

Multiomics analysis reveals a distinct response mechanism in multiple primary lung adenocarcinoma after neoadjuvant immunotherapy

Chao Zhang,^{1,2} Kai Yin,¹ Si-Yang Liu,¹ Li-Xu Yan,³ Jian Su,¹ Yi-Long Wu ,¹ Xu-Chao Zhang,¹ Wen-Zhao Zhong,¹ Xue-Ning Yang¹

To cite: Zhang C, Yin K, Liu S-Y, *et al*. Multiomics analysis reveals a distinct response mechanism in multiple primary lung adenocarcinoma after neoadjuvant immunotherapy. *Journal for ImmunoTherapy of Cancer* 2021;**9**:e002312. doi:10.1136/jitc-2020-002312

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jitc-2020-002312>).

CZ, KY and S-YL contributed equally.

Accepted 24 February 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Guangdong Lung Cancer Institute, Guangdong Provincial People's Hospital, Guangzhou, Guangdong, China

²School of Medicine, South China University of Technology, Guangzhou, Guangdong, China

³Department of Pathology, Guangdong Provincial People's Hospital, Guangzhou, Guangdong, China

Correspondence to

Dr Xue-Ning Yang;
yangxuening@gdph.org.cn

Dr Wen-Zhao Zhong;
syzhongwenzhao@scut.edu.cn

ABSTRACT

Multiple primary lung cancer (MPLC) remains a tough challenge to diagnose and treat. Although neoadjuvant immunotherapy has shown promising results in early stage non-small cell lung cancer, whether such modality can benefit all primary lesions remains unclear. Herein, we performed integrated multiomics analysis in one patient with early stage MPLC with remarkable tumor shrinkage in a solid nodule and no response in two subsolid nodules after treatment with three cycles of neoadjuvant pembrolizumab. Genomic heterogeneity was observed among responding nodules with high levels of infiltrating CD8⁺ and CD68⁺ immune cells. Substantially downregulated human leukocyte antigen (HLA)-related genes and impaired T lymphocyte function were observed in non-responding nodules. A larger proportion of infiltrating tissue resident memory T cells (Trm) along with high T cell receptor repertoire clonality in responding nodules were validated as predictive and prognostic biomarkers in multiple cancer types using external public datasets. These results suggested that neoadjuvant programmed death 1 (PD-1)/programmed death ligand 1 inhibitors alone may not be an optimal therapeutic strategy for MPLC due to disparities in genomic alterations and immune microenvironment among different lesions. Additionally, we postulate that increased infiltration of Trm may be a unique marker of early immune responses to PD-1 blockade.

BACKGROUND

Multiple primary lung cancer (MPLC) is a lung cancer subtype with a unique diagnosis and therapeutic strategy. Distinguishing a second primary lung cancer from intrapulmonary metastasis can be clinically challenging and thus, influences subsequent treatment.^{1 2} A previous study showed extremely low incidence of intrapulmonary metastasis among ground-glass or subsolid nodules.³ Although surgery is considered the optimal choice for MPLC treatment, the extent of resection and the potential of other local treatment approaches has not been fully determined.⁴⁻⁶ Immune checkpoint

inhibitors blocking negative key regulator of T cells have shifted the management of multiple advanced cancers^{7 8}; however, only a fraction of patients with advanced disease respond to immunotherapy aside from long-term benefits.^{9 10}

Unlike advanced disease, early stage disease may harbor distinct tumor microenvironments. Immunoediting, a concept used to illustrate tumor immunogenicity over time,^{11 12} has demonstrated correlations between disease progression and metastasis.^{13 14} Initial disease, known as precancerous lesions, is accompanied by early immune sensing through the activation of resident immune cells and the activation of innate and adaptive immune responses.¹⁵ However, very low expression of programmed death 1/programmed death ligand 1 (PD-1/PD-L1) protein in either tumor cells or immune cells during precancerous and early stage disease can indicate the existence of other mechanisms of immune escape beyond PD-1/PD-L1 modulation.¹⁶ Conversely, relatively high PD-L1 expression has been observed in late-stage lung cancer.^{17 18} Since the development of neoadjuvant immunotherapy for early stage non-small cell lung cancer (NSCLC),¹⁹ multiple trials have been initiated to evaluate the efficacy of different checkpoint blockades and modalities and have shown encouraging preliminary results. However, no reported trials have assessed the feasibility of neoadjuvant immunotherapy in MPLC. Here, we report an integrated multiomics analysis in one patient with MPLC receiving three cycles of neoadjuvant pembrolizumab, and showing mixed response.

CASE PRESENTATION

A woman aged 72 years admitted to outer hospital was accidentally found with

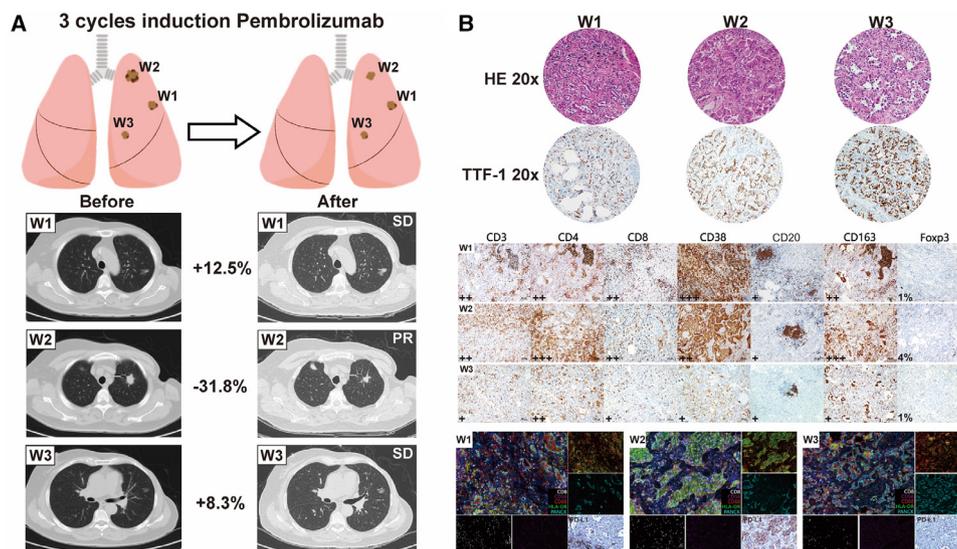


Figure 1 Patient clinical and pathological evaluation following neoadjuvant immunotherapy. (A) Radiological evaluation of three resected nodules after three cycles of neoadjuvant immunotherapy showing notable tumor shrinkage of the solid nodule; however, the other nodules remained stable. (B) Comprehensive pathological evaluation integrated with mIHC of three surgically resected nodules. The extent of staining and the percentage of Foxp3-positive cells with nucleus staining are listed on the lower left corner of each slide. mIHC and PD-L1 images using the Dako 22C3 assay of nodules are obtained. mIHC panels are displayed in color on the lower right corner of the IHC slides: CD8 (white), CD56 (purple), CD68 (red), HLA-DR (green); PANCK (cyan). The expression level of PD-L1 is listed on the lower right corner of IHC slides. CNV, copy number variation; HLA, human leukocyte antigen; IC, immune cell; mIHC, multiplex immunohistochemistry; PD-L1, programmed death ligand 1; PR, partial response; RECIST, Response Evaluation Criteria in Solid Tumors; SD, stable disease; TMB, tumor mutation burden; TTF-1, thyroid transcription factor 1.

multiple pulmonary nodules on a CT scan. One nodule in the left upper lobe was a solid lesion, whereas all others were mixed or pure ground-glass nodules. PET scan showed no distant metastasis or enlarged lymph nodes with increased uptake values. Percutaneous lung biopsy of the major solid nodule confirmed lung adenocarcinoma. Next-generation sequencing showed pan-negative results for all common driver mutations. Immunohistochemistry (IHC) using the Dako 22C3 assay revealed a PD-L1 expression level of 40%. However, due to the presence of bilateral pulmonary nodules, intrapulmonary metastasis and multiple primary lung cancer were indistinguishable. Given the negative results for driver mutations and the relatively high PD-L1 expression, the patient was prescribed the ‘first-line’ pembrolizumab, with informed consent. After three cycles (21 days/cycle) of pembrolizumab treatment, the patient was scheduled for radical treatment at our hospital. Chest CT showed significant shrinkage (31.8%) of the solid nodule; however, mixed and pure ground-glass nodules remained stable or slightly increased in size (figure 1A). Surgical resection of the major nodule and other nodules in the upper left lobe was considered due to the potential of early stage disease and radiological surveillance was recommended for mixed or pure ground-glass nodules in other lobes. IHC and multiplex immunohistochemistry (mIHC) revealed significantly high levels of infiltrating CD8⁺ lymphocytes and activated CD68⁺HLA-DR⁺ macrophages in responding nodules (W2). Notably,

CD20 staining showed tertiary lymphoid structures (TLS) formation in all three nodules, which has been verified its potential of predicting immunotherapy efficacy.²⁰ As expected, the extent of TLS observed in non-responding nodules (W1/W3) was smaller than that in W2 (figure 1B).

Heterogeneous genomic features and T cell receptor repertoire profile among nodules

Multomics sequencing was performed on three resected nodules as described in the online supplemental materials. Whole exome sequencing (WES) and T cell receptor repertoire (TCR) in resected nodules identified highly heterogeneous genomic features among different nodules. Both W1 and W3, which were unresponsive to immunotherapy, harbored the *EGFR* exon 21 L858R mutation, whereas W2 harbored an oncogenic driver mutation in *KRAS*. Arm-level somatic copy number alterations indicated relatively fewer events in W1 and W2, likely leading to lower immune evasion.²¹ The tumor mutation burden of W2 was notably higher but showed the lowest genomic instability (figure 2A). Additionally, HLA phenotypes predicted nearly separate neoantigen distribution, indicating that the different nodules induced a heterogeneous microenvironment (figure 2B). TCR-beta chain sequencing identified a markedly amplified complementarity-determining region 3 (CDR3) alignment (CAISLKSSGFTGELFE) in the W2 cell type, contributing to over 10% of the total repertoire. In contrast, the most predominant CDR3

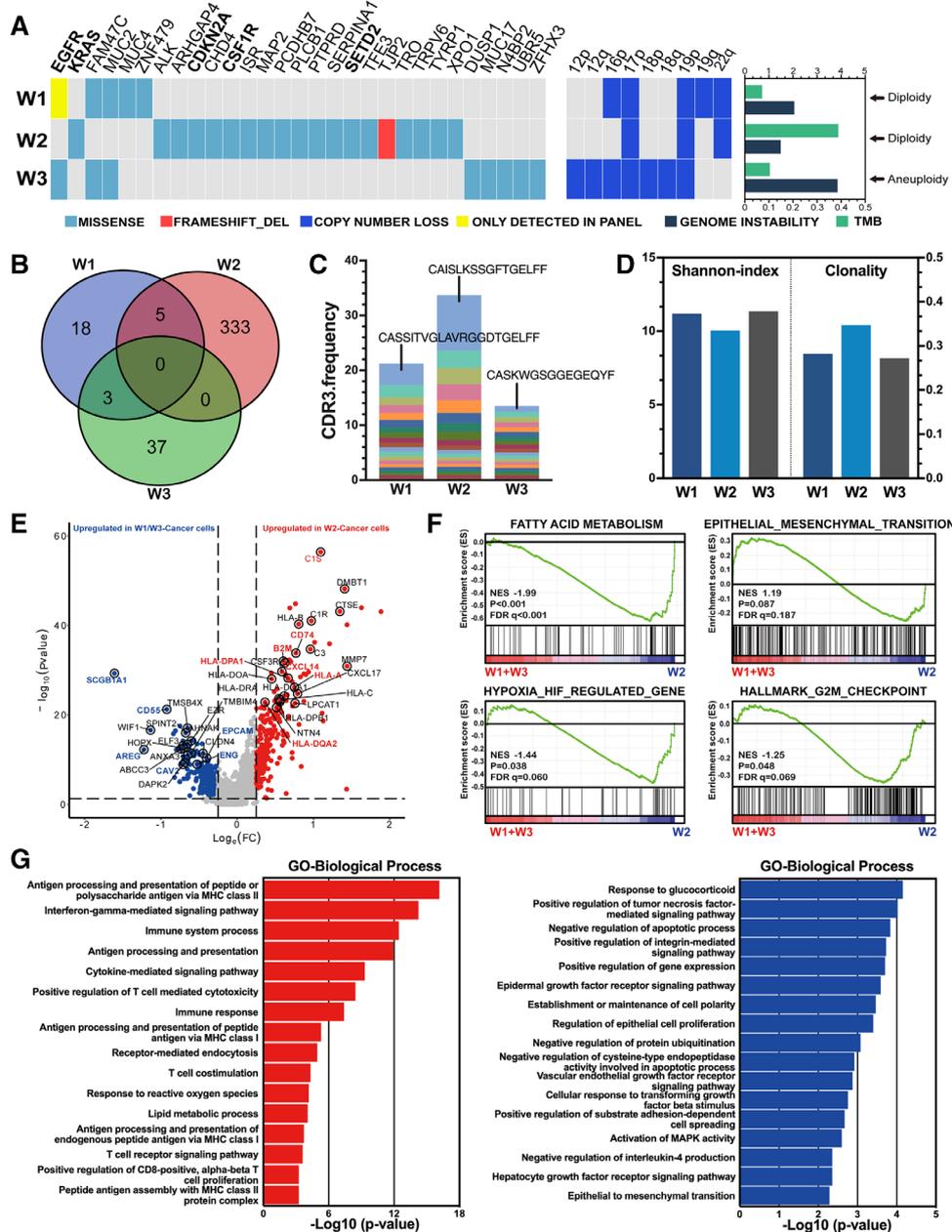


Figure 2 Heterogeneity of multiple primary lung cancer (MPLC) and single-cell RNA sequencing (scRNA-seq) analysis of differentially enriched pathways. (A) Genomic alteration of MPLC and somatic copy number variants. Bar chart on the right shows genomic instability, tumor mutation burden (TMB) and ploidy for each corresponding lesion. (B) A Venn diagram showing highly inconsistent human leukocyte antigen (HLA) types among nodules. (C) Stacked column chart of CDR3 frequency among nodules highlighting high frequency of CDR3 alignment in responding nodules. (D) Shannon index and CDR3 clonality of T cell receptor repertoire. (E) A volcano plot of differentially expressed genes between W2 and W1/W3 cancer cells (EPCAM+) with immune modulation (red) and tumor growth (blue) associated genes highlighted in corresponding color. (F) Gene set enrichment analysis (GSEA) between W2 and W1/W3 cancer cells among four cancer-related pathways, including fatty acid metabolism, epithelial mesenchymal transition, hypoxia-regulated genes and G2/M checkpoint. (G) Gene Ontology (GO) analysis of differentially enriched biological processes between W2 (red) and W1/W3 (blue) cancer cells.

clonotypes in W1 and W3 accounted for only 3.9% and 1.1%, respectively, suggesting a specific T cell proliferation in W2 (figure 2C). In line with the presence of a dominant clone, TCR clonality was relatively higher in W2; however, the W2 clonotype had lower TCR diversity compared with the other two nodules even under the same systemic immunity (figure 2D) (original data of

mIHC, WES as well as HLA phenotype was summarized in online supplemental file 2).

Differentially enriched pathways and major histocompatibility complex expression modulation

We then analyzed gene expression in tumor cells selected using InferCNV from single-cell RNA sequencing

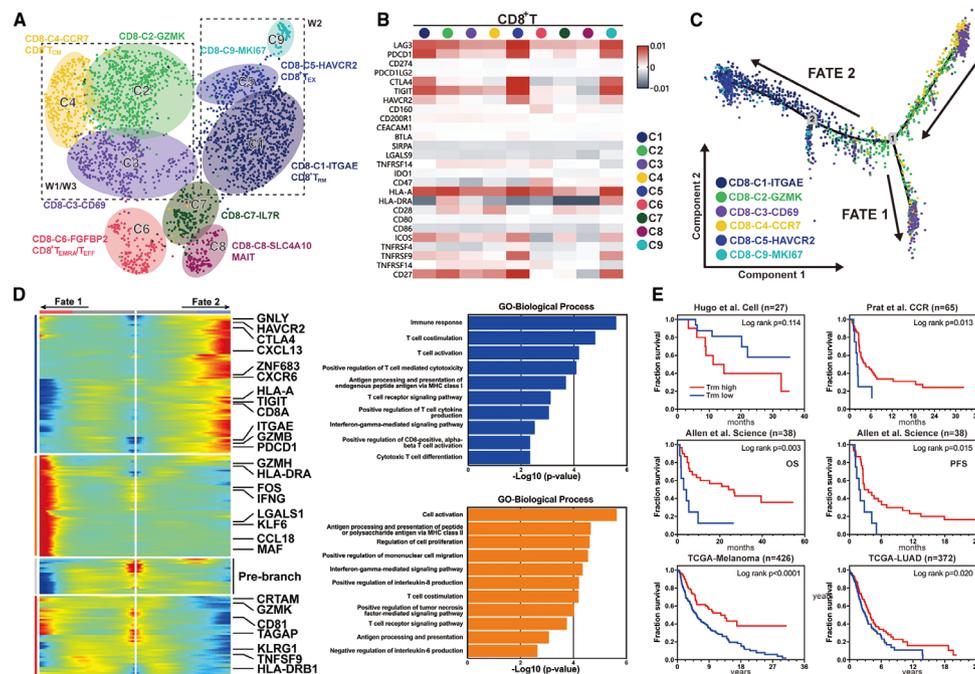


Figure 3 Diverse differentiation trajectories of CD8⁺ T lymphocytes and prognostic value of tissue resident memory T cells (Trm). (A) UMAP projection of 2755 CD8⁺ T lymphocytes from three nodules clustered into nine groups by feature markers. (B) Heatmap of average gene expression value of immune checkpoint and costimulation-related genes among different CD8 T clusters. (C) Pseudo-time analysis of six major CD8⁺ T lymphocyte clusters showing two differentiation trajectory fates. (D) Radiation plot showing significant differential expression for each fate. Highly expressed genes are colored in red and genes with low expression in blue. Bar charts on the right indicate highly involved processes in fate 1 (orange) and fate 2 (blue), respectively. (E) Confirmation of the predictive and prognostic value of the Trm signature using external validation cohorts including two TCGA datasets and three retrospective studies. Log-rank p value is presented for each independent survival analysis. TCGA, The Cancer Genome Atlas; UMAP, Uniform manifold approximation and projection.

(scRNA-seq) and determined differentially expressed genes in W2 and W1/W3 cancer cells. Genes correlated with HLA, complement activation and immune modulation were significantly upregulated in W2 cancer cells, whereas genes related to cell growth and adhesion were upregulated in W1/W3 cancer cells (figure 2E). We further examined several gene signatures that could influence the efficacy of antitumor treatment. The fatty acid metabolism gene signature was significantly enriched in W2 ($p < 0.001$, false discovery rate (FDR) $q < 0.001$). Both hypoxia-related ($p = 0.038$, FDR $q = 0.060$) and G2/M checkpoint signatures ($p = 0.048$, FDR $q = 0.069$) were markedly enriched in the W2. However, there was no significant enrichment in the epithelial mesenchymal transition signature ($p = 0.087$, FDR $q = 0.187$) among nodules (figure 2F). W2 cancer cells were significantly involved in antigen processing and presentation through major histocompatibility complex class II ($p < 0.001$, FDR < 0.001). Additionally, W2 cancer cells were involved in positive regulation of T cell-mediated cytotoxicity and other T cell activation processes. In contrast, W1/3 cancer cells were significantly involved in growth factor receptor signaling ($p < 0.001$, FDR = 0.058) and cell proliferation ($p < 0.001$, FDR = 0.060) (figure 2G), which could partially explain the limited response to immunotherapy.

Impaired lymphocyte immunity with fewer infiltrating tissue resident memory T cells

Given the major therapeutic mechanism of PD-1 checkpoint inhibitors on T lymphocytes, we evaluated differential T cell interactions between responding and non-responding nodules. T cell subtypes were distinguished by specific markers in CD4 (CD3⁺CD8⁺CD45⁺) and CD8 (CD3⁺CD8⁺CD45⁺) lymphocytes (figure 3A), respectively. To investigate how immunotherapy affected the expression of checkpoints and costimulator molecules in CD8 T lymphocytes, we examined the expression of 26 relevant genes in different cell types. CD8-C1-ITGAE, CD8-C5-HAVCR2 and CD8-C9-MKI67 showed relatively high expression of immune checkpoint proteins, such as LAG3, PDCD1, cytotoxic T-lymphocyte antigen 4 (CTLA-4), TIGIT and HLA-A, whereas other cell types had weak or no expression of these proteins (figure 3B). To map the evolutionary trajectory of CD8⁺ lymphocytes under immunotherapy, we performed pseudo-time analysis and identified two fates. W2 cells tended toward fate 2, characterized by high expression of HAVCR2, TIGIT, PDCD1, GNLV, HLA-A and GZMB, which correlated with T cell exhaustion and cytotoxic signals (figure 3C). W1/W3 cells tended toward fate 1, which was characterized by the high expression of GZMH, HLA-DRA and IFNG (figure 3C). These results were corroborated by Gene Ontology

analysis, showing that fate 2 was enriched for the induction of antigen-specific cytotoxicity in CD8 lymphocytes. However, CD8 lymphocytes in fate 1 also exhibited activated T cell functions, such as T cell costimulation and TCR signaling (figure 3D). Trm, previously reported to reside in epithelia, connective tissues and other tissues,^{22,23} have been associated with superior long-term prognosis in various cancers.^{24,25} In this study, Trm subsets were prominently enriched in W2 cancer cells, accounting for nearly 50% of T lymphocytes. Therefore, we established a Trm signature by screening specific genes using the Wilcoxon rank-sum test, as described in the online supplemental materials. Using three studies^{26–28} and two datasets from The Cancer Genome Atlas (<https://gdc.cancer.gov/>) as external validation cohorts to assess the predictive immunotherapy and prognostic value of Trm, we found that cells with high Trm infiltration showed significantly prolonged overall survival in both lung adenocarcinoma and melanoma. Except in one study, patients with high Trm signature treated with either PD-1 or CTLA-4 inhibitors showed appreciable survival compared with patients with low Trm signature indicating that Trm was both a predictive and prognostic factor for immune infiltration in multiple cancer types (figure 3E).

DISCUSSION

In this study, we showed that neoadjuvant immunotherapy unlikely enhances systemic antitumor immunity against all lesions in patients with early stage MPLC. A recent study revealed that 45% of advanced NSCLC had mixed progression after resistance to immunotherapy²⁹ despite moderate tumor heterogeneity. Genomics analyses have shown that unlike metastatic lesions, multiple primary lesions are much more heterogeneous,³⁰ and may explain the diverse HLA-phenotypes and TCR expansion observed in this study. Therefore, a single biopsy of one lesion may not be sufficient as a guide for subsequent treatment of the remaining lesions. Furthermore, multiple subsolid lesions revealed impaired antigen presentation and cytotoxicity induction by CD4 and CD8 lymphocytes, which might be a potential mechanism of insensitivity to immunotherapy, given their involvement in early phase antitumor immunity.³¹ Notably, impaired T lymphocyte immunity was not driven by the expression of PD-1/PD-L1 or other immune checkpoint proteins, as shown by our findings that non-responding nodules were negative for PD-L1 and exhibited low expression of immune checkpoints in lymphocytes. These results might be explained by differential distribution of Trm in responding and non-responding lesions of immune-modulated and immune-surveillant CD8 T cell types,³² indicating effective durable stimulation by tumor antigens. In contrast, CD8 T lymphocytes of non-responding nodules preferentially displayed naïve or early activated T cell genotypes, concordant with early stage disease. The TCR repertoire showed higher clonality in responding nodules with a specific amplified CDR3 alignment (CAISLKSSGFTGELFE). Furthermore,

survival analysis demonstrated the prognostic value of Trm with or without immunotherapy in advanced-stage disease. We propose, therefore, that Trm might be a more sophisticated marker as an early response signal for immunotherapy than CD8 infiltrating T cells.

There are several caveats to our study. First, despite our multiomics approach, this is a case-based study; thus, the results should be carefully interpreted and a larger sample size study is warranted to establish clinical efficacy. Second, it should be noted that pseudo-progression might also explain the diverse responses in different lesions. However, two to four cycles of neoadjuvant treatment may not be sufficient to determine whether the diverse responses can be attributed to pseudo-progression. Besides, we have performed in-depth pathological evaluation including pathological regression showing relatively consistent results with radiological shrinkage (pathological regression –80% for W2 and –25%/–10% for W1/W3, respectively), which may in a way suggest unlikelihood of pseudo-progression. Third, due to the lack of pretreated specimens, changes in different infiltrating immune cells after immunotherapy could not be reliably determined. Moreover, the acquisition of all small nodules in MPLC, especially for subsolid lesions, is technically impracticable before surgery. Nevertheless, by comparing nodules with different responses to a specific treatment, it may be possible to analyze how immunotherapy influences the tumor immune microenvironment, leading to diverse clinical outcomes.

In summary, our results revealed that early stage MPLC harbored diverse genomic phenotypes and exhibited distinct neoantigen peptides and downregulated HLA genes, suggesting lower immunogenicity in subsolid nodules. Moreover, we showed that impaired T lymphocyte immunity could be attributed to an inferior response to immunotherapy within subsolid nodules, which accounted for the largest proportion of multiple pulmonary nodules. Several trials (NCT04047186, NCT04026841) evaluating the efficacy of neoadjuvant immunotherapy for MPLC are still ongoing. Recently, a retrospective study showed limited efficacy of PD-1/PD-L1 checkpoint blockade alone in multiple ground-glass nodules of advanced NSCLC.³³ Altogether, these data suggested that neoadjuvant immunotherapy alone might not be an ideal option for patients with MPLC due to insufficient antitumor response for all subsolid lesions. However, given the genomic and immune microenvironment heterogeneity among MPLCs, neoadjuvant immunotherapy for MPLC with subsolid nodules should be deliberately considered for future clinical trials. In addition, Trm might be incorporated as an extensive immune biomarker in the design of future immunotherapy trials.

Acknowledgements The authors would like to thank Geneseq technology for providing high-throughput whole exome sequencing plus T cell receptor repertoire sequencing and NovolBio for providing scRNA-seq and bioinformatics analysis support. The authors would like to thank Editage (www.editage.cn) for English language editing. The authors would also like to thank the patient and her family for consenting to our publishing her clinical case.

Contributors Conceptualization: CZ, W-ZZ, Y-LW, X-NY. Resources: CZ, W-ZZ and S-YL. Methodology: CZ, S-YL, JS, L-XY and X-CZ. Bioinformatic analysis: CZ, S-YL, KY. Writing—original draft: CZ and S-YL. Writing—review and editing: CZ, S-YL, KY, W-ZZ, X-NY and Y-LW. All authors approved the submission of the manuscript.

Funding This study was supported by National Natural Science Foundation (grant 81872510 to W-ZZ) and High-level Hospital Construction Project (grant DFJH201801 to W-ZZ).

Competing interests W-ZZ has received honoraria from AstraZeneca and Roche outside the submitted work; Y-LW has received honoraria from AstraZeneca, Eli Lilly, Roche, Pierre Fabre, Pfizer and Sanofi; consulting or advisory role with AstraZeneca, Roche, Merck and Boehringer Ingelheim and Roche outside the submitted work.

Patient consent for publication Obtained.

Ethics approval Written informed consent of tumor acquisition for research has been obtained before surgery and approved by internal review board from Guangdong Lung Cancer Institute and Guangdong Provincial People's Hospital (Guangzhou, China, IRB approved protocol number GDREC2016175H).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Single-cell RNA sequencing datasets generated in this study are available on the GEO database under the accession number GSE146100. Detailed mIHC data, WES MAF files, and HLA-phenotype data are summarized in supplemental data. External datasets used in this study were accessed using the accession numbers: GEO, GSE93157 (Prat et al.); GEO, GSE78220 (Hugo et al.). All other relevant data are available from the corresponding author of this study (Xue-Ning Yang, yangxuening@gdph.org.cn) upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Yi-Long Wu <http://orcid.org/0000-0002-3611-0258>

REFERENCES

- Girard N, Deshpande C, Lau C, et al. Comprehensive histologic assessment helps to differentiate multiple lung primary non-small cell carcinomas from metastases. *Am J Surg Pathol* 2009;33:1752–64.
- Detterbeck FC, Franklin WA, Nicholson AG, et al. The IASLC lung cancer staging project: background data and proposed criteria to distinguish separate primary lung cancers from metastatic foci in patients with two lung tumors in the forthcoming eighth edition of the TNM classification for lung cancer. *J Thorac Oncol* 2016;11:651–65.
- Suh YJ, Lee H-J, Sung P, et al. A novel algorithm to differentiate between multiple primary lung cancers and intrapulmonary metastasis in multiple lung cancers with multiple pulmonary sites of involvement. *J Thorac Oncol* 2020;15:203–215.
- Lee BE, Port JL, Stiles BM, et al. Tnm stage is the most important determinant of survival in metachronous lung cancer. *Ann Thorac Surg* 2009;88:1100–5.
- Ishikawa Y, Nakayama H, Ito H, et al. Surgical treatment for synchronous primary lung adenocarcinomas. *Ann Thorac Surg* 2014;98:1983–8.
- De Leyn P, Moons J, Vansteenkiste J, et al. Survival after resection of synchronous bilateral lung cancer. *Eur J Cardiothorac Surg* 2008;34:1215–22.
- Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science* 2018;359:1350–5.
- Hargadon KM, Johnson CE, Williams CJ. Immune checkpoint blockade therapy for cancer: an overview of FDA-approved immune checkpoint inhibitors. *Int Immunopharmacol* 2018;62:29–39.
- Schadendorf D, Hodi FS, Robert C, et al. Pooled analysis of long-term survival data from phase II and phase III trials of ipilimumab in unresectable or metastatic melanoma. *J Clin Oncol* 2015;33:1889–94.
- Garon EB, Hellmann MD, Rizvi NA, et al. Five-year overall survival for patients with advanced Non-Small-Cell lung cancer treated with pembrolizumab: results from the phase I KEYNOTE-001 study. *J Clin Oncol* 2019;37:2518–27.
- Dunn GP, Bruce AT, Ikeda H, et al. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol* 2002;3:991–8.
- Mittal D, Gubin MM, Schreiber RD, et al. New insights into cancer immunoediting and its three component phases—elimination, equilibrium and escape. *Curr Opin Immunol* 2014;27:16–25.
- Angelova M, Mlecnik B, Vasaturo A, et al. Evolution of metastases in space and time under immune selection. *Cell* 2018;175:751–65.
- Matsushita H, Vesely MD, Koboldt DC, et al. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* 2012;482:400–4.
- Mascaux C, Angelova M, Vasaturo A, et al. Immune evasion before tumour invasion in early lung squamous carcinogenesis. *Nature* 2019;571:570–5.
- Zhang C, Zhang J, Xu F-P, et al. Genomic landscape and immune microenvironment features of preinvasive and early invasive lung adenocarcinoma. *J Thorac Oncol* 2019;14:1912–23.
- Tsunoda A, Morikawa K, Inoue T, et al. A prospective observational study to assess PD-L1 expression in small biopsy samples for non-small-cell lung cancer. *BMC Cancer* 2019;19:546.
- Yu H, Boyle TA, Zhou C, et al. Pd-L1 expression in lung cancer. *J Thorac Oncol* 2016;11:964–75.
- Forde PM, Chaft JE, Smith KN, et al. Neoadjuvant PD-1 blockade in resectable lung cancer. *N Engl J Med* 2018;378:1976–86.
- Helmink BA, Reddy SM, Gao J, et al. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 2020;577:549–55.
- Davoli T, Uno H, Wooten EC, et al. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* 2017;355. doi:10.1126/science.aaf8399. [Epub ahead of print: 20 Jan 2017].
- Kumar BV, Ma W, Miron M, et al. Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Rep* 2017;20:2921–34.
- Mueller SN, Mackay LK. Tissue-resident memory T cells: local specialists in immune defence. *Nat Rev Immunol* 2016;16:79–89.
- Djenidi F, Adam J, Goubar A, et al. CD8+CD103+ tumor-infiltrating lymphocytes are tumor-specific tissue-resident memory T cells and a prognostic factor for survival in lung cancer patients. *J Immunol* 2015;194:3475–86.
- Savas P, Virassamy B, Ye C, et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat Med* 2018;24:986–93.
- Hugo W, Zaretsky JM, Sun L, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* 2016;165:35–44.
- Prat A, Navarro A, Paré L, et al. Immune-related gene expression profiling after PD-1 blockade in non-small cell lung carcinoma, head and neck squamous cell carcinoma, and melanoma. *Cancer Res* 2017;77:3540–50.
- Van Allen EM, Miao D, Schilling B, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 2015;350:207–11.
- Osorio JC, Arbour KC, Le DT, et al. Lesion-level response dynamics to programmed cell death protein (PD-1) blockade. *J Clin Oncol* 2019;37:3546–55.
- Asmar R, Sonett JR, Singh G, et al. Use of oncogenic driver mutations in staging of multiple primary lung carcinomas: a single-center experience. *J Thorac Oncol* 2017;12:1524–35.
- Galon J, Bruni D. Immunology T. Tumor immunology and tumor evolution: intertwined histories. *Immunity* 2020;52:55–81.
- van der Leun AM, Thommen DS, Schumacher TN. CD8+ T cell states in human cancer: insights from single-cell analysis. *Nat Rev Cancer* 2020;20:218–32.
- Wu F, Li W, Zhao W. Synchronous ground-glass nodules showed limited response to anti-PD-1/PD-L1 therapy in patients with advanced lung adenocarcinoma. *Clin Transl Med* 2020;10:e149.

Supplementary Material

Materials and Methods

Human tumor samples

By performing left upper lobectomy, we have collected 3 nodules including one solid nodule and two mixed ground glass nodules (mGGN). Each resected nodule contained samples of freshly resected, frozen-section and formalin-fixed and paraffin-embedded (FFPE). Written informed consent of tumor acquisition for research has been obtained before surgery and approved by internal review board from Guangdong Lung Cancer Institute and Guangdong Provincial People's Hospital (Guangzhou, China, IRB approved protocol number GDREC2016175H).

Pathology evaluation and immunohistochemistry analysis

Comprehensive pathological evaluation was performed in 3 resected nodules. 4-um non-stained formalin-fixed and paraffin-embedded slides were set for immune staining to assess infiltrating immune cell subsets. The antibodies used in this study included: anti-CD3, anti-CD4, anti-CD20, anti-CD38, anti-CD163, anti-Foxp3. Staining results were manually evaluated and confirmed by specialized pathologist. Expression score was assessed based on staining quantity and intensity as follows: -, absence; +, low intensity; ++, intermediate intensity; +++, high intensity. Negative controls were used for all antibodies.

Expression of PD-L1 (22C3 pharmDx; Agilent Technologies Inc., Santa Clara, CA, USA) of tissue section was scored by specialized pathologist. PD-L1 expression was evaluated by defining the proportion of PD-L1 positive cells among tumor cells. Positive expression of PD-L1 was defined as $\geq 1\%$ for tumor cell while negative expression was opposite.

Multiplex Immunohistochemistry Staining and Analysis

5 slides of 5-um non-stained formalin-fixed and paraffin-embedded samples were obtained for multiplex immunohistochemistry staining. To identify the cell subsets expressing PD-L1 in the TME and multiplex immunofluorescence staining, we used PANO 7-plex IHC kit (cat 0004100100, Panovue, Beijing, China) including CD8, CD56, CD68, HLA-DR and PANCK. Different primary antibodies were sequentially applied, followed by horseradish peroxidase-conjugated secondary antibody incubation and tyramide signal amplification. The slides were microwave heat-treated after each TSA operation. Nuclei were stained with 4'-6'-diamidino-2-phenylindole (DAPI, SIGMA-ALDRICH) after all the human antigens had been labelled. To obtain multispectral images, the stained slides were scanned using the Mantra System (PerkinElmer, Waltham, Massachusetts, US), which captures the fluorescent spectra at 20-nm wavelength intervals from 420 to 720 nm with identical exposure time; the scans were combined to build a single stack image.

Images of unstained and single-stained sections were used to extract the spectrum of autofluorescence of tissues and each fluorescein, respectively. The extracted images were further used to establish a spectral library required for multispectral unmixing by inForm image analysis software 2.2.1. (PerkinElmer, Waltham, Massachusetts, US). Using this spectral library, we obtained reconstructed images of sections with the autofluorescence removed.

Fresh Tissues Processing for scRNA-seq

Fresh resected tissues within half hour after surgical resection were kept in MACS Tissue Storage Solution (Miltenyi Biotec) until processing. Samples were first washed with phosphate-buffered

saline (PBS), minced into small pieces (approximately 1mm³) on ice and enzymatically digested with 250U/mL collagenase I (Worthington), 100U/mL collagenase IV (Worthington) and 30U/mL DNase I (Worthington) for 45 min at 37°C, with agitation. After digestion, samples were sieved through a 70µm cell strainer, and centrifuged at 300g for 5 min. After the supernatant was removed, the pelleted cells were suspended in red blood cell lysis buffer (Miltenyi Biotec) to lyse red blood cells. After washing with PBS containing 0.04% BSA, the cell pellets were re-suspended in PBS containing 0.04% BSA and re-filtered through a 35µm cell strainer. Dissociated single cells were then stained with AO/PI for viability assessment using Countstar Fluorescence Cell Analyzer for quality assessment.

Single-Cell RNA Sequencing

The scRNA-Seq libraries were generated using the 10X Genomics Chromium Controller Instrument and Chromium Single Cell 3' V3 Reagent Kits (10X Genomics, Pleasanton, CA). Briefly, cells were concentrated to 1000 cells/uL and approximately 8,000 cells were loaded into each channel to generate single-cell Gel Bead-In-Emulsions (GEMs), which results into expected mRNA barcoding of 5,000 single-cells for each sample. After the RT step, GEMs were broken and barcoded-cDNA was purified and amplified. The amplified barcoded cDNA was fragmented, A-tailed, ligated with adaptors and index PCR amplified. The final libraries were quantified using the Qubit High Sensitivity DNA assay (Thermo Fisher Scientific) and the size distribution of the libraries were determined using a High Sensitivity DNA chip on a Bioanalyzer 2200 (Agilent). All libraries were sequenced by HiSeq Xten (Illumina, San Diego, CA) on a 150 bp paired-end run.

Single-Cell RNA Data Analysis

Fastq data was applied with default parameter filtering the adaptor sequence and removed the low-quality reads to achieve the clean data. Then the feature-barcode matrices were obtained by aligning reads to the human genome (GRCh38) using CellRanger v3.1.0. Cells contained over 200 expressed genes and mitochondria UMI rate below 20% passed the cell quality filtering. Mitochondria genes were removed in the expression table but used for cell expression regression to avoid the effect of the cell status for clustering analysis and marker analysis of each cluster.

Seurat package (version: 2.3.4, <https://satijalab.org/seurat/>) was used for cell normalization and regression based on the expression table according to the UMI counts of each sample and percent of mitochondria rate to obtain the scaled data. PCA was constructed based on the scaled data with all high variable genes and top 10 principals were used for tSNE construction. The fastMNN function ($k = 5$, $d = 50$, `approximate = TRUE`, `auto.order = TRUE`) from R package scran (v1.10.2) was used to apply mutual nearest neighbor method to correct for batch effect among samples. Utilizing graph-based cluster method, we acquired the unsupervised the cell cluster result based the PCA top 10 principal and we calculated the marker genes by FindAllMarkers function with wilcox rank sum test algorithm under following criteria: 1. $\log_{2}FC > 0.25$; 2. $P \text{ value} < 0.05$; 3. $\text{min.pct} > 0.1$. In order to identify the cell type detailed, the clusters of same cell type were selected for re-tSNE analysis, graph-based clustering and marker analysis.

Single-Cell Trajectories analysis was calculated with Monocle2 (<http://cole-trapnell-lab.github.io/monocle-release>) using DDR-Tree and default parameter. Before Monocle analysis, we select marker genes of the Seurat clustering result and raw expression counts of the cell passed filtering. Based on the pseudo-time analysis, branch expression analysis modeling (BEAM Analysis)

was applied for branch fate determined gene analysis. The function FindMarkers with wilcox rank sum test algorithm was used under following criteria: 1. $\log_{2}FC > 0.25$; 2. $P \text{ value} < 0.05$; 3. $\text{min.pct} > 0.1$ was used to identify differentially expressed genes among samples.

Targeted Panel Sequencing and Whole Exome Sequencing

1 μg of genomic DNA from 3 resected nodules were used for whole-exome sequencing. Genomic DNA from tumor frozen sections or the whole blood control sample was extracted using QIAamp DNA FFPE Tissue Kit (Qiagen) and DNeasy Blood and tissue kit (Qiagen), respectively, and fragmented by M220 Focused-ultrasonicator (Covaris) into approximately 250bp. Whole genome library was prepared using KAPA Hyper Prep Kit (KAPA Biosystems). For WES, exome capture was performed using the Illumina Rapid Capture Extended Exome Kit (Illumina Inc). For the targeted-panel, capture-based enrichment was performed using a customized panel targeting 425 cancer-related genes with Dynabeads M-270 (Life Technologies) and xGen Lockdown hybridization and Wash Reagent kit (Integrated DNA Technologies) according to the manufacturers' protocols. The captured libraries were then amplified with Illumina p5 (50 AAT GAT ACG GCG ACC ACC GA 30) and p7 primers (50 CAA GCA GAA GAC GGC ATA CGA GAT 30) in KAPA HiFi HotStart ReadyMix (KAPA Biosystems), and subsequently purified using Agencourt AMPure XP beads. Sequencing libraries were quantified by qPCR using KAPA Library Quantification kit (KAPA Biosystems). Library fragment size was determined by Bioanalyzer 2100 (Agilent Technologies). Both WES and target-enriched libraries were sequenced using the Illumina HiSeq 4000 platform (Illumina) according to the manufacturer's instructions.

Variant Filtering and Mutation Calling

Trimmomatic was used to trim adapters and remove low quality reads (quality reading below 20) or N bases from FASTQ files. Burrows-Wheeler Aligner (BWA) was then used to align clean paired-end reads to the reference human genome (hs37d5). PCR deduplication was performed using Picard and indel realignment and base quality score recalibration were performed using Genome Analysis Toolkit (GATK 3.4.0 (McKenna et al.,2010)). Matching of tumor and whole blood control samples was assessed using VCF2LR (GeneTalk) for same SNP fingerprint. Cross sample contamination was estimated using ContEst (Broad Institute) by evaluating the likelihood of detecting alternate alleles of SNPs reported in the1000g database. Somatic SNV and insertion/deletions (INDELs) were called using Vardict (V 1.5.4). SNVs and INDELs were further filtered using previously reported criteria (Zhang et al. 2019). Final mutations were annotated using vcf2maf.

HLA Alleles Typing

Typing for Class I HLA alleles was done through alignment with the IGMT/HLA database running Opitype (V1.3.1) with default parameters. Neoantigen calling was performed using the identified HLA types with netMHCpan (V4.0) for binding prediction under the NeoPrePipe pipeline.

Copy-Number Variants Analysis

Copy number analysis was performed using FACETS (Ver 0.5.13). Somatic CNA events were assigned based on sample-ploidy values calculated in the FACETS algorithm (Riaz et al., 2017). Chromosome arm-level CNA gain ($>$ sample average ploidy +1) was defined if segments of amplification and deep amplification events account for more than 60% of total segments for the

corresponding chromosome arm. Similarly, arm-level CNA loss ($<$ sample average ploidy -1) was identified if segments of deletion and deep deletion events account for more than 60% of total segments for the given chromosome.

TCR library preparation and sequencing

Genomic DNA from frozen section was extracted and quantified as described above. Minimum 1 μ g of DNA was required for subsequent experiments. Multiplex PCR reaction was prepared using QIAGEN Multiplex PCR Plus Kit with customized TCR primer mixture comprising 51 forward primers complementary to the V gene segments and 13 reverse primers complementary to the J gene segment. To correct amplification bias from multiplex PCR primers, 663 barcoded synthetic templates (i.e. a synthetic repertoire of all possible V-J combinations) were used to calibrate PCR efficiency. These templates contain universal P5 and P7 ends for standard primer recognition, barcodes and V gene and J gene segments flanking barcoded internal markers. Amplified synthetic products and tumor samples were then purified using AxyPrep MAG FragmentSelect-I Kit (Axygen). Subsequently, tumor TCR library was prepared with KAPA Hyper Prep Kit (KAPA Biosystems). Briefly, A-trailing and end-repair of fragments was performed before ligation of index adaptors in the TruSeq DNA PCR-free Library Prep Kit (Illumina). Purified ligation product was then amplified with Illumina p5 and p7 primers in the KAPA HiFi HotStart ReadyMix (KAPA Biosystems), followed by a final purification step using the Axygen beads. TCR library was sequenced using the Illumina HiSeq 4000 platform according to the manufacturer's instructions.

TCR analysis and profiling

Trimmomatic was used to remove adaptors and filter low quality reads from FASTQ files. Non-V-J paired reads was further removed using Cutadapt (V 1.18). Next, paired-end read merger (PEAR, V 0.9.10) was employed to merge paired reads and non-barcoded reads were removed for synthetic standards, while barcoded reads were filtered for tumor samples. Clean reads were subsequently assembled using MiXCR (V 2.1.11). Reads were aligned to the reference V or J gene segments according to the international ImMunoGeneTics (IMGT) database. Clonotypes were then built from alignments using the assemble pipeline of the software. For final repertoire profiling, sample V/J counts, CDR3 counts and clonality counts were calculated with normalization using corresponding counts from synthetic standards. Statistical analyses and graphical visualization were performed with R packages tcR and Vdjtools.

Signature Establishment and External Datasets

In this study, we established two signatures related to Trm cluster, MDSC-like macrophage cluster and TAM-like macrophage cluster using principal component analysis (PCA) to identify specific genes for different signatures. Trm signature contained CXCR6, ZNF683 and ITGAE. MDSC-like macrophages signature contained CCL2, PLA2G7, AQP9, VCAN, CD36 and FCN1. TAM-like macrophages signature contained GPNMB, C1QC, CCL13, SLC40A1 and FOLR2.

To validate prognostic value of the signatures we established in this study, we collected five different cohorts which have already published (referenced in the figure) or available from The Cancer Genome Atlas (TCGA) datasets (<https://gdc.cancer.gov/>).

Statistical analysis

All statistical analysis was performed in SPSS version 25.0 software (SPSS, Inc., Chicago, IL, USA). t-test was used for two group comparison and ANOVA or Kruskal–Wallis test was used for more than two group comparisons. Fisher’s exact test was used to compare two categorical variables. Kaplan Meier analysis was used to compare survival between groups and log-rank test p value was calculated. All reported p-value were 2-tailed, and the statistical significance was defined as $p < 0.05$.