanton, C. Intratumor heterogeneity: seeing the wood for the trees. *Science translational medicine* **4**, 127ps110-127ps110 (2012).
13. Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer research* **72**, 4875-4882 (2012).
14. Fisher, R., Pusztai, L. & Swanton, C. Cancer heterogeneity: implications for targeted therapeutics. *British journal of cancer* **108**, 479 (2013).
15. Gerlinger, M.*, et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature genetics* **46**, 225 (2014).
16. Gerlinger, M.*, et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England journal of medicine* **366**, 883-892 (2012).
17. Zhang, J.*, et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* **346**, 256-259 (2014).
18. Reuben, A.*, et al.* TCR repertoire intratumor heterogeneity in localized lung adenocarcinomas: an association with predicted neoantigen heterogeneity and postsurgical recurrence. *Cancer discovery* **7**, 1088-1097 (2017).
19. Jamal-Hanjani, M.*, et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *The New England journal of medicine* **376**, 2109-2121 (2017).
20. Lee, W.C.*, et al.* Multiregion gene expression profiling reveals heterogeneity in molecular subtypes and immunotherapy response signatures in lung cancer. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **31**, 947-955 (2018).
21. Quek, K.*, et al.* DNA methylation intratumor heterogeneity in localized lung adenocarcinomas. *Oncotarget* **8**, 21994-22002 (2017).
22. Rosenthal, R.*, et al.* Neoantigen-directed immune escape in lung cancer evolution. *Nature* **567**, 479-485 (2019).

23. van Meerbeeck, J.P., Fennell, D.A. & De Ruysscher, D.K. Small-cell lung cancer. *Lancet (London, England)* **378**, 1741-1755 (2011).
24. Hendriks, L.E., Menis, J. & Reck, M. Prospects of targeted and immune therapies in SCLC. *Expert review of anticancer therapy* **19**, 151-167 (2019).
25. Bonanno, L.*, et al.* The role of immune microenvironment in small-cell lung cancer: Distribution of PD-L1 expression and prognostic role of FOXP3-positive tumour infiltrating lymphocytes. *European journal of cancer (Oxford, England : 1990)* **101**, 191-200 (2018).
26. Reuben, A.*, et al.* Comprehensive T cell repertoire characterization of non-small cell lung cancer. *Nat Commun* **11**, 603 (2020).
27. Peifer, M.*, et al.* Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet* **44**, 1104-1110 (2012).
28. George, J.*, et al.* Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47-53 (2015).
29. Rudin, C.M.*, et al.* Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat Genet* **44**, 1111-1116 (2012).
30. Nong, J.*, et al.* Circulating tumor DNA analysis depicts subclonal architecture and genomic evolution of small cell lung cancer. *Nature communications* **9**, 3114 (2018).
31. Roth, A.*, et al.* PyClone: statistical inference of clonal population structure in cancer. *Nature methods* **11**, 396-398 (2014).
32. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Campbell, P.J. & Stratton, M.R. Deciphering signatures of mutational processes operative in human cancer. *Cell reports* **3**, 246-259 (2013).
33. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519-525 (2012).
34. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543-550 (2014).
35. Schumacher, T.N. & Schreiber, R.D. Neoantigens in cancer immunotherapy. *Science* **348**, 69-74 (2015).
36. Strønen, E.*, et al.* Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337-1341 (2016).
37. McGranahan, N.*, et al.* Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* **171**, 1259-+ (2017).
38. Tran, E.*, et al.* T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. *N Engl J Med* **375**, 2255-2262 (2016).
39. Davoli, T., Uno, H., Wooten, E.C. & Elledge, S.J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**(2017).
40. Xie, F.*, et al.* Multifactorial Deep Learning Reveals Pan-Cancer Genomic Tumor Clusters with Distinct Immunogenomic Landscape and Response to Immunotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* (2020).

41. Yang, C.-F.J., *et al.* Role of Adjuvant Therapy in a Population-Based Cohort of Patients With Early-Stage Small-Cell Lung Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **34**, 1057-1064 (2016).
42. Li, G.H., Wu, Y., Zhang, X.J. & Cui, Y.S. [A comparative study of survival time of surgery combined with chemotherapy and non-surgical chemotherapy in SCLC]. *Zhonghua yi xue za zhi* **90**, 2212-2214 (2010).
43. Brock, M.V., *et al.* Surgical resection of limited disease small cell lung cancer in the new era of platinum chemotherapy: Its time has come. *The Journal of thoracic and cardiovascular surgery* **129**, 64-72 (2005).
44. Jin, K., *et al.* 从早期小细胞肺癌患者中选择以手术作为局部治疗的候选者: 基于人群的分析. *癌症*, 3 (2018).
45. Devarakonda, S., *et al.* Tumor Mutation Burden as a Biomarker in Resected Non-Small-Cell Lung Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **36**, 2995-3006 (2018).
46. Orr, H.A. Fitness and its role in evolutionary genetics. *Nature Reviews Genetics* **10**, 531-539 (2009).
47. Wagner, A.H., *et al.* Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. *Nat Commun* **9**, 3787 (2018).
48. Kim, R., Emi, M. & Tanabe, K. Cancer immunoediting from immune surveillance to immune escape. *Immunology* **121**, 1-14 (2007).
49. Alexandrov, L.B., *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421 (2013).
50. McGranahan, N., *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463-1469 (2016).
51. Hellmann, M.D., *et al.* Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. *Cancer cell* **33**, 853-861. e854 (2018).
52. Liu, L., *et al.* Combination of TMB and CNA Stratifies Prognostic and Predictive Responses to Immunotherapy Across Metastatic Cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **25**, 7413-7423 (2019).
53. Hutchinson, L. Biomarkers: Aneuploidy and immune evasion - a biomarker of response. *Nature reviews. Clinical oncology* **14**, 140 (2017).
54. Kawakami, M., Liu, X. & Dmitrovsky, E. New Cell Cycle Inhibitors Target Aneuploidy in Cancer Therapy. *Annual review of pharmacology and toxicology* **59**, 361-377 (2019).
55. Birkbak, N.J., *et al.* Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer research* **71**, 3447-3452 (2011).
56. Beroukhim, R., *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905 (2010).
57. Eymin, B. & Gazzeri, S. Role of cell cycle regulators in lung carcinogenesis. *Cell Adh Migr* **4**, 114-123 (2010).
58. Tian, H., *et al.* DNA damage response–a double-edged sword in cancer prevention and cancer therapy. *Cancer letters* **358**, 8-16 (2015).

59. Curtin, N.J. DNA repair dysregulation from cancer driver to therapeutic target. *Nature reviews. Cancer* **12**, 801-817 (2012).
60. Zhang, H.*, et al.* CDK7 Inhibition Potentiates Genome Instability Triggering Anti-tumor Immunity in Small Cell Lung Cancer. *Cancer Cell* **37**, 37-54.e39 (2020).
61. Daniel, D.*, et al.* Trilaciclib (T) decreases myelosuppression in extensive-stage small cell lung cancer (ES-SCLC) patients receiving first-line chemotherapy plus atezolizumab. *Annals of Oncology* **30**, v713 (2019).
62. Hu, X.*, et al.* Multi-region exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. *Nat Commun* **10**, 2978 (2019).
63. Cibulskis, K.*, et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology* **31**, 213-219 (2013).
64. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic acids research* **47**, D886-d894 (2019).
65. Favero, F.*, et al.* Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol* **26**, 64-70 (2015).
66. Felsenstein, J. *PHYLIP (phylogeny inference package), version 3.5 c*, (Joseph Felsenstein., 1993).
67. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S. & Swanton, C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol* **17**, 31-31 (2016).
68. Talevich, E., Shain, A.H., Botton, T. & Bastian, B.C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol* **12**, e1004873 (2016).
69. Lundegaard, C.*, et al.* NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic acids research* **36**, W509-512 (2008).
70. Lundegaard, C., Lund, O. & Nielsen, M. Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics (Oxford, England)* **24**, 1397-1398 (2008).
71. Nielsen, M.*, et al.* NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PloS one* **2**, e796 (2007).
72. Bai, Y., Wang, D. & Fury, W. PHLAT: Inference of High-Resolution HLA Types from RNA and Whole Exome Sequencing. *Methods in molecular biology (Clifton, N.J.)* **1802**, 193-201 (2018).
73. McGranahan, N.*, et al.* Allele-specific HLA loss and immune escape in lung cancer evolution. *Cell* **171**, 1259-1271. e1211 (2017).
74. Favero, F.*, et al.* Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Annals of oncology : official journal of the European Society for Medical Oncology* **26**, 64-70 (2015).
75. Roh, W.*, et al.* Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med* **9**(2017).

671

672

673 **Figure Legends**

674 **Figure 1. Phylogenetic trees of 18 SCLC tumors with multiregional whole exome**
675 **sequencing (WES).** Blue, brown and red lines represent trunk, branch, and private
676 mutations, respectively. The length of trunk, branch and private branch are proportional
677 to the numbers of mutations. Commonly mutated cancer genes *TP53, RB1* and are
678 mapped to the phylogenetic trees as indicated. Patient ID: pink = alive; green = expired.

679

680 **Figure 2. Genomic intra-tumor heterogeneity of small cell lung cancers. (A)**
681 Proportion of trunk (blue), branch (brown), and private (red) mutations representing
682 mutations detected in all tumor regions, some but not all and only in one single tumor
683 region from any given tumor. Purple patient IDs = patients who were alive; Green
684 patient IDs = patients who were decreased. **(B)** Proportion of clonal *versus* subclonal
685 mutations defined by PyClone in 50 SCLC tumor specimens. Mutations were classified
686 as clonal (estimated cancer cell fraction = 1, indicating mutations presenting in all
687 cancer cells, blue) or subclonal (estimated cancer cell fraction < 1, indicating mutations
688 only present in a subset of cancer cells, orange) in each tumor specimen.

689

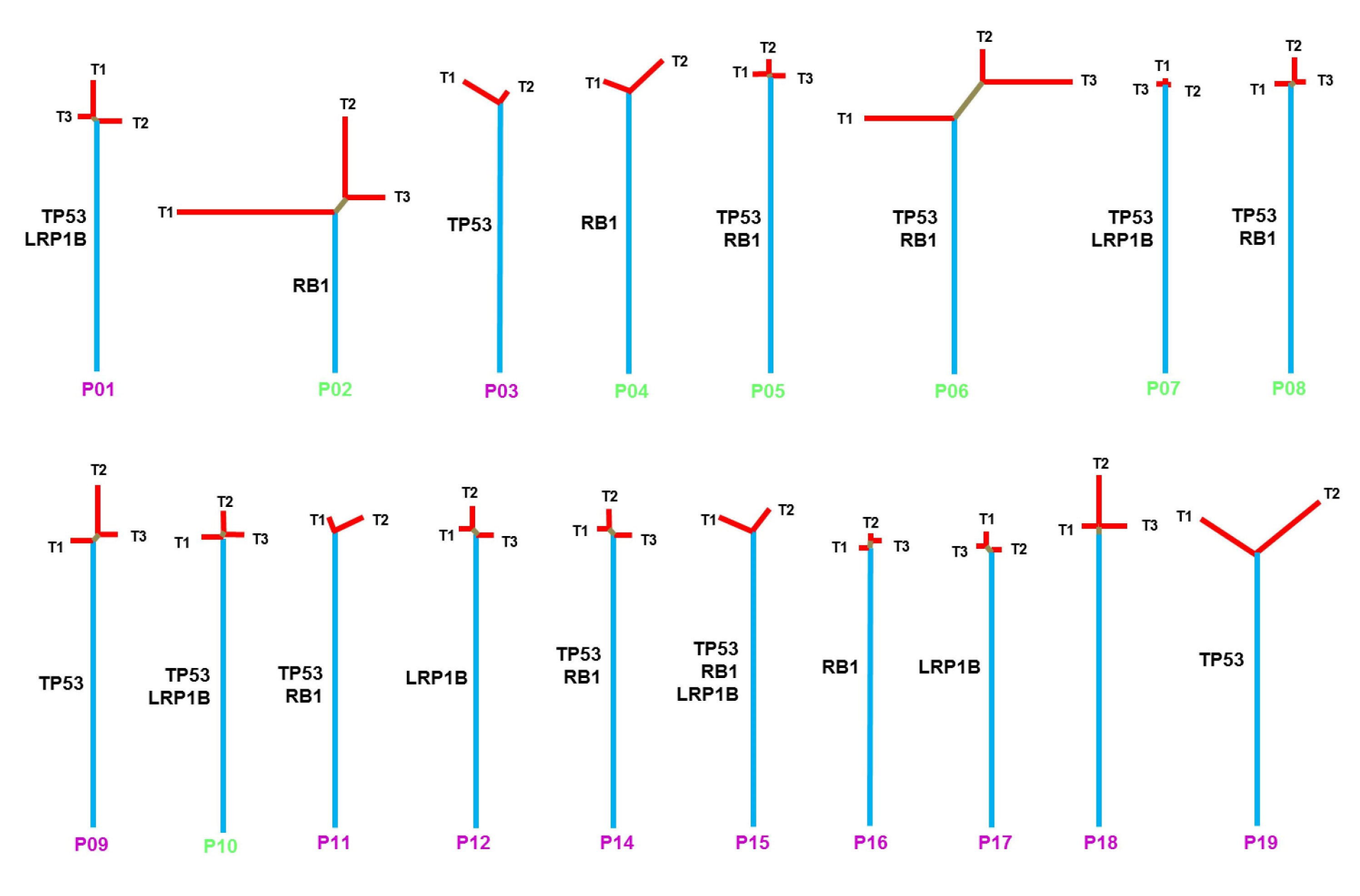
690 **Figure 3. The mutational processes in small cell lung cancers. (A)** The top
691 COSMIC mutational signatures in 50 SCLC specimens. **(B)** The top COSMIC
692 mutational signatures associated with trunk mutations. Bar chart on the left: top
693 signatures associated with trunk mutations in each patient. Pie chart on the right: the
694 average of contribution of each signature across the 19 patients. **(C)** The top COSMIC
695 mutational signatures associated with non-trunk mutations. Bar chart on the left: top
696 signatures associated with non-trunk mutations in each patient. Pie chart on the right:
697 the average of contribution of top signatures across the 19 patients. Patient ID: pink =
698 alive; green = expired.

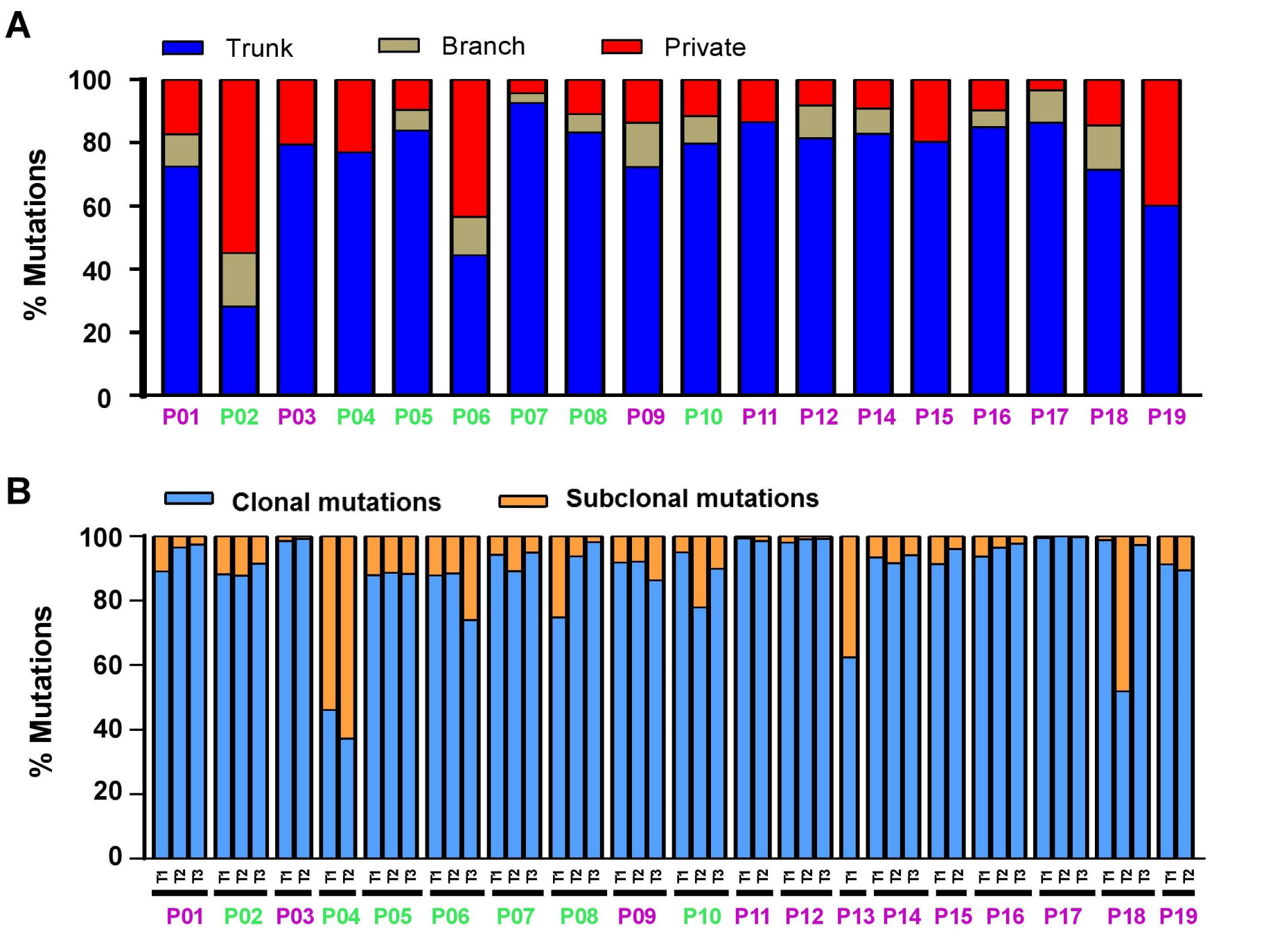
699

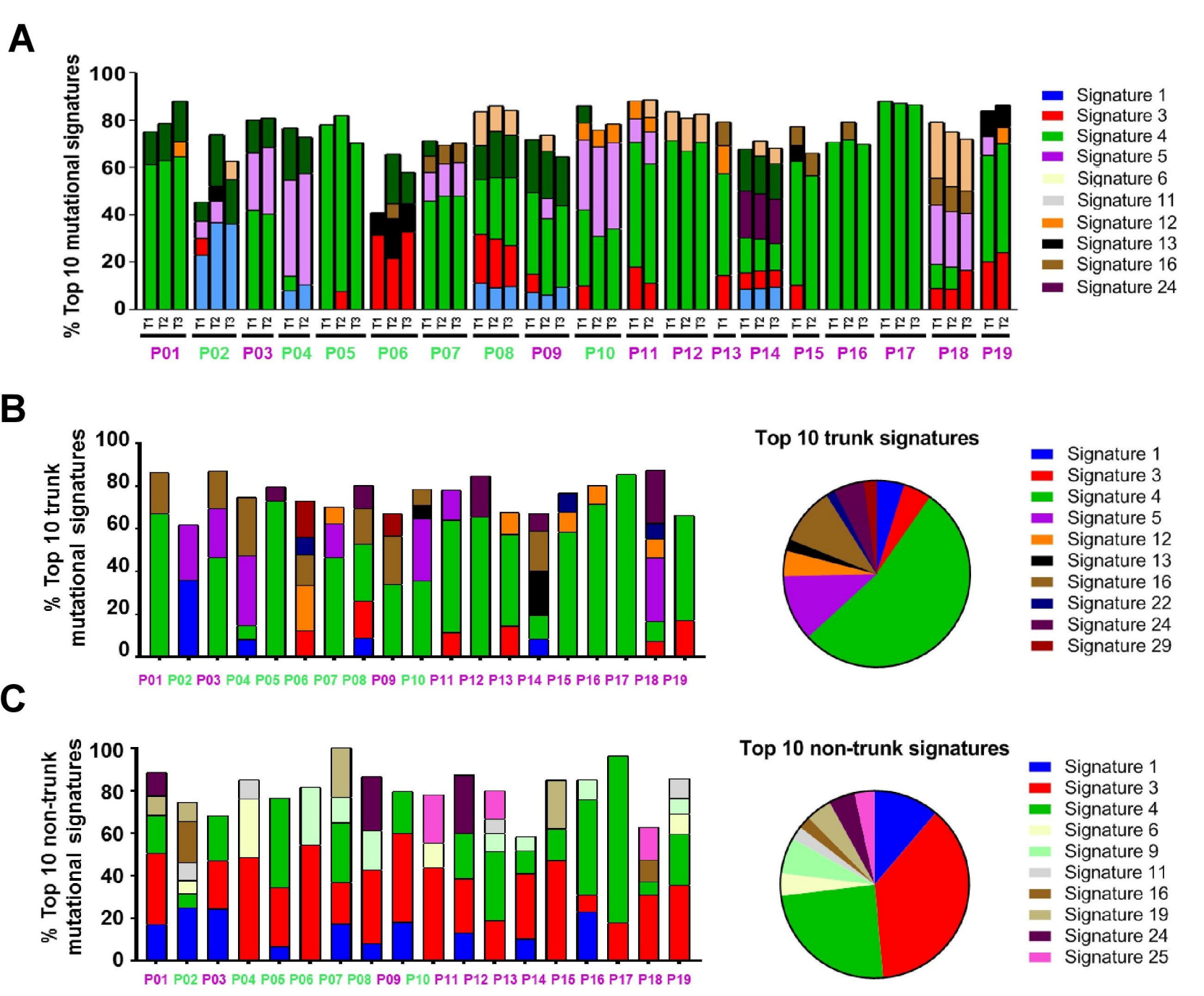
700 **Figure 4. Substantial TCR repertoire intratumor heterogeneity (ITH) in small cell**
701 **lung cancer. (A)** Quantification of T cell receptor (TCR) ITH by Jaccard index (JI), a

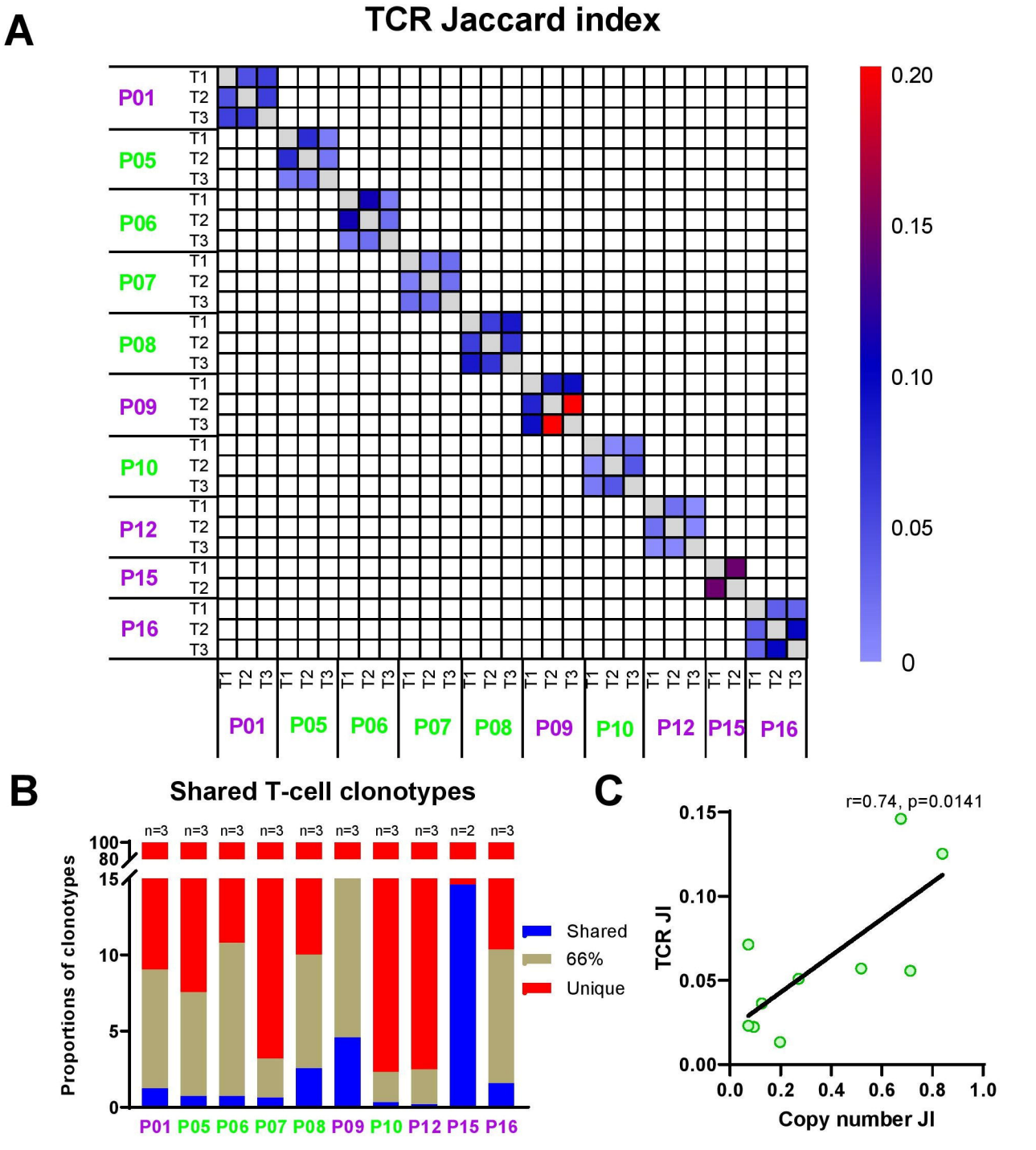
metric representing the proportion of shared T-cell clonotypes between two samples. **(B)** Proportions of T-cell clonotypes detected in all regions (shared, blue), in 2/3 (brown) and restricted to a single region (red) from the same tumors. Patient ID: pink = alive; green = expired. **(C)** Correlations between TCR ITH and TCR ITH by JI.

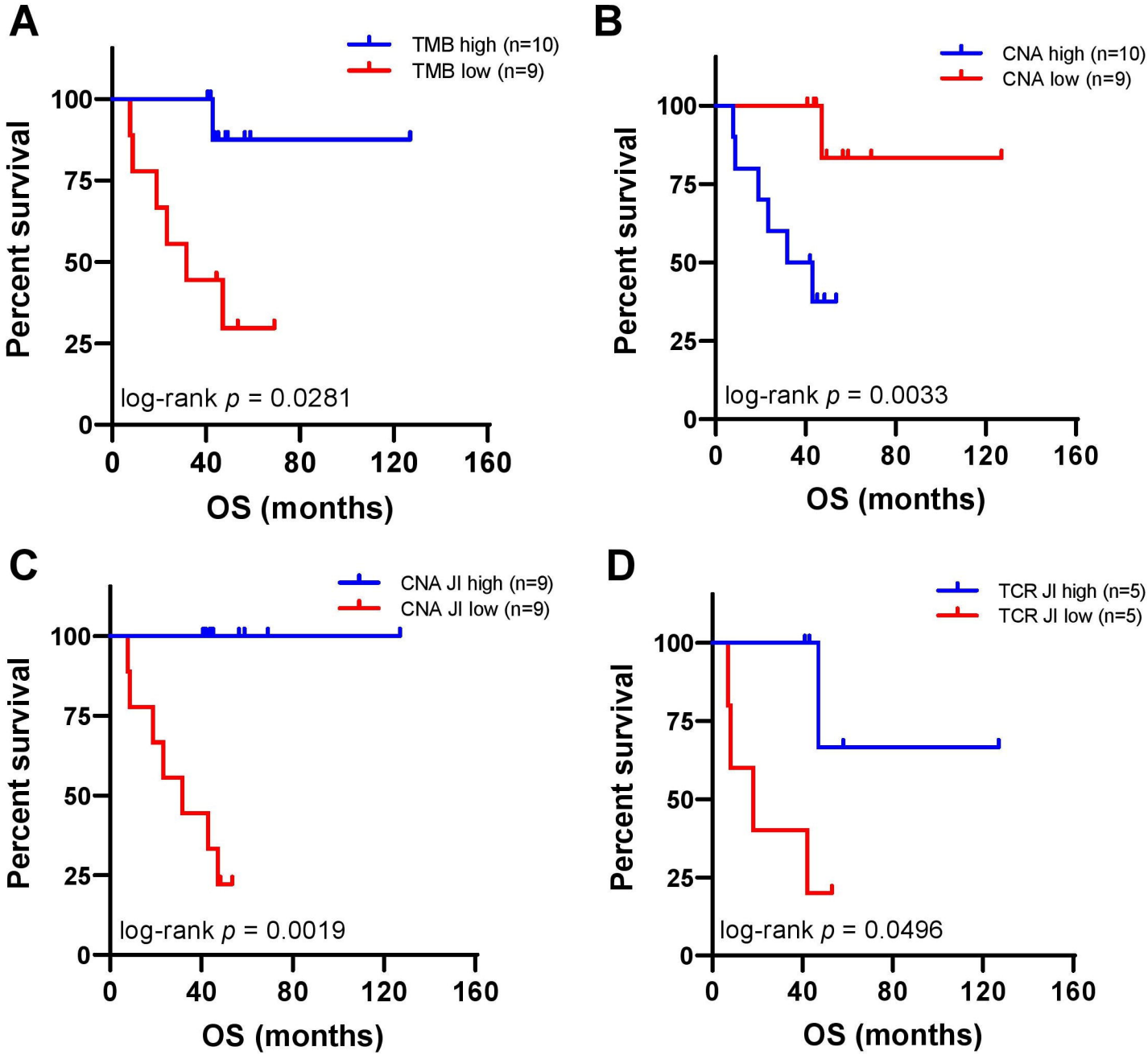
Figure 5. Association of overall survival (OS) with genomic and TCR landscape. **(A)** OS is longer in patients with higher (above median, blue) TMB than patients with lower (below median, red) TMB. **(B)** OS is shorter in patients with higher (above median, blue) CNA burden than patients with lower (below median, red) CNA burden. **(C)** OS is longer in patients with less CNA ITH (higher CNA JI, blue) than patients with higher level of CNA ITH (lower CNA JI, red). **(D)** OS is longer in patients with more homogenous TCR repertoire (higher above median TCR JI, blue) than patients with more heterogeneous TCR repertoire (lower below median TCR JI, red).


T1
T3
T2
TP53
LRP1B
P01
T2
T1
T3
RB1
P02
T1
T2
TP53
P03
T1
T2
RB1
P04
T2
T1
T3
TP53
RB1
P05
T2
T1
T3
TP53
RB1
P06
T1
T3
T2
TP53
LRP1B
P07
T2
T1
T3
TP53
RB1
P08
T2
T1
T3
TP53
P09
T2
T1
T3
TP53
LRP1B
P10
T1
T2
T3
TP53
RB1
P11
T2
T1
T3
LRP1B
P12
T2
T1
T3
TP53
RB1
P14
T2
T1
T3
TP53
RB1
LRP1B
P15
T2
T1
T3
RB1
P16
T1
T3
T2
LRP1B
P17
T2
T1
T3
P18
T1
T2
TP53
P19


A
Trunk
Branch
Private
% Mutations
100
80
60
40
20
0
P01 P02 P03 P04 P05 P06 P07 P08 P09 P10 P11 P12 P14 P15 P16 P17 P18 P19
B
Clonal mutations
Subclonal mutations
% Mutations
100
80
60
40
20
0
T1 T2 T3
P01 P02 P03 P04 P05 P06 P07 P08 P09 P10 P11 P12 P13 P14 P15 P16 P17 P18 P19


A
% Top 10 mutational signatures
100
80
60
40
20
0
T1 T2 T3
P01 P02 P03 P04 P05 P06 P07 P08 P09 P10 P11 P12 P13 P14 P15 P16 P17 P18 P19
Signature 1
Signature 3
Signature 4
Signature 5
Signature 6
Signature 11
Signature 12
Signature 13
Signature 16
Signature 24
B
% Top 10 trunk mutational signatures
P01 P02 P03 P04 P05 P06 P07 P08 P09 P10 P11 P12 P13 P14 P15 P16 P17 P18 P19
Top 10 trunk signatures
Signature 1
Signature 3
Signature 4
Signature 5
Signature 12
Signature 13
Signature 16
Signature 22
Signature 24
Signature 29
C
% Top 10 non-trunk mutational signatures
P01 P02 P03 P04 P05 P06 P07 P08 P09 P10 P11 P12 P13 P14 P15 P16 P17 P18 P19
Top 10 non-trunk signatures
Signature 1
Signature 3
Signature 4
Signature 6
Signature 9
Signature 11
Signature 16
Signature 19
Signature 24
Signature 25


A
TCR Jaccard index
P01
P05
P06
P07
P08
P09
P10
P12
P15
P16
T1
T2
T3
0.20
0.15
0.10
0.05
0
B
Shared T-cell clonotypes
n=3
n=2
100
80
15
10
5
0
Proportions of clonotypes
Shared
66%
Unique
C
r=0.74, p=0.0141
TCR JI
Copy number JI
0.15
0.10
0.05
0.00
0.0
0.2
0.4
0.6
0.8
1.0


A
TMB high (n=10)
TMB low (n=9)
Percent survival
100
75
50
25
0
log-rank p = 0.0281
0
40
80
120
160
OS (months)
B
CNA high (n=10)
CNA low (n=9)
Percent survival
100
75
50
25
0
log-rank p = 0.0033
0
40
80
120
160
OS (months)
C
CNA JI high (n=9)
CNA JI low (n=9)
Percent survival
100
75
50
25
0
log-rank p = 0.0019
0
40
80
120
160
OS (months)
D
TCR JI high (n=5)
TCR JI low (n=5)
Percent survival
100
75
50
25
0
log-rank p = 0.0496
0
40
80
120
160
OS (months)