Dysregulation of $T_{FH}$-B-$T_{RM}$ lymphocyte cooperation is associated with unfavorable anti-PD-1 responses in EGFR-mutant lung cancer

Jae-Won Cho¹, Seyeon Park², Gamin Kim³, Heonjong Han¹, Hyo Sup Shim⁴, Sunhye Shin³, Yong-Soo Bae⁵, Seong Yong Park⁶, Sang-Jun Ha², Insuk Lee¹ & Hye Ryun Kim³

Patients with non-small cell lung cancer (NSCLC) with epidermal growth factor receptor (EGFR) mutations exhibit an unfavorable response to PD-1 inhibitor through unclear mechanisms. Hypothesizing that EGFR mutations alter tumor-immune interactions, we compare tumor-infiltrating lymphocytes between EGFR mutant (EGFR-MT) and wild type (EGFR-WT) tumors through single-cell transcriptomic analysis. We find that B cells, CXCL13-producing follicular helper CD4+ T ($T_{FH}$)-like cells, and tissue-resident memory CD8+ T ($T_{RM}$)-like cells decreased in EGFR-MT tumors. The NOTCH-RBPJ regulatory network, which is vital for persistence of $T_{RM}$ state, is perturbed, and the interactions between $T_{FH}$ and B cells through the CXCL13-CXCR5 axis disappear in EGFR-MT tumors. Notably, the proportion of $T_{RM}$-like cells is predictive for anti-PD-1 response in NSCLC. Our findings suggest that the impairment of $T_{FH}$-B-$T_{RM}$ cooperation in tertiary lymphoid structure formation, accompanied by the dysregulation of $T_{RM}$ homeostasis and the loss of $T_{FH}$-B crosstalk, underlies unfavorable anti-PD-1 response in EGFR-MT lung tumors.

¹Department of Biotechnology, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Korea. ²Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Korea. ³Division of Medical Oncology, Department of Internal Medicine, Yonsei Cancer Center, Yonsei University College of Medicine, Seoul 03722, Korea. ⁴Department of Pathology, Yonsei University College of Medicine, Seoul 03722, Korea. ⁵Department of Biological Sciences, Science Research Center (SRC) for Immune Research on Non-lymphoid Organ (CIRNO), Sungkyunkwan University, Jangan-gu, Suwon, Gyeonggi-do 16419, Korea. ⁶Department of Thoracic and Cardiovascular Surgery, Yonsei University College of Medicine, Seoul 03722, Korea. ✉email: syparkcs@yuhs.ac; sjha@yonsei.ac.kr; insuklee@yonsei.ac.kr; nobelg@yuhs.ac
Epidermal growth factor receptor (EGFR) mutations are major drivers of oncogenic alterations that account for an important molecular subtype of lung adenocarcinoma, comprising 40–60% of cases in the South-East Asian population and 10–20% among Caucasian patients. Patients with lung cancer with sensitizing-EGFR mutations clinically benefit from EGFR-tyrosine kinase inhibitors (TKIs) as first-line therapies. However, acquired resistance inevitably develops within 9–18 months, which is the main limitation to the effectiveness of EGFR-TKIs.

Programmed death receptor-1 (PD-1) inhibitors improve survival outcomes by offering a durable response in patients with non-small cell lung cancer (NSCLC)1.2. Accordingly, these inhibitors are used as first- and second-line therapies for NSCLC3.4. With the increase in the number of checkpoint inhibitor-based clinical trials being conducted, interest in integrating immunotherapy for patients with EGFR-mutant type (EGFR-MT) lung cancer has also increased. However, recent large-scale clinical trials and meta-analysis data have revealed unfavorable treatment outcomes of these EGFR-MT patients compared to those with EGFR wild-type (EGFR-WT), indicating that patients with EGFR-MT do not benefit from PD-1 inhibitors5.6. In addition, pembrolizumab for the treatment of PD-L1-positive EGFR-TKI naive patients with sensitizing-EGFR mutation did not show objective response benefit7. A recent clinical trial study reported that a combination of an anti-PD-L1 inhibitor, atezolizumab, with bevacizumab and chemotherapy showed a survival benefit in pretreated EGFR-MT patients who had failed TKI8. Determining the mechanisms underlying the lower efficacy of PD-1 inhibitors in patients with EGFR-MT than EGFR-WT can guide the development of treatment strategies to improve the survival outcomes of these patients.

We predicted that the TME of EGFR-MT tumors differs from that of EGFR-WT tumors in patients with NSCLC. However, previous studies using immunohistochemistry (IHC) approaches did not comprehensively compare the TMEs of the two lung cancer subtypes. Single-cell RNA sequencing (scRNA-seq) enables high-resolution dissection of cellular heterogeneity of solid tissues, including tumors. A systematic comparison of single-cell transcriptome profiles between EGFR-WT and EGFR-MT tumors can reveal immune cell subsets that differentially accumulate between groups of molecular subtypes in NSCLC tumors, some of which may provide important clues regarding their difference in clinical outcomes. Additionally, identifying biomarkers and therapeutic targets can improve anti-PD-1 immunotherapy. Although multiple scRNA-seq studies have depicted immune cell landscapes of lung cancers9–12, EGFR-WT and EGFR-MT tumors have not been systematically compared.

In this work, we compare the single-cell transcriptome of tumor-infiltrating immune cells between the two molecular subtypes of lung cancer and seek key cell subsets and their molecular features associated with EGFR mutation. We find that B cells, CXCL13-producing follicular helper CD4+ T (Tfh)-like cells, and tissue-resident memory CD8+ T (TRM)-like cells are depleted in EGFR-MT compared to EGFR-WT. Homeostasis of TRM cells and CXCL13-CXCR5 axis-mediated crosstalk between Tfh and B cells are dysregulated in EGFR-MT tumors. Finally, the proportion of cytotoxic TRM-like cells estimated by the expression signature reveals a highly predictive value for the anti-PD-1 response. Altogether, these results suggest the importance of Tfh–B–TRM lymphocyte cooperation in the formation of tertiary lymphoid structure (TLS) for antitumor immunity. Furthermore, its impairment in EGFR-MT tumor results in an unfavorable response to anti-PD-1 immunotherapy.

**Results**

Comparative scRNA-seq analysis reveals immune cell subsets differentially accumulated between EGFR-mutant and wild-type tumors in NSCLC. To compare the cellular landscape between two different molecular subtypes of lung cancer based on the EGFR mutation status, we prepared tumor specimens surgically resected from 10 patients with NSCLC in stages IA–IIIA: five EGFR-WT and five EGFR-MT (Supplementary Data 1). Focusing on the landscape of tumor-infiltrating immune cells, we sorted CD45+ leukocytes from tumor tissues and conducted scRNA-seq using 10X Genomics technology for individual patients (Fig. 1a and Supplementary Fig. S1). A total of 33,719 cells were sequenced, 31,887 of which passed quality control (Supplementary Data 1). We preprocessed scRNA-seq data using Seurat v.3.0 package13 and visualized them through UMAP dimension reduction14. Systematic bias was observed among data sets across different patients (Fig. 1b). To correct batch effects, scRNA-seq data were integrated using Seurat v.3.0, which is based on canonical correlation analysis and anchoring data sets between samples15. After batch correction, the homogeneous population structures of patients and mutation status were reduced in the latent space by UMAP dimensions (Fig. 1c). Using the Louvain algorithm of network community detection, we defined 18 clusters (C0–C17) corresponding to distinct cell subpopulations (Fig. 1d). Based on the expression of canonical marker genes (Supplementary Fig. S2), we annotated clusters for major cell types such as CD8+ T cell (C0, C2, C11, and C14) (Supplementary Figs. S3 and 4), CD4+ T cell (C1, C3, C5, C6, and C9) (Supplementary Figs. S3 and 4), regulatory T cell (C4) (Supplementary Figs. S3 and 4), natural killer (NK) cell (C13) (Supplementary Fig. S5), natural killer T (NKT) cell (C7, C10, and C12) (Supplementary Fig. S5), B cell (C8) (Supplementary Fig. S6), macrophage and monocyte (C15) (Supplementary Fig. S7), dendritic cell (DC) (C16) (Supplementary Fig. S8), and innate lymphoid cell 2 (ILC2) (C17) (Supplementary Fig. S9). Many clusters apparently had unequal cell distributions between EGFR-WT and EGFR-MT (Fig. 1e), indicating differences in the composition of tumor-infiltrating immune cells between the two molecular subtypes of NSCLC. The majority of the clusters were composed of cells derived from most of the 10 patients without severe bias toward a specific patient. However, the three smallest clusters—C15 (macrophage and monocyte), C16 (DC), and C17 (ILC2)—were mostly derived from patient 6 (P6) (Fig. 1f and Supplementary Data 2). Thus, we excluded these clusters from downstream analyses to avoid formulating hypotheses based on patient-biased data.

To investigate proportional changes across the 15 cell subsets (C0–C14) between EGFR-WT and EGFR-MT, we evaluated deviations in the observed cell count from the expected counts for each subset using Pearson residual ($\chi^2$). As the size of the residual indicates the direction of the difference in the observed count from the expected count, $\chi^2 > 0$ and $\chi^2 < 0$ represent augmentation and depletion compared to the expected count, respectively. We considered only subsets with $\chi^2 > 3.5$ and $\chi^2 < -3.5$ as augmentation and depletion with high confidence, respectively, for follow-up functional interpretation (see “Methods” section for a detailed description of $\chi^2$ threshold selection). We visualized the results of the goodness of fit test using a mosaic plot, in which subsets with a deviation of observed from expected cell counts are indicated in blue (augmentation), red (depletion), or gray (no significant change) (Fig. 1g). The goodness of fit for all subsets was also evaluated by the chi-square statistic ($p$-value of Fig. 1g). We observed above-threshold level augmentation or depletion ($\chi^2 > 3.5$ or $\chi^2 < -3.5$) of the cell count for EGFR-WT and EGFR-MT from eight subsets (Supplementary Data 2). We expected that a
successful interpretation of differentially accumulated subsets between EGFR-WT and EGFR-MT would provide clues to understanding the mechanisms underlying unfavorable anti-PD-1 response in patients with EGFR-MT NSCLC. To characterize individual subsets for each major cell type (e.g., CD8\(^+\) T, CD4\(^+\) T, Treg, NK, NKT, and B cells), we identified differentially expressed genes (DEGs) of a subset compared to all other subsets for the same major cell type (subset-DEG) and DEGs between
EGFR-WT and EGFR-MT for each subset (group-DEG) (Supplementary Data 3).

**Proportion of cytotoxic NK cells decreases, whereas that of NKT cells with low cytotoxicity increases in EGFR-MT tumors.** NK and NKT cells are cytotoxic lymphocytes of the innate immune system that provide the first line of defense in the body before activating the adaptive immune system. Although these cells have a limited range of antigen specificity and lower capacity for memory-generation, they can kill cancer cells directly via cytotoxic molecules. The subset of cytotoxic NK cells (C13) was depleted in EGFR-MT, although its significance was not above the pre-determined threshold ($r = -3.34$) (Fig. 2a). Among the three subsets for NKT cells (C7, C10, and C12), the proportion of NKT-C7 and NKT-C10 was significantly augmented in EGFR-MT ($r = 6.35$ and 10.57, respectively; Fig. 2a). The subset-DEG analysis revealed that NK-C13 and NKT-C12, which were depleted and unchanged in EGFR-MT, respectively, expressed more genes involved in activation (TNFRSF18, IL2RB, CD7, CD44, TYROBP, and TNFRSF1B) and cytotoxicity (PRF1 and GZMB) than other subsets. In contrast, NKT-C7 and NKT-C10 proportion increased in EGFR-MT and these subsets expressed genes involved in cytotoxicity at lower levels than in other subsets. Overall, NKT-C10 is likely to have the lowest cytotoxic activity among all NK and NKT cell subsets, as the expression of genes involved in activation and cytotoxicity was lower than that in all other subsets (Fig. 2b, c), which is consistent with the highest augmentation in EGFR-MT ($r = 10.57$) (Fig. 2a). In summary, single-cell transcriptome analysis revealed the augmentation of NKT cell subsets with relatively low expression of genes involved in activation and cytotoxicity in EGFR-MT.

**Proportion of B cells decreases with reduced antigen presentation in EGFR-MT tumors.** B cells secrete antibodies to neutralize antigens or tag tumor cells for attack by other immune cells. Recently, tumor-infiltrated B cells have been reported to promote the response to immunotherapy and survival in cancer patients through the formation of TLSs, in which B cells can present a tumor-derived neoantigen to T cells to activate them. We observed a significant depletion of B cells in EGFR-MT.
EGFR-MT compared to EGFR-WT \((r = -6.23,\) Fig. 2d), suggesting that EGFR-MT tumor has a lower proportion of tumor-infiltrating B cells than EGFR-WT. In addition, we observed that B cells from EGFR-MT showed lower expression levels of HLA genes (MHC class I: HLA-C and MHC class II: HLA-DRB5, and HLA-DQA2) and TAP1 (antigen peptide transporter involved in processing antigens that will be presented to MHC class I) (Fig. 2e). These results suggest that B cells in EGFR-MT tend to less effectively activate both CD8+ and CD4+ T cells than EGFR-WT, due to the lower expression of MHC class I and II, respectively.

Proportion of CXCL13-producing T\textsubscript{FH}-like cells decreases, whereas that of immunosuppressive CD4+ T cells increases in EGFR-MT tumors. Along with CD8+ T cells, CD4+ T cells are major lymphocytes governing adaptive immunity; they differentiate into diverse types of helper T cells that activate other immune cells, including cytotoxic T cells. Accordingly, various conventional CD4+ T cells are involved in antitumor or tumor-promoting processes, whereas CD4+ Treg suppresses antitumor immunity in the TME. We obtained six subsets of CD4+ T cells (Fig. 3a), which were characterized by subset-DEG analysis for various categories of immune signature genes (Fig. 3b). Among the six CD4+ subsets, the two largest ones showed significant augmentation in EGFR-MT (C1 and C3 by \(r = 8.48\) and 7.66, respectively). We found that these subsets showed a high expression of genes involved in CD4+ T cell-mediated immune suppression (\(PD29,30,\) ANXA1,132, ANXA233, and CXC434). The C5 subset, for which the marker-based cellular identity was relatively unclear, showed augmentation in EGFR-MT with a subthreshold level of significance \((r = 2.74)\). The C6 subset, annotated as naive CD4+ T cells based on the high expression of \(CCR7\) and \(SELL\), showed no significant difference in relative abundance between EGFR-WT and EGFR-MT.

Notably, the C9 subset was CD4+ T cells with the largest depletion in EGFR-MT \((r = -7.37)\). We found several genes related to T\textsubscript{FH} cells (CXCL335, ICOS36, MAP37, PDCD138, and SH2D1A39) among the subset-DEGs for C9 (Fig. 3b, c and Supplementary Data 3). T\textsubscript{FH} cells reportedly help B cells by recruiting B cells with the function of antibody affinity in germinal centers (GC)35. PD-1 regulates GC positioning and helps with the function of T\textsubscript{FH} cells38. Human T\textsubscript{FH} cells produce CXCL13, a major chemoattractant in homing B cells to follicles35. CXCL13 has been associated with TLS development40,41, CXCR5, a receptor for CXCL13 and another hallmark of T\textsubscript{FH} cells, was not detected as a subset-DEG of C9. Moreover, the C9 subset does not express BCL6, a lineage-defining transcription factor of T\textsubscript{FH} cells35. These CXCR5 and BCL6-negative T\textsubscript{FH}-like cells that produce CXCL13 are found in local inflammatory sites41, where TLS development frequently occurs42. Given that the major role of the T\textsubscript{FH} cells is to recruit B cells, depletion of CXCL13-producing T\textsubscript{FH}-like cells in EGFR-MT may result in the inefficient formation of TLS. The observed depletion of tumor-infiltrating B cells (Fig. 2d) also suggests lower TLS development in EGFR-MT. Recently, B cells and TLS have been shown to be associated with responses to immunotherapy26-28. We also observed a substantially lower level of CXCL13 expression in EGFR-MT than in EGFR-WT tumor (Fig. 3d), which further supports the hypothesis that CXCL13-producing T\textsubscript{FH}-like cells contribute to the anti-PD-1 response through recruiting B cells to functional TLS in tumors. To confirm the reduced abundance of CXCL13+CD4+ T cells in EGFR-MT compared with EGFR-WT tumor, inferred from the single-cell transcriptome analysis, we performed the multiplexed immunofluorescence (IF) staining in archival tissue specimens of the same samples. We could obtain results of multiplexed IF for nine of the prior samples, because the sample quantity for the P3 patient was not sufficient for the downstream analysis. The results of multiplexed IF confirmed the reduced number of CD4+ and CXCL13+CD4+ T cells in EGFR-MT compared with EGFR-WT tumors (Fig. 3e, f and Supplementary Figs. S10, S11, Supplementary Data 4a). Although not statistically significant due to the small sample size, these results are congruous with the observations from the single-cell transcriptional analysis. To validate these findings further, we performed a similar analysis on a validation set of tumor samples from 19 independent NSCLC patients with EGFR-WT \((n = 9)\) or EGFR-MT \((n = 10)\) (Supplementary Data 5a). For this validation set with a larger sample size, we could observe significantly reduced frequencies of CXCL13+CD4+ T cells in EGFR-MT compared with EGFR-WT tumors for the validation set (Fig. 3g and Supplementary Fig. S12, Supplementary Data 4c).

The C4 subset was annotated as Treg cells based on the high expression signature genes of tumor-infiltrating Tregs (FOXP3, IL2RA, IL2RB, IL2RG, IL10RA, MAGEH1, LAYN, and GATA3). They also showed a high expression of genes involved in T cell activation (CD27, CD28, ICOS, and TNFRSF4 encodingOX40, TNFRSF9 encoding RBB, and TNFRSF18 encoding GITR) and checkpoint molecules (TIGIT and CTLA4). Counterintuitively, we observed a depletion of the Treg subset in EGFR-MT (Fig. 3a). However, we also found that Treg cells of EGFR-MT tumors expressed genes known to promote Treg activity (TSC22D3 encoding GILZ43 and NR4A244) at levels significantly higher than those in EGFR-WT (Fig. 3h and Supplementary Data 3), which may compensate for its disadvantages in abundance.

Proportion of cytotoxic CD8+ T\textsubscript{RM}-like cells decreases in EGFR-MT tumors. CD8+ T cells were clustered into four different subsets (C0, C2, C11, and C14), among which, only C2 exhibited significant differences in its proportion between EGFR-WT and EGFR-MT, particularly depletion in EGFR-MT \((r = -6.44)\) (Fig. 4a). We characterized each subset of CD8+ T cells based on the expression of various categories of signature genes (Fig. 4b) and found that the C2 subset had the highest cytotoxicity and expressed ITGAE (also known as CD103), a hallmark of T\textsubscript{RM} cells45. In addition, the C2 subset showed a significantly higher effector signature \((CD2, IFNG)\) and memory signature \((CD7)\). T\textsubscript{RM} cells are a subset of effector memory T cells in the peripheral tissue, and their roles in antitumor immunity have recently gained considerable attention46,47. We also found that the C2 subset highly expressed the signature genes for tumor-infiltrating T\textsubscript{RM}-like CD8+ T cells such as PDCDI1, HAVCR2, and LAG3, but rarely expressed TCF7, a lung T\textsubscript{RM} suppressor48-50 (Fig. 4b). Previously, HAVCR2 (also known as TIM-3) expression has been reported as a feature of lung tumor T\textsubscript{RM} cells that are not involved in exhaustion but rather reflect a functional state49. In this regard, we defined the C2 subset as T\textsubscript{RM}-like cells. As a fast-reactive memory T cell subset, T\textsubscript{RM} cells constitutively express perforin (PRF1), granzyme B (GZMB), and IFN-gamma (IFNG). We found that the tumor-infiltrating T\textsubscript{RM}-like subset expressed a high level of GZMB, whereas other subsets showed much lower expression, indicating the T\textsubscript{RM}-like subset as a major cytolytic lymphocyte in the tumor (Fig. 4c). High expression of ENTPD1 (also known as CD39) indicates that the C2 subset is composed of tumor-reactive T\textsubscript{RM}-like cells rather than bystander T cells51 (Fig. 4d). The expression of effector molecules and markers for naive/memory states suggest that C0 and C11 are effector CD8+ T cell subsets with lower cytotoxicity, and C14 is a naive T cell subset (Fig. 4e, f). Notably, we observed a higher expression of GZMB and genes involved in T cell
activation in T\textsubscript{RM}-like cells from EGFR-WT than in those from EGFR-MT (Fig. 4g), indicating that T\textsubscript{RM}-like cells in EGFR-WT tumors have the highest antitumor activity for the direct killing of cancer cells.

We evaluated the protein expression of selected molecules (CD8, CD103, and GZMB) using multiplexed IF for the whole slides to confirm the presence of T\textsubscript{RM}-like cells in an archival tissue specimen. The abundance of CD8\textsuperscript{+}, CD103\textsuperscript{+}CD8\textsuperscript{+}, and GZMB\textsuperscript{+} CD103\textsuperscript{+}CD8\textsuperscript{+} T cells tended to be lower in EGFR-MT than EGFR-WT (Fig. 4h, i and Supplementary Figs. S10, 11, Supplementary Data 4a), consistent with the results based on the transcriptional level. For the validation set of tumor samples from...
19 independent NSCLC patients (Supplementary Data 5a), we could observe a significantly lower abundance of CD8^+ and CD103^+ CD8^+ T cells in EGFR-MT compared with EGFR-WT (Fig. 4j and Supplementary Fig. S12, Supplementary Data 4c). For further validation, we performed a flow cytometry analysis using multiple antibodies, including anti-CD8, anti-CD103, and anti-GZMB for another validation set that consist of tumors from independent 35 NSCLC patients with EGFR-WT (n = 16) or EGFR-MT (n = 19) (Supplementary Data 5b). The T RM-like cell subset was highly enriched for cells with a high expression of both CD103 and GZMB (area within the solid circle) among CD8^+ cells (area within the dotted circle) in EGFR-WT compared to EGFR-MT (Fig. 4k). The proportion of CD103^+ cells among CD8^+ cells and that of GZMB^+ cells among CD8^+ cells was significantly lower in EGFR-MT than EGFR-WT (Fig. 4l, m). Therefore, the depletion of CD8^+ T RM-like cells in EGFR-MT lung tumors was confirmed by protein-based immunophenotyping analysis.

**Homeostasis of T RM-like cells is impaired in EGFR-MT tumors.** Hitherto our comparative single-cell transcriptome analysis between EGFR-WT and EGFR-MT revealed three lymphocyte subsets significantly depleted in lung tumors with EGFR mutations: B cells, CD4^+ T HF-like cells, and CD8^+ T RM-like cells. Given that patients with lung cancer with EGFR mutations exhibit unfavorable anti-PD-1 responses, regulation of cooperation among these lymphocyte subsets may be important for antitumor immunity in lung tumors. Based on the constitutively high expression of major cytolytic enzymes, such as perforin and GZMB, in T RM-like cells, the persistence of T RM cell state in a tumor may be a major determinant of the effective killing of cancer cells.

To evaluate the homeostasis of T RM-like cells in EGFR-WT and EGFR-MT tumors, we measured the transition probability between CD8^+ subsets (C0, C2, C11, and C14; Fig. 5a) based on RNA velocity analysis using velocyto52, which can be estimated from the ratio of unspliced to spliced mRNA. Based on the probability of future mRNA abundance for all genes, we can predict the future state (subset) of each cell (Fig. 5b). For example, the mean transition probability from C2 to C0 was higher in EGFR-MT than in EGFR-WT, whereas that from C2 to C2 (that is, the persistence of C2 state) was higher in EGFR-WT than in EGFR-MT (Fig. 5c). We compared and calculated the transition probability among four subsets, C0, C2, C11, and C14, between EGFR-WT and EGFR-MT tumors to identify the changes in homeostasis. Notably, C2 (T RM-like cells) derived from EGFR-WT and EGFR-MT showed probabilities of 0.57 and 0.41 for persistence, respectively (Fig. 5d and Supplementary Data 6). In contrast, all other subsets showed similar probabilities for persistence between EGFR-WT and EGFR-MT: 0.77 vs. 0.76 for C0, 0.16 vs. 0.16 for C11, and 0.042 vs. 0.034 for C14. We also performed scVeLo analysis that enables estimation of RNA velocity without steady-state assumption and observed consistent results (Supplementary Fig. S13). These results suggest that the persistence of the T RM cell state is impaired in EGFR-MT lung tumors.

**NOTCH-RBPJ regulatory network for maintaining T RM cells is perturbed in EGFR-MT tumors.** Transcriptional regulatory networks tightly control differentiation and functional maintenance of immune cells. We hypothesized that the dysregulation of the transcriptional regulatory network for homeostasis in the T RM cell state occurs in EGFR-MT lung tumors. To reconstruct the key regulatory network, we first identified differentially upregulated transcription factors (TFs) in the C2 subset compared to other CD8^+ cell subsets (Fig. 5e). Among the upregulated TFs, ID2, RBPI, and BHLLHE40 are known to be involved in T RM cell development and homeostasis54–56. Particularly, RBPI is associated with NOTCH to form a transcriptional activator required for maintaining T RM cells in humans55. Next, we compared the co-regulatory network of RBPI between EGFR-WT and EGFR-MT to identify the gain or loss of regulatory targets for RBPI between the two molecular subtypes of lung tumors. To improve the robustness of gene-gene correlation analysis for highly sparse scRNA-seq data, we employed the MetaCell57 analysis method, which partitions single-cell transcriptome profiles into disjoint and homogeneous groups of profiles. As expected, we observed a high correlation of expression between RBPI and many known genes for T RM cell differentiation and maintenance45–47 in C2 subset cells from EGFR-WT tumors (Fig. 5f). Additionally, NOTCH1 showed a high correlation with the same regulatory target genes, although it was not significantly upregulated in the C2 subset compared with other CD8^+ cell subsets. Interestingly, many transcriptional correlations with RBPI disappeared in the C2 subset cells of EGFR-MT. Particularly, the expression of RGS2, NOTCH1, RGS1, PDCDI1, CRTAM, CXCR6, TIA1, HIF1A, NAB1, and ITGAE positively correlated with RBPI in the C2 subset of EGFR-WT, but the gene-gene correlations were lost or even inversely correlated in the C2 subset of EGFR-MT (Fig. 5g and Supplementary Data 7). The loss of the transcriptional correlation in EGFR-MT was more marked with NOTCH1 (Fig. 5f and Supplementary Data 7). These results indicate that the persistence of the tumor-infiltrating T RM cell state depends on the NOTCH-RBPJ regulatory network, which is dysregulated in EGFR-MT lung tumors.

**Proportion of tumor-infiltrating T RM-like cells is highly predictive for anti-PD-1 response in NSCLC.** Given that the proportion of tumor-infiltrating T RM-like cells decrease in EGFR-MT lung tumors, which tends to show poor response to anti-PD-1 immunotherapy, we hypothesized that the relative abundance of T RM-like cells in the tumor by itself could predict the anti-PD-1 response. The current gold-standard estimation of the proportion of each immune cell type is flow cytometry analysis using canonical markers specific for each cell type. However, this approach is valid for only major cell types or subtypes with well-
characterized marker genes. Single-cell biology enables discovering an increasing number of cell subsets, and identifying specific markers for the subsets would be a daunting task. Therefore, quantifying cell subsets from a heterogeneous cell population using cell-type-specific expression signature has recently been employed. To infer the expression signature of the 18 subsets of immune cells from lung tumor samples, we applied CIBERSORTx, a deconvolution method developed for inferring gene expression signature for a specific cell type, to the single-cell transcriptome profiles and generated a signature matrix of 18 subsets (Supplementary Data 8a). We then estimated the relative abundance of TRM-like cells in The Cancer Genome Atlas.
(TCGA) cohort of lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC)\textsuperscript{59}. Consistent with the single-cell transcriptome results, the signature score for T\textsubscript{RM}-like cells was significantly higher in TCGA tumors with wild-type EGFR than in the mutant type ($p = 0.0173$, one-tailed Mann–Whitney U test) (Fig. 5h and Supplementary Data 8b). This result confirms that the inferred expression signature is a suitable estimator of the relative abundance of T\textsubscript{RM}-like cells in the tumor.

Next, we evaluated the expression signature of T\textsubscript{RM}-like cells to predict the response to anti-PD-1 immunotherapy using bulk RNA-seq data and clinical information available for two independent public cohorts of patients with NSCLC (GSE126045 and GSE135222)\textsuperscript{60,61}. Notably, we observed a positive correlation between the signature score, indicating an abundance of tumor-infiltrating T\textsubscript{RM}-like cells and anti-PD-1 response in both cohorts (Fig. 5i and Supplementary Data 8c, d). In the two NSCLC cohorts, the signature score of responders was significantly higher than that of non-responders ($p = 0.0114$ and 0.0404, one-tailed Mann–Whitney U test). Furthermore, we prioritized patients exhibiting the highest signature score. This stratification resulted in a large area under the receiver operating characteristic curve (AUROC) scores, indicating that a high abundance of tumor-infiltrating T\textsubscript{RM}-like cells is a predictive marker of the anti-PD-1 response (Fig. 5j). Notably, only 2 of the 18 patients in the cohort recruited by Yonsei Cancer Center (YCC)\textsuperscript{60} were EGFR-MT; one was a responder, whereas the other was not (EGFR mutation status information was not available for the cohort by Jung et al.\textsuperscript{61}). Therefore, the identified expression signature for T\textsubscript{RM}-like cells by itself can be used as a biomarker to stratify patients with lung cancer for anti-PD-1 immunotherapy regardless of their EGFR mutation status.

CXCL13–CXCR5-mediated T\textsubscript{FH}–B and T\textsubscript{RM}–B interactions are dysregulated in EGFR-MT tumors. Previous studies also demonstrated that T\textsubscript{FH} cells release CXCL13 to recruit B cells into TLS\textsuperscript{35}, which then present tumor antigens to T cells\textsuperscript{62}. As demonstrated in our study, the abundance of CXCL13-producing T\textsubscript{FH}-like cells was higher in EGFR-WT than in EGFR-MT patients (Fig. 3e, f). These observations suggest that recruiting B cells into the TLS of the tumor become less effective in EGFR-MT, likely due to a loss of CXCL13–CXCR5-mediated interactions, resulting in reduced cytotoxic T cell activation. To test this hypothesis, we evaluated cell-to-cell interactions among T\textsubscript{FH}-like cells (C9), B cells (C8), and cytotoxic T\textsubscript{RM}-like cells (C2) using CellPhoneDB\textsuperscript{63}. We found more significant cell–cell interactions in EGFR-WT than EGFR-MT (Fig. 6a). We found that T\textsubscript{FH}–B cell interaction through the CXCL13–CXCR5 axis was substantially weaker in EGFR-MT (lower mean expression of ligand and receptor) than in EGFR-WT. Notably, we could also detect CXCL13–CXCR5-mediated interactions between T\textsubscript{RM}–B cells in EGFR-WT but not in EGFR-MT. Indeed, T\textsubscript{RM}-like cells also express CXCL13, albeit not as much as in T\textsubscript{FH}-like cells.

Recently, a PD-1\textsuperscript{+}CD8\textsuperscript{+} T cell subset that secretes CXCL13 has been detected close to the B cell area of the TLS in lung cancer\textsuperscript{64}. B cells of EGFR-WT appeared to present more molecules involved in antigen processing and presentation than EGFR-MT (Fig. 3d). These molecules will promote the activation of CD4\textsuperscript{+} T\textsubscript{FH}-like cells (through MHC class II) and CD8\textsuperscript{+} T\textsubscript{RM} cells (through MHC class I). Subsequently, the enhanced antigen-presenting ability might boost B cell expansion and stimulate tumor-specific CD8\textsuperscript{+} T\textsubscript{RM} cells. In EGFR-WT, the abundance of CD8\textsuperscript{+} T\textsubscript{RM}-like cells and their GZMB production were higher than those in EGFR-MT (Fig. 4h, i, k, m). Consequently, the increased proportion of CD8\textsuperscript{+} T\textsubscript{RM}-like cells can effectively lyse cancer cells by producing cytolytic materials such as GZMB.

To test our hypothesis, we first defined the area of TLS-like lesion (see “Methods” section) and then evaluated the abundance of T\textsubscript{FH}-like (CXCL13\textsuperscript{+}CD4\textsuperscript{+}), B (CXCR5\textsuperscript{+}CD20\textsuperscript{+}), and T\textsubscript{RM}-like (CD103\textsuperscript{+}CD8\textsuperscript{+}) cells within the TLS-like lesions of archival tumor tissue samples using multiplexed IF. Interestingly, we found that the total area of TLS-like lesion was significantly smaller in EGFR-MT than EGFR-WT (Fig. 6b, left and right upper panels and Supplementary Data 4a). For the validation set of 19 independent tumor samples used for multiplexed IF, we also observed a smaller area of TLS-like lesion in EGFR-MT than EGFR-WT (Fig. 6b, right lower panel and Supplementary Data 4c). In overall TLS-like lesion, the much less number of CD4\textsuperscript{+} and CD8\textsuperscript{+} T cells were found in EGFR-MT than EGFR-WT (Fig. 6c, d and Supplementary Figs. S10, 11). Notably, the numbers of T\textsubscript{FH}-like (CXCL13\textsuperscript{+}CD4\textsuperscript{+}) and T\textsubscript{RM}-like (CD103\textsuperscript{+}CD8\textsuperscript{+}) cells in EGFR-MT were significantly lower than EGFR-WT (Fig. 6c, d and Supplementary Figs. S10, 11, Supplementary Data 4b). We observed consistent results from the validation set of 19 independent tumor samples (Fig. 6e and Supplementary Fig. S12, Supplementary Data 4d). These results altogether suggest that CXCL13–CXCR5-mediated interactions between T\textsubscript{FH}–B and T\textsubscript{RM}–B cells scaffold T\textsubscript{FH}–B–T\textsubscript{RM} lymphocyte cooperation in TLS for antitumor immunity in lung cancer as summarized in Fig. 6f.
Discussion
In the present study, we performed comparative single-cell transcriptome analysis between EGFR-WT and EGFR-MT lung tumors across subsets of CD45+ immune cells using scRNA-seq to elucidate the underlying mechanisms of the different response rates to anti-PD-1 immunotherapy. Recently, accumulating evidence has indicated that activating the oncogenic pathway in tumor cells could diminish the antitumor immune response by modulating immune cell differentiation and recruitment65. Our study reveals the significant depletion of tumor-infiltrating B cells, CD4+ T_FH-like cells, and CD8+ T_RM-like cells in EGFR-MT tumors. In addition, T_FH-like, B, and T_RM-like cells of EGFR-MT
showed a lower expression of genes involved in TLS formation, antigen presentation, and cytotoxicity, respectively, than EGFR-WT. EGFR-WT and EGFR-MT patient groups showed bias towards males with smoking history and female non-smokers, respectively, in both discovery and validation cohorts. Previous studies also reported a higher prevalence of EGFR-mutant NSCLC in female non-smokers than male smokers. To test whether gender or smoking status was a confounder of the observed depletion of TFH-like, B, and TRM-like cells in EGFR-MT tumors, we compared CIBERSORTx scores for the three cell types between different smoking status or genders in EGFR-WT and EGFR-MT groups of the NSCLC tumors of TCGA cohort, separately. In neither EGFR-WT nor EGFR-MT group, we could observe a significant decrease of the three cell types in non-smokers and females (Supplementary Fig. S14a), which indicates that neither smoking status nor gender was a confounder of the observed association between EGFR mutation status and TME phenotypes. In addition, we observed only a weak correlation of TRM-like cell abundance or TLS area with PD-L1 expression (Supplementary Fig. S14b), which suggests that PD-L1 was not a confounder of our observed association between EGFR mutation status and TME phenotypes. Altogether, these results suggest that a higher intrinsic resistance of EGFR-mutant tumors to anti-PD-1 immunotherapy is primarily due to the depletion of these three lymphocyte subsets, major functional components of TLS for antitumor immunity, and their dysfunctional modulation in tumors. Furthermore, we found the disappearance of TFH-like and TRM-like interactions mediated by the CCL13–CXCR5 axis in EGFR-MT tumors, which may negatively affect the formation of TLS in the tumor. Thus, targeting genes involved in these intra- and inter-cellular networks may improve the anti-PD-1 response in EGFR-MT patients.

Previous studies have demonstrated a correlation between the intratumoral abundance of TRM cells and patient prognosis in lung cancer. In this study, we demonstrate the feasibility of predicting the response to anti-PD-1 immunotherapy based on the proportion of TRM-like cells. Quantification of a specific intratumoral cell subset would be a useful biomarker for predicting clinical outcomes. Using the recently developed CIBERSORTx for gene expression deconvolution with single-cell transcriptome data, we extracted expression signatures for each cell subset in the lung tumor and found that CD8+ TRM-like cells were highly predictive of the anti-PD-1 response in two independent cohorts of NSCLC. Notably, we tested the expression signature of TRM-like cells constructed from lung tumors to identify anti-PD-1 responders in melanoma cohorts but observed no prediction power. Given that TRM cells permanently reside in the peripheral tissue, these results are attributable to the tissue-specific nature of TRM cells by adaptation to different local microenvironment.

In the present study, we employed Pearson residual analysis with mosaic plot visualization to systematically compare the cell subset distribution between sample groups (EGFR-WT and EGFR-MT tumors). Using the chi-square significance test with Bonferroni adjustment for multiple hypothesis tests, we could retrieve subsets showing proportional changes that were consistent with the observed phenotype differences between sample groups. The same analytical framework can be applied to various comparative single-cell transcriptome analyses between sample groups with distinct genotypes and phenotypes, and the interpretation of results is highly intuitive. The Pearson residual method can be applied in a comparative analysis between more than two groups, if not too many. Therefore, we anticipate expanded the applicability of the Pearson residual framework to single-cell transcriptome analysis in the future.

There are several limitations to this study. First, although we identified three cell subsets associated with antitumor immunity in the TLS area and their functional regulation, whether they are regulated independently or hierarchically coordinated remains unknown. For example, if CXCL13 produced by TFH-like cells promotes the persistence of TRM-like cells by interacting with recruited B cells, the transcriptional regulator of CXCL13 expression in T cells could be the primary target to improve an anti-PD-1 response in EGFR-MT tumors. Second, this study lacks data for myeloid cells (e.g., dendritic cells and myeloid-derived suppressor cells) and stromal cells (e.g., cancer-associated fibroblasts), which also contribute to cancer–immune interactions. Although they are not directly involved in killing cancer cells, their activity enhances or suppresses major cytotoxic lymphocytes through cytokine-mediated cell-to-cell communication and the presentation of tumor antigens. Therefore, a comparison of subsets of myeloid and stromal cells would reveal additional cellular components for determining the response to anti-PD-1 immunotherapy. Third, our study cohort has a bias toward the adenocarcinoma subtype of NSCLC (it includes only a single patient with the squamous carcinoma subtype). Thus, some findings from the present study are possibly relevant to the adenocarcinoma subtype only. Fourth, two patients with G719A (P5) and G719S (P8) were included in the present study. Although G719X are considered sensitizing mutations, these were not included in a phase 3 trial comparing osimertinib with other EGFR-TKIs for EGFR-mutant NSCLC patients as first line (FLAURA) and are under-represented in clinical trials compared to exon 19del and L858R. Fifth, although we demonstrated that...
smoking status and genders are not confounders of the observed decrease of TLS in EGFR mutants, we may not rule out other confounding factors. Much large sample sets of genomically and clinically annotated tumors will provide a more comprehensive assessment of our findings. Finally, although we demonstrated the importance of homeostasis of T_{RM}-like cells for antitumor immunity in TME of NSCLC through single-cell RNA velocity and co-regulatory network analysis, we could not provide functional validations in the present study. Thus, we need to conduct further experimental analyses in vitro and in vivo conditions to confirm our proposed mechanisms in future study.
In conclusion, we performed a comparative single-cell transcriptome analysis between EGFR-mutant and wild-type lung tumors and found several key lymphocyte subsets involved in the mechanisms underlying the anti-PD-1 response. In particular, we elucidated that T_{FH}-B–T_{R M} lymphocyte cooperation is critical to antitumor immunity in lung cancer. Dysregulation of the homeostasis of T_{R M} cell state and T_{FH}-B and T_{R M}-B crosstalk involved in T_{LS} development in EGFR-MT lung tumor may account for unfavorable responses to the antitumor immunotherapy with PD-1 blockades. Our findings provide guidelines for developing strategies to enhance anti-PD-1 efficacy in patients with EGFR mutations.

Methods

Tumor specimen collection and preparation. Patients diagnosed with NSCLC and who underwent surgical resection at Severance Hospital (Seoul, Republic of Korea) between June 2018 and August 2018, were prospectively enrolled. We consecutively enrolled five patients with EGFR-MT and five patients with EGFR-WT lung cancer. As depicted in Supplementary Data 1, six patients were in pathological stage I, two in stage II, and two in stage III; all patients with EGFR-MT were female never smokers, and those with EGFR-WT were males with a smoking history. This bias regarding the contribution of gender and smoking history to EGFR mutation has been previously reported in Asian patients with lung cancer.

No patient had undergone chemotherapy or radiation therapy before surgery (46.4 ± 4.7 months). This study was approved by the Institutional Review Board of Severance Hospital (IRB No. 4-2018-1210, 4-2014-0775, and 4-2020-0945). We have obtained written informed consents from all patients in this study.

We collected fresh tumor tissues on the day of the surgery, which were mechanically and enzymatically dissociated using a gentleMACS dissociator (Miltenyi Biotec, Gladbach Bergisch, Germany; Cat# 130-093-235) and the Human Tumor Dissociation Kit (Miltenyi Biotec, Cat# 130-095-929) following the manufacturer’s instructions. After incubation for 1 h at 37 °C, the suspended samples were filtered through a 70-μm MACS SmartStrainer (Miltenyi Biotec, Cat# 130-098-462) into RPMI-1640 medium (Corning, Inc., Corning, NY, USA) supplemented with 10% fetal bovine serum (Biowest, Riverside, MO, USA) and centrifuged at 300 × g for 10 min. The pellet was suspended in PBS containing 2% fetal bovine serum. For fluorescence-activated cell sorting (FACS) analysis, single-cell suspensions were stained for 20 min at 4 °C with the following fluorescent-dye-conjugated antibodies: LIVE/DEAD® Fixable Red Dead Cell Stain Kit (Invitrogen, Carlsbad, CA, USA; Cat# 3300–05), CD3 (Biolegend, San Diego, CA, USA; clone 2D1.1:100). The samples were sorted using a FACS Aria III sorter (BD Biosciences, Franklin Lakes, NJ, USA) into live and single CD45+ cells. The sorted cell population was analyzed to ensure successful cell sorting, as shown in Supplementary Fig. S1. Cell viability was assessed by the Luna FL Fluorescence Cell Counter (Logos Biosystems, Anyang, Korea, Cat#LP2001) and confirmed to be 75–98% in all samples.

10× single-cell RNA sequencing. Single-cell samples for constructing libraries and sequencing were prepared using a Chromium platform according to the 10× Single Cell 3’ v2 protocol (10× Genomics, Pleasanton, CA, USA). To acquire 10,000 targeted cells, the cells were diluted in nuclease-free water and mixed with the master mix. The single-cell suspensions were loaded with Single Cell 3’ Gel Beads and Partitioning Oil (GEM) into a Single Cell 3’ Chip. To generate the GEMs, single cells were uniaxially barcoded, and cDNAs were used to construct libraries according to the manufacturer’s instructions. The libraries were sequenced using a HiSeq 2500 sequencing system (llumina, San Diego, CA, USA).

Single-cell RNA sequencing data analysis. From the raw FASTQ files of scRNA-seq data (forward/reverse strands and barcode sequence), we used Cell Ranger analysis pipelines (v2.1.1) to obtain a gene-count matrix. We used the Seurat version3.0.19 pipeline to analyze scRNA-seq data. For quality control, we filtered cells with a higher percentage of mitochondrial genes or those showing outlier behavior in feature-mRNA and nCount RNA from the “FeatureScatter” plot for each sample as recommended. We then normalized the data using the “NormalizeData” function with the recommended setting (log-normalization scale factor of 10,000). The normalized samples were integrated with 7431 highly variable genes by integration functions (“FindIntegrationAnchors” and “IntegrationData”) with default settings. For dimension reduction and visualization of the integrated data, “ScaleData,” “RunPCA,” and “RunUMAP” functions were used with default settings. Unsupervised clustering was performed by “FindNeighbors,” which conducts shared-nearest-neighbor-based clustering with 2000 highly variable genes and “FindClusters,” which conducts Louvain clustering (dimension of 20 and resolution of 0.65). The parameters for UMAP plotting before batch correction were the same as those for the integrated one.

Comparative single-cell transcriptome analysis based on Pearson residual. To investigate proportional changes between sample groups (e.g., EGFR-WT and EGFR-MT) across cell subsets defined by clustering, we evaluated deviations in the observed cell count for each group from the expected count for a subset using Pearson residual:

$$ \text{Pearson residual} (r) = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}} $$

where $i$ and $j$ represent indices for each group and cell subsets, respectively, and $O$ and $E$ represent the observed and expected cell counts, respectively. The expected cell count ($E$) for a group $i$ of a subset $j$ was calculated by the following equation:

$$ E_{ij} = \frac{T_i T_j}{T_{ij}} $$

where $T_i$, $T_j$, and $T_{ij}$ represent total cell count for the entire data set, total cell count for a group $i$, and total cell count for a subset $j$, respectively. The advantage of the Pearson residual is that the sign of the residual indicates the direction of the difference of the observed count from the expected count (i.e., positive for augmentation and negative for depletion compared to expected count). Pearson residual ($r$) follows an approximately normal distribution; thus, scores larger than 2 or smaller than −2 are significant by $p < 0.05$. For a highly conservative statistical test along with Bonferroni correction for multiple hypothesis test adjustment, we used a significance threshold of p-value = 0.00056 (p-value of 0.01 was divided by the number of subsets, 18, for Bonferroni correction). Thus, we considered only cell subsets with $r > 3.5$ (augmentation) and $r < −3.5$ (depletion) for follow-up functional interpretation. We visualized the results of the goodness of fit testing using a mosaic plot, in which subsets with deviation in observed cell counts from the expected cell counts are indicated by blue (augmentation), red (depletion), or gray (no significant change) colors. The goodness of fit for all subsets was also evaluated by the chi-square statistic (p-value).

Categorizing cell subsets into major immune cell types. We used marker gene expression to categorize cell subsets into immune cell types. The following list of marker genes was adapted from the Reference Gene Expression Profiles (REGEP)71: CD3D, CD3E, CD3G, and CD28 for T cells; CD4 for CD4+ T cells; CD8A and CD8B for CD8+ T cells; FOXP3 and IL2RA for Treg cells; FCGR3A, FCGR3B, NCAM1, KLRB1, KLRK1, KLRD1, KLRF1, and KLRK1 for NK and NKT cells; CD69, CD163, and EGFR for macrophages and monocytes; CD19, MS4A1, CD79A, CD79B, and BLNK for B cells; IL3RA, CLEC4C, and NRP1 for dendritic cells; and ITGALL for ILC2 cells. The expression of these marker genes was...
visualized as violin plots (Supplementary Fig. S2) and validated on UMAP (Supplementary Figs. S3–9).

Differentially expressed gene analysis with single-cell transcriptome data. We performed two types of DEG analysis using the “FindMarkers” function. The “group-DEG” was evaluated by comparing the EGF-R WT group to the EGF-R MT group for each subset. The “subset-DEG” was evaluated by comparing one subset to all other subsets within a major cell type (i.e., CD8+ T cells, NK cells, CD4+ T cells). The DEGs were accepted if their respective FDR was smaller than 0.05 and |avgLogFC| > 0.25 (default threshold by Seurat package). We excluded genes with high expression but a low ratio of non-zero cells.

RNA velocity-based estimation of cellular transition probability. We performed RNA velocity analysis using velocyto24 using the default pipeline. We processed single-cell transcriptome data for EGF-R WT and EGF-R MT tumors separately. We used the pre-built UMAP coordinate of CD8+ T cells from the Seurat package as the embedding coordinates for plotting. The transition probability was obtained from the “transition_prob” variable. RNA velocity predicts the transcriptional abundance of each gene from the ratio of spliced and unspliced mRNA levels based on a steady-state model. For ScVelo25 (ver 0.2.3) analysis which enables estimation of RNA velocity without steady-state assumption, we used the default pipeline. Loom files for splicing and unsplicing count information for each cell were obtained from RNA velocity analysis. The transition probability was obtained from the “scvelo.tl.transition_matrix” class. The future state of each cell is predicted based on all genes. For a given subset of cells, the probability of transition to another subset can be calculated for each cell. The mean values of transition probability for each subset of cells were obtained by averaging the corresponding cell transition to other cell subsets.

Immunofluorescence staining. We cut 4 μm-thick tissue sections from formalin-fixed paraffin-embedded blocks, and then performed multiplex immunofluorescence staining with a Leica Bond Rx (Leica Biosystems, Germany). Brieﬂy, the slides were heated for 30 min and dewaxed with Leica Bond Dewax solution (#AR9222), followed by antigen retrieval with Bond Epitope Retrieval 2 (#AR9640, Leica Biosystems) in pH 9.0 solution for 30 min. Each sample was down-sampled to 7000 randomly selected CD8+ T cells from the Seurat package embedding coordinates for plotting. The transition probability was obtained from the “transition_prob” variable. RNA velocity predicts the transcriptional abundance of each gene from the ratio of spliced and unspliced mRNA levels based on a steady-state model. For ScVelo53 (ver 0.2.3) analysis which enables estimation of RNA velocity without steady-state assumption, we used the default pipeline. Loom files for splicing and unsplicing count information for each cell were obtained from RNA velocity analysis. The transition probability was obtained from the “scvelo.tl.transition_matrix” class. The future state of each cell is predicted based on all genes. For a given subset of cells, the probability of transition to another subset can be calculated for each cell. The mean values of transition probability for each subset of cells were obtained by averaging the corresponding cell transition to other cell subsets.

MetaCell-based analysis of gene expression correlations. The correlation coefficient calculated from single-cell gene expression data was markedly low, mainly because of the zero-inflation property of scRNA-seq data. The sparsity of single-cell expression profiles can be compensated by imputation approaches, but they generate many false gene-gene correlations22. For more sensitive and robust detection of co-expression between two genes, we employed MetaCell57, which pools k nearest neighbors. We collected the raw count matrix from the Seurat object of single-cell data for EGF-R WT and EGF-R MT groups, separately. Each data set was processed using default parameter settings in the MetaCell pipeline without cell filtering step. T_k = 0.08 for mccell.get.transition_matrix(), T_top3 = 100 and T_top5 = 2 for mccell.get.transition_matrix_cov(); K = 100 and T_top3 = 100 and T_top5 = 2 for mccell.get.transition_matrix_cov(); K = 30, min_mc_size = 20, p = r = 0.75, and n = 500 for mccell_cooccur_graph(); K = 30, min_mc_size = 30, and alpha = 2 for mccell_mc_from_graph(); and T_lfc = 3 for mccell_mc_split_flip(). The pooled or “metacell” was generated from the corrected expression matrix calculated by the regularized geometric mean of each “metacell.” We calculated Spearman’s rank correlation coefficient between two genes based on the corrected expression matrix.

Flow cytometry analysis. Flow cytometry was performed using CytoFLEX (Beckman Coulter, Brea, CA, USA). Data were analyzed using FlowJo software (Tree Star, Ashland, OR, USA). The apoptosis and necrosis of immunologically unaltered monomacrophages against the following proteins were used: CD8 (RPA-T8, Cat# 301048, Biolegend, San Diego, CA, USA, 1:50), CD1 (SK7, Cat# 344808, Biolegend, San Diego, CA, USA, 1:100), PE-1 (EHL1.217, Cat# 329933, Biolegend, San Diego, CA, USA, 1:20), CD80 (Bev-3AT, Cat# 3501, Biolegend, San Diego, CA, USA, 1:20), CD39 (Cat# 328210, Biolegend, San Diego, CA, USA, 1:20), CD73 (QA16A2, Cat# 372214, Biolegend, San Diego, CA, USA, 1:50), and CD4 (RPA-T4, Cat# 560837, BD Biosciences, San Diego, CA, USA, 1:50). The LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit was from Invitrogen (L34973, Waldbam, Massachusetts, USA, 1:100). The sequential gating strategies used for analysis for data obtained from flow cytometry are described in Supplementary Fig. S15.

Analysis of receptor-ligand-mediated cell-cell interactions. We used CellPhoneDB26 to analyze the receptor–ligand-mediated cell–cell interactions. We used the raw count matrix with the 7431 highly variable genes based on the scRNA-seq data for cells of C0–14 clusters as input data. We ran the EGF-R WT and EGF-R MT data separately with the default pipeline and parameters.
Received: 12 December 2020; Accepted: 29 September 2021; Published online: 18 October 2021

References
1. Hsu, W. H., Yang, J. C., Mok, T. S. & Loong, H. H. Overview of current systemic management of EGFR-mutant NSCLC. Ann. Oncol. 29, 13–19 (2018).
2. Mok, T. S. et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. N. Engl. J. Med. 376, 629–640 (2017).
3. Soria, J. C. et al. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. N. Engl. J. Med. 378, 113–125 (2018).
4. Brahmer, J. et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N. Engl. J. Med. 373, 123–135 (2015).
5. Borghaei, H. et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N. Engl. J. Med. 373, 1627–1639 (2015).
6. Herbst, R. S. et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet 387, 1540–1550 (2016).
7. Reck, M. et al. Updated analyses of KEYNOTE-026 pembrolizumab versus platinum-based chemotherapy for advanced non-small-cell lung cancer with PD-L1 tumor proportion score of 50% or greater. J. Clin. Oncol. 37, 537–546 (2019).
8. Gainor, J. F. et al. EGFR mutations and ALK rearrangements are associated with low response rates to PD-1 pathway blockade in non-small-cell lung cancer: a retrospective analysis. Clin. Cancer Res. 22, 4585–4593 (2016).
9. Dong, Z. Y. et al. EGFR mutation correlates with unin...
10. Lee, C. K. et al. Checkpoint inhibitors in metastatic EGFR-mutated non-small cell lung cancer. J. Thorac. Oncol. 12, 403–407 (2017).
11. Lisberg, A. et al. A phase II study of pembrolizumab in EGFR-mutant, PD-L1+...
This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation funded by the Ministry of Science and ICT (NRF-2019M3A9B6065231, 2017M3A9E802971, 2017M3A9E9072669 to H.R.K., 2018R1A5A2025879, 2018M3C9A5064709, 2019M3A9B6065192 to I.L., 2019M3A9B6065221, 2018R1A2A1A05076997, 2017R1A5A1014560 to S.-J.H.). The work was supported in part by Brain Korea 21(BK21) FOUR program.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-26362-0.

Correspondence and requests for materials should be addressed to Seong Yong Park, Sang-Jun Ha, Inseuk Lee or Hye Ryun Kim.

Peer review information Nature Communications thanks Dwight H. Owen, Manu Setty and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information are available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements
This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation funded by the Ministry of Science