

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

Ming Chen

Zhejiang Cancer Hospital <https://orcid.org/0000-0003-2570-0858>

Runzhe Chen

The University of Texas MD Anderson Cancer Center

Ying Jin

Zhejiang Cancer Hospital

Jun Li

MD Anderson Cancer Center

Jiexin Zhang

MD Anderson Cancer Center <https://orcid.org/0000-0001-9526-6730>

Junya Fujimoto

MD Anderson Cancer Center

Won-Chul Lee

MD Anderson Cancer Center

Xin Hu

MD Anderson Cancer Center

Shawna Hubert

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0002-7907-9609>

Bo Zhu

MD Anderson Cancer Center

Yanhua Tian

State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College

Julie George

University of Cologne <https://orcid.org/0000-0002-4272-3683>

Xiao Hu

Zhejiang Cancer Hospital

Yamei Chen

Zhejiang Cancer Hospital

Meijuan Wu

Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital)

Carmen Behrens

MD Anderson

Chi-Wan Chow

The University of Texas MD Anderson Cancer Center

Hoa Pham

Nagasaki University <https://orcid.org/0000-0002-0998-4486>

Junya Fukuoka

Toyama University Hospital/Nagasaki University

Jia Wu

MD Anderson Cancer Center

Edwin Roger Parra

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0001-9068-1636>

Carl Gay

University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0002-4907-0718>

Latasha Little

The University of Texas MD Anderson Cancer Center

Curtis Gumbs

The University of Texas MD Anderson Cancer Center

Xingzhi Song

The University of Texas MD Anderson Cancer Center

Lixia Diao

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0001-6995-4492>

Qi Wang

University of Texas M. D. Anderson Cancer Center <https://orcid.org/0000-0003-0858-9393>

Robert Cardnell

The University of Texas M D Anderson Cancer Center

Jianhua Zhang

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0001-5412-9860>

Jing Wang

UT M.D. Anderson Cancer Center <https://orcid.org/0000-0002-5398-0802>

Don Gibbons

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0003-2362-3094>

John Heymach

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0001-9068-8942>

J. Jack Lee

The University of Texas M D Anderson Cancer Center <https://orcid.org/0000-0001-5469-9214>

William William Jr.

The University of Texas MD Anderson Cancer Center

Bonnie Glisson

MD Anderson Cancer Center

Ignacio Wistuba

M. D. Anderson Cancer Center

Andrew Futreal

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0001-8663-2671>

Roman Thomas

University of Cologne

Alexandre Reuben

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0003-4510-0382>

Lauren Byers

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0002-0780-2677>

Jianjun Zhang (✉ jzhang20@mdanderson.org)

The University of Texas MD Anderson Cancer Center <https://orcid.org/0000-0001-7872-3477>

Article

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

Posted Date: January 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-148314/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

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, using multi-region whole exome/T cell receptor (TCR) sequencing as well as immunohistochemistry (IHC), we revealed a rather homogeneous mutational landscape but extremely cold and heterogeneous TCR repertoire in limited-stage SCLC (LS-SCLC) tumors. Compared to localized non-small cell lung cancers (NSCLCs), LS-SCLCs had similar predicted neoantigen burden and genomic ITH, but significantly colder and more heterogeneous TCR repertoire associated with higher chromosomal copy number aberration (CNA) burden. Higher mutational burden, higher T cell infiltration and positive PD-L1 expression were associated with longer overall survival (OS), while higher CNA burden were associated with shorter OS in patients with LS-SCLC.

Introduction

Small-cell lung cancer (SCLC) accounts for ~15% of all newly diagnosed lung cancers leading to ~30,000 deaths in the United States annually¹. 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 (LS) that can be treated with chemotherapy combined with radiotherapy or surgical resection, while the remaining patients present with extensive-stage disease (ES) 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) below 7% for ES-SCLC^{1,7,8}. 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 survival⁶. National Cancer Institute (NCI) has identified SCLC as a recalcitrant malignancy¹. Translational studies to understand the mechanisms underlying recurrence and therapeutic resistance remain an unmet need to design novel therapeutic strategies.

Tumors are composed of cancer cells and stromal cells of distinct molecular and phenotypic features, a phenomenon termed intratumor heterogeneity (ITH). Evolutionary theory suggests populations of high genetic variation have survival advantages⁹. Similarly, tumors of complex ITH may be difficult to eradicate. Higher levels of molecular ITH have been demonstrated to associate with inferior outcome of cancer patients¹⁰⁻¹². 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^{10,12-17}. It has been speculated that SCLC has an extremely complex ITH architecture that lead to poor prognosis¹⁸. Another plausible explanation for the poor outcome is that SCLC is associated with an immunosuppressive tumor

microenvironment, particularly T cell responses^{19,20}. In localized NSCLC, our recent work has revealed that a cold and heterogeneous T cell receptor (TCR) repertoire is associated with inferior survival^{13,21}.

The genomic and TCR ITH architecture of SCLC and their potential clinical impact have not been well studied, largely due to lack of adequate tumor specimens²²⁻²⁵ as surgical resection is not the standard of care for the majority of SCLC patients. Through international collaboration, we conducted multi-region whole exome sequencing (WES) and TCR sequencing of 50 tumor samples from 19 resected LS-SCLCs with adequate tissues available (**Supplementary Fig. 1**) to depict the immunogenomic ITH architecture of SCLC. We then compared these LS-SCLCs to the cohorts 100 early-stage NSCLCs with multi-region WES data available (TRACERx cohort)¹⁴ as well as 242 localized NSCLCs (PROSPECT cohort) with WES and TCR repertoire data available²¹. We further conducted immunohistochemistry (IHC) to validate our findings in TCR repertoire on 66 LS-SCLCs and 68 localized NSCLC tumors (**Supplementary Data 1**) matched for clinical characteristics (**Supplementary Table 1**) and assessed the impact of immunogenomic attributes on patient survival.

Results

Overall homogeneous genomic landscape in SCLC

A total of 50 spatially separated tumor regions (hereafter referred as region) from 19 resected LS-SCLC tumors (hereafter referred as tumor) with adequate tissue available were subjected to WES (**Supplementary Fig. 1, Supplementary Fig. 2, Supplementary Data 1**). All patients underwent upfront surgery without preoperative chemotherapy or radiation. In total, 3,773 nonsilent (nonsynonymous, frameshift, stop-gain and stop-loss) mutations were identified from these 50 tumor regions (**Supplementary Data 2**) with a median nonsilent tumor mutational burden (TMB) of 4.69/Mb (interquartile (25th-75th) range: 2.66-6.43/Mb). TMB varied substantially between patients, but was similar between different regions within the same tumors (**Supplementary Fig. 3a**).

We constructed phylogenetic trees of 18 SCLCs for which multi-region WES data was 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 previously¹². A median of 80.1% (28%-93%) (interquartile (25th-75th) range: 72%-84%) 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 tumors, compared to 72% (8%-99.6%) (interquartile (25th-75th) range: 55%-84%) trunk mutations in 100 early-stage NSCLC in TRACERx cohort ($p=0.218$)¹⁴. Considering the number of regions per tumor may introduce bias into the proportion of trunk mutations, we also calculated the average mutational Jaccard index (JI) between each pair of specimens within the same tumors in 18 SCLCs (**Supplementary Fig. 4a**) and compared that with NSCLC tumors from the TRACERx cohort. The results demonstrated that genomic JI was comparable in the cohort of SCLC and NSCLC (**Supplementary Fig. 4b**, $p=0.6058$).

Next, we assessed the chromosomal copy number aberrations (CNAs) in this cohort of LS-SCLC using a gene-based CNA analysis algorithm²⁶ for WES data that compares the CNA burden between different samples. The TMB, CNA burden was similar between different regions within the same tumors whilst varied substantially between different patients (**Supplementary Fig. 3b**). The median CNA JI was 0.50 (range: 0.02 to 0.93) (interquartile (25th-75th) range: 0.27-0.85) in these 18 SCLC tumors with multi-region WES data available (**Supplementary Fig. 4c**) significantly higher than that of NSCLC tumors in TRACERx cohort (median: 0.21, range: 0 to 0.99) (interquartile (25th-75th) range: 0.06-0.35, **Supplementary Fig. 4d**).

Furthermore, using PyClone²⁷, we classified mutations as clonal, defined as estimated cancer cell fraction (CCF) =1 (indicating mutations present in all cancer cells) or subclonal, defined as estimated CCF<1 (indicating mutations only present in a subset of cancer cells) in each tumor specimen after adjusting for tumor purity and copy number changes. A median of 92.8% (37%-99.9%) (interquartile (25th-75th) range: 88%-97.8%) of mutations were clonal in these 50 SCLC tumor regions (**Figure 2b**). Taken together, these data suggest the genomic landscape is overall homogeneous in these SCLCs, comparable to NSCLCs.

A total of 134 cancer gene alterations were identified in these tumors including 84 point mutations, 8 copy number gains of oncogenes and 42 copy number loss of tumor suppressors (**Supplementary Fig. 5**). As expected, the most frequently altered cancer genes were *RB1* and *TP53*, identified in 16 and 15 of the 19 patients, respectively (**Supplementary Fig. 5**). Of note, 86 of 134 cancer gene alterations were detected in all regions within the same tumors (**Figure 1** and **Supplementary Fig. 5**) suggesting these cancer gene alterations were early genomic events during evolution of SCLC and the genomic landscape of these SCLC tumors was homogeneous not only quantitatively, but also qualitatively in key cancer genes alterations.

Different mutational processes associated with early clonal mutagenesis versus later subclonal mutagenesis in SCLC

Understanding how mutational processes shape cancer evolution may inform mechanisms underlying tumor adaptation²⁸. We next analyzed the mutational spectrum and signatures in these SCLCs²⁹. C>A transversions were the most common nucleotide substitutions (**Supplementary Fig. 6**) and Cosmic Signature 4 (associated with cigarette smoking) was the predominant mutational signature (**Supplementary Fig. 7a**) as expected, given all 19 patients were smokers or heavy second-hand 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, respectively²⁸. As shown in **Supplementary Fig. 7b-c**, among the top 5 mutational signatures, Cosmic Signature 4 remained as the predominant signature in trunk mutations contributing to 62% of the top 5 signatures, consistent with previous reports that smoking associated mutational processes play critical roles during early

mutagenesis of lung cancers^{12,14,30}. On the other hand, the contribution of Signature 4 was significantly reduced (29% for non-trunk mutational signatures versus 62% for trunk mutational signatures among the top 5 signatures, $p=0.002$), while Cosmic Signature 3 (associated with defect of DNA double-strand break-repair) emerged as the predominant signature for non-trunk mutations (45% for non-trunk mutational signatures versus 6% for trunk mutational signatures among the top 5 signatures, $p<0.0001$). Taken together, these results highlight the dynamic nature of mutagenesis at different time points 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 defects.

Cold TCR repertoire in SCLC tumors

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% (interquartile (25th-75th) range: 0.57%-3.97%) (**Supplementary Fig. 8a**). T-cell richness, a measure of T-cell diversity, ranged from 38 to 8,286 unique T-cells (median: 510) (interquartile (25th-75th) range: 181-1034) per specimen (**Supplementary Fig. 8b**) and T-cell clonality, a metric indicating T-cell expansion and reactivity, ranged from 0.002 to 0.139 (median=0.009) (interquartile (25th-75th) range: 0.005-0.0025) (**Supplementary Fig. 8c**). 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; **Supplementary Fig. 8d-f**), indicating a reduced T cell infiltration, diversity and proliferation in tumor tissues similar to NSCLC²¹. As shown in **Supplementary Fig. 8**, P13 is clearly an outlier in TCR parameters. We therefore reran this analysis excluding P13, which didn't lead to significant changes (**Supplementary Fig. 9**). Of particular interest, compared with the TCR repertoire data generated using the same method from NSCLCs (the PROSPECT cohort), all three TCR metrics were significantly lower than those from NSCLCs (T-cell density 0.014 vs. 0.21, $p<0.0001$ (**Figure 3a**); diversity 510 vs. 3,246, $p<0.0001$ (**Figure 3b**); and T-cell clonality 0.009 vs. 0.14, $p<0.0001$ (**Figure 3c**)²¹. Similarly, the total TCR templates were also significantly lower in SCLC tumors compared to NSCLC tumors (**Supplementary Fig. 10**).

Higher levels of immune cell infiltration can lead to low tumor purity. *Vice versa*, lower tumor purity often indicates a higher level of immune cell infiltration. We estimated tumor purity by Sequenza³¹ in this cohort of LS-SCLCs and compared that to localized NSCLCs from the PROSPECT cohort. As shown in **Figure 4a**, significantly higher tumor purity was observed in these SCLC compared with NSCLC from PROSEPT cohort supporting possible lower immune infiltration in SCLC.

Furthermore, we derived the immune score by quantifying the density of all immune cells within tumors and estimated the fraction of T cell infiltration by deconvoluting previously published RNA sequencing (RNA-seq) data of 81 SCLCs²³ and compared those to 1,027 NSCLCs from TCGA^{30,32}. In line with TCR

repertoire findings, both immune scores and estimated T-cell fractions were significantly lower in SCLCs than NSCLCs ($p < 0.0001$, $p < 0.0001$, respectively; **Figure 4b-c**).

To orthogonally assess tumor infiltrated T cells, we performed IHC by T cell markers CD3 and PD-L1 (**Supplementary Fig. 11**) on 66 SCLC tumors including the same 18/19 tumors in the current study (one tumor tissue was exhausted) and 68 NSCLC tumors (**Supplementary Data 1**) with matched clinical characteristics including gender, age, smoking status and tumor size (**Supplementary Table 1**). As shown in **Figure 4d**, SCLC tumors had significantly lower T cell infiltration than NSCLC tumors ($p < 0.0001$). Importantly, the T cell infiltration was negatively associated with tumor purity ($r = -0.47$, $p = 0.0405$, **Figure 4e**) indicating lower T cell infiltration was one potential reason underlying observed higher tumor purity in SCLCs than NSCLCs. Moreover, the PD-L1 expression was also significantly lower in SCLCs than NSCLCs (**Figure 4f**, $p < 0.0001$). Taken together, these data suggest that SCLC had an overall lower T cell infiltration rather than higher level of adaptive immunosuppression compared to NSCLCs.

Substantial TCR repertoire heterogeneity in SCLC

To gain further insights into TCR heterogeneity, we calculated the Jaccard index (JI), a metric measuring the proportion of shared T-cell clonotypes between two samples. Substantial TCR heterogeneity was evident across all SCLC tumors, with a median JI of 0.05 (0.02 to 0.15) in the 10 SCLCs with multi-region TCR data available (**Figure 5a**), significantly lower than the 11 localized NSCLCs¹³ (median 0.05 in SCLC vs. 0.16 in NSCLC, $p < 0.0001$) (**Figure 3d**) with multi-region TCR data available¹³. Furthermore, 79.9%-97.7% of T-cell clones were restricted to individual tumor regions while only 0.2%-14.6% (median: 0.98%, interquartile (25th-75th) range: 0.54%-3.06%) were identified in all regions within the same tumors (**Figure 5b**), significantly lower than NSCLC (median: 5.7%, range: 1.6% to 14.5%) (interquartile (25th-75th) range: 3.6%-7.4%) ($p = 0.0048$)¹³ demonstrating profound TCR ITH in SCLC even beyond NSCLC. Of note, TCR JI was positively correlated with T-cell density, richness and clonality although some correlations did not reach statistical significance (**Supplementary Fig. 12a-c**).

To compare shared TCR clonotypes considering T cell expansion, we further used another two TCR metrics: Morisita index (MOI), a metric taking into consideration of not only the composition of T cell clonotypes but also the abundance of individual T cell clonotypes, and the proportion of shared top (most expanded/abundant) TCR clonotypes. Similar to TCR JI, compared to the NSCLCs¹³, SCLC tumors showed significantly lower TCR MOI and shared top TCR clonotypes (**Figure 3e-f**) suggesting higher level of TCR ITH in SCLC tumors for not only the overall TCR repertoire but also for the most expanded/active T cells.

High copy number burden associated with cold TCR repertoire in SCLC

To identify genomic aberrations that could associate with the cold TCR repertoire 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^{33,34}. We performed *in silico* prediction of human leukocyte antigen (HLA)-A-, -B-, and -C-presented neoantigens. A median of 78 (26-463) (interquartile (25th-75th) range: 67-123) predicted neoantigens ($IC_{50} < 500$ nmol/L) per tumor were detected (**Supplementary Fig. 13a**), which was similar to NSCLCs from the PROSPECT cohort (median: 72/tumor, 2-801, interquartile (25th-75th) range: 35-125, $p=0.31$). Similar to somatic mutations, 81% (43%-93%, interquartile (25th-75th) range: 64%-86%) of predicted neoantigens were present across different regions within the same tumors (**Supplementary Fig. 13b**), and 90% of predicted neoantigens (51%-100%, interquartile (25th-75th) range: 81%-94%) were associated with clonal mutations in individual tumor regions (**Supplementary Fig. 13c**). These results suggest that the cold TCR repertoire in SCLC was not likely to be due to the low clonal neoantigen burden.

As higher chromosomal copy number aberration (CNA) burden has been reported to associate with cold immune microenvironment in multiple cancer types^{35,36}, we next assessed whether the CNA burden is associated with the observed cold TCR repertoire in SCLC. A median of 2,310 CNA events per tumor (26-8044, interquartile (25th-75th) range: 444-4318) were identified from these SCLCs (**Figure 6a**), significantly higher than 111 per tumor (range: 0-7741, interquartile (25th-75th) range: 0-561) in NSCLCs from the PROSPECT cohort ($p < 0.0001$)²¹, which have shown significantly hotter TCR repertoire than this cohort of SCLCs (**Figure 3a-c**). Importantly, the CNA burden was negatively associated with T-cell density, richness and clonality in this cohort of SCLCs (**Figure 6a-c**, $r = -0.40$, $p = 0.0157$; $r = -0.36$, $p = 0.0321$; $r = -0.33$, $p = 0.0490$; respectively) suggesting that higher CNA burden could be important genomic alterations associated with profoundly cold TCR repertoire in these SCLCs.

Next, we explored possible links between CNA and immune evasion. One potential impact of CNA on immune evasion is loss of neoantigens^{35,37}. We therefore investigated whether predicted neoantigen-associated genes were more frequently lost. With the caveat of imperfect performance of neoantigen prediction algorithms, a significantly higher fraction of copy number loss was observed in predicted neoantigen-associated genes than that in other mutated genes ($p = 0.0019$) (**Figure 6d**). We further calculated the fraction of mutated genes showing neutral LOH (CNN-LOH) or copy number loss LOH (CNL-LOH). Similar to that for overall copy number losses, both CNN-LOH and CNL-LOH were significantly more common in predicted neoantigen-associated genes than mutated genes that not associated with neoantigens (**Figure 6e-f**). These results suggest loss of neoantigen via copy number loss could contribute to underlying cold immune microenvironment in SCLC.

Different genomic alterations associated with cold immune repertoire in different tumor regions within the same tumors

Evidence of loss of heterozygosity (LOH) of HLA, a potential immune evasion mechanism in cancer^{38,39} was revealed in 9 of 19 SCLC tumors, numerically higher than in NSCLCs (9/19 vs. 60/216, $p=0.11$) suggesting there was once strong immune pressure during carcinogenesis of these SCLC and HLA loss could be a common genomic alteration underlying immune evasion in SCLC. Eventually, these SCLC tumors were able to survive the immune pressure leading to overall cold immune infiltration as evidenced by low T cell infiltration and low T cell expansion. Interestingly, among the 9 SCLC tumors, 5 tumors had HLA LOH in all regions (P04, 09, 17, 18, 19), while 4 tumors had HLA LOH in 1 or 2 regions (P01, 02, 07, 14). Furthermore, in P09, all 3 TCR metrics were much lower in region T1 compared to regions T2 and T3 (**Supplementary Fig. 14**) reflecting ITH of T cell response. Of note, the TMB was almost identical among these 3 tumor regions. However, T1 had significantly higher fraction of predicted neoantigens associated with copy number loss than T2 or T3 (6/115 predicted neoantigens with copy number loss in T1 versus 0/124 in T2 and 0/129 in T3, $p=0.00085$) (**Supplementary Table 2**). This is in line with the overall cohort that a significantly higher fraction of predicted neoantigens were associated with copy number loss. Taken together, these data indicate that different immune escaping mechanisms could be in play not only between different patients but also in different subclones within the same tumors.

Genomic and TCR ITH associated with survival in patients with SCLC

We next attempted to assess whether the genomic and T cell features impact the clinical outcome. Given the small sample size and the fact that recurrence status was unavailable for some patients, for the 19 SCLC patients with WES and TCR data available, we focused on only overall survival (OS). Higher TMB was associated with significantly longer OS (**Figure 7a**, HR=0.85, $p=0.036$), consistent with previous reports in NSCLC⁴⁰. Conversely, significantly shorter OS was observed in patients with higher CNA burden (**Figure 7b**, HR=1.21, $p=0.014$). No TCR parameter (T-cell density, richness and clonality) was associated with OS (**Supplementary Fig. 12d-f**), however, patients with a more heterogeneous TCR repertoire (lower TCR JI) exhibited significantly shorter OS (**Figure 7c**, HR=0.87, $p=0.037$).

Interestingly, lower tumor purity, presumably with more immune infiltration was associated with longer OS (**Figure 7d**, HR=1.30, $p=0.001$). This finding is further supported by the IHC data on the 66 SCLC patients, where significantly longer OS was observed in patients with higher level of CD3 (**Figure 7e**, HR=0.32, $p=0.0167$) or positive PD-L1 expression (**Figure 7f**, HR=0.35, $p=0.0249$). Prolonged DFS was also observed in patients with higher CD3 infiltration (**Supplementary Fig. 15a**, HR=0.30, $p=0.0059$) or positive PD-L1 expression (**Supplementary Fig. 15b**, HR=0.54, $p=0.1411$) in their tumors suggesting increased intratumor T cell infiltration was associated with superior disease control and survival in these SCLC patients.

Discussion

Although pioneering studies have revealed pivotal molecular features^{22-25,41}, the genomic ITH architecture of SCLC has not been defined, primarily due to the lack of adequate tumor specimens for multi-region 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^{7,42}. Surprisingly, all SCLCs in the current study demonstrated homogeneous genomic landscape for both mutations and CNAs. Additionally, previous work from our group⁴³ and others⁴¹ has demonstrated striking similarity of the mutational landscape between primary and relapsed SCLC. Taken together, these data indicate that complex genomic 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 surveillance⁴⁴. Our study delineates for the first time, the intratumor TCR repertoire of SCLC and demonstrates an extremely cold TCR repertoire quantitatively (density) and qualitatively (richness and clonality), compared to not only matched normal lung tissues but also compared to NSCLC tumors despite the fact SCLC and NSCLC tumors had similar TMB and comparable mutational homogeneity. Comparing the immune score and estimated T-cell fraction by deconvolution of transcriptomic data from a previously published larger SCLC cohort (n=81)²³ to TCGA NSCLC cohorts (n=1,027) also revealed significantly lower immune infiltration in SCLCs than NSCLCs. These findings were further validated orthogonally by IHC using anti-CD3 on 66 SCLC tumors and 68 NSCLC tumors matched for clinicopathological characteristics. Furthermore, IHC data also revealed significantly lower PD-L1 expression in SCLC than NSCLC tumors. Taken together, these results suggested an overall cold immune infiltration rather than adaptive immune suppression in SCLC tumors.

In addition to the cold intratumor TCR repertoire, SCLC tumors also demonstrated an 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 NSCLC¹³, which may further impair the efficacy of anti-tumor immune response. SCLC is among the cancers with high TMB²⁸ and our study also demonstrated a homogenous mutational landscape, both of which have been reported to associate with benefit from ICB⁴⁵. However, compared to NSCLC and other tumor types, fewer SCLC patients benefit from ICB⁴⁶. The cold and heterogeneous TCR repertoire may be one potential reason underlying suboptimal response to immunotherapy.

In subsequent attempts to identify genomic features that may account for the cold TCR repertoire in SCLC, a significantly higher CNA burden was observed in SCLC compared with that in NSCLC. Moreover, the CNA burden was negatively associated with both T cell quantity (density) and quality (richness and clonality). These results suggest that high CNA burden may be one of the genomic changes underlying the cold TCR repertoire in SCLC. High CNA burden has been reported to correlate with cold microenvironment and inferior benefit from ICB across different cancer types^{35,47,48}. The mechanisms

underlying the association between high CNA burden and immunosuppression are not well understood. In this cohort of SCLC tumors, genes associated with predicted neoantigens were observed to be more frequently lost than other mutated genes not associated with neoantigens indicating loss of neoantigens through chromosomal aberrations could be one potential mechanism underlying immune evasion in SCLC. Functional studies are eventually warranted to determine how exactly CNA impacts anti-tumor immune response.

Compared to many malignancies, our understanding of the molecular landscape of SCLC is rudimentary, primarily due to the lack of adequate tumor tissues as the majority of SCLC patients are not treated with surgical resection. As such, our study was limited by the small sample size and therefore, these intriguing findings have to be validated by larger cohorts in future studies that may need multi-institutional collaboration. Nevertheless, WES and TCR data from multiregional specimens make these data unique to understand the ITH architecture of SCLC tumors. Additionally, one of the main findings regarding cold T cell infiltration was validated orthogonally by IHC on a larger cohort of SCLC tumors and NSCLC tumors matched for clinicopathologic features. Another potential limitation is that it is unclear whether the biology of these resected early-stage SCLC tumors resemble that of advanced-stage SCLCs. On the other hand, we⁴³ and Wagner *et al*⁴¹ have demonstrated similar genomic landscape between treatment-naïve resected SCLC tumors and recurrent disease suggesting that some of the findings in the current study may be applicable to advanced SCLCs.

To the best of our knowledge, this is the first study on genomic and TCR ITH of SCLC. In summary, we demonstrate that despite a homogeneous genomic landscape, SCLC exhibits a cold and heterogeneous T cell infiltration that could potentially lead to ineffective anti-tumor immune surveillance. From a therapeutic standpoint, these findings demand further investigation into approaches capable of overcoming the cold and heterogeneous intratumor T cell repertoire to improve the efficacy of immunotherapy for SCLC patients. Enhancement of T cell trafficking into the SCLC tumors as well as within the tumors may have the potential to promote T cell infiltration and make the immune infiltrate more consistent across the tumors, therefore generating more effective antitumor responses. T cell trafficking is a delicate and complicated process that can be impaired by tumor vascular structure⁴⁹, aberrant tumor extracellular matrix architecture⁵⁰, presence of myeloid-derived suppressor cells (MDSCs)⁵¹, hypoxia⁵², lack of T cell recruitment associated chemokines such as CCL5 and CXCL9⁵³ etc. Targeting these pathways to reprogram T cell motility may provide opportunities to improve efficacy of immunotherapy in tumors with cold and heterogeneous immune infiltration such as SCLC.

Methods

Patients

A total of 67 patients with LS SCLC who underwent surgical resection at Zhejiang Cancer Hospital, Hangzhou, China from 2008 to 2020 were enrolled. All the enrolled patients underwent upfront surgery without preoperative chemotherapy or radiation and no patients received immunotherapy prior to or post

surgery. As a control group, 68 patients with localized NSCLC, who underwent upfront surgery without preoperative chemotherapy or radiation were included. The clinical information of all 135 patients is included in the **Supplementary Data 1**. The SCLC group and NSCLC group were matched for clinical characteristics including gender, age, smoking status and tumor size (**Supplementary Table 1**). 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 (H&E) 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. An H&E slide from each tumor region was first reviewed by experienced lung cancer pathologists to assess the percentage of tumor versus adjacent normal tissues. Three regions from each tumor FFPE block, were collected by a “grid” approach representing the spatial heterogeneity of the primary tumors. Only tumor regions with enough tumor cells were selected for DNA extraction and sequencing (**Supplementary Fig. 2**). DNA was extracted using the AllPrep® DNA/RNA FFPE Kit (Qiagen, Hilden, Germany) from 50 spatially separated tumor regions from 19 LS-SCLC tumors with adequate tissues available (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 described⁵⁴.

Whole exome sequencing (WES)

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)⁵⁵ 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 $\geq 20\times$ in tumor DNA and $\geq 10\times$ in germline DNA; and 2) variant allele 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 ⁵⁶.

Clonal and subclonal analysis

Tumor contents and major/minor copy number changes were estimated by Sequenza (v2.1.2)³¹. The cancer cell fraction (CCF) and mutant allele copy number for each SNV was inferred using PyClone 12.3²⁷. 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 phylogenetic tree⁵⁷.

Mutational signature analysis

The R package “DeconstructSigs” package⁵⁸ 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)⁵⁹, 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. CNA burden was quantified using a gene-based CNA analysis algorithm²⁶ for exome sequencing data that allows comparing the CNA between different samples.

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 alleles⁶⁰⁻⁶². Candidate peptides were considered HLA binders when IC50<500 nM.

TCR β sequencing and comparison parameters

Immunosequencing of the CDR3 regions of human TCR β chains was performed using the protocol of ImmunoSeq (Adaptive Biotechnologies, hsTCR β Kit) ⁶³⁻⁶⁵. Two sets of PCRs were performed on DNA extracted from the tissues collected. The initial PCR used a mix of multiplexed V- and J-gene primers which amplify all possible recombined receptor sequences from the DNA sample, and then a second PCR designing to add unique DNA barcodes to each PCR product was followed. After that, samples were pooled together with a negative and a positive control. The pools were then sequenced on an Illumina MiSeq platform using a 150 cycles paired end protocol and sequence-ready primers. After sequencing, the raw data were transferred to Adaptive Biotechnologies and processed into a report including those passed quality-check samples and a normalized and annotated TCR β profile repertoire accordingly.

T-cell density was calculated by normalizing TCR- β template counts to the total amount of DNA usable for TCR sequencing, where the amount of usable DNA was determined by PCR-amplification and sequencing of housekeeping genes expected to be present in all nucleated cells. T-cell richness is a metric of T cell diversity, and it is calculated by on the T-cell unique rearrangements. T-cell clonality is a metric of T cell proliferation and reactivity, and it is defined as 1-Peilon's evenness and is calculated on productive rearrangements by:

$$1 + \frac{\sum_i^N p_i \log_2(p_i)}{\log_2(N)}$$

where p_i is the proportional abundance of rearrangement i , and N is the total number of rearrangements. Clonality ranges from 0 to 1: values approaching 0 indicate a very even distribution of frequency of different clones (polyclonal), whereas values approaching 1 indicate a distinct asymmetric distribution in which a few activated clones are present at high frequencies (monoclonal). Statistical analysis was performed in R version 3.2. The immunoSEQ Assay is for research use only and not for use in diagnostic procedures. TCR Jaccard index (JI) is conceptually a percentage of how many objects of two sets have in common out of how many objects they have in total. JI was calculated by the number of rearrangements shared/sum of total number of rearrangements between any two specimens. Morisita index (MOI) is a measure of the similarity in the T cell repertoire between samples ranging from 0 to 1, taking into account the specific rearrangements and their respective frequencies, with an MOI of 1 being an identical T cell repertoire.

Human leukocyte antigen loss of heterozygosity analysis

For Human Leukocyte Antigen Loss Of Heterozygosity (HLA LOH) analysis, we first performed HLA typing using PHLAT⁶⁶. 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, LOHHLA³⁸ using purity and ploidy information estimated by Sequenza³¹. 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 $p < 0.01$.

Copy number LOH analysis

For each sample, we calculated the fraction of copy number losses for those genes associated with predicted neoantigens versus other mutated genes not associated with predicted neoantigens, respectively. The segment copy number file (predicted by Sequenza³¹) was used to evaluate whether a gene showing evidence of copy-number-neutral LOH (CNN-LOH) or copy-number-loss LOH (CNL-LOH). Those candidate genes within the chromosomal segments with the count of A allele equaling to 2 and B allele equaling to 0 were considered as CNN-LOH, while count of A allele equaling to 1 and B allele equaling to 0 were considered as CNL-LOH⁶⁷. For those mutated genes not associated with predicted neoantigens, the fraction was calculated by the total number of mutated genes with CNN-LOH/CNL-LOH divided by the total number of mutated genes. For those predicted neoantigen-associated genes, the

fraction was calculated by the total number of predicted neoantigen-associated genes with CNN-LOH/CNL-LOH divided by the total number of genes associated with predicted neoantigens.

Immunohistochemistry

Immunohistochemistry (IHC) of CD3 and PD-L1 was performed on 66 LS-SCLC tumors (including the same 18/19 tumors that underwent multi-region WES and TCR sequencing) and 68 localized NSCLC tumors with matched clinical characteristics including gender, age, smoking status and tumor size (**Supplementary Data 1** and **Supplementary Table 1**). Representative IHC figures is shown in **Supplementary Fig. 11**. Serial sections with a thickness of 4 μm from FFPE samples were cut onto glass slides, followed by IHC staining. PD-L1 IHC testing was performed using the PD-L1 clone 22C3 pharmDx kit and Dako Automated Link 48 platform (Agilent Technologies/Dako, Carpinteria, CA, USA). Slides stained with CD3 were labeled by a mouse anti-CD3 monoclonal antibody (clone LN10, Leica Biosystems, Newcastle Upon Tyne, UK) at a working solution and incubated on an autostainer (Leica BOND-III, Leica Biosystems, Newcastle Upon Tyne, UK). Two independent observers examined the stained slides in a blinded fashion. The PD-L1 tumor proportion score (TPS) values were determined by the formula (PD-L1 positive tumor cells/total tumor cells*100%) in the slides. The PD-L1 positive tumor cells was defined by clear membrane staining tumor cells with or without plasma staining at any extant. For each slide, CD3 positive tumor infiltrating lymphocytes (TILs) infiltrating were indicated by the values (CD3+ TILs/ tumor cells) in each 10 typical high-power phases (20 \times).

Analysis of published data

RNA sequencing data from 81 SCLCs²³ and 1,027 NSCLCs from TCGA^{30,32} were downloaded. The estimated T-cell fraction was calculated by MCPcounter⁶⁸. The immune score was calculated with the R package "ESTIMATE" (<https://sourceforge.net/projects/estimateproject/>), which infers the infiltration of immune cells by 141 gene expression signatures⁶⁹. Copy number comparison of this study makes use of data generated by The TRacking Non-small Cell Lung Cancer Evolution Through Therapy (Rx) (TRACERx) Consortium and provided by the UCL Cancer Institute and The Francis Crick Institute. The TRACERx study is sponsored by University College London, funded by Cancer Research UK and coordinated through the Cancer Research UK and UCL Cancer Trials Centre.

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 clinical characteristics and TCR metrics. Mann-Whitney test was used to compare differences

between two independent groups. Chi-squared test was used to compare categorical variables in two groups. Log-rank test was used for survival analysis.

Data availability

The whole exome sequencing (WES) dataset has been deposited at European Bioinformatics Institute European Genome–phenome Archive (EGA) (accession number: pending) through controlled access. To protect patient privacy, interested researchers need to apply via data access committee (DAC), which will grant all reasonable requests. All other data may be found within the main manuscript or supplementary information or available from the authors upon request.

Declarations

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 from Merck, Johnson and Johnson, and consultant fees from BMS, Johnson and Johnson, AstraZeneca, Geneplus, OrigMed, Innovent outside the submitted work. The remaining authors declare no competing interests.

Acknowledgements

This work was supported by 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 Project of Zhejiang Provincial Health Science Foundation jointly Constructed by Province and Ministry (grant No. 2017211789).

Author contributions

M.C. and Jianjun Z. conceived the study. R.C., J.L. and Jiexin Z. led the data analysis. Y.J., J.F, H.P, CW.C. and J.F. led the pathological assessment, multi-region sample preparation and DNA extraction. Y.J., Y.C., M.W. and Xiao 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 Xin H. performed downstream bioinformatics analyses. R.C., S.H., Y.T., B.Z., 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., J.W., A.R., L.A.B. and Jianjun Z. interpreted the data for clinical and pathological correlation. L.D., Q.W., J.W., and J.J.L. performed statistical analyses. R.C., A.R., C.G. and Jianjun 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, doi:10.1200/jco.2005.04.4859 (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, doi:10.1056/NEJMoa1809064 (2018).
- 7 Pietanza, M. C. & Ladanyi, M. Bringing the genomic landscape of small-cell lung cancer into focus. *Nature genetics* **44**, 1074 (2012).
- 8 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, doi:10.1016/S0140-6736(19)32222-6 (2019).
- 9 Orr, H. A. Fitness and its role in evolutionary genetics. *Nature Reviews Genetics* **10**, 531-539 (2009).

- 10 Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England journal of medicine* **366**, 883-892 (2012).
- 11 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).
- 12 Zhang, J. *et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* **346**, 256-259 (2014).
- 13 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).
- 14 Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *The New England journal of medicine* **376**, 2109-2121, doi:10.1056/NEJMoa1616288 (2017).
- 15 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, doi:10.1038/s41379-018-0029-3 (2018).
- 16 Quek, K. *et al.* DNA methylation intratumor heterogeneity in localized lung adenocarcinomas. *Oncotarget* **8**, 21994-22002, doi:10.18632/oncotarget.15777 (2017).
- 17 Rosenthal, R. *et al.* Neoantigen-directed immune escape in lung cancer evolution. *Nature* **567**, 479-485, doi:10.1038/s41586-019-1032-7 (2019).
- 18 van Meerbeeck, J. P., Fennell, D. A. & De Ruyscher, D. K. Small-cell lung cancer. *Lancet (London, England)* **378**, 1741-1755, doi:10.1016/s0140-6736(11)60165-7 (2011).
- 19 Hendriks, L. E., Menis, J. & Reck, M. Prospects of targeted and immune therapies in SCLC. *Expert review of anticancer therapy* **19**, 151-167 (2019).
- 20 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, doi:10.1016/j.ejca.2018.06.023 (2018).
- 21 Reuben, A. *et al.* Comprehensive T cell repertoire characterization of non-small cell lung cancer. *Nat Commun* **11**, 603, doi:10.1038/s41467-019-14273-0 (2020).
- 22 Peifer, M. *et al.* Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet* **44**, 1104-1110, doi:10.1038/ng.2396 (2012).

- 23 George, J. *et al.* Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47-53, doi:10.1038/nature14664 (2015).
- 24 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, doi:10.1038/ng.2405 (2012).
- 25 Nong, J. *et al.* Circulating tumor DNA analysis depicts subclonal architecture and genomic evolution of small cell lung cancer. *Nature communications* **9**, 3114 (2018).
- 26 Lee, W. C. *et al.* Multiomics profiling of primary lung cancers and distant metastases reveals immunosuppression as a common characteristic of tumor cells with metastatic plasticity. *Genome Biol* **21**, 271, doi:10.1186/s13059-020-02175-0 (2020).
- 27 Roth, A. *et al.* PyClone: statistical inference of clonal population structure in cancer. *Nature methods* **11**, 396-398, doi:10.1038/nmeth.2883 (2014).
- 28 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421, doi:10.1038/nature12477 (2013).
- 29 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, doi:10.1016/j.celrep.2012.12.008 (2013).
- 30 Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519-525, doi:10.1038/nature11404 (2012).
- 31 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, doi:10.1093/annonc/mdu479 (2015).
- 32 Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543-550, doi:10.1038/nature13385 (2014).
- 33 Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* **348**, 69-74 (2015).
- 34 Strønen, E. *et al.* Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337-1341 (2016).
- 35 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**, doi:10.1126/science.aaf8399 (2017).

- 36 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*, doi:10.1158/1078-0432.ccr-19-1744 (2020).
- 37 McGrail, D. J. *et al.* Multi-omics analysis reveals neoantigen-independent immune cell infiltration in copy-number driven cancers. *Nature Communications* **9**, 1317, doi:10.1038/s41467-018-03730-x (2018).
- 38 McGranahan, N. *et al.* Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* **171**, 1259+, doi:10.1016/j.cell.2017.10.001 (2017).
- 39 Tran, E. *et al.* T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. *N Engl J Med* **375**, 2255-2262, doi:10.1056/NEJMoa1609279 (2016).
- 40 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, doi:10.1200/jco.2018.78.1963 (2018).
- 41 Wagner, A. H. *et al.* Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. *Nat Commun* **9**, 3787, doi:10.1038/s41467-018-06162-9 (2018).
- 42 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).
- 43 Negro, M. V. *et al.* Exome sequencing of paired primary and relapsed small cell lung cancers reveals increased copy number aberration complexity to be associated with disease relapse (AACR, 2018).
- 44 Kim, R., Emi, M. & Tanabe, K. Cancer immunoediting from immune surveillance to immune escape. *Immunology* **121**, 1-14 (2007).
- 45 McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463-1469, doi:10.1126/science.aaf1490 (2016).
- 46 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).
- 47 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, doi:10.1158/1078-0432.ccr-19-0558 (2019).
- 48 Hutchinson, L. Biomarkers: Aneuploidy and immune evasion - a biomarker of response. *Nature reviews. Clinical oncology* **14**, 140, doi:10.1038/nrclinonc.2017.23 (2017).

- 49 Griffioen, A. W., Damen, C. A., Blijham, G. H. & Groenewegen, G. Tumor angiogenesis is accompanied by a decreased inflammatory response of tumor-associated endothelium. *Blood* **88**, 667-673 (1996).
- 50 Salmon, H. *et al.* Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J Clin Invest* **122**, 899-910, doi:10.1172/jci45817 (2012).
- 51 Engelhardt, J. J. *et al.* Marginating dendritic cells of the tumor microenvironment cross-present tumor antigens and stably engage tumor-specific T cells. *Cancer Cell* **21**, 402-417, doi:10.1016/j.ccr.2012.01.008 (2012).
- 52 Manaster, Y. *et al.* Reduced CTL motility and activity in avascular tumor areas. *Cancer immunology, immunotherapy : CII* **68**, 1287-1301, doi:10.1007/s00262-019-02361-5 (2019).
- 53 Dangaj, D. *et al.* Cooperation between Constitutive and Inducible Chemokines Enables T Cell Engraftment and Immune Attack in Solid Tumors. *Cancer Cell* **35**, 885-900.e810, doi:10.1016/j.ccell.2019.05.004 (2019).
- 54 Hu, X. *et al.* Multi-region exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. *Nat Commun* **10**, 2978, doi:10.1038/s41467-019-10877-8 (2019).
- 55 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213-219, doi:10.1038/nbt.2514 (2013).
- 56 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, doi:10.1093/nar/gky1016 (2019).
- 57 Felsenstein, J. *PHYLIP (phylogeny inference package), version 3.5 c.* (Joseph Felsenstein., 1993).
- 58 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, doi:10.1186/s13059-016-0893-4 (2016).
- 59 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, doi:10.1371/journal.pcbi.1004873 (2016).
- 60 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, doi:10.1093/nar/gkn202 (2008).
- 61 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, doi:10.1093/bioinformatics/btn128 (2008).

- 62 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, doi:10.1371/journal.pone.0000796 (2007).
- 63 Robins, H. S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099-4107, doi:10.1182/blood-2009-04-217604 (2009).
- 64 Carlson, C. S. *et al.* Using synthetic templates to design an unbiased multiplex PCR assay. *Nature communications* **4**, 2680, doi:10.1038/ncomms3680 (2013).
- 65 Robins, H. *et al.* Ultra-sensitive detection of rare T cell clones. *Journal of immunological methods* **375**, 14-19, doi:10.1016/j.jim.2011.09.001 (2012).
- 66 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, doi:10.1007/978-1-4939-8546-3_13 (2018).
- 67 Ryland, G. L. *et al.* Loss of heterozygosity: what is it good for? *BMC medical genomics* **8**, 45 (2015).
- 68 Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* **17**, 218, doi:10.1186/s13059-016-1070-5 (2016).
- 69 Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature communications* **4**, 1-11 (2013).

Figures

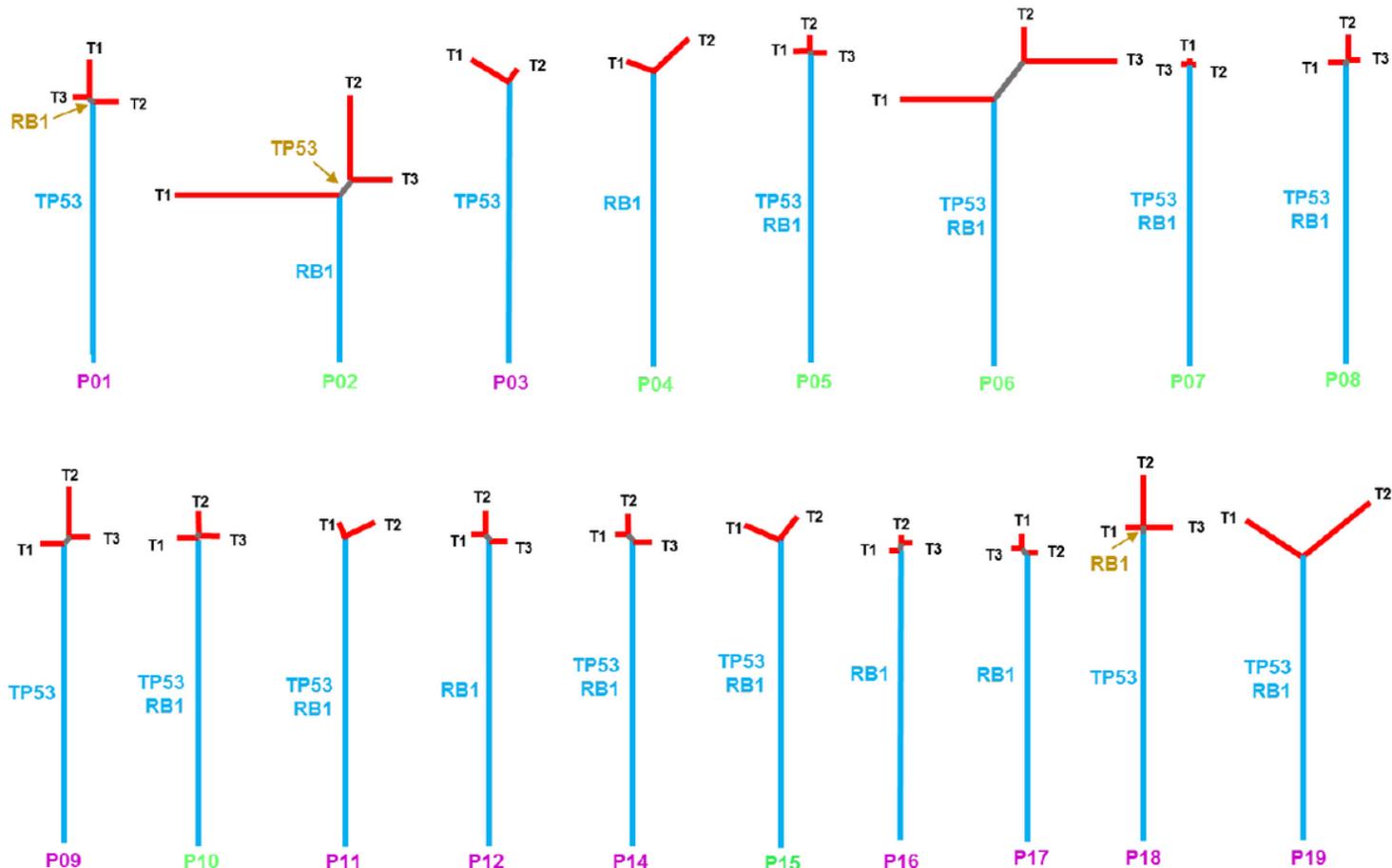


Figure 1

Phylogenetic trees of 18 limited stage small-cell lung cancer (SCLC) tumors with multi-region whole exome sequencing (WES). Blue, brown and red lines represent trunk, branch, and private mutations, respectively. The length of trunk, branch and private branch are proportional to the numbers of mutations shared by 3, 2 or 1 tumor regions. Mutated cancer genes TP53 and RB1 are mapped to the phylogenetic trees as indicated. Purple patient IDs = patients who were alive; Green patient IDs = patients who were deceased.

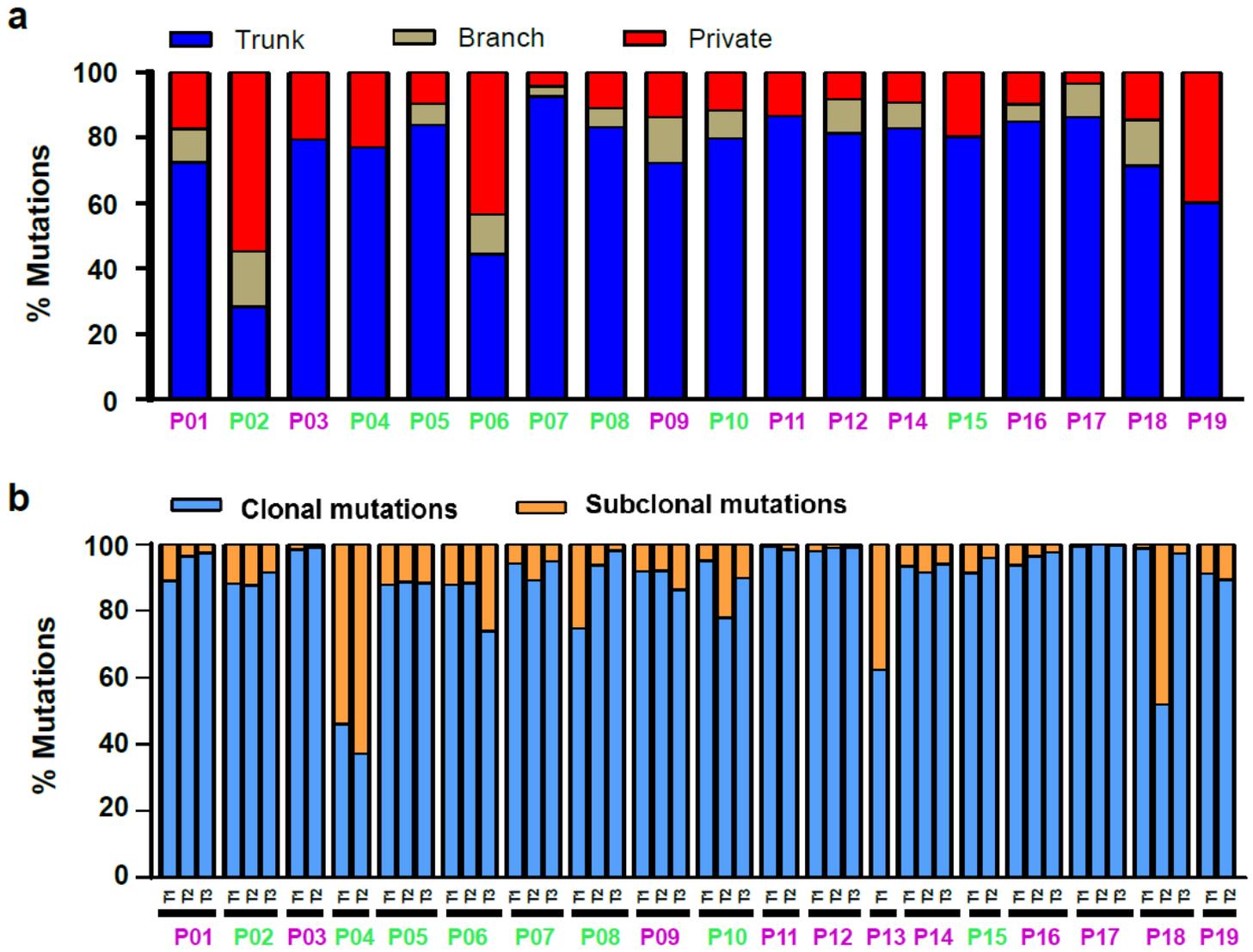


Figure 2

Genomic intratumor heterogeneity of small-cell lung cancers (SCLCs). (a) Proportion of trunk (blue), branch (brown), and private (red) mutations representing mutations detected in all 3 tumor regions, 2 tumor regions and only in one single tumor region from any given tumor. (b) Proportion of clonal versus subclonal mutations defined by PyClone in 50 SCLC tumor specimens. Mutations were classified as clonal (estimated cancer cell fraction = 1, indicating mutations presenting in all cancer cells, blue) or subclonal (estimated cancer cell fraction < 1, indicating mutations only present in a subset of cancer cells, orange) in each tumor specimen. Purple patient IDs = patients who were alive; Green patient IDs = patients who were deceased.

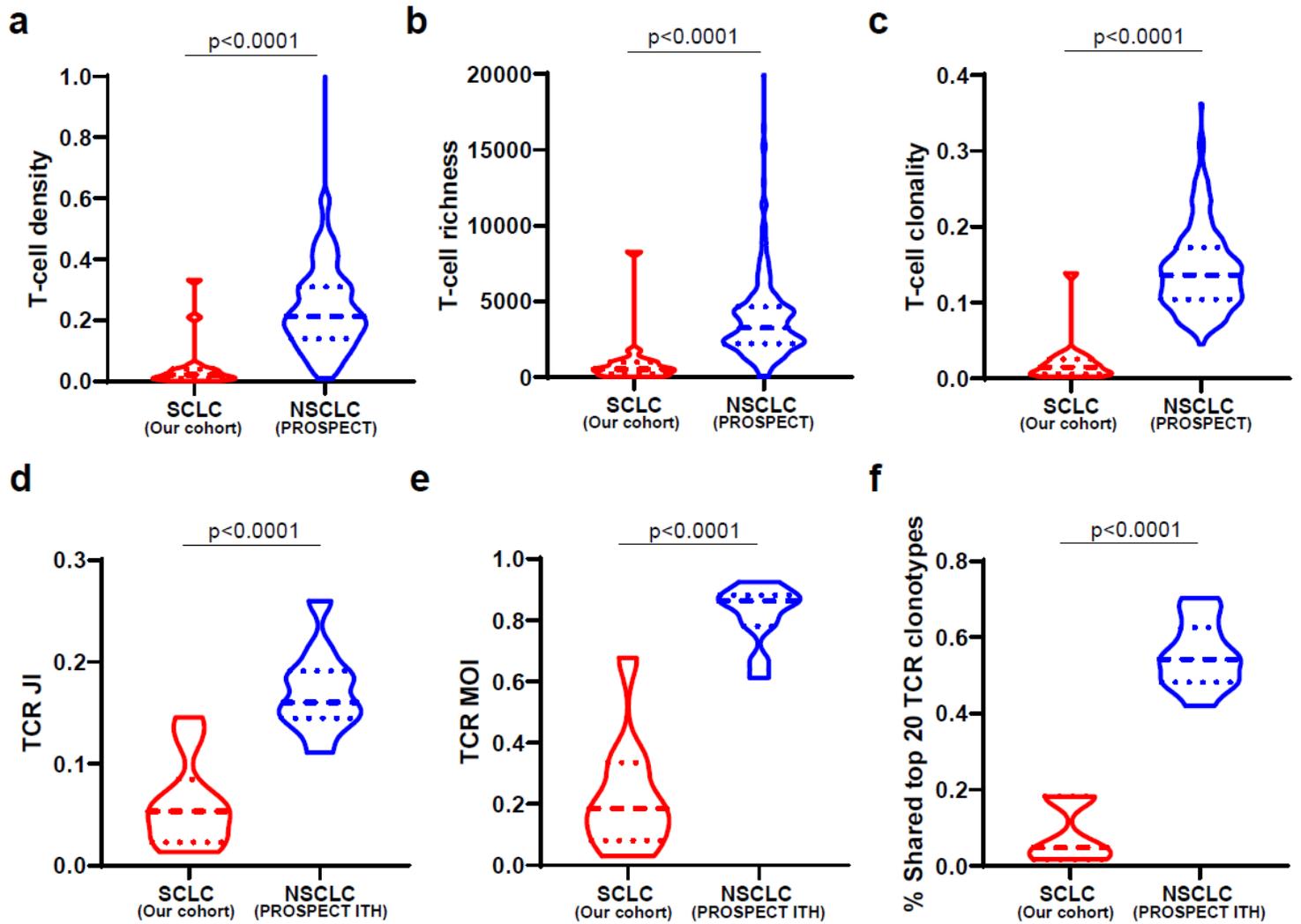


Figure 3

TCR metrics in small-cell lung cancer (SCLC) versus non-small cell lung cancer (NSCLC) tumors (the PROSPECT cohort). (a) T-cell density - an estimate of the proportion of T cells in a specimen, (b) T-cell richness - a measure of T-cell diversity and (c) T-cell clonality - a metric indicating T-cell expansion and reactivity. TCR intratumor heterogeneity (ITH) in SCLC versus NSCLC using (d) the average Jaccard index (JI), a metric representing the proportion of shared T-cell clonotypes, (e) Morisita index (MOI), a metric taking into consideration the not only the composition of T cell clonotypes but also the abundance of individual T cell clonotypes and (f) Proportion of shared top 20 TCR clonotypes between any paired samples within the same tumors in SCLC (blue) versus NSCLC tumors (red) with multiregional TCR data available.

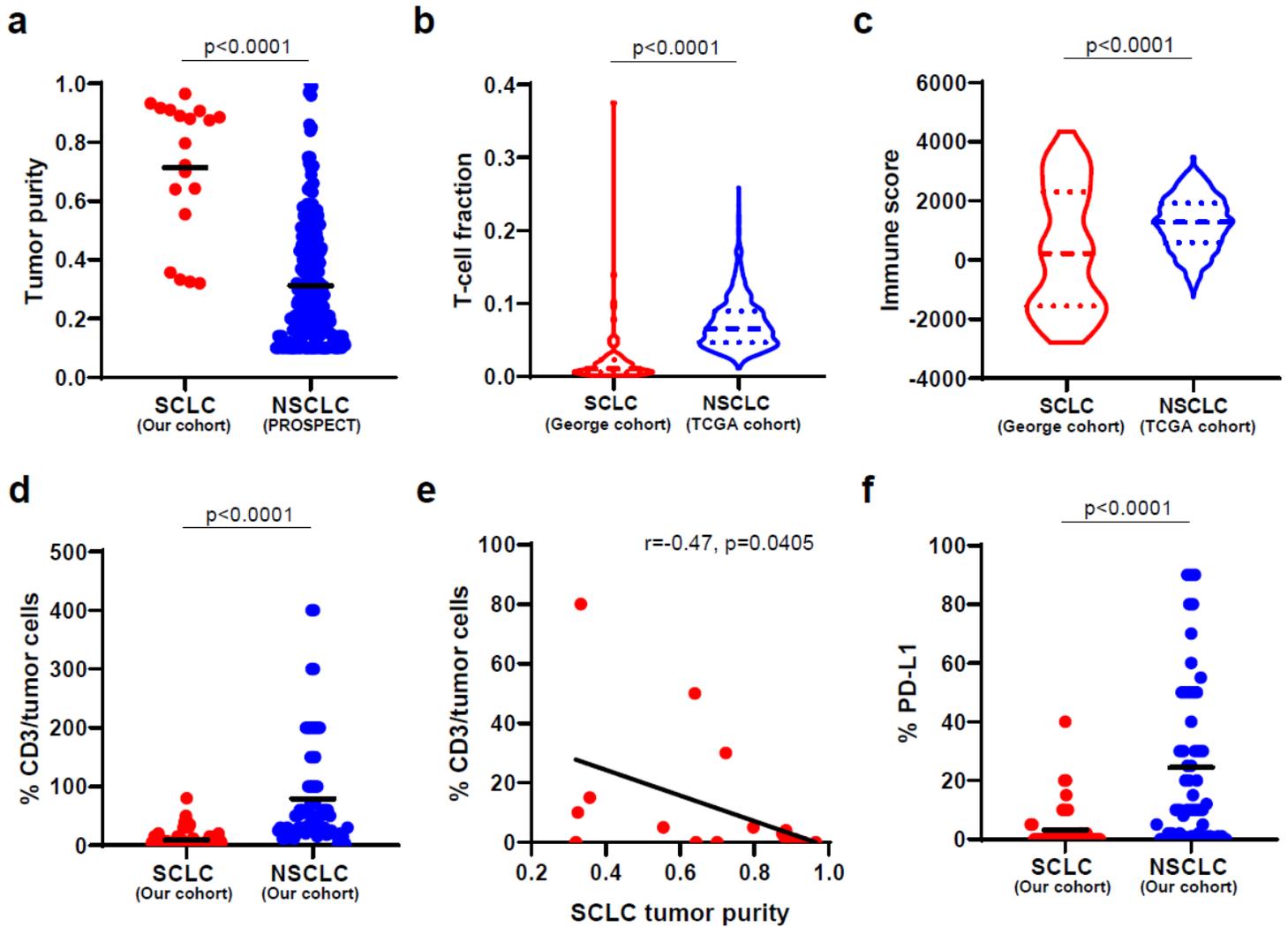
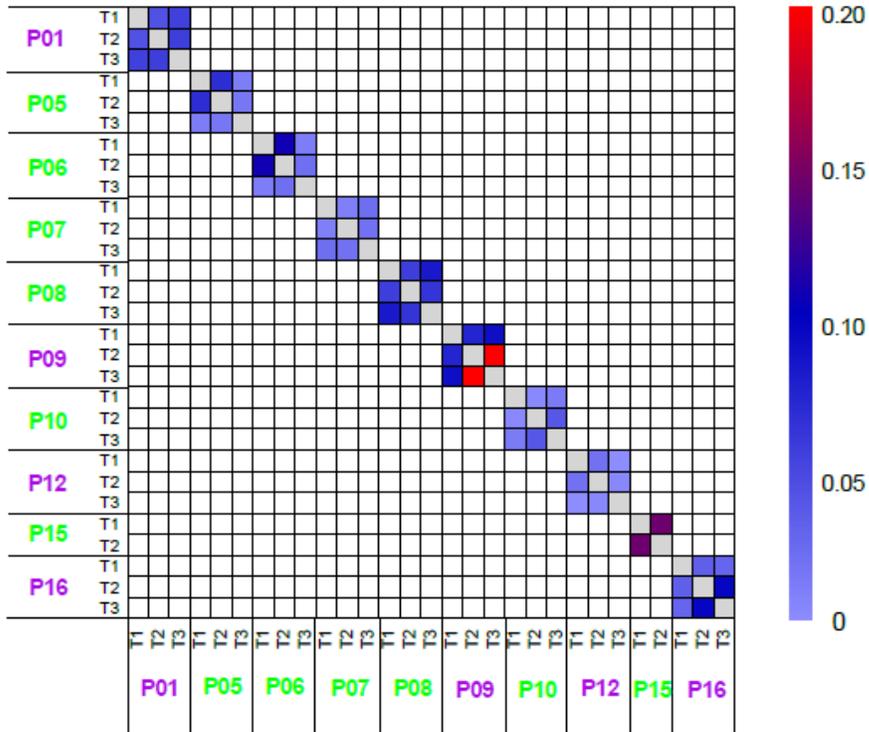
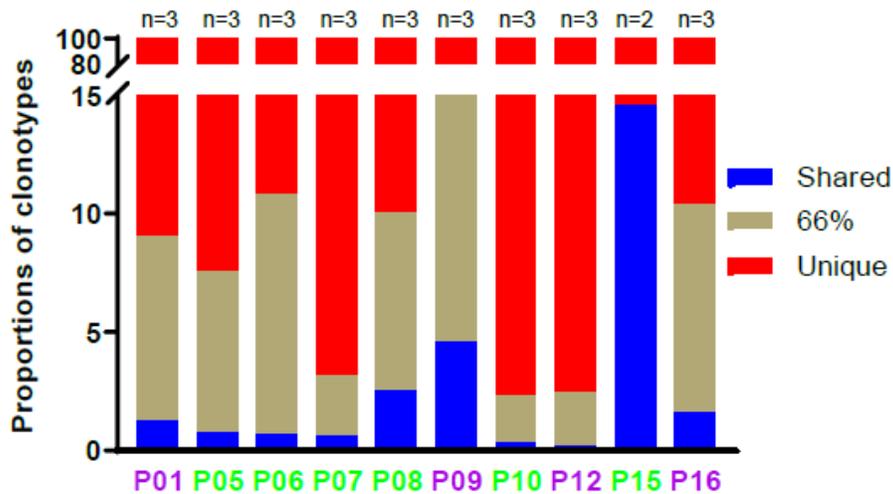


Figure 4

Comparison of immune features in SCLCs versus NSCLCs. (a) Tumor purity in SCLCs versus NSCLCs. Tumor purity were derived from WES data from 19 SCLC tumors (blue) versus 242 NSCLC tumors (red) from the PROSPECT cohort. (b) T-cell infiltration in SCLCs compared to NSCLCs. T-cell infiltration was derived by deconvolution of RNA-seq data of 81 SCLC tumors (George cohort) versus 1,027 NSCLC tumors from TCGA. (c) Immune scores in SCLCs compared to NSCLCs. The immune scores were calculated from RNA-seq data to quantify all immune cells within the tumors from 81 SCLC tumors (George cohort) versus 1,027 NSCLC tumors from TCGA. (d) Expression of CD3 in tumor infiltrating lymphocytes (TILs) of SCLC versus NSCLC tumors. (e) Expression of PD-L1 in SCLC versus NSCLC tumors. (f) Association of tumor purity with CD3+ TIL in SCLCs.

a**TCR Jaccard index****b****Shared T-cell clonotypes****Figure 5**

Substantial TCR repertoire intratumor heterogeneity (ITH) in small-cell lung cancer (SCLC). (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. Purple patient IDs = patients who were alive; Green patient IDs = patients who were deceased.

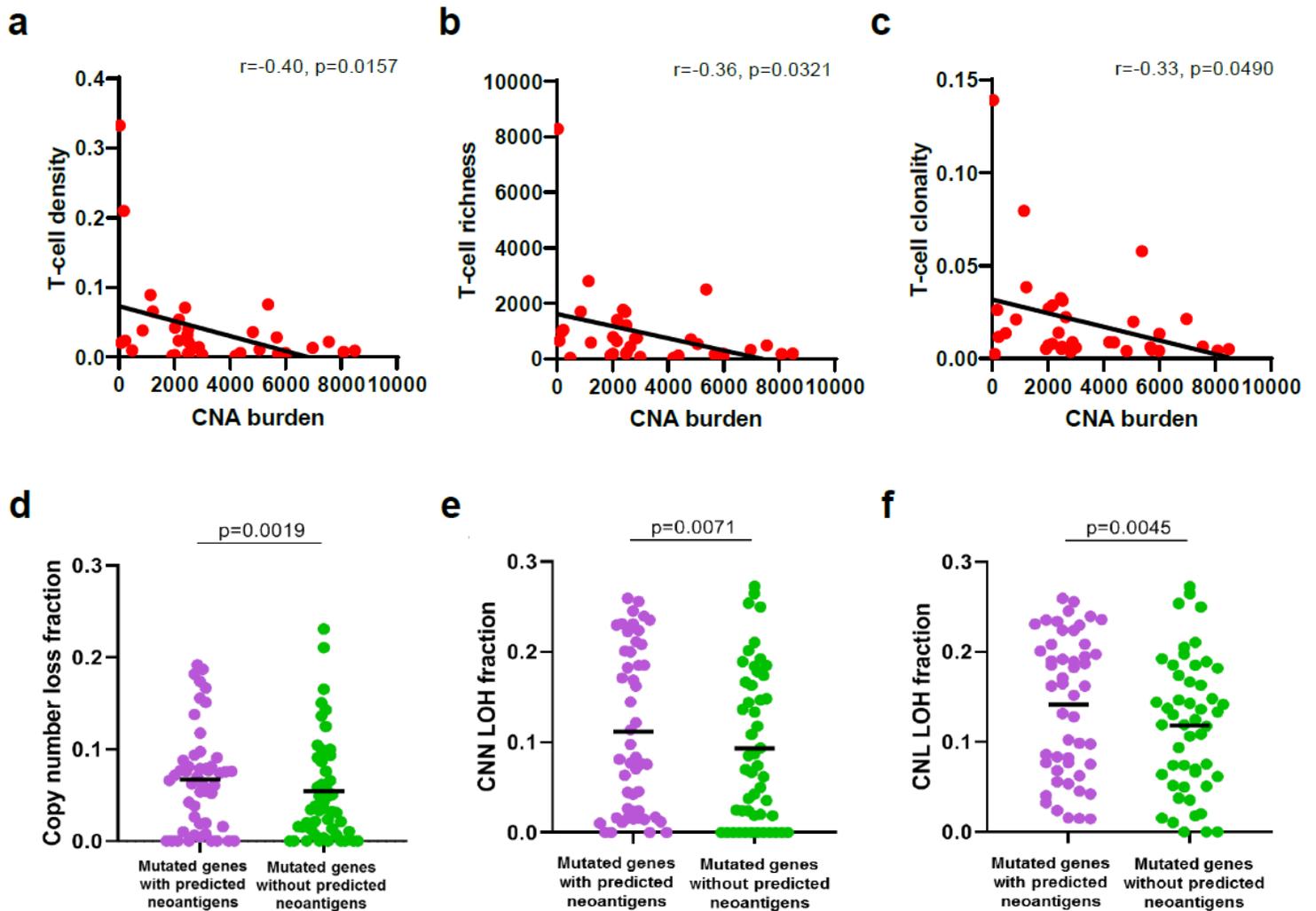


Figure 6

Associations of chromosomal copy number aberrations (CNAs) with TCR repertoire and loss of neoantigens. CNA burden and its negative correlations with T-cell (a) density, (b) richness and (c) clonality. (d) copy number loss fraction, (e) copy number neutral LOH (CNN-LOH) fraction and (f) copy number loss LOH (CNL-LOH) fraction in predicted neoantigen-associated genes (purple) versus other mutated genes not associated with predicted neoantigens (green) in SCLC tumors.

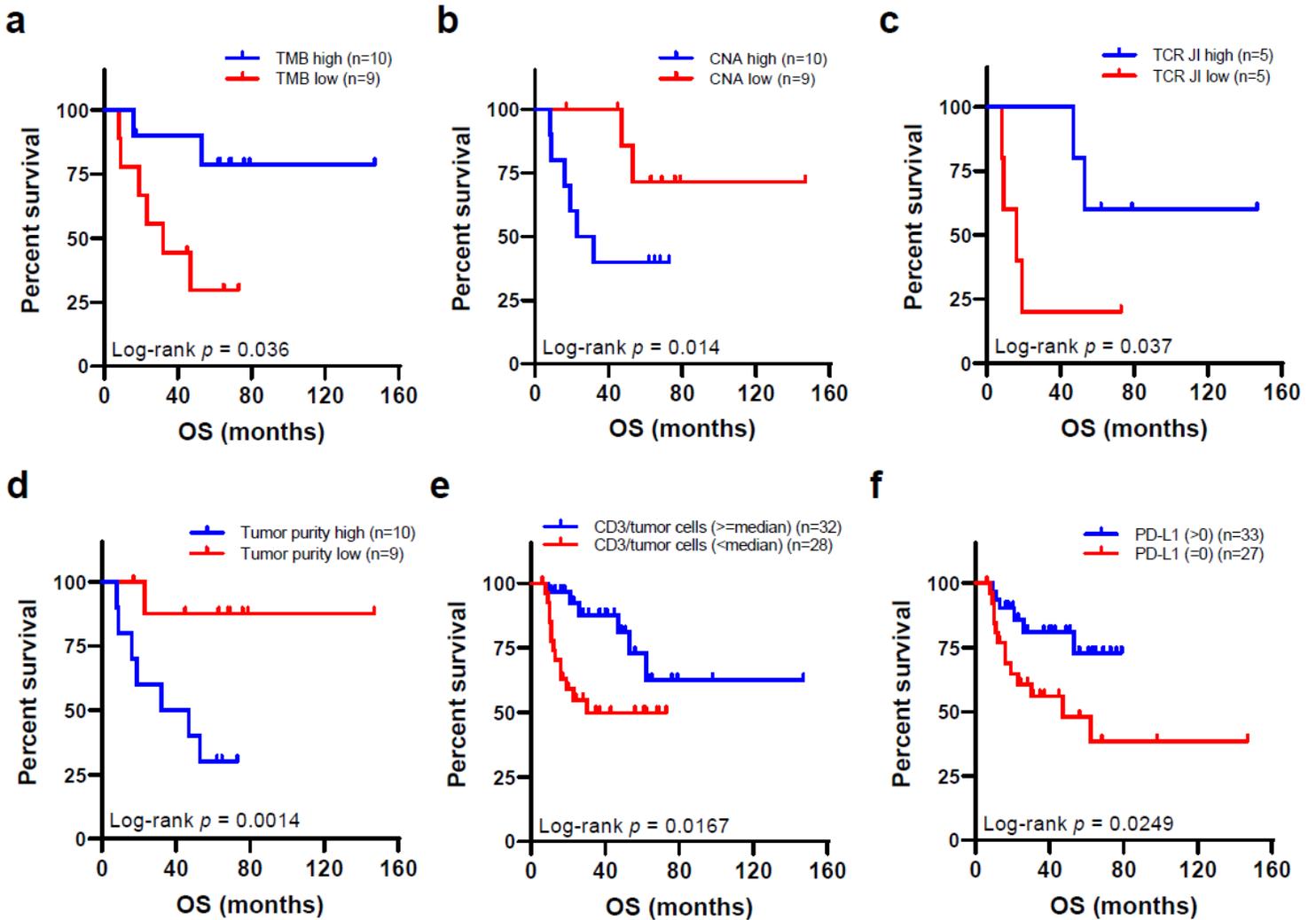


Figure 7

Association of overall survival (OS) with immunogenomic landscape. (a) OS of patients with higher (above median, blue) TMB versus patients with lower (below median, red) TMB. (b) OS of patients with higher (above median, blue) CNA burden versus patients with lower (below median, red) CNA burden. (c) OS of patients with more homogenous TCR repertoire (higher above median TCR JI, blue) versus patients with more heterogeneous TCR repertoire (lower below median TCR JI, red). (d) OS of patients with tumors of higher (above median, blue) tumor purity versus patients with tumors of lower (below median, red) tumor purity. (e) OS of patients with tumors of higher (no less than median, blue) CD3+ tumor infiltrating lymphocytes (TILs) versus patients with tumors of lower (below median, red) CD3+ TILs. (f) OS of patients with tumors of positive (above 0, blue) PD-L1 expression versus patients with tumors of negative (equal 0, red) PD-L1 expression.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [2021.01.15SCLCSupplementaryinfo.docx](#)
- [2021.01.15supplementarydata1.xlsx](#)
- [SupplementaryData2mutectt.xlsx](#)