

Single-cell dissection reveals the role of aggrephagy patterns in tumor microenvironment components aiding predicting prognosis and immunotherapy on lung adenocarcinoma

Xinti Sun^{1,*}, Fei Meng^{1,*}, Minyu Nong^{2,*}, Hao Fang¹, Chenglu Lu³, Yan Wang¹, Peng Zhang¹

¹Department of Cardiothoracic Surgery, Tianjin Medical University General Hospital, Tianjin 300052, China

²School of Clinical Medicine, Youjiang Medical University for Nationalities, Baise, Guangxi, China

³Department of Pathology, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China

*Equal contribution

Correspondence to: Peng Zhang; email: pengzhang01@tmu.edu.cn

Keywords: aggrephagy, lung cancer, bioinformatics, tumor microenvironment, scRNA

Received: July 17, 2023

Accepted: November 6, 2023

Published: December 13, 2023

Copyright: © 2023 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Lung adenocarcinoma (LUAD) is one of the leading malignant cancers. Aggrephagy plays a critical role in key genetic events for various cancers; yet, how aggrephagy functions within the tumor microenvironment (TME) in LUAD remains to be elucidated.

Methods: In this study, by sequential non-negative matrix factorization (NMF) algorithm, pseudotime analysis, cell-cell interaction analysis, and SCENIC analysis, we have shown that aggrephagy genes demonstrated various patterns among different cell types in LUAD TME. LUAD and Immunotherapy cohorts from public repository were used to determine the prognosis and immune response of aggrephagy TME subtypes. The aggrephagy-deprived prognostic score (ADPS) was quantified based on machine learning algorithms.

Results: The cancer-associated fibroblasts (CAFs), tumor-associated macrophages (TAMs), and CD8+ T cells have various aggrephagy patterns, which enhance the intensity of intercellular communication and transcription factor activation. Furthermore, based on the signatures of the newly defined aggrephagy cell subtypes and expression profiles of large cohorts in LUAD patients, we determine that DYNC1I2+CAF-C1, DYNLL1+CAF-C2, PARK7+CAF-C3, VIM+Mac-C1, PARK7+Mac-C2, VIM+CD8+T_cells-C1, UBA52+CD8+T_cells-C2, TUBA4A+CD8+T_cells-C3, and TUBA1A+CD8+T_cells-C4 are crucial prognostic factors for LUAD patients. The developed ADPS could predict survival outcomes and immunotherapeutic response across ten cohorts ($n = 1838$), and patients with low ADPS owned a better prognosis, lower genomic alterations, and are more sensitive to immunotherapy. Meanwhile, based on PRISM, CTRP, and CMAP databases, PLK inhibitor BI-2536, may be a potential agent for patients with high ADPS.

Conclusions: Taken together, our novel and systematic single-cell analysis has revealed the unique role of aggrephagy in remodeling the TME of LUAD. As a newly demonstrated biomarker, the ADPS facilitates the clinical management and individualized treatment of LUAD.

INTRODUCTION

Non-small cell lung cancer (NSCLC) is a malignancy with a notable prevalence and mortality rate, wherein the lung adenocarcinoma (LUAD) represents the foremost

prevalent pathological subclass [1]. The amassing of molecular insights through burgeoning technologies has facilitated the development of targeted therapeutics alongside traditional interventions, such as surgery and chemotherapy. The 5-year relative survival rate of LUAD

patients, at 21%, mirrors the difficulties encountered in managing this aggressive malignancy, frequently detected at an advanced stage, whereby the disease has already metastasized to other regions of the body [2]. The difficulty in treating LUAD is often due to its resistance to traditional chemotherapy and radiation therapy, which could lead to different individual therapeutic responses and contribute to its poor survival rate. Tumor heterogeneity, including characterized by diverse TME patterns as well as cancer cell types, represents a significant contributor to reduced response rates and drug resistance [3]. Thankfully, advances in molecular biology, genomics, and proteomics might help unravel the complexity as well as heterogeneity at the molecular level, leading to new individualized strategies for LUAD treatment.

Autophagy represents a highly conserved eukaryotic cellular recycling process that plays an indispensable role in the degradation of cytoplasmic organelles, proteins, as well as macromolecules, essential for the survival and maintenance of cells [4]. Significantly, autophagy serves as a sustainable source of biomolecules and energy to maintain homeostasis under stressful conditions, including those encountered within the tumor microenvironment [5]. Misfolded proteins accumulate in cells to form protein aggregates, which interfere with the normal physiological activities of cells and cause various human diseases [6, 7]. Aggrephagy is a specific clearance pathway for intracellular protein aggregates, which is a type of selective autophagy. It undertakes most of the tasks involved in degrading protein aggregates and plays a decisive role in the abnormal protein quality control system. This pathway has potential applications in the treatment of various diseases, like neurodegenerative diseases, muscular atrophy, and cancer [8]. Previous studies have indicated that protein misfolding and aggregation can impact the function of the p53 protein in cancer. When the p53 protein is mutated or aggregated, it can lose its functionality, leading to tumor progression [9–11]. Consequently, comprehending the mechanisms of aggrephagy has arisen as a promising approach for targeted therapy as well as holds potential for further research in drug development. Single-cell transcriptomics has uncovered the intricate intercellular crosstalk between diverse subtypes of cells in the TME and cancer cells, revealing a complex network of intercellular signaling pathways that regulate tumor growth and progression [12]. The TME consists of a diverse range of cellular components, including CAFs, TAMs, T cells, and tumor cells. It is defined by a tumor-promoting and immunosuppressive phenotype, characterized by complex intercellular signaling networks that regulate the progression and growth of the tumor [13]. Meanwhile, TME has long been known as a nutrient-

depleted environment, study indicated that the autophagy of cancer cells rescued itself from T cell-mediated cytotoxicity by blocking cytokine-induced apoptosis [14]. Zhao et al. found that autophagy loss impedes CAFs activation via downregulation proline biosynthesis [15]. Additionally, wang et al. discovered that elevated TUBA1A, a classical aggrephagy markers, lead the worse clinical outcomes of gastric cancer patients, and be associated with the infiltration of macrophages in the TME [16]. PARK7, alternatively referred to as DJ-1, is overexpressed in a substantial 86% of individuals diagnosed with NSCLC. 72.2% predominantly exhibit PARK7 expression in the cytoplasm, and the heightened expression of this gene is strongly correlated with unfavorable clinical outcomes, including increased risk of recurrence and reduced overall survival rates for LUAD patients [17, 18]. However, there has been limited research investigating the cell-cell interactions between TME cell subtypes and prognosis associated with aggrephagy modification in LUAD.

We explored the influence of aggrephagy on the main TME cells based on LUAD single-cell RNA sequencing data. Through NMF analysis of 44 aggrephagy genes, as previously described [19], it was discovered that distinct expression patterns of aggrephagy mRNA in various subpopulations of LUAD TME cells demonstrated extensive and diverse intercellular communication with epithelial cells, and associated with different immune characteristics, metabolic pathways, as well as transcription characteristics. Moreover, upon integration with the bulk RNA-seq data of sizable LUAD patient cohorts, we substantiated that multiple aggrephagy cell subtypes exerted a substantial influence on both the prognosis and response to immune checkpoint blockade (ICB) therapy. Based on this, we created and multi-center assessed a 32-gene combined ADPS via machine learning algorithms. ADPS has shown strong predictive ability for survival outcomes, immunotherapy response, and drug efficacy in multiple multicenter cohorts. To our knowledge, this extensive single-cell analysis is the first to unveil the potential role of aggrephagy mRNA in mediating intercellular communication between TME cells and tumor cells, thus impeding the progression of LUAD. Furthermore, our findings provide crucial insights for early detection, prognostic assessment, risk stratification, and personalized therapeutic interventions in clinical settings.

MATERIALS AND METHODS

Data acquisition

scRNA-seq data were obtained from the GEO database under accession number: GSE149655 [20]. LUAD bulk RNA-seq data including clinical traits were acquired

from the GDC portal of TCGA and GEO databases: TCGA-LUAD, GSE3141, GSE31210, GSE37745, GSE50081, and GSE68465. The six datasets of 1512 patients were integrated and batch effects were adjusted by the Combat algorithm using the “sva” package [21]. Normal lung tissue data were available from the GTEx and TCGA databases. Two datasets treated with PD-(L)1 and containing clinical traits were downloaded: IMvigor210 [22], and GSE78220 [23]. Supplementary Table 1 summarizes the data sources and details of this study. A total of 44 aggrephagy genes were downloaded from https://www.gsea-msigdb.org/gsea/msigdb/cards/REACTOME_AGGREPHAGY.

scRNA-seq data process

The normalization, integration, dimension reduction, and clustering were performed stepwise using the Seurat pipeline with the R package “Seurat” [24]. Cell annotation was performed by referring to common tumor microenvironment cell markers, published studies [25, 26], the CellMarker website (<http://xteam.xbio.top/CellMarker/>), and the PanglaoDB website (<https://panglaoDB.se/>). Normalization data from the Seurat object were analyzed by single-cell NMF based on aggrephagy gene expression [27]. Cells that expressed no aggrephagy-related genes and aggrephagy-related genes that had no expression in all cell types were removed during the analysis. The NMF method was set to snmf/r and a maximum of ten clusters was used. The FindAllMarkers function was employed to identify the markers of each NMF cluster for every cell type in LUAD. Clusters with aggrephagy genes with log₂ (fold change) exceeding 1.0 were termed (“Gene + Cell type”). Clusters with no characteristics of aggrephagy genes were termed “Non-Aggrephagy-Cell type”. Clusters with characteristics aggrephagy genes with log₂ (fold change) less than 1 were termed “Unclear-Cell type”.

Pseudotime trajectory, cell-cell interaction analysis, and transcription factor analysis

To explore the correlation between aggrephagy genes and cell pseudotime trajectories, we utilized the R package “Monocle2” to analyze scRNA data for all cell types in LUAD [28, 29]. In short, the Monocle object underwent size factor and dispersion estimation, followed by the identification of highly variable features. Afterward, dimensionality reduction was performed and cell ordering was carried out for the purpose of pseudotime visualization. Besides, cell-cell interactions were conducted using “CellChat”, an R package that identifies and quantifies intercellular communication between different cell types within a single-cell dataset [30]. Secreted signaling in humans

was included in the cell-cell interaction analysis [31]. We utilized the “pySCENIC” package, which is a Python-based implementation of the SCENIC pipeline, to explore the transcription factor (TF) gene regulatory network in LUAD [32]. The scRNA-seq data of LUAD were subjected to analysis using two gene-motif rankings (hg19-tss-centered-10 kb and hg19-500 bp-upstream) obtained from the RcisTarget database, with the aim of detecting transcription start sites (TSS) and gene regulatory networks. TFs with a Benjamini-Hochberg false discovery rate (BH-FDR) <0.05 were selected for further investigation.

Gene set scoring

The package “GSVA” was utilized to perform single-sample gene set enrichment analysis (ssGSEA) for gene set scoring in bulk RNA sequencing data [33, 34]. GSVA was also employed to assess the previously established CAF-subtype in the single-cell RNA sequencing data [35]. The “AddModuleScore” function was utilized to evaluate the expression of M1-like/M2-like polarization markers derived from published studies [36]. The metabolic scores of different aggrephagy cell subtypes were calculated using “scMetabolism” package [37].

Survival analyses and ICB response prediction

We utilized the “GSVA” package to calculate the gene signature scores of the aggrephagy cell subtypes across all LUAD cohorts. The relationship between aggrephagy-related NMF signatures and patients’ prognosis was explored using the log-rank test and Cox proportional hazard regression. The “survminer” package was applied to plot Kaplan-Meier curves and determine the cutoff values of different NMF cell signatures in the different LUAD cohorts. Furthermore, we utilized the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm available at <http://tide.dfci.harvard.edu/> to obtain TIDE scores, which enabled us to predict the clinical response to ICB in LUAD cohorts. Prognosis-associated genes were identified by performing univariate Cox regression analysis with a significance threshold of $P < 0.05$ on the aforementioned aggrephagy cell subtypes that were associated with prognosis. To compress the gene number and identify the variables that have the greatest impact on the target variable, we employed a two-step approach to analyze the data, beginning with a LASSO Cox regression analysis to shrink the coefficient estimates. Subsequently, a multivariate Cox regression analysis with stepwise regression method was performed to identify the most significant predictors of the prognosis [38]. Based on the results of the multivariate Cox model, ADPS was calculated using the following formula: $ADPS = \sum \beta \times \text{Exp}_i$.

Having established that i denotes a gene in the ADPS, exp_i represents the expression level of gene i , as well as β_i stands for the coefficients of gene i . Then, we performed zero-mean normalization on the ADPS of the patients and categorized them into high ADPS and low ADPS groups.

Development of potential therapeutic agents

We followed the protocol outlined by Yang et al. [39] to identify potential agents for LUAD patients with high ADPS: (1) We downloaded drug sensitivity data for cancer cell lines (CCLs) from the Cancer Therapeutic Response Portal (CTRP, includes 481 compounds over 835 CCLs, <https://portals.broadinstitute.org/ctrp>) and profiling relative inhibition simultaneously in mixtures (PRISM, includes 1448 compounds over 482 CCLs, <https://www.theprismlab.org/>) datasets, as well as expression data of CCLs from the Cancer Cell Line Encyclopedia (CCLE, <https://portals.broadinstitute.org/ccle/>) database. (2) The CTRP and PRISM datasets possess AUC values, and decreased AUC values indicate heightened responsiveness to this compound. (3) Using the Wilcoxon rank-sum test, we conducted a differential analysis of drug response between the top 10% (high ADPS group) and bottom 10% (low ADPS group) of samples. We set a threshold of $\log_2FC > 0.1$ to identify compounds with significantly lower AUC values in the high ADPS group. (4) Furthermore, we employed Spearman correlation analysis to detect compounds with AUC values that showed negative correlation coefficients (using a threshold of $R < -0.4$) for subsequent screening. (5) We further identified potential agents by taking the intersection of the compounds obtained from steps (3) and (4). In the end, according to differential expression analysis, we identified potential agents using Connectivity Map (CMap, <https://clue.io/>) [40] to further verify the agents derived from the CTRP and PRISM databases.

Multi-omics alteration characteristics targeting ADPS

GISTIC 2.0 analysis (<https://gatk.broadinstitute.org>) was applied to distinguish recurrently amplified and deleted regions for genomic alterations. TCGA GDC data Portal was used to download “maf” data for LUAD samples (VarScan2Variant Aggregation and Masking; <https://portal.gdc.cancer.gov>). The TMB score with high ADPS and low ADPS groups was further calculated according to the “maftools” package [41]. The fraction of genome alteration (FGA), the fraction of genome gained (FGG), as well as the fraction of genome lost (FGL), were calculated as follows: $FGA = \text{total CNV/all bases}$, $FGG = \text{gain bases/all bases}$, and $FGL = \text{loss bases/all bases}$. These metrics were used to evaluate the extent of genomic alterations in the samples.

Statistical analysis

A Wilcoxon rank-sum test was utilized to compare the continuous variables. Cox regression was performed using the “survival” package to analyze the relationship between variables and survival outcomes. The glm function was utilized for logistic regression to examine the relationship between variables and binary outcomes. The log-rank P test was utilized for the Kaplan-Meier analysis. The ROC curve was plotted via the “timeROC” package and the calibration curve was plotted using the “rms” package. Statistical significance was set at $P < 0.05$.

Availability of data and materials

The original contributions presented in the study are included in the Article/Supplementary Materials; further inquiries can be directed to the corresponding authors.

RESULTS

Heterogeneity of aggrephagy genes across cells

The workflow chart illustrates the general design of this study (Figure 1A). We found that the aggrephagy score was significantly increased in normal samples compared LUAD samples analyzed via the ssGSEA algorithm (Figure 1B, $P < 0.001$). We then examined the landscape of aggrephagy on each cell type by using the scRNA-seq data of LUAD. After using the Seurat pipeline, a total of 12,554 cells were divided into 18 clusters (Figure 1C) and were annotated with major cell types based on classical marker genes, including mast cells, macrophages, plasma cells, T cells, stromal cells, epithelial cells, fibroblasts, and smooth muscle cells (Figure 1D). Marker genes for each cell population showed a clear separation between each cell type (Figure 1E). To further investigate whether aggrephagy activity was dynamic among TME in LUAD scRNA-seq level, we quantified aggrephagy score using ssGSEA, AUCcell, Ucell, addmodule score, and singscore algorithms (Figure 1F). Interestingly, the outcomes demonstrated relatively heightened aggrephagy score in smooth muscle cells, fibroblasts, epithelial, stromal, and macrophage, while lower in T, plasma, and mast cells. Moreover, a comparison between average aggrephagy score in LUAD and normal tissues unveiled intriguing observations: macrophage, plasma, T, epithelial, and fibroblasts within tumors exhibited notably low average aggrephagy score, while stromal cells exhibited notably high average aggrephagy score (Figure 1J). As a result of comparing the expression patterns of aggrephagy genes across cell types in LUAD, we found that aggrephagy genes also showed high heterogeneity

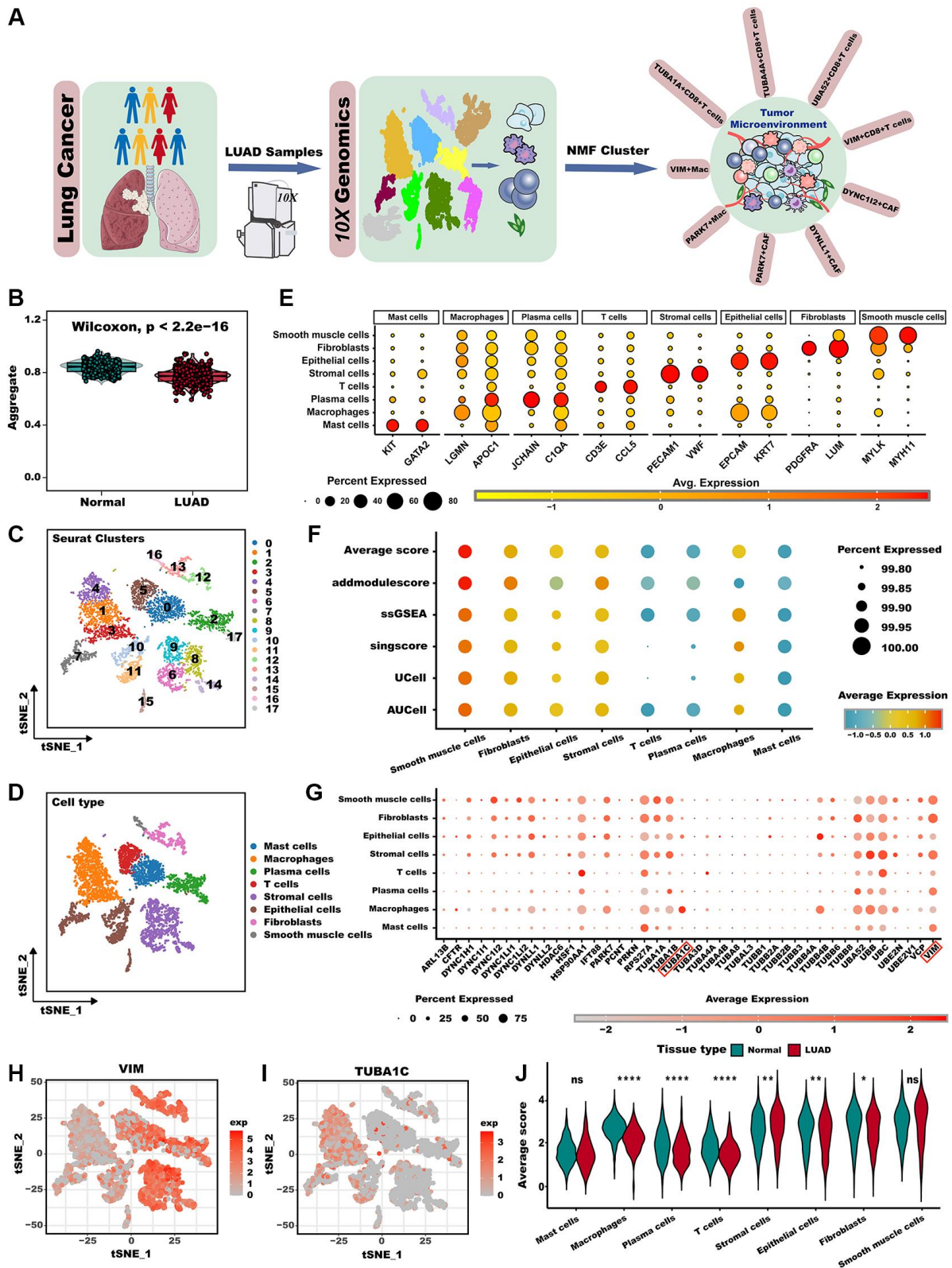


Figure 1. Overview of aggrephagy gene expression in the scRNA-seq for LUAD. (A) The overall design of this study. (B) The GSVA scores based on the aggrephagy gene set were compared between normal and LUAD samples. (C) t-SNE plot colored by 18 cluster of cells. (D) Cell type annotations clustered by Seurat t-SNE in the scRNA-seq data. (E) Dot plot showing representative marker genes for each cell type. (F) Bubble plot showing aggrephagy scores for each cell type. (G) Dot plot showing the expression level of aggrephagy genes in all cell types. (H, I) Expression of example genes VIM (H), and TUBA1C (I) in different cell types. (J) The difference in average aggrephagy score in each cell type in normal and LUAD samples. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.

among cell types (Figure 1G). For instance, VIM exhibited higher expression levels in fibroblasts, and plasma cells, while showed lower expression levels in T cells and macrophages (Figure 1H). TUBA1C is highly expressed in macrophages and epithelial cells, but shared lower expression levels in other cells (Figure 1I), indicating the necessity for further investigation of aggrephagy genes by targeting specific cell types.

Novel aggrephagy-mediated fibroblasts contribute to the TME of LUAD

Recent research has recognized CAFs as a key factor in TME, and they are emerging as a research hotspot [42]. Advancements in single-cell analysis techniques have led to a more comprehensive comprehension of the diversity and role of CAFs. For instance, according to one of the most well-established categorizations of CAFs, CAFs can be categorized as myCAF (myfibroblastic CAF), dCAF (development CAF), iCAF (inflammatory CAF), as well as pCAF (Pdpn CAF). Distinctive characteristics differentiate various subgroups of CAFs. Specifically, myCAF subset is situated in proximity to cancer cell nests and is characterized by elevated expression levels of both FAP+ and α -SMA. Conversely, the iCAF subset is localized far from cancerous cells and is marked by low expression levels of α -SMA but high expression levels of IL-6; the dCAF subset is distinguished by the activation of genes associated with diverse types of stem cells [43]. Similarly, to determine whether aggrephagy expression has an effect on CAFs, we performed a dimension reduction analysis. Pseudotime analysis reveals that aggrephagy genes are expressed at various developmental stages (Figure 2A). For instance, the early development stages of CAFs were characterized by the significant expression of TUBA4A, TUBA1C, and ARL13B, whereas TUBB4B, VIM, and TUBA1B were the feature of late development. Based on the NMF algorithm, CAFs were further divided into four clusters, and we identified that DYNC1I2+CAF-C1, DYNLL1+CAF-C2, PARK7+CAF-C3, as well as Non-Aggrephagy-CAF-C4 subtypes (Figure 2B, Supplementary Table 2). The developing status of NMF-based CAF clusters varied greatly as shown in pseudotime analysis (Figure 2C, 2D). Interestingly, using the Cellchat analysis, DYNC1I2+CAF-C1, DYNLL1+CAF-C2, and PARK7+CAF-C3 presented more and tighter connections with other cell types than Non-Aggre-CAF-C4 (Figure 2E–2G). GSVA was employed to derive scores for established classical CAF phenotype markers, with the aim of uncovering the possible phenotypes and functions of the aggrephagy CAF subtypes [44, 45] (Figure 2I). DYNC1I2+CAF-C1 showed the most prominent

scores of pan-myCAF, pan-dCAF, and pan-pCAF, while PARK7+CAF-C3 exhibited the prominent scores of pan-iCAF. In contrast, Non-Aggrephagy-CAF-C4 had the lowest score among all classical CAF subtypes. In addition, we also investigated important CAF phenotype markers, like pro-inflammatory genes, neo-angiogenic genes, and MMPs. From the pathway heatmap (Figure 2H), aggrephagy-CAF subtypes had a significantly different expression of common pathway genes. PARK7+CAF-C3 had significantly higher expression of MMPs, ECM, Neo-Angio, and pro-inflammatory pathways genes, DYNC1I2+CAF-C1 had significantly higher expression of TGFb and RAS pathways genes, while Non-Aggre-CAF-C4 had lower expression of above pathways genes. Ultimately, we compared TF regulation features and subtype-specific TFs between aggrephagy-related CAFs as well as non-aggrephagy-related CAFs. We found that the average activities of TFs including FOS and ATF3 were exclusively higher in DYNC1I2+CAF-C1, while downregulated in DYNLL1+CAF-C2 except for FOXO3, NFIA, and CEBPB (Figure 2J). In aggregate, these discoveries imply that aggrephagy could exert a significant influence on CAF remodeling within the TME of LUAD.

Aggrephagy participate in TAMs metabolism and polarization

Subsequently, we investigated the potential impact of aggrephagy on the phenotypes and functions of TAMs. Throughout the pseudotime analysis, aggrephagy genes were observed to exhibit diverse expression patterns at different stages of macrophage development, implying their complex and dynamic roles in regulating the differentiation and function of macrophages (Figure 3A). NMF with aggrephagy genes separated macrophages into VIM+Mac-C1 and PARK7+Mac-C2 (Figure 3B, Supplementary Table 3). The developing status of NMF-based TAM clusters varied greatly as shown in pseudotime analysis (Figure 3C, 3D). Further, VIM+Mac-C1 and PARK7+Mac-C2 had comparatively stronger interactions with other cellular components, especially with epithelial cells (Figure 3E, 3F). We inferred the specific pathways of intercellular communication and found that mainly outgoing (secreting) of VIM+Mac-C1 and PARK7+Mac-C2 were EGF and VISFATIN, and incoming (target) signal pathways were PTN and MK (Figure 3G). A tumor microenvironment induces metabolic reprogramming of macrophages, leading to protumor macrophages having an overactive metabolism [37], we conducted scMetabolism algorithm to explore the metabolic heterogeneity in different aggrephagy TAM subtypes (Figure 3H). Interestingly, the result revealed that the two TAM subtypes exhibit unique metabolic activation

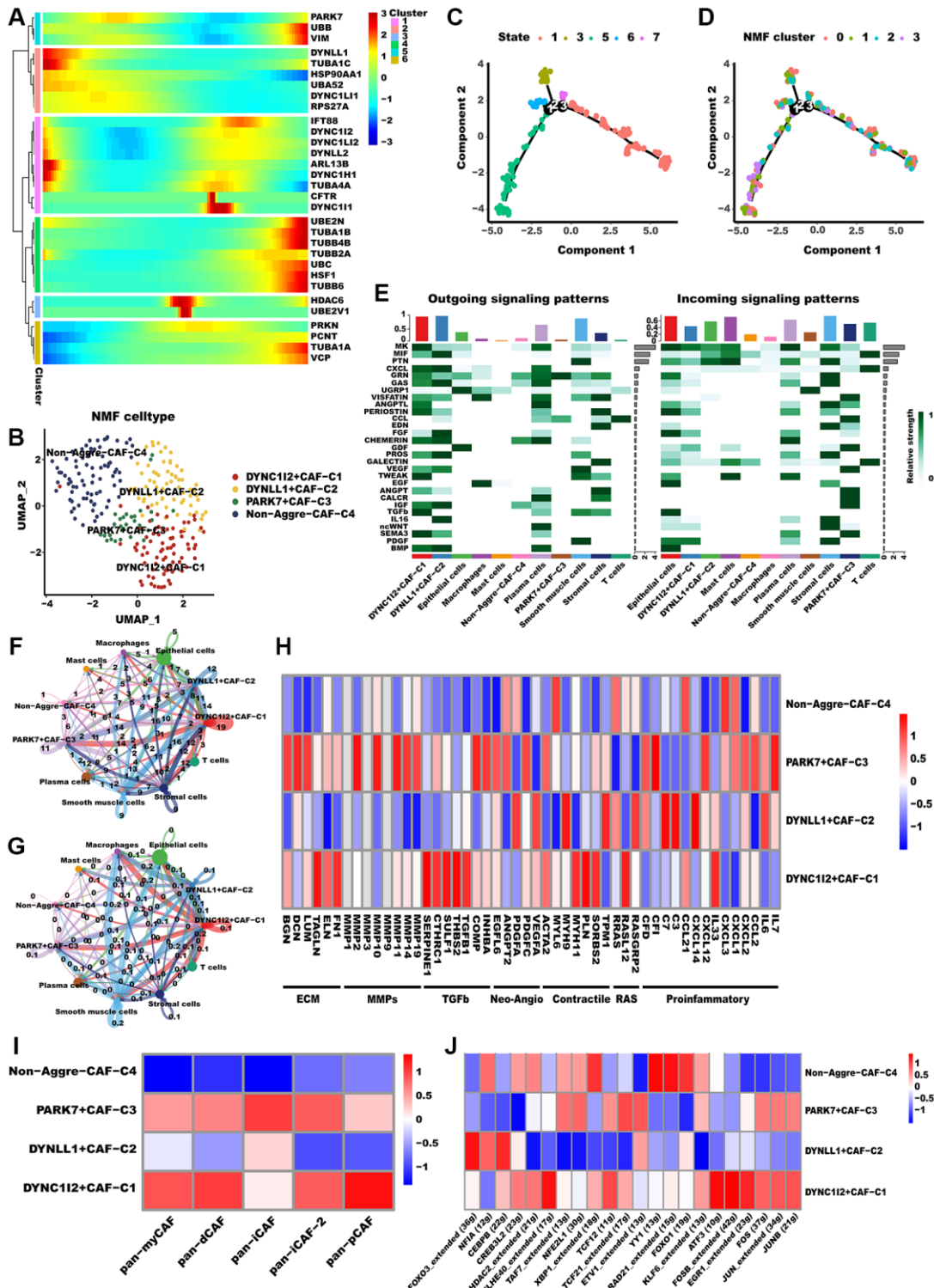


Figure 2. Aggrephagy genes modified the features of CAF. (A) Pseudotime trajectory analysis of aggrephagy genes in CAFs. (B) NMF clustering and annotation in CAFs classified by aggrephagy gene expression features. (C, D) The developing status of NMF-based CAF clusters obtained in pseudotime analysis. (E) A heat map summarizing the outgoing (secreting) and incoming (target) signal pathways of NMF-based aggrephagy-related CAF subtypes and other cell types. (F, G) The number (F, number of interactions) and weight (G, interaction weights/strength) of cell-cell interactions between Agg-related CAF subtypes and other cell types. (H) Heatmap showing the different average expression of common signaling pathway genes in the aggrephagy-related CAF subtypes, including collagens, ECM, MMPs, TGFb, Neo-Angio, Contractile, RAS and Proinflammatory. (I) Different aggrephagy-related CAF subtypes were correlated with the previous signatures. (J) Heatmap showing the significantly different activities of TFs among aggrephagy-related CAF subtypes by comparing the average AUC using pySCENIC in Python.

pathways, indicating that aggregophagy could be a critical modulator of metabolic regulation. For instance, PARK7+Mac-C2 was distinguished by the

upregulation of multiple metabolic pathways such as glycolysis, TCA cycle, fatty acid degradation, as well as oxidative phosphorylation, while VIM+Mac-C1 was

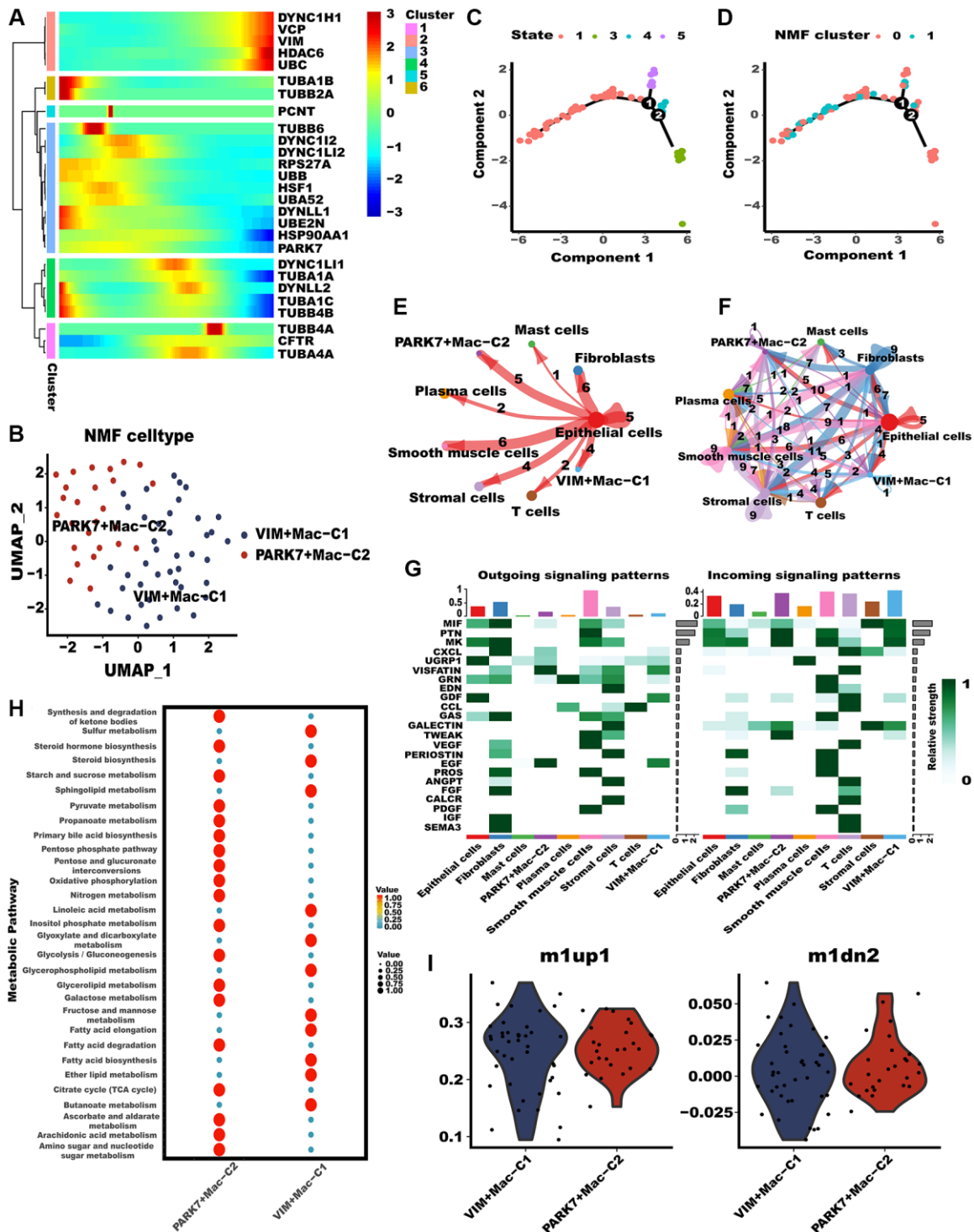


Figure 3. Tumor-associated macrophages (TAMs) differed in metabolism and polarization during aggregophagy. (A) Pseudotime trajectory analysis of aggregophagy genes in TAMs. (B) NMF clustering and annotation in TAMs classified by aggregophagy gene expression features. (C, D) The developing status of NMF-based TAM subtypes obtained in pseudotime analysis. (E, F) The number and weight of cell-cell interactions between aggregophagy-related TAM subtypes and other cell types. (G) A heat map summarizing the outgoing (secreting) and incoming (target) signal pathways of NMF-based aggregophagy-related TAM subtypes and other cell types. (H) Metabolic status of aggregophagy-related TAM subtypes analyzed by scMetabolism. (I) M1-like/M2-like phenotype scoring among different aggregophagy-related TAM subtypes.

distinguished by the upregulation of fatty acid elongation and ether lipid metabolism. In addition, we computed the M1-like/M2-like polarization scores of different aggrephagy TAMs subtypes [46]. Notably, PARK7+Mac-C2 and VIM+Mac-C1 exhibited a preference for expressing genes that were upregulated in M1 macrophages, suggesting their enhanced M1-like anti-tumor functions (Figure 3I).

Aggrephagy contributed to the subgroup formation in CD8+T cells

As CD8+T cells are known to significantly impact tumor metastasis and treatment outcomes, we conducted a comprehensive analysis of these cells [47]. Likewise, Monocle2 revealed that aggrephagy genes occupy different developmental stages (Figure 4A). TUBA1A, concordant with CAFs, was highly expressed during late development, while PARK7 and UBA52 were expressed during early development. Furthermore, CD8+T cells were categorized as four clusters via the NMF algorithm, and the developing status of NMF-based CD8+T cells varied greatly (Figure 4C, 4D, Supplementary Table 4), as well as named them as VIM+CD8+T cells C1, UBA52+CD8+T cells C2, TUBA4A+CD8+T cells C3, and TUBA1A+CD8+T cells C4, respectively (Figure 4B). Similarly, the aforementioned aggrephagy CD8+ T cells subtypes were found to have more extensive and robust interactions with other components of the TME, indicating their potential role in orchestrating the tumor immune microenvironment (Figure 4E). Furthermore, we inferred the specific pathways of intercellular communication and found that VIM+CD8+T cells C1 and TUBA1A+CD8+T cells C4 presented stronger activity in the CD40 pathways compared to UBA52+CD8+T cells C2, TUBA4A+CD8+T cells C3 (Figure 4F). We computed the T exhaustion and T cytotoxic scores for these aggrephagy related CD8+ T cell subtypes, utilizing a previously published panel for calculating scores that evaluate overall functions and phenotypes. These scores play a crucial role in determining the effects of ICB and its significance in LUAD progression [48]. Interestingly, we found that UBA52+CD8+T cells C2 showed higher T cytotoxic scores, TUBA4A+CD8+T cells C3 demonstrated higher T exhaustion scores but lower T cytotoxic scores, while VIM+CD8+T cells C1 and TUBA1A+CD8+T cells C4 presented lower both scores (Figure 4G), indicating TUBA4A likely to be involved in CD8+T cell exhaustion, result in tumor immune escape. In addition, we performed a comparison of the mean expression levels of genes related to T cell function and immune checkpoint inhibitors across these subtypes of CD8+ T cells. Consequently, VIM+CD8+T cells C1, UBA52+CD8+T cells C2, TUBA4A+CD8+T cells C3,

and TUBA1A+CD8+T cells C4 exhibited different immune function-related terms (Figure 4H). Ultimately, we investigated the regulatory mechanisms through network regulatory analysis. TUBA1A+CD8+T cells C4 demonstrated apparent activation of TFs including FOS, JUNB, FOSB, BCLAF1, and BATF (Figure 4H). In conclusion, our investigation has illustrated the role of aggrephagy in the restructuring of CD8+ T cells within the TME of LUAD.

Aggrephagy-mediated TME remodeling contributes to prognosis and immunotherapy response in LUAD

We have set out to investigate whether the newly defined aggrephagy subtypes have an impact on the survival rates of patients diagnosed with LUAD. Initially, we depicted the intercellular communication among all the aggrephagy-related subtypes, which elucidated all the possible connections between these subtypes (Figure 5A). Then we utilized ssGSEA algorithm to compute the enrichment score of each aggrephagy subtype based on the corresponding differentially expressed genes (Supplementary Table 5) and investigated their prognostic significance in LUAD patients across multiple cohorts. Through the implementation of univariate Cox regression analysis, we acquired the hazard ratio of each aggrephagy cell subtype in TCGA, GSE68465, GSE50081, GSE37745, GSE3141, GSE31210, Meta-GEO, and Meta cohorts, which showed that PARK7-Mac-C2, UBA52+CD8+T_cells-C2, PARK7+CAF-C3, and VIM+Mac-C1 were poor prognostic factors (Figure 5B, Supplementary Table 6). In addition, certain aggrephagy-related cell subtypes could distinguish the OS in TCGA, GSE68465, GSE50081, GSE37745, GSE3141, GSE31210, and Meta-GEO cohorts (Supplementary Figure 1A–1AJ). Remarkably, all cell subtypes related to aggrephagy were able to differentiate overall survival (OS) among meta-cohorts, which included 1512 patients with LUAD (Figure 5D–5L). Having established the potential role of aggrephagy in shaping the TME, we proceeded to investigate whether the aggrephagy-based patterns of TME had an impact on the response to ICB therapy. We employed the TIDE algorithm to predict the response of each patient to ICB therapy across multiple cohorts. Additionally, we observed that in the TCGA cohort, DYNC1I2+CAF-C1, DYNLL1+CAF-C2, and PARK7+CAF-C3, TUBA4A+CD8+T_cells-C3 were downregulated in responders, suggesting that these cells may be linked to ICB resistance (Figure 5M), which was verified in other cohorts (Supplementary Figure 2). Utilizing logistic regression, we obtained the odds ratio value for response prediction, unveiling the detrimental impact of these cells (Figure 5C, Supplementary Table 7). To verify the presence of the above cell types, we

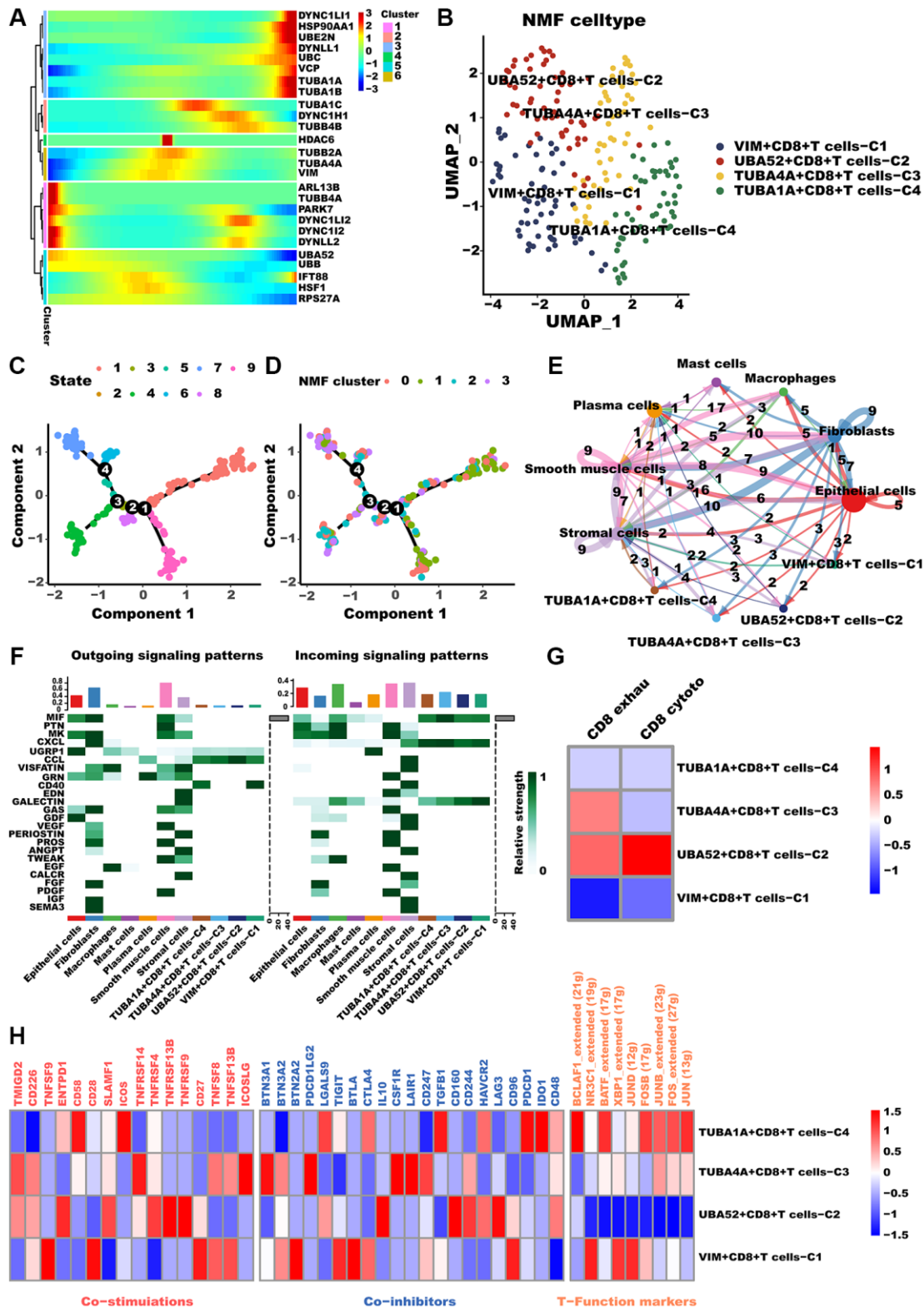


Figure 4. CD8+T cells differed in metabolism and polarization during aggregaphy. (A) Pseudotime trajectory analysis of aggregaphy genes in CD8+T cells. (B) NMF clustering and annotation in CD8+T cells classified by aggregaphy gene expression features. (C, D) The developing status of NMF-based CD8+T cells subtypes. (E) The number and weight of cell-cell interactions between aggregaphy-related CD8+T cells subtypes and other cell types. (F) A heatmap summarizing the outgoing (secretory) and incoming (target) signal pathways of NMF-based aggregaphy-related CD8+T cells subtypes and other cell types. (G) Heatmap showing the comparison of CD8+T cell function signatures (exhaustion score and T cytotoxic score) between aggregaphy-related CD8+T cells subtypes. (H) Heatmap showing significantly different features among aggregaphy-related CD8+T cells subtypes, including Co-stimulations (left), Co-inhibitors (middle), and TFs (right).

used IHC to observe cellular localization of DYNC112, TUBA1A, TUBA4A, UBA52 and VIM. We discovered that some of the fibroblasts in stromal

did express DYNC112. (Figure 6A, first panel). Similarly, TUBA1A+ T cells were observed, showing significant positivity, from which immune cells were

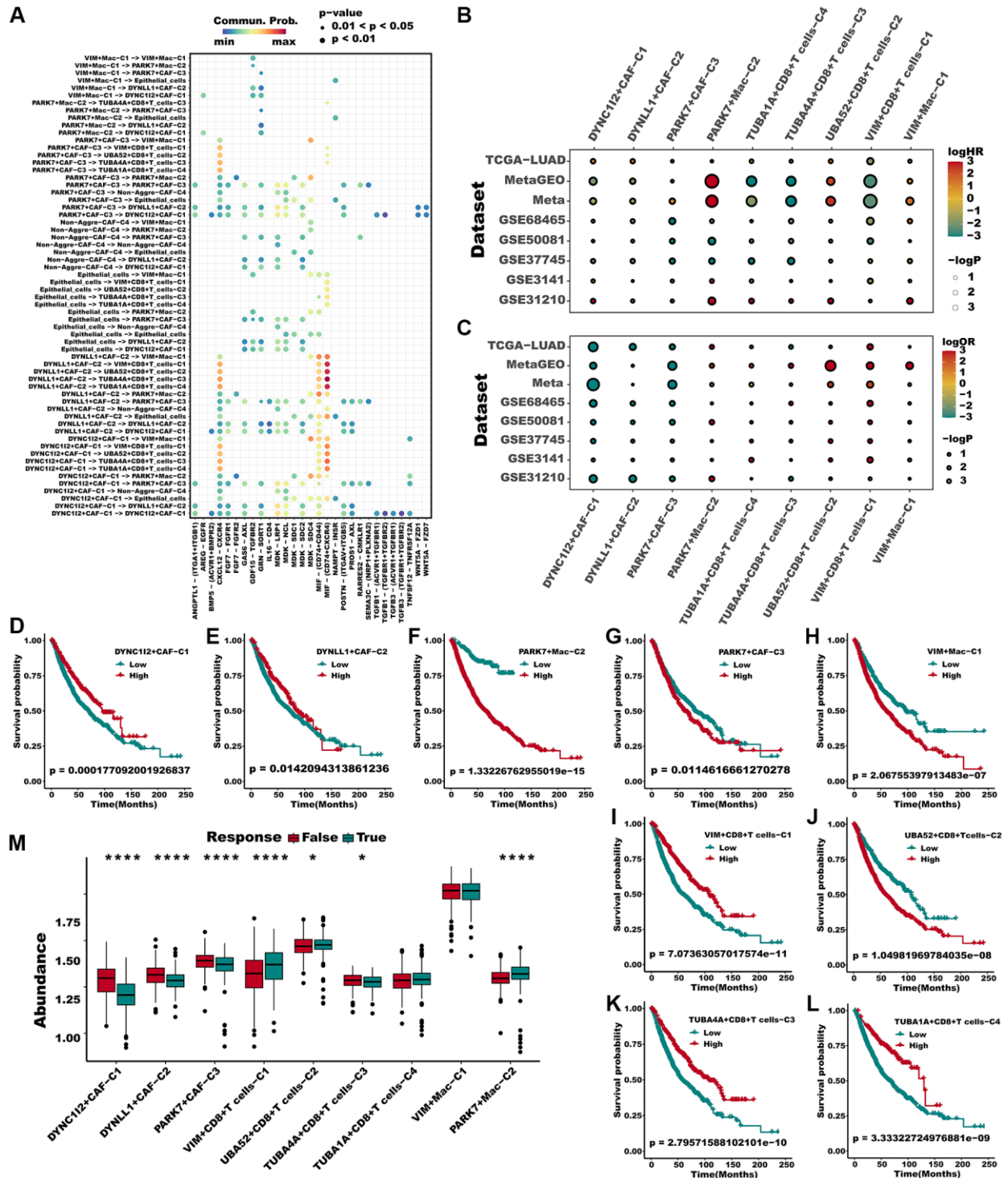


Figure 5. Multiple aggregraphy cell subtypes influenced the prognosis and immunotherapy response of LUAD patients. (A) The landscape of all cell-cell communication within all identified aggregraphy subtypes. **(B)** A hazard ratio of different aggregraphy subtypes in 8 LUAD cohorts. **(C)** Odd ratio produced by logistic regression of different subtypes in 8 cohorts. **(D–L)** All aggregraphy cell subtypes could distinguish the survival of patients in Meta cohort. **(M)** Various aggregraphy-related subtypes demonstrated significantly different infiltration in responders and non-responders of ICB in TCGA cohorts, as predicted by TIDE. Mann Whitney-Wilcoxon test was applied between responders and non-responders. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.

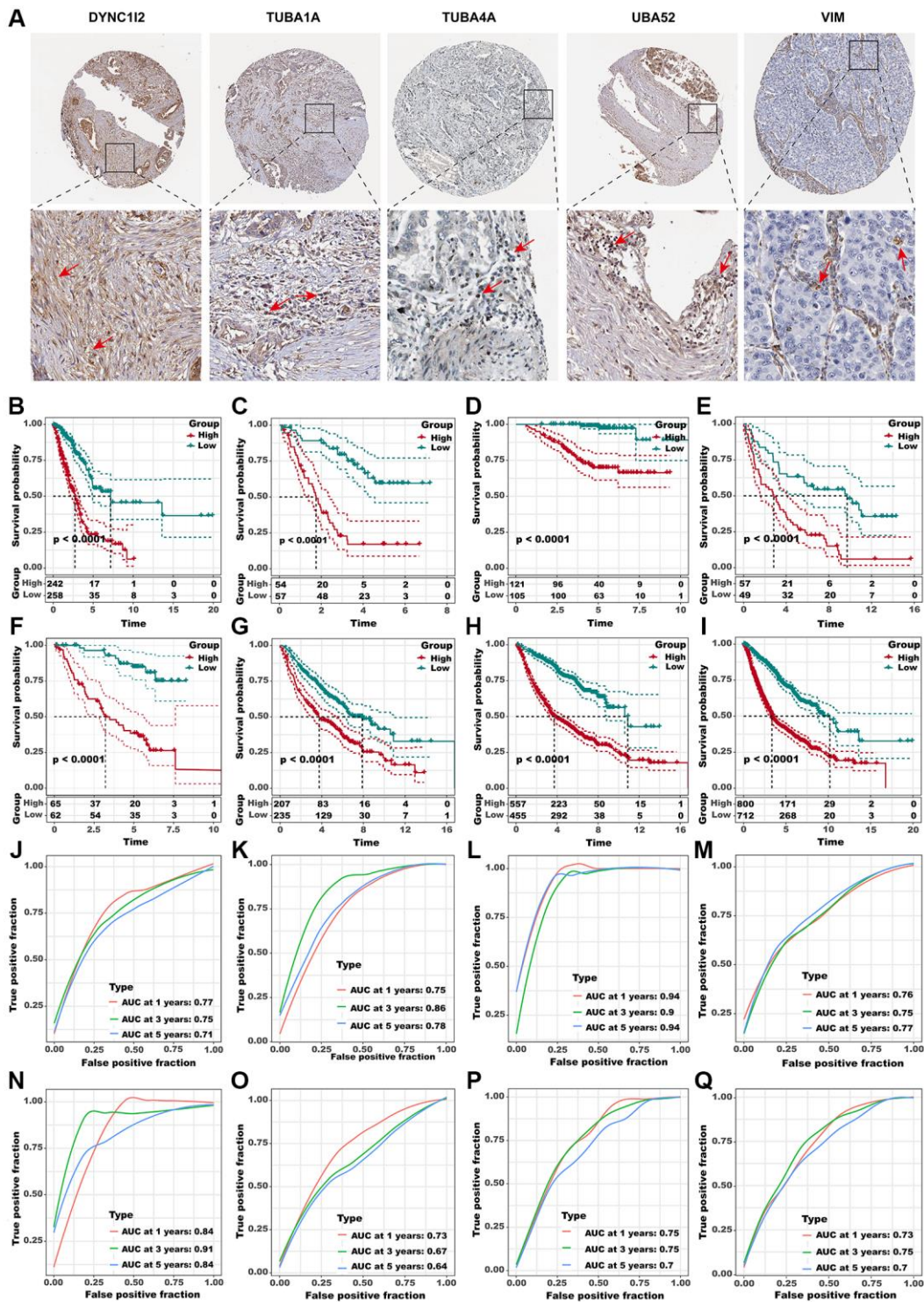


Figure 6. Identification of the hub predictive genes to construct ADPS. (A) IHC staining images of five crucial aggregophagy genes. Panel1: DYNC1I2+ fibroblasts were marked with a red arrow. Protein was mainly expressed on cell membrane (source HPA. Patient ID: 448, female, age 76 years). Panel2: TUBA1A+ T cells were marked with a red arrow. Protein was mainly expressed on cytoplasmic/membranous (source HPA. Patient ID: 448, female, age 76 years). Panel3: TUBA4A+ T cells were marked with a red arrow. Protein was mainly expressed on cytoplasmic/membranous (source HPA. Patient ID: 2403, female, age 65 years). Panel4: UBA52+ T cells were marked with red arrows. Protein was mainly expressed in the nucleus (source HPA. Patient ID: 2393, female, age 54 years). Panel5: VIM + Macrophages were marked with red arrows. Protein was mainly expressed in the cytoplasmic/membranous (source HPA. Patient ID: 1421, female, age 76 years). (B–I) Kaplan-Meier survival curves of the ADPS regarding OS in the TCGA (B), GSE3141 (C), GSE31210 (D), GSE37745 (E), GSE50081 (F), GSE68465 (G), MetaGEO (H), and Meta cohorts (I). (J–Q) Time-dependent ROC curves of the ADPS regarding 1-, 3-, and 5-year OS in the TCGA (J), GSE3141 (K), GSE31210 (L), GSE37745 (M), GSE50081 (N), GSE68465 (O), MetaGEO (P), and Meta cohorts (Q).

seen to exude (Figure 6A, second panel). TUBA4A was expressed on the cell nuclear of some T cells, with stromal cell nearby (Figure 6A, third panel). As a highly conserved nuclear and cytoplasmic protein, UBA52 was expressed in some T cells. As an immune cells, we also observed a high abundance of fibroblasts around UBA52+ T cells in pathological sections (Figure 6A, fourth panel). As we known, VIM was an epithelial-mesenchymal transition biomarker, highly expressed in stromal cell and tumor cell. In pathological sections, we found VIM was expressed in some macrophages, with tumor cell nearby. (Figure 6A, fifth panel).

Aggrephagy-derived prognostic score (ADPS)

Subsequently, considering the concordant performances of aggrephagy subtypes in anticipating the overall survival outcomes as well as immunotherapy response in multiple LUAD cohorts, we retired marker genes and a total of 571 genes from aggrephagy-related subtypes were obtained (Supplementary Table 8). A total of 114 genes were found to have prognostic value through the assessment of each gene using univariate Cox regression analysis (Supplementary Figure 3A). To streamline the number of genes, we utilized LASSO Cox regression analysis, which resulted in 32 remaining genes at a lambda value of 0.0252 (Supplementary Figure 3B, 3C). The ADPS was ultimately generated by performing multivariate Cox regression analysis (Supplementary Figure 3D, Supplementary Table 9). We calculated the ADPS for each patient according to the expression as well as weighted regression coefficients of the ADPS-related genes, and then partitioned them to high ADPS and low ADPS groups after z-mean normalization. K-M survival analyses revealed that the mortality rate in the high ADPS group was significantly higher than the low ADPS group in the training cohort (TCGA-LUAD, $n = 500$, $P < 0.005$), and other seven validation cohorts GSE3141 ($n = 111$, $P < 0.0001$), GSE31210 ($n = 226$, $P < 0.0001$), GSE37745 ($n = 106$, $P < 0.0001$), GSE50081 ($n = 126$, $P < 0.0001$), GSE68465 ($n = 442$, $P < 0.0001$), MetaGEO ($n = 1012$, $P < 0.0001$), and Meta cohort ($n = 1512$, $P < 0.0001$) (Figure 6B–6I). Moreover, in the TCGA cohort, the ADPS exhibited excellent performance, with time-dependent AUCs of 0.77, 0.75, and 0.71 at 1, 3, and 5 years (Figure 6J). Comparable results were also acquired across the validation cohorts GSE3141 (0.75/0.86/0.78), GSE31210 (0.94/0.90/0.94), GSE37745 (0.76/0.75/0.77), GSE50081 (0.84/0.91/0.84), GSE68465 (0.73/0.67/0.64), MetaGEO (0.75/0.75/0.70), as well as Meta cohort (0.73/0.75/0.7), respectively (Figure 6K–6Q).

Predictive value of ADPS for immunotherapy

In addition, we evaluated the prognostic significance of ADPS about immunotherapy, utilizing the real-

world cohort including IMvigor210 and GSE78220 datasets. Within the IMvigor210 cohort, which consisted of 298 patients, responses to anti-PD-L1 receptor blockers ranged from the complete response (CR) and partial response (PR) to stable disease (SD) and progressive disease (PD). A significant clinical benefit was observed in the low ADPS group of the IMvigor210 cohort compared with the high ADPS group (Figure 7A). CR/PR patients presented lower ADPS than SD/PD patients (Figure 7B, $P < 0.0001$). The low ADPS group displayed a higher percentage of CR/PR compared to the high ADPS group (Figure 7C). Survival differences between different ADPS groups were significant specifically in patients with Stage I+II (Figure 7D, $P < 0.001$), and in Stage III+IV patients (Figure 7E, $P < 0.001$). Our finding also indicated that patients with a low ADPS had a significantly better overall survival outcome compared to those with a high ADPS in the GSE78220 cohort (Figure 7F, $P < 0.001$). Besides, CR/PR patients also presented lower ADPS than SD/PD patients (Figure 7G, $P < 0.01$). Additionally, a higher percentage of SD/PD was observed in the high ADPS group as compared to the low ADPS group (Figure 7H). In addition to the abundance of immune checkpoints, we also noted a significantly higher expression of these checkpoints in the low ADPS group as compared to the high ADPS group (Figure 7I). Utilizing the TIDE web tool, we discovered that the low ADPS group had notably lower TIDE scores and higher rates of immunotherapy response compared to the high ADPS group (Figure 7J). There was a significant positive correlation observed between ADPS and TIDE score (Figure 7K). Responders to ICIs were found to be more prevalent among patients belonging to the high ADPS group, as predicted by the TIDE algorithm (Figure 7L, $P < 0.001$). Parallel to these findings, the subclass mapping analysis (Submap, modules in GenePattern, <https://cloud.genepattern.org>) also revealed similarities in expression patterns between patients with low ADPS and those with melanoma who responded to immunotherapy [49, 50] (A melanoma dataset that responded to immunotherapy was selected as the reference, while default settings were applied) responding to ICB (Figure 7M).

Immune landscape for ADPS

To explore the immune characteristics reflected by the ADPS, we investigated the association between the ADPS and immune cell type, immune scores, stromal scores, as well as critical immune checkpoint scores. Figure 8A, 8B demonstrated that the group with low ADPS exhibited elevated levels of immune infiltrating cells as well as immune modulators. This suggests the presence of an inflamed yet relatively immune-supportive

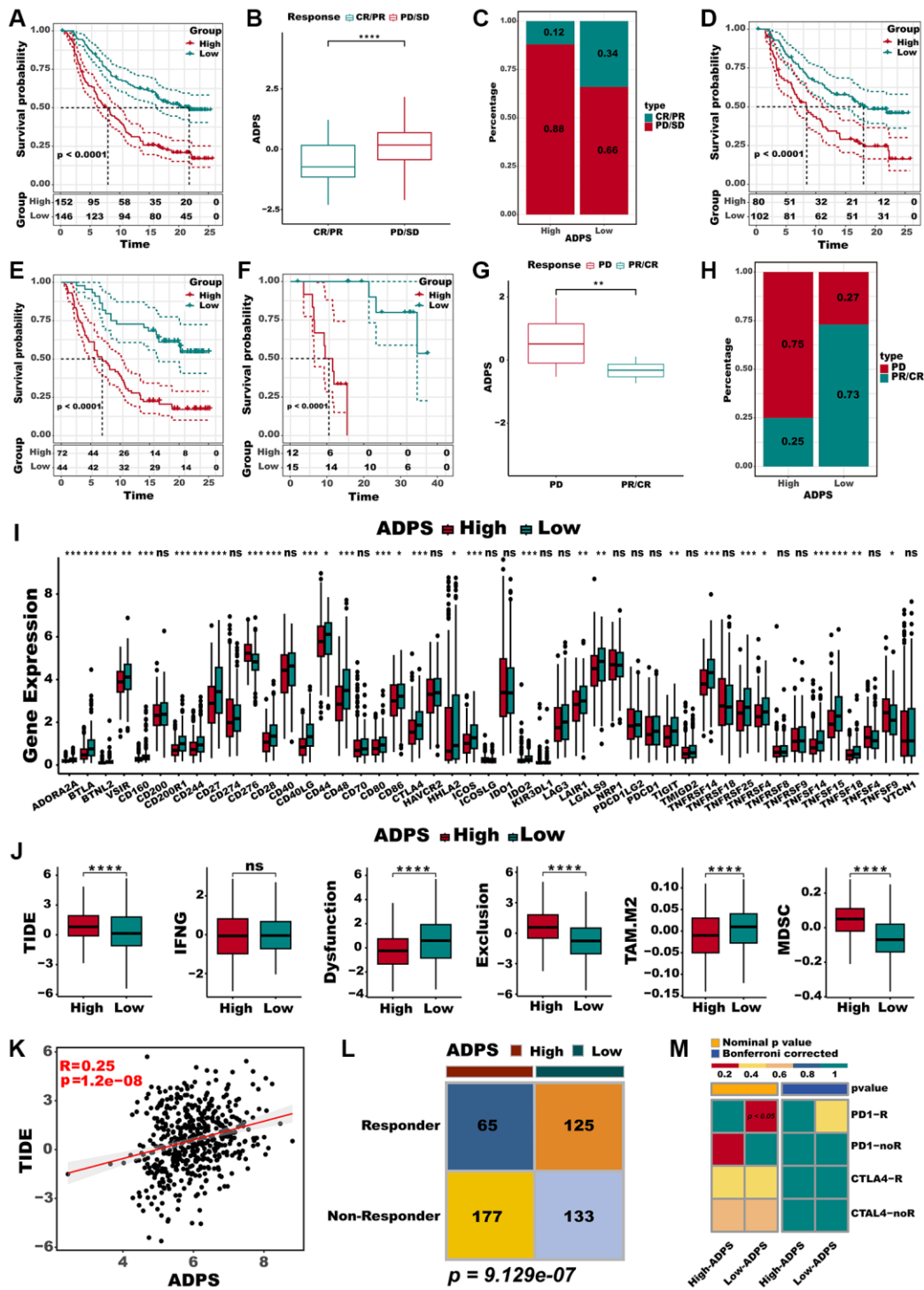


Figure 7. Immunotherapy response prediction of the ADPS. (A) Kaplan-Meier survival curves of the ADPS regarding OS in the IMvigor210 cohort. (B) Differences in ADPS among immunotherapy responses in the IMvigor210 cohort. (C) Distribution of immunotherapy responses among ADPS groups in the IMvigor210 cohort. (D) Prognostic differences between ADPS groups in early stage (stage I+II) patients in the IMvigor210 cohort. (E) Prognostic differences between ADPS groups in advanced stage (stage III+IV) patients in the IMvigor210 cohort. (F) Kaplan-Meier survival curves of the ADPS regarding OS in the GSE78220 cohort. (G) Differences in ADPS among immunotherapy responses in the GSE78220 cohort. (H) Distribution of immunotherapy responses among ADPS groups in the GSE78220 cohort. (I) Boxplot of relative expression levels at immune checkpoints between the high and low ADPS patients. (J) Boxplot of TIDE score between the high and low ADPS patients. (K) The relationship between the TIDE score and ADPS. (L) Contingency table between immunotherapy responses and ADPS groups based on TIDE algorithm. (M) Submap analysis of the two groups and 47 pretreated patients with comprehensive immunotherapy annotations. In submap analysis, a smaller p -value indicated a greater similarity of paired expression profiles. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.

microenvironment, which may be more conducive to benefiting from immunotherapy [51]. The immune score, stromal score, and ESTIMATE score had a

negative correlation to ADPS, respectively (Figure 8C–8E). In parallel, patients in the low ADPS group had higher stromal, immune, and ESTIMATE scores

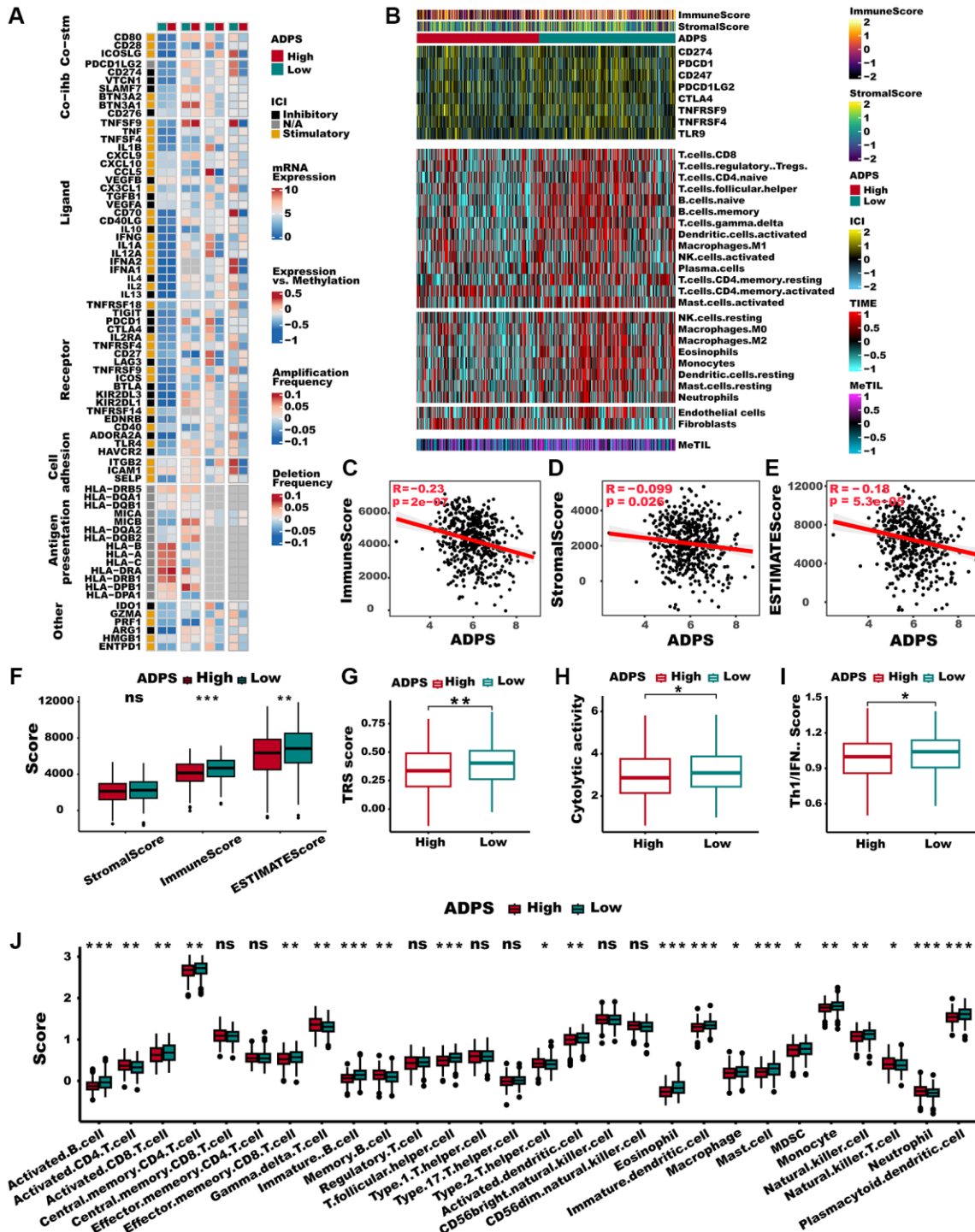


Figure 8. Immune characteristics of the ADPS in the TCGA dataset. (A) The correlation between the ADPS and immune modulators. (B) Heatmap exhibiting the immune score, stromal score, critical checkpoints, and cell types calculated through CIBERSORT analysis of the high and low ADPS groups. (C–E) Correlations between ADPS and immune score (C), stromal score (D), and ESTIMATE score (E). (F) Boxplot of relative stromal, immune, and ESTIMATE score between high and low ADPS groups. (G) Boxplot of relative TRS score between high and low ADPS groups. (H) Boxplot of relative cytolytic activity between high and low ADPS groups. (I) Boxplot of relative Th1/IFN score between high and low ADPS groups. (J) Boxplot of relative infiltrate abundance of 28 immune cell types between high and low ADPS groups. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.

compared to the high ADPS group (Figure 8F). Besides, we compared the status of the TRS score, Cytolytic activity, and Th1/IFN Score, which was more related to a more immunoreactive microenvironment between the two ADPS score groups [52, 53]. The results revealed that all of these indicators were significantly upregulated in the low ADPS group (Figure 8G–8I). According to 28 immune cells infiltration [54] assessed by ssGSEA (Figure 8J) and CIBERSORT algorithm (Supplementary Figure 4A), we further confirmed that the low ADPS group had markedly higher overall infiltration abundance than the high ADPS group. Thus, we defined the high ADPS group as “immune-cold” tumors and the low ADPS group as “immune-hot” tumors. Considering the upregulation of immune-related characteristics observed within the group of patients with a low ADPS, we sought to delve deeper into the biological mechanisms responsible for this phenomenon. KEGG-based GSVA analysis indicated the ADPS had a strong positive correlation with many pathways facilitating tumor growth, such as the P53 signaling pathway, cell cycle, glycolysis, and WNT signaling, while the immunological pathways, such as B/T cell receptor signaling pathways, cytokine receptor interaction pathway showed a weaker correlation with ADPS (Supplementary Figure 4B). Additionally, we further explored the immune characteristics from gene level, as displayed in Figure 9A, 9B, most ADPS-related genes were positively correlated with immune score as well as immune cell types.

Searching for potential therapeutic agents for the high ADPS patients

The CTRP and PRISM databases comprise gene expression as well as drug sensitivity profiles from hundreds of cancer cell lines (CCLs), providing an opportunity to construct a model for predicting drug response. After removing duplicates, a total of 1770 unique compounds were obtained in the two datasets, with 160 compounds being shared between them (Figure 9C). Figure 9E demonstrates our strategic approach to developing potential agents for high ADPS patients. Subsequently, we implemented this methodology to pinpoint promising agent candidates for the high ADPS patients and generated three CTRP-derived agents (BI-2536, paclitaxel, and SB-743921), as well as five PRISM-derived agents (cabazitaxel, epothilone-b, gemcitabine, ispinesib, and SNS-314). The AUC values of these agents, as estimated, showed a statistically significant negative correlation with ADPS scores and were significantly lower in the high ADPS group (Figure 9F–9I). Furthermore, we leveraged the differential expression profiles of LUAD patients as well as normal samples to identify potential candidate compounds using the Cmap tool. Specifically, we aimed

to identify agents whose gene expression patterns were opposite to the LUAD-specific expression patterns, indicating potential efficacy in suppressing LUAD tumor growth (i.e., gene expression increased in tumor tissues but decreased by treatment of certain compounds). Following the cross-referencing of the results from CTRP and PRISM, we were left with a trio of prospective agents: BI-2536, a PLK inhibitor, paclitaxel, and gemcitabine (Figure 9D). Among them, BI-2536, with a CMap score of -87.58 , exhibited high sensitivity in LUAD patients, implying its potential as a therapeutic agent for those with high ADPS.

Multi-omics alteration characteristics targeting ADPS

The differences in frequently altered chromosomes were detected in two ADPS groups (Figure 10A, 10B). We conducted an integrated analysis of mutations and copy number alterations (CNA, Figure 10C) to examine the genomic heterogeneity of the high and low ADPS groups. In the high ADPS group, we observed a higher frequency of mutations in classical tumor suppressor genes TP53 and CSMD3, as well as the oncogene KRAS, compared to the low ADPS group (Figure 10C). Additionally, compared to the low ADPS group, the high ADPS group showed significantly higher levels of amplification or deletion at the focal and chromosome arm levels, such as the amplification of 8q24.21, 12p12.1, as well as the deletion of 11p15.5, and 9p21.3 (Figure 10C). This finding was further supported at the gene level by the clear amplification of the oncogene MYC located at 8q24.21 and the distinct deletion of the tumor suppressor genes CDKN2A located at 9p21.3 (Figure 10C). Furthermore, we explored the mutation frequency of 10 major oncogenic pathways [55] between the high and low ADPS groups, and result found that both groups had detectable mutations in most of the oncogenic pathways, such as RTK-RAS, PI3K, TP53, NOTCH, and Hippo pathways (Figure 10D, 10E). We further explored the relationship between ADPS-related genes and LUAD by analyzing their correlations with different molecular signatures (Figure 10F). Our findings showed that FHL1 and ADGRD1 were strongly and positively correlated with molecular signatures associated with genomic instability in LUAD, including Aneuploidy Score, Homologous Recombination Defects, Fraction Altered, Number of Segments, as well as Nonsilent Mutation Rate. Additionally, the high ADPS group exhibited a higher TMB score in comparison to the low ADPS group (Figure 10G, $P < 0.01$). Only a small fraction of samples exhibited copy number variation (CNV) in the ADPS-related genes (Figure 10H). Additionally, we explored the single-nucleotide variant (SNV) mutations of the ADPS-related genes. It showed that COL4A1,

COL1A2, and LUZP2 had SNV mutations in more samples (Supplementary Figure 4C). Finally, we investigated the co-occurrence probability between ADPS-related genes and the top 10 genes with the

highest mutation frequency. Supplementary Figure 4D reveals significant co-occurrence probabilities between CACYBP and ENU1, as well as between RELA and AP2M.

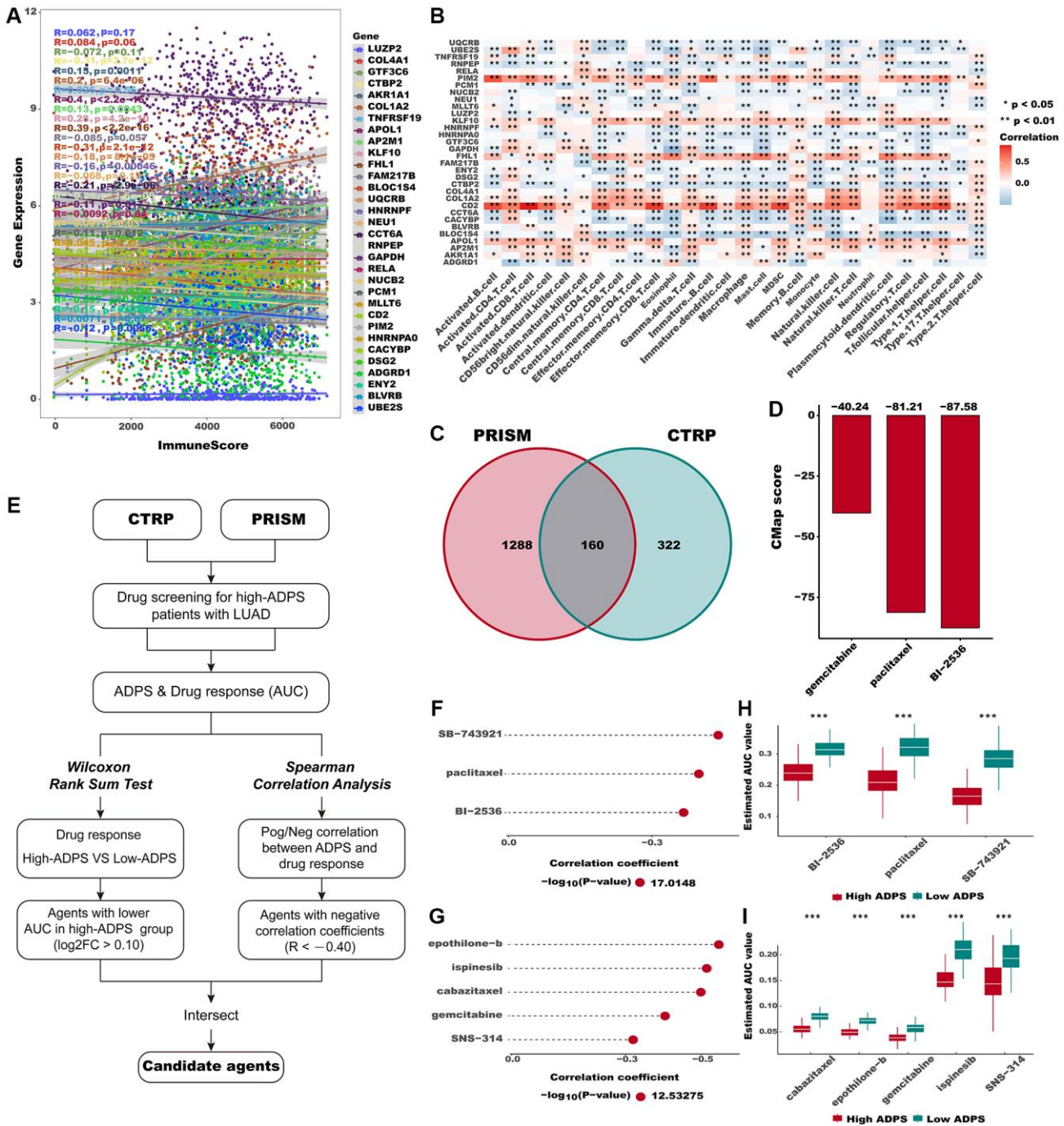


Figure 9. Identification of candidate agents with higher drug sensitivity in high ADPS group. (A) The correlation of gene expression and immune score. (B) The correlation of gene expression and immune cell types. (C) A Venn diagram for summarizing included compounds from CTRP and PRISM datasets. (D) Barplot of gemcitabine, paclitaxel, and BI-2536 CMap scores in patients with high ADPS. (E) Schematic outlining the strategy to identify agents with higher drug sensitivity in high ADPS patients. (F, G) The result of Spearman's correlation analysis and differential drug response analysis of CTRP-derived compounds (F) and profiling relative inhibition simultaneously in mixtures (PRISM)-derived compounds (G). (H, I) The results of differential drug response analysis of CTRP -derived compounds (H) and PRISM -derived compounds (I), The lower the value of the y-axis, the greater the drug sensitivity. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.

Nomogram based on ADPS and clinical features

To better apply ADPS in a clinical setting and optimize its predictive performance, we performed

univariate and multivariate Cox regression analysis to integrate clinicopathological characteristics as well as ADPS. ADPS was identified as the most significant independent prognostic factor of LUAD according to

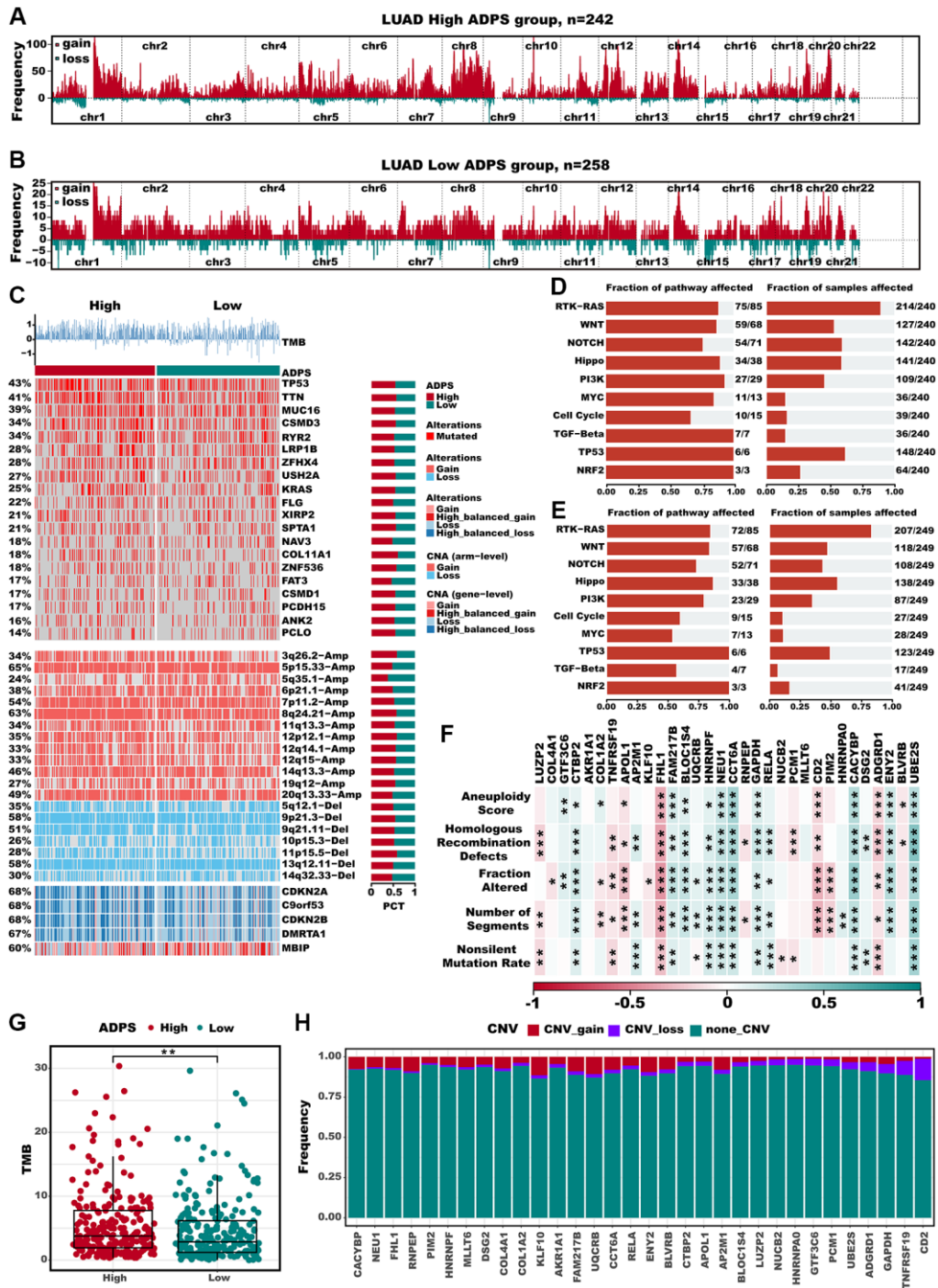


Figure 10. Multi-omics alteration characteristics of the ADPS in the TCGA dataset. (A, B) GISTIC 2.0-based chromosome amplifications and deletions in high (A) and low (B) ADPS groups. (C) Genomic alteration landscape according to ADPS. Tumor mutation burden (TMB), relative contribution of four mutational signatures, top 20 mutated genes and broad-level CNA (>20%). The proportion of the high and low ADPS groups in each alteration is presented in the right bar charts. (D, E) The mutation frequencies of nine common oncogenic pathways in the high ADPS (D) and low (E) groups. (F) Correlation heatmap of ADPS-related genes with Aneuploidy Score, Homologous Recombination Defects, Fraction Altered, Number of Segments, and Nonsilent Mutation Rate. (G) Difference of TMB score between high and low ADPS group. (H) CNV mutations (gain, loss, none) of ADPS-related genes. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.

the results of univariate Cox analysis (hazard ratio (HR) = 2.718, 95% confidence interval (CI): 2.25–3.284, $p < 0.001$), as well as multivariate Cox regression analysis (hazard ratio (HR) = 2.687, 95% confidence interval (CI): 2.127–3.394, $p < 0.001$), respectively (Figure 11A,

11B). The circo plot illustrated a significant correlation between ADPS and survival status, tumor stage, as well as TNM staging system in the TCGA cohort (Figure 11C). Moreover, the ADPS showed superior predictive accuracy, as indicated by the higher C-index values,

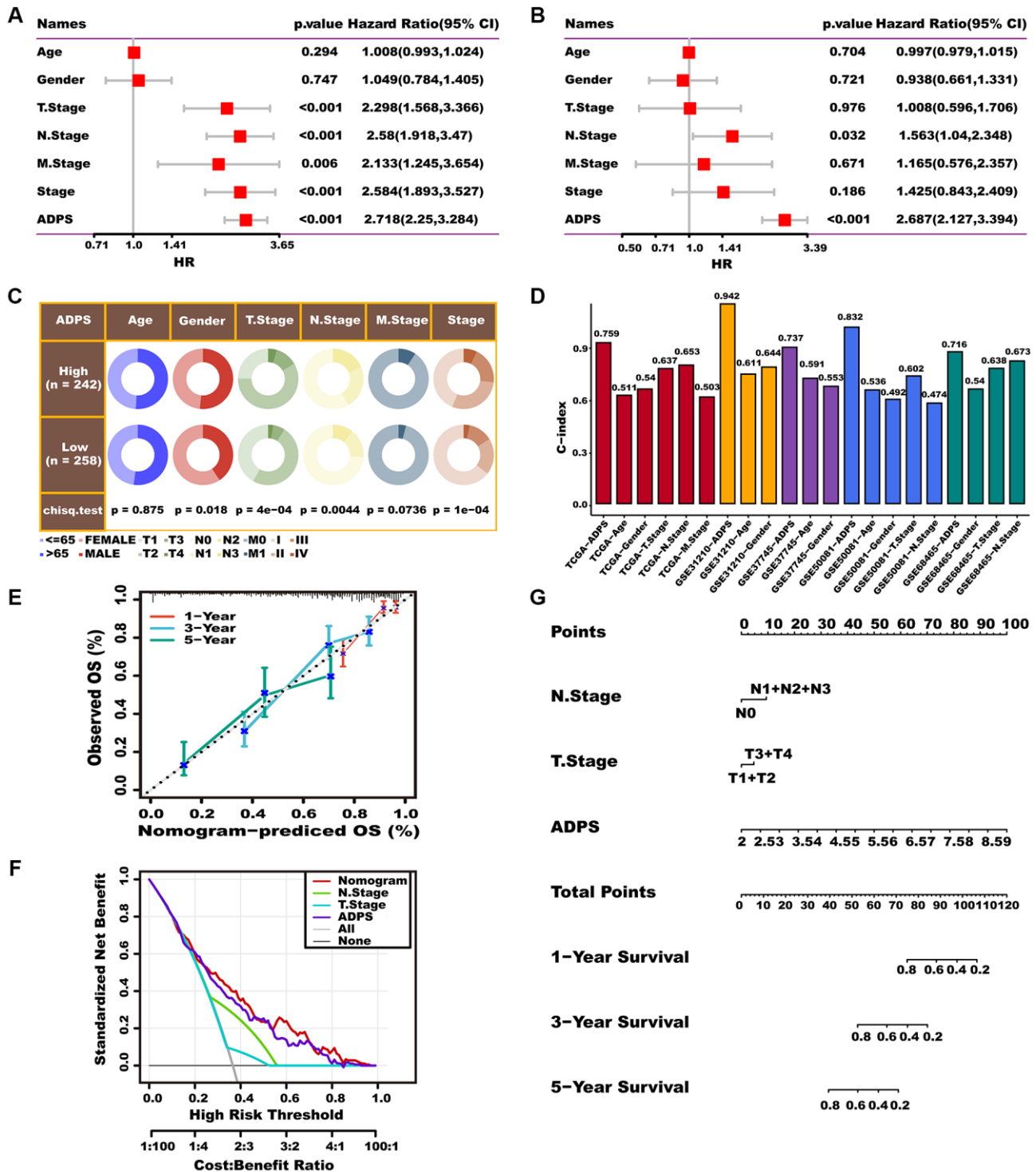


Figure 11. Prognostic value of the ADPS and nomogram construction. (A, B) Univariate Cox (A) and multivariate Cox (B) analysis of the ADPS and clinicopathological characteristics. (C) Circo plot of different clinical factors in two ADPS groups. (D) The C-index of the ADPS and various clinical factors in the TCGA, GSE31210, GSE37745, GSE50081, and GSE68465 datasets. (E) Calibration curves for 1, 3, and 5 years of nomogram. (F) Decision curve for nomogram. (G) Nomogram model integrating the ADPS and stage was constructed.

compared to traditional clinical factors including age, gender, tumor stage, and TNM staging system, in various independent LUAD cohorts including TCGA, GSE31210, GSE37745, GSE50081, as well as GSE68465 (Figure 11D). Thus, a nomogram was created by integrating both the stage and ADPS, as illustrated in Figure 11G. The calibration plot illustrated the effective ability of the nomogram in predicting the actual survival outcomes (Figure 11E). Furthermore, the decision curve analysis (DCA) indicated that the nomogram had a superior ability to identify high ADPS patients compared to both the ADPS as well as stage alone, as illustrated in Figure 11F. TimeROC analysis indicated that the ADPS and nomogram had higher AUC values than other indicators (Supplementary Figure 4E). Our research further revealed that ADPS was significantly increased in the T stage (Supplementary Figure 4F), N stage (Supplementary Figure 4G), and tumor stage (Supplementary Figure 4I), without significant differences in M stage (Supplementary Figure 4H).

DISCUSSION

To date, aggrephagy has escalatingly emerged as a hotspot in anti-tumor research, and several studies have revealed the correlation between aggrephagy modification and the pathogenesis of tumorigenesis [6, 56, 57]. Understanding the mechanisms behind this process could contribute to the development of new treatments for diseases associated with protein aggregation. However, no related study focuses on the potential tumorigenic role of aggrephagy-modified single cells. Multiple algorithms such as AUCell and addmodulescore suggested the entire aggrephagy score was significantly activated in various TME cells in LUAD, and aggrephagy classical genes also shared high heterogeneity among TME. So, comprehensively exploring how aggrephagy shapes and reprograms TME components is necessary. We further dissected the intricate intercellular interactions among aggrephagy-associated TME subtypes in LUAD at the scRNA-seq level and validated the cell types through IHC. This study provides a novel perspective for understanding how the cell-specific expression patterns of aggrephagy genes shape the TME, thus affecting the prognosis and outcomes of immune checkpoint blockade therapy in individual LUAD patients.

CAF, as a highly adaptable and dynamic constituent of the TME, exerts a crucial function in promoting cancer progression via intricate crosstalk with other cellular and non-cellular constituents within the TME [42]. According to their distinct molecular features, CAFs can be categorized into various subtypes, including pan-iCAFs, pan-myCAFs, pan-dCAFs, pan-nCAFs, as

well as pan-pCAFs [35]. Until now, there has been no report on the expression patterns of aggrephagy in CAFs. This study identified four subtypes of CAFs: DYNC1I2+CAF-C1, DYNLL1+CAF-C2, PARK7+CAF-C3, and Non-Aggre-CAF-C4, and further explored the interactions between these subtypes and other components in the TME. We found that aggrephagy-related CAFs exhibited more extensive interaction with other components compared to non-aggrephagy-related CAFs. Additionally, we highlighted DYNC1I2+CAF-C1 and PARK7+CAF-C3 because of their high correlation with pan-dCAF, pan-iCAF, and pan-pCAF, along with elevated expression of well-recognized TGF β genes such as TGF β 1, SULF1, and THBS2 in DYNC1I2+CAF-C1, as well as pro-inflammatory genes such as IL7, CCL2 and CXCL12 in PARK7+CAF-C3 [58, 59]. We emphasized DYNC1I2+CAF-C1, DYNLL1+CAF-C2, as well as PARK7+CAF-C3 subtypes due to their exceptional ability to differentiate LUAD patient survival. It is worth noting that the high DYNC1I2+CAF-C1, and DYNLL1+CAF-C2 score presented better survival, but the high PARK7+CAF-C3 score presented dismal survival. Thus, we postulated that alterations in aggrephagy could potentially impact the functional and phenotypic properties of CAFs, leading to significant changes in the immunosuppressive TME that may have a consequential impact on the malignant progression and metastasis of LUAD.

Recently, there has been growing interest in studying the role of aggrephagy in immune cell components of the TME, with particular attention focused on TAMs [60]. The TAMs were classified into two subclusters based on NMF clustering, and both subclusters demonstrated broad communication with other TME constituents. The metabolic activity of TAMs has emerged as a critical factor influencing cancer progression and immune responses, with glucose, glutamine, and fatty acid metabolism being among the key metabolic pathways involved [61]. Interestingly, we showed that the two aggrephagy TAM subtypes have distinct metabolic statuses. Obvious activation of pathways such as oxidative phosphorylation, glycolysis/gluconeogenesis, as well as TCA cycle were presented in PARK7+Mac-C2 subtype. We revealed that the functional and metabolic characteristics of TAMs are modulated by aggrephagy genes, pointing towards a potential mechanism of immune evasion facilitated by TAMs in the context of LUAD. In addition, aggrephagy-related subtypes of the four main T cell subtypes demonstrated varying degrees of T cell activity and inactivity. TF analysis revealed that aggrephagy-related subtype all manifested distinct TFs characteristics. For CAFs, TFs such as BHLHE40, FOXO3, ATF3, JUNB, FOSB, as well as CEBPB are notably activated in different aggrephagy-related CAF

subtypes. FOSB and JUNB, as constituents of the Activator Protein-1 (AP-1) family, hold a pivotal position in transcriptional control of multiple genes that govern a variety of cellular activities encompassing cell proliferation, differentiation, migration, immune surveillance, and survival [62]. Furthermore, for CD8 T cells, TUBA1A+CD8+ T cells C4 exhibited a unique TF gene signature, such as BATF, BCLAF1, FOSB, and JUNB. Zhang et al. demonstrated that depletion of BATF in diverse chimeric antigen receptor T cell models and mouse OT-1 cells results in enhanced resistance to exhaustion and superior efficacy in tumor eradication [63]. In conclusion, the modulation of distinct transcription factor (TF) regulatory networks by aggrephagy-mediated cell subtypes may result in a reshaped and reprogrammed TME. Furthermore, our cell network analysis demonstrated the close connectivity and communication between these aggrephagy-mediated TME cells and tumor cells. Remarkably, aggrephagy-mediated CAFs and immune cell subtypes exhibited increased crosstalk with cancer epithelial cells, implicating that aggrephagy regulation may contribute to the establishment of an immunosuppressive microenvironment.

Given the intricate intrinsic patterns of aggrephagy in TME cells, we conducted a comprehensive analysis to assess the correlation between the scores of these aggrephagy-related cell subtypes and both prognosis and immune response based on RNA-seq data from multiple centers. Clearly, the degree of aggrephagy gene dominance in TME cells exhibited significant prognostic differences in LUAD patients, and highly distinguished the immune response in patients treated with ICB therapy. These findings highlight the critical role of TME aggrephagy in LUAD, which warrants further investigation. Given the prognostic values of aggrephagy-related cell subtypes, we established a risk model named ADPS with 32 genes. It consisted of 14 protective genes and 18 risk genes. Additionally, ADPS exhibited superior stability in the stratification of patients with different prognoses in multiple cohorts; hence, targeted clinical interventions for patients with varying levels of ADPS are necessary. A lower ADPS was found to be a significant predictor of increased sensitivity to immunotherapy in both the IMvigor210 and GSE91061 cohorts, and this result was further validated by TIDE and Submap analyses.

Furthermore, the comprehensive exploration of immune infiltration from multiple perspectives revealed that the low ADPS group demonstrated a greater abundance of immune cell types, such as activated CD8 T cells and CD4 T cells. It is widely acknowledged that the abundance of effector immune cells, like activated CD8 T cells and CD4 T cells may augment the anti-

tumor immunity and confer better immunotherapeutic outcomes [64, 65]. Consistently, the low ADPS group demonstrated an elevated expression of immune modulators, immune checkpoints, as well as biomarkers that reflect the presence of an immunoreactive microenvironment, like CYT, TCR, and IFN- γ . Furthermore, a decreased ADPS was associated with an activated cancer immunity cycle and several immunological pathways, indicating its potential utility in predicting response to immunotherapy. The concept of precision medicine necessitates the early identification of patients who would be responsive to diverse treatments for further personalized interventions. Due to the high sensitivity of the low ADPS patients to immunotherapy, we explored the integration of CTRP, PRISM, as well as CMap databases to create personalized drugs for patients with high ADPS [39, 40, 66]. Finally, BI-2536, a PLK inhibitor, caught our attention. Zhou et al. found BI2536 inhibited lung cancer growth and promoted activation of T cells and DC cells [67]. Going forward, further clinical trials are needed to validate the potential of BI2536 in LUAD, particularly with high ADPS patients. Using a multi-omics approach, we conducted an in-depth investigation into the mutation and copy number variation (CNV) characteristics associated with ADPS. Our findings indicate that the high ADPS group exhibited elevated tumor mutational burden (TMB) and a higher frequency of mutations in the classical tumor suppressor gene TP53 and oncogene KRAS. These mutations have been previously linked to increased invasion and immune evasion in patients with LUAD, leading to a poorer prognosis [68, 69].

After conducting a thorough review of the literature, we discovered that CTBP2 expression was markedly elevated in LUAD tissues when compared to normal lung tissue, and that high levels of CTBP2 expression were linked to a poorer prognosis among LUAD patients [70]. The expression of CCT6A was found to be significantly correlated with both relapse-free and overall survival in patients with LUAD. Moreover, overexpression of CCT6A was observed to enhance cell growth and invasion in LUAD [71]. He et al. proposed a mechanism in which miR-3613-5p expression is induced by RELA through its direct interaction with JUN, thus activating the AKT/mitogen-activated protein kinase (MAPK) pathway and promoting oncogenesis in LUAD [72]. HNRNPF, as a critical alternative splicing regulator, was associated with worse survival of LUAD [73]. DSG2, a member of the cadherin superfamily, has been implicated in cell-cell adhesion and tumorigenesis. Jin et al. demonstrated that high expression of DSG2 is associated with poor prognosis in LUAD patients and promotes cell proliferation and migration, as well

as increases resistance to the EGFR tyrosine kinase inhibitor Osimertinib [74]. The overexpression of CACYBP has been shown to enhance the proliferative, invasive, and migratory capacities of LUAD cells, and it may represent a novel therapeutic target for advanced LUAD [75]. Furthermore, our study revealed robust biological associations between ADPS and mutations/TME in LUAD, underscoring the critical role of ADPS in both prognosis and the immune microenvironment. These findings highlight the potential of ADPS as a valuable clinical tool for the precision management and treatment of LUAD.

As an initial investigation, our analysis has some notable limitations including the relatively low sequencing depth of scRNA-seq data and the limited sample size. As a result, further validation in larger patient cohorts is necessary to confirm our conclusions. Compared to bulk RNA-seq, the scRNA-seq analysis of aggrephagy in LUAD is characterized by low coverage and a higher proportion of zero counts, which could potentially introduce bias to the NMF clustering method employed in our study. Additionally, it is regrettable that based on the HPA database, we only found the IHC results for some key aggrephagy-related genes. Further experiments are still required to validate other aggrephagy-related genes, and to further investigate their mechanism of affecting the TME. Quantitative ADPS scoring system included the 32 genes were found to be associated with multiple prognostic features of LUAD, suggesting their potential prognostic value, most of their precise roles in LUAD are still unclear and also require further functional experimental validation in the future.

CONCLUSIONS

We have identified specific aggrephagy cell subtypes of TME cells for the first time using the single-cell sequencing analysis method. Our findings reveal the aggrephagy-mediated intercellular communication within the tumor microenvironment, which plays a crucial role in regulating tumor growth and modulating the antitumor immune response. Based on machine-learning algorithm, we have developed a robust and powerful signature that accurately predicts the prognosis and immune response of individual LUAD patients, allowing for optimized decision-making and surveillance protocols.

AUTHOR CONTRIBUTIONS

XTS and FM designed the study. XTS, MYN, and YW analyzed the data, participated in data collection, and prepared the manuscript. HF, CLL, and PZ helped

the analysis with constructive discussions. All authors critically revised the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

FUNDING

This work was supported by the Beijing-Tianjin-Hebei Basic Research Cooperation Project (20JCZXJC00190) and Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-061B).

REFERENCES

1. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, Petrella F, Spaggiari L, Rosell R. Non-small-cell lung cancer. *Nat Rev Dis Primers*. 2015; 1:15009. <https://doi.org/10.1038/nrdp.2015.9> PMID:27188576
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin*. 2021; 71:7–33. <https://doi.org/10.3322/caac.21654> PMID:33433946
3. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol*. 2018; 15:81–94. <https://doi.org/10.1038/nrclinonc.2017.166> PMID:29115304
4. Parzych KR, Klionsky DJ. An overview of autophagy: morphology, mechanism, and regulation. *Antioxid Redox Signal*. 2014; 20:460–73. <https://doi.org/10.1089/ars.2013.5371> PMID:23725295
5. Kocaturk NM, Akkoc Y, Kig C, Bayraktar O, Gozuacik D, Kutlu O. Autophagy as a molecular target for cancer treatment. *Eur J Pharm Sci*. 2019; 134:116–37. <https://doi.org/10.1016/j.ejps.2019.04.011> PMID:30981885
6. Miller DR, Thorburn A. Autophagy and organelle homeostasis in cancer. *Dev Cell*. 2021; 56:906–18. <https://doi.org/10.1016/j.devcel.2021.02.010> PMID:33689692
7. Malampati S, Song JX, Chun-Kit Tong B, Nalluri A, Yang CB, Wang Z, Gopalkrishnashetty Sreenivasamurthy S, Zhu Z, Liu J, Su C, Krishnamoorthi S, Iyaswamy A, Cheung KH, et al. Targeting Aggrephagy for the Treatment of Alzheimer's Disease. *Cells*. 2020; 9:311. <https://doi.org/10.3390/cells9020311> PMID:32012902

8. Ma X, Lu C, Chen Y, Li S, Ma N, Tao X, Li Y, Wang J, Zhou M, Yan YB, Li P, Heydari K, Deng H, et al. CCT2 is an aggregophagy receptor for clearance of solid protein aggregates. *Cell*. 2022; 185:1325–45.e22. <https://doi.org/10.1016/j.cell.2022.03.005> PMID:35366418
9. Gong H, Yang X, Zhao Y, Petersen RB, Liu X, Liu Y, Huang K. Amyloidogenicity of p53: a hidden link between protein misfolding and cancer. *Curr Protein Pept Sci*. 2015; 16:135–46. PMID:25692950
10. Ishimaru D, Andrade LR, Teixeira LS, Quesado PA, Maiolino LM, Lopez PM, Cordeiro Y, Costa LT, Heckl WM, Weissmüller G, Foguel D, Silva JL. Fibrillar aggregates of the tumor suppressor p53 core domain. *Biochemistry*. 2003; 42:9022–7. <https://doi.org/10.1021/bi034218k> PMID:12885235
11. Kanapathipillai M. Treating p53 Mutant Aggregation-Associated Cancer. *Cancers (Basel)*. 2018; 10:154. <https://doi.org/10.3390/cancers10060154> PMID:29789497
12. Obradovic A, Chowdhury N, Haake SM, Ager C, Wang V, Vlahos L, Guo XV, Aggen DH, Rathmell WK, Jonasch E, Johnson JE, Roth M, Beckermann KE, et al. Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages. *Cell*. 2021; 184:2988–3005.e16. <https://doi.org/10.1016/j.cell.2021.04.038> PMID:34019793
13. Devalaraja S, To TKJ, Folkert IW, Natesan R, Alam MZ, Li M, Tada Y, Budagyan K, Dang MT, Zhai L, Lobel GP, Ciotti GE, Eisinger-Mathason TSK, et al. Tumor-Derived Retinoic Acid Regulates Intratumoral Monocyte Differentiation to Promote Immune Suppression. *Cell*. 2020; 180:1098–114.e16. <https://doi.org/10.1016/j.cell.2020.02.042> PMID:32169218
14. Young TM, Reyes C, Pasnikowski E, Castanaro C, Wong C, Decker CE, Chiu J, Song H, Wei Y, Bai Y, Zambrowicz B, Thurston G, Daly C. Autophagy protects tumors from T cell-mediated cytotoxicity via inhibition of TNF α -induced apoptosis. *Sci Immunol*. 2020; 5:eabb9561. <https://doi.org/10.1126/sciimmunol.abb9561> PMID:33443027
15. Bai J, Liu T, Tu B, Yuan M, Shu Z, Fan M, Huo S, Guo Y, Wang L, Wang H, Zhao Y. Autophagy loss impedes cancer-associated fibroblast activation via downregulating proline biosynthesis. *Autophagy*. 2023; 19:632–43. <https://doi.org/10.1080/15548627.2022.2093026> PMID:35786294
16. Wang D, Jiao Z, Ji Y, Zhang S. Elevated TUBA1A Might Indicate the Clinical Outcomes of Patients with Gastric Cancer, Being Associated with the Infiltration of Macrophages in the Tumor Immune Microenvironment. *J Gastrointest Liver Dis*. 2020; 29:509–22. <https://doi.org/10.15403/jgld-2834> PMID:33331338
17. Jin W. Novel Insights into PARK7 (DJ-1), a Potential Anti-Cancer Therapeutic Target, and Implications for Cancer Progression. *J Clin Med*. 2020; 9:1256. <https://doi.org/10.3390/jcm9051256> PMID:32357493
18. Chang YH, Lee SH, Chang HC, Tseng YL, Lai WW, Liao CC, Tsay YG, Liao PC. Comparative secretome analyses using a hollow fiber culture system with label-free quantitative proteomics indicates the influence of PARK7 on cell proliferation and migration/invasion in lung adenocarcinoma. *J Proteome Res*. 2012; 11:5167–85. <https://doi.org/10.1021/pr300362g> PMID:22985211
19. Garcia-Mata R, Gao YS, Sztul E. Hassles with taking out the garbage: aggravating aggregates. *Traffic*. 2002; 3:388–96. <https://doi.org/10.1034/j.1600-0854.2002.30602.x> PMID:12010457
20. Dost AFM, Moya AL, Vedaie M, Tran LM, Fung E, Heinze D, Villacorta-Martin C, Huang J, Hekman R, Kwan JH, Blum BC, Louie SM, Rowbotham SP, et al. Organoids Model Transcriptional Hallmarks of Oncogenic KRAS Activation in Lung Epithelial Progenitor Cells. *Cell Stem Cell*. 2020; 27:663–78.e8. <https://doi.org/10.1016/j.stem.2020.07.022> PMID:32891189
21. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012; 28:882–3. <https://doi.org/10.1093/bioinformatics/bts034> PMID:22257669
22. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE III, Koepfen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, et al. TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature*. 2018; 554:544–8. <https://doi.org/10.1038/nature25501> PMID:29443960
23. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hui-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G, Seja E, Lomeli S, Kong X, et al. Genomic and Transcriptomic Features of Response to Anti-PD-1

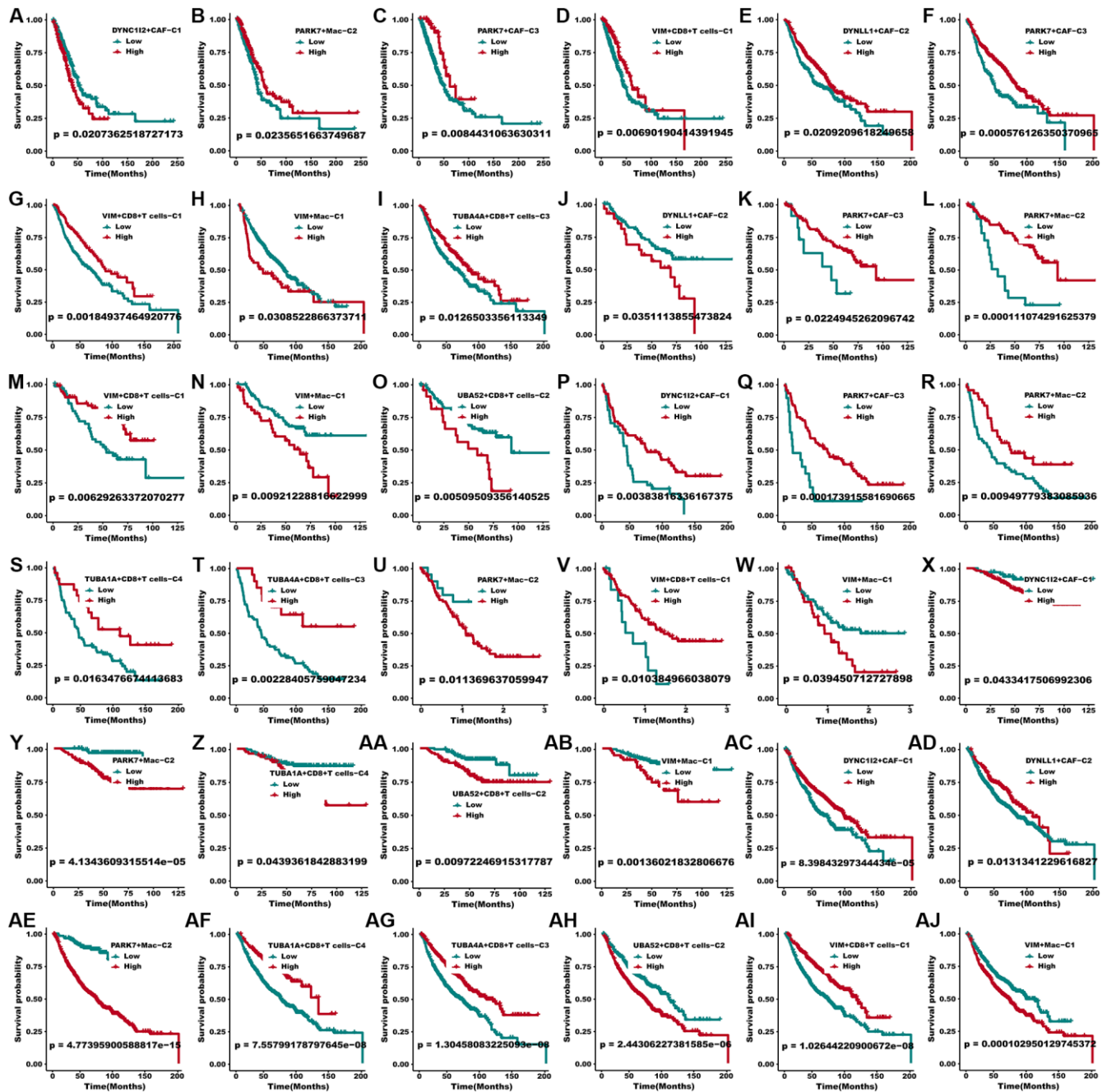
- Therapy in Metastatic Melanoma. *Cell*. 2016; 165:35–44.
<https://doi.org/10.1016/j.cell.2016.02.065>
PMID:26997480
24. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015; 33:495–502.
<https://doi.org/10.1038/nbt.3192>
PMID:25867923
25. Mo S, Shen X, Wang Y, Liu Y, Sugawara T, Yang Z, Gu W, Nakajima T. Systematic single-cell dissecting reveals heterogeneous oncofetal reprogramming in the tumor microenvironment of gastric cancer. *Hum Cell*. 2023; 36:689–701.
<https://doi.org/10.1007/s13577-023-00856-z>
PMID:36662371
26. Gao Y, Wang H, Chen S, An R, Chu Y, Li G, Wang Y, Xie X, Zhang J. Single-cell N⁶-methyladenosine regulator patterns guide intercellular communication of tumor microenvironment that contribute to colorectal cancer progression and immunotherapy. *J Transl Med*. 2022; 20:197.
<https://doi.org/10.1186/s12967-022-03395-7>
PMID:35509079
27. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018; 36:411–20.
<https://doi.org/10.1038/nbt.4096>
PMID:29608179
28. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016; 34:525–7.
<https://doi.org/10.1038/nbt.3519>
PMID:27043002
29. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017; 14:979–82.
<https://doi.org/10.1038/nmeth.4402>
PMID:28825705
30. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polański K, Goncalves A, Gardner L, Holmqvist S, Henriksson J, et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature*. 2018; 563:347–53.
<https://doi.org/10.1038/s41586-018-0698-6>
PMID:30429548
31. Guan X, Polesso F, Wang C, Sehrawat A, Hawkins RM, Murray SE, Thomas GV, Caruso B, Thompson RF, Wood MA, Hipfinger C, Hammond SA, Graff JN, et al. Androgen receptor activity in T cells limits checkpoint blockade efficacy. *Nature*. 2022; 606:791–6.
<https://doi.org/10.1038/s41586-022-04522-6>
PMID:35322234
32. Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, Seurinck R, Saelens W, Cannoodt R, Rouchon Q, Verbeiren T, De Maeyer D, Reumers J, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc*. 2020; 15:2247–76.
<https://doi.org/10.1038/s41596-020-0336-2>
PMID:32561888
33. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013; 14:7.
<https://doi.org/10.1186/1471-2105-14-7>
PMID:23323831
34. Sun X, Li Z, Meng F, Huang X, Wang J, Song J, Sun L, Zhang P. Cuproptosis associated genes affect prognosis and tumor microenvironment infiltration characterization in lung adenocarcinoma. *Am J Cancer Res*. 2022; 12:4545–65.
PMID:36381320
35. Galbo PM Jr, Zang X, Zheng D. Molecular Features of Cancer-associated Fibroblast Subtypes and their Implication on Cancer Pathogenesis, Prognosis, and Immunotherapy Resistance. *Clin Cancer Res*. 2021; 27:2636–47.
<https://doi.org/10.1158/1078-0432.CCR-20-4226>
PMID:33622705
36. Bischoff P, Trinks A, Obermayer B, Pett JP, Wiederspahn J, Uhlitz F, Liang X, Lehmann A, Jurmeister P, Elsner A, Dziondzio T, Rückert JC, Neudecker J, et al. Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. *Oncogene*. 2021; 40:6748–58.
<https://doi.org/10.1038/s41388-021-02054-3>
PMID:34663877
37. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, Cheng Y, Huang S, Liu Y, Jiang S, Liu J, Huang X, Wang X, et al. Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. *Cancer Discov*. 2022; 12:134–53.
<https://doi.org/10.1158/2159-8290.CD-21-0316>
PMID:34417225
38. Yu L, Shen N, Shi Y, Shi X, Fu X, Li S, Zhu B, Yu W, Zhang Y. Characterization of cancer-related fibroblasts (CAF) in hepatocellular carcinoma and construction of CAF-based risk signature based on single-cell RNA-seq and bulk RNA-seq data. *Front Immunol*. 2022; 13:1009789.
<https://doi.org/10.3389/fimmu.2022.1009789>
PMID:36211448

39. Yang C, Huang X, Li Y, Chen J, Lv Y, Dai S. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: an in silico strategy towards precision oncology. *Brief Bioinform.* 2021; 22:bbaa164.
<https://doi.org/10.1093/bib/bbaa164>
PMID:32789496
40. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell.* 2017; 171:1437–52.e17.
<https://doi.org/10.1016/j.cell.2017.10.049>
PMID:29195078
41. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 2018; 28:1747–56.
<https://doi.org/10.1101/gr.239244.118>
PMID:30341162
42. Mao X, Xu J, Wang W, Liang C, Hua J, Liu J, Zhang B, Meng Q, Yu X, Shi S. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol Cancer.* 2021; 20:131.
<https://doi.org/10.1186/s12943-021-01428-1>
PMID:34635121
43. Kanzaki R, Pietras K. Heterogeneity of cancer-associated fibroblasts: Opportunities for precision medicine. *Cancer Sci.* 2020; 111:2708–17.
<https://doi.org/10.1111/cas.14537>
PMID:32573845
44. Kasashima H, Duran A, Martinez-Ordoñez A, Nakanishi Y, Kinoshita H, Linares JF, Reina-Campos M, Kudo Y, L'Hermitte A, Yashiro M, Ohira M, Bao F, Tauriello DVF, et al. Stromal SOX2 Upregulation Promotes Tumorigenesis through the Generation of a SFRP1/2-Expressing Cancer-Associated Fibroblast Population. *Dev Cell.* 2021; 56:95–110.e10.
<https://doi.org/10.1016/j.devcel.2020.10.014>
PMID:33207226
45. Affo S, Nair A, Brundu F, Ravichandra A, Bhattacharjee S, Matsuda M, Chin L, Filliol A, Wen W, Song X, Decker A, Worley J, Caviglia JM, et al. Promotion of cholangiocarcinoma growth by diverse cancer-associated fibroblast subpopulations. *Cancer Cell.* 2021; 39:883.
<https://doi.org/10.1016/j.ccell.2021.05.010>
PMID:34129825
46. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015; 1:417–25.
<https://doi.org/10.1016/j.cels.2015.12.004>
PMID:26771021
47. Cui C, Wang J, Fagerberg E, Chen PM, Connolly KA, Damo M, Cheung JF, Mao T, Askari AS, Chen S, Fitzgerald B, Foster GG, Eisenbarth SC, et al. Neoantigen-driven B cell and CD4 T follicular helper cell collaboration promotes anti-tumor CD8 T cell responses. *Cell.* 2021; 184:6101–18.e13.
<https://doi.org/10.1016/j.cell.2021.11.007>
PMID:34852236
48. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, Kang B, Liu Z, Jin L, Xing R, Gao R, Zhang L, Dong M, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med.* 2018; 24:978–85.
<https://doi.org/10.1038/s41591-018-0045-3>
PMID:29942094
49. Hubble J, Demeter J, Jin H, Mao M, Nitzberg M, Reddy TB, Wymore F, Zachariah ZK, Sherlock G, Ball CA. Implementation of GenePattern within the Stanford Microarray Database. *Nucleic Acids Res.* 2009; 37:D898–901.
<https://doi.org/10.1093/nar/gkn786>
PMID:18953035
50. Lu X, Jiang L, Zhang L, Zhu Y, Hu W, Wang J, Ruan X, Xu Z, Meng X, Gao J, Su X, Yan F. Immune Signature-Based Subtypes of Cervical Squamous Cell Carcinoma Tightly Associated with Human Papillomavirus Type 16 Expression, Molecular Features, and Clinical Outcome. *Neoplasia.* 2019; 21:591–601.
<https://doi.org/10.1016/j.neo.2019.04.003>
PMID:31055200
51. Wu W, Liu Y, Zeng S, Han Y, Shen H. Intratumor heterogeneity: the hidden barrier to immunotherapy against MSI tumors from the perspective of IFN- γ signaling and tumor-infiltrating lymphocytes. *J Hematol Oncol.* 2021; 14:160.
<https://doi.org/10.1186/s13045-021-01166-3>
PMID:34620200
52. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 2015; 160:48–61.
<https://doi.org/10.1016/j.cell.2014.12.033>
PMID:25594174
53. Danilova L, Ho WJ, Zhu Q, Vithayathil T, De Jesus-Acosta A, Azad NS, Laheru DA, Fertig EJ, Anders R, Jaffee EM, Yarchoan M. Programmed Cell Death Ligand-1 (PD-L1) and CD8 Expression Profiling Identify an Immunologic Subtype of Pancreatic Ductal Adenocarcinomas with Favorable Survival. *Cancer Immunol Res.* 2019; 7:886–95.

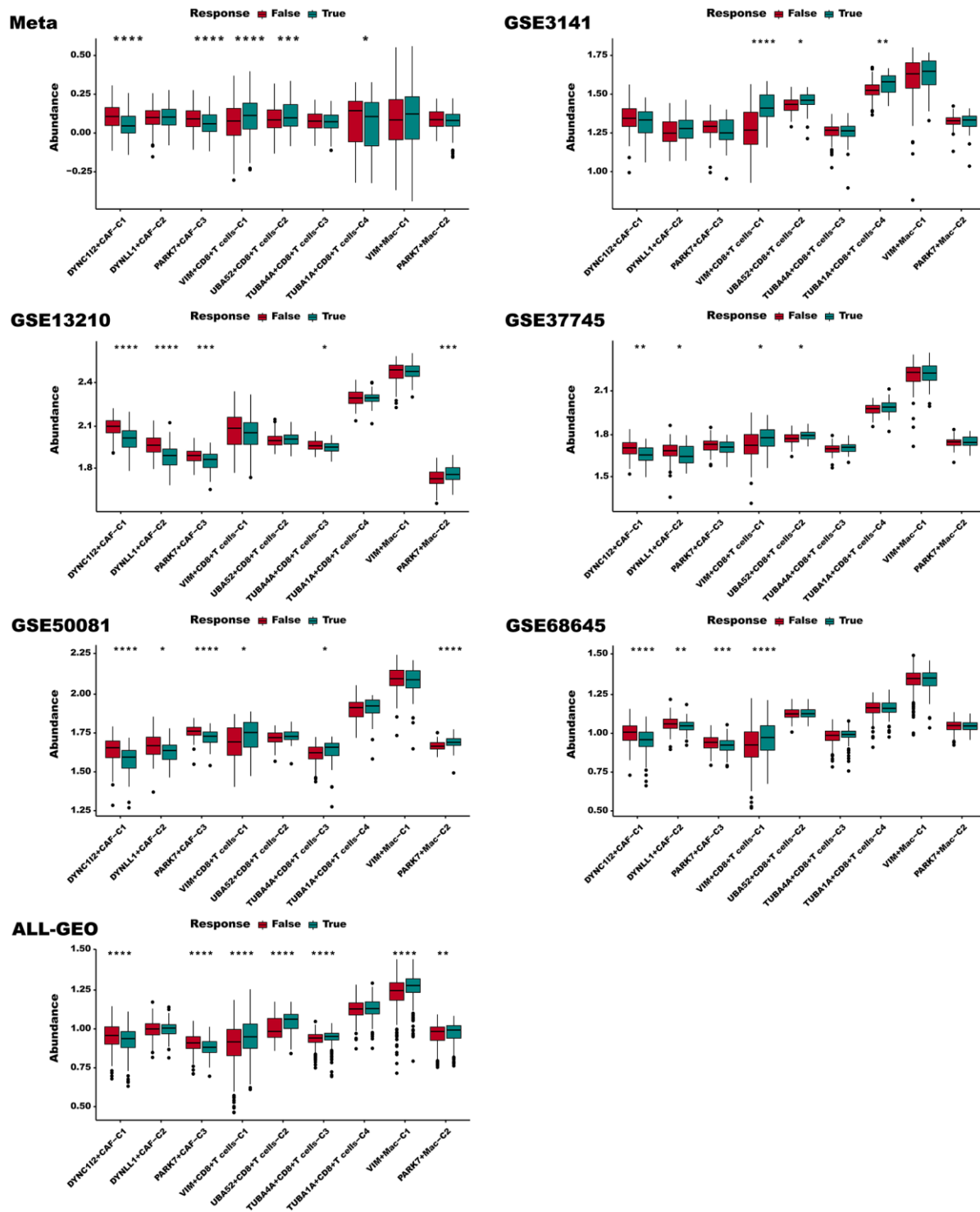
- <https://doi.org/10.1158/2326-6066.CIR-18-0822>
PMID:[31043417](https://pubmed.ncbi.nlm.nih.gov/31043417/)
54. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* 2017; 18:248–62.
<https://doi.org/10.1016/j.celrep.2016.12.019>
PMID:[28052254](https://pubmed.ncbi.nlm.nih.gov/28052254/)
55. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, Dimitriadou S, Liu DL, Kantheti HS, Saghafinia S, Chakravarty D, Daian F, Gao Q, et al, and Cancer Genome Atlas Research Network. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell.* 2018; 173:321–37.e10.
<https://doi.org/10.1016/j.cell.2018.03.035>
PMID:[29625050](https://pubmed.ncbi.nlm.nih.gov/29625050/)
56. Marsh T, Debnath J. Autophagy suppresses breast cancer metastasis by degrading NBR1. *Autophagy.* 2020; 16:1164–5.
<https://doi.org/10.1080/15548627.2020.1753001>
PMID:[32267786](https://pubmed.ncbi.nlm.nih.gov/32267786/)
57. Poillet-Perez L, White E. Role of tumor and host autophagy in cancer metabolism. *Genes Dev.* 2019; 33:610–9.
<https://doi.org/10.1101/gad.325514.119>
PMID:[31160394](https://pubmed.ncbi.nlm.nih.gov/31160394/)
58. Janssens R, Struyf S, Proost P. The unique structural and functional features of CXCL12. *Cell Mol Immunol.* 2018; 15:299–311.
<https://doi.org/10.1038/cmi.2017.107>
PMID:[29082918](https://pubmed.ncbi.nlm.nih.gov/29082918/)
59. Marković I, Savvides SN. Modulation of Signaling Mediated by TSLP and IL-7 in Inflammation, Autoimmune Diseases, and Cancer. *Front Immunol.* 2020; 11:1557.
<https://doi.org/10.3389/fimmu.2020.01557>
PMID:[32849527](https://pubmed.ncbi.nlm.nih.gov/32849527/)
60. Mulder K, Patel AA, Kong WT, Piot C, Halitzki E, Dunsmore G, Khalilnezhad S, Irac SE, Dubuisson A, Chevrier M, Zhang XM, Tam JKC, Lim TKH, et al. Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. *Immunity.* 2021; 54:1883–900.e5.
<https://doi.org/10.1016/j.immuni.2021.07.007>
PMID:[34331874](https://pubmed.ncbi.nlm.nih.gov/34331874/)
61. Vitale I, Manic G, Coussens LM, Kroemer G, Galluzzi L. Macrophages and Metabolism in the Tumor Microenvironment. *Cell Metab.* 2019; 30:36–50.
<https://doi.org/10.1016/j.cmet.2019.06.001>
PMID:[31269428](https://pubmed.ncbi.nlm.nih.gov/31269428/)
62. Barrett CS, Millena AC, Khan SA. TGF- β Effects on Prostate Cancer Cell Migration and Invasion Require FosB. *Prostate.* 2017; 77:72–81.
<https://doi.org/10.1002/pros.23250>
PMID:[27604827](https://pubmed.ncbi.nlm.nih.gov/27604827/)
63. Zhang X, Zhang C, Qiao M, Cheng C, Tang N, Lu S, Sun W, Xu B, Cao Y, Wei X, Wang Y, Han W, Wang H. Depletion of BATF in CAR-T cells enhances antitumor activity by inducing resistance against exhaustion and formation of central memory cells. *Cancer Cell.* 2022; 40:1407–22.e7.
<https://doi.org/10.1016/j.ccell.2022.09.013>
PMID:[36240777](https://pubmed.ncbi.nlm.nih.gov/36240777/)
64. Borst J, Ahrends T, Bąbała N, Melief CJM, Kastenmüller W. CD4⁺ T cell help in cancer immunology and immunotherapy. *Nat Rev Immunol.* 2018; 18:635–47.
<https://doi.org/10.1038/s41577-018-0044-0>
PMID:[30057419](https://pubmed.ncbi.nlm.nih.gov/30057419/)
65. Wu F, Zhang W, Shao H, Bo H, Shen H, Li J, Liu Y, Wang T, Ma W, Huang S. Human effector T cells derived from central memory cells rather than CD8(+)T cells modified by tumor-specific TCR gene transfer possess superior traits for adoptive immunotherapy. *Cancer Lett.* 2013; 339:195–207.
<https://doi.org/10.1016/j.canlet.2013.06.009>
PMID:[23791878](https://pubmed.ncbi.nlm.nih.gov/23791878/)
66. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, Gevaert O, Colaprico A, Czerwińska P, Mazurek S, et al, and Cancer Genome Atlas Research Network. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell.* 2018; 173:338–54.e15.
<https://doi.org/10.1016/j.cell.2018.03.034>
PMID:[29625051](https://pubmed.ncbi.nlm.nih.gov/29625051/)
67. Zhou J, Yang Q, Lu L, Tuo Z, Shou Z, Cheng J. PLK1 Inhibition Induces Immunogenic Cell Death and Enhances Immunity against NSCLC. *Int J Med Sci.* 2021; 18:3516–25.
<https://doi.org/10.7150/ijms.60135>
PMID:[34522178](https://pubmed.ncbi.nlm.nih.gov/34522178/)
68. Uras IZ, Moll HP, Casanova E. Targeting KRAS Mutant Non-Small-Cell Lung Cancer: Past, Present and Future. *Int J Mol Sci.* 2020; 21:4325.
<https://doi.org/10.3390/ijms21124325>
PMID:[32560574](https://pubmed.ncbi.nlm.nih.gov/32560574/)
69. Mogi A, Kuwano H. TP53 mutations in nonsmall cell lung cancer. *J Biomed Biotechnol.* 2011; 2011:583929.
<https://doi.org/10.1155/2011/583929>
PMID:[21331359](https://pubmed.ncbi.nlm.nih.gov/21331359/)

70. Li B, Xiang Z, Xiong F, Yan B, Huang Q. C-terminal binding protein-2 is a prognostic marker for lung adenocarcinomas. *Medicine (Baltimore)*. 2020; 99:e20076.
<https://doi.org/10.1097/MD.00000000000020076>
PMID:[32756072](https://pubmed.ncbi.nlm.nih.gov/32756072/)
71. Wang H, Wang X, Xu L, Lin Y, Zhang J. CCT6A and CHCHD2 Are Coamplified with EGFR and Associated with the Unfavorable Clinical Outcomes of Lung Adenocarcinoma. *Dis Markers*. 2022; 2022:1560199.
<https://doi.org/10.1155/2022/1560199>
PMID:[35937942](https://pubmed.ncbi.nlm.nih.gov/35937942/)
72. He T, Shen H, Wang S, Wang Y, He Z, Zhu L, Du X, Wang D, Li J, Zhong S, Huang W, Yang H. MicroRNA-3613-5p Promotes Lung Adenocarcinoma Cell Proliferation through a RELA and AKT/MAPK Positive Feedback Loop. *Mol Ther Nucleic Acids*. 2020; 22:572–83.
<https://doi.org/10.1016/j.omtn.2020.09.024>
PMID:[33230458](https://pubmed.ncbi.nlm.nih.gov/33230458/)
73. Li H, Liu J, Shen S, Dai D, Cheng S, Dong X, Sun L, Guo X. Pan-cancer analysis of alternative splicing regulator heterogeneous nuclear ribonucleoproteins (hnRNPs) family and their prognostic potential. *J Cell Mol Med*. 2020; 24:11111–9.
<https://doi.org/10.1111/jcmm.15558>
PMID:[32915499](https://pubmed.ncbi.nlm.nih.gov/32915499/)
74. Jin R, Wang X, Zang R, Liu C, Zheng S, Li H, Sun N, He J. Desmoglein-2 modulates tumor progression and osimertinib drug resistance through the EGFR/Src/PAK1 pathway in lung adenocarcinoma. *Cancer Lett*. 2020; 483:46–58.
<https://doi.org/10.1016/j.canlet.2020.04.001>
PMID:[32272148](https://pubmed.ncbi.nlm.nih.gov/32272148/)
75. Zhang P, Pei S, Gong Z, Feng Y, Zhang X, Yang F, Wang W. By integrating single-cell RNA-seq and bulk RNA-seq in sphingolipid metabolism, CACYBP was identified as a potential therapeutic target in lung adenocarcinoma. *Front Immunol*. 2023; 14:1115272.
<https://doi.org/10.3389/fimmu.2023.1115272>
PMID:[36776843](https://pubmed.ncbi.nlm.nih.gov/36776843/)

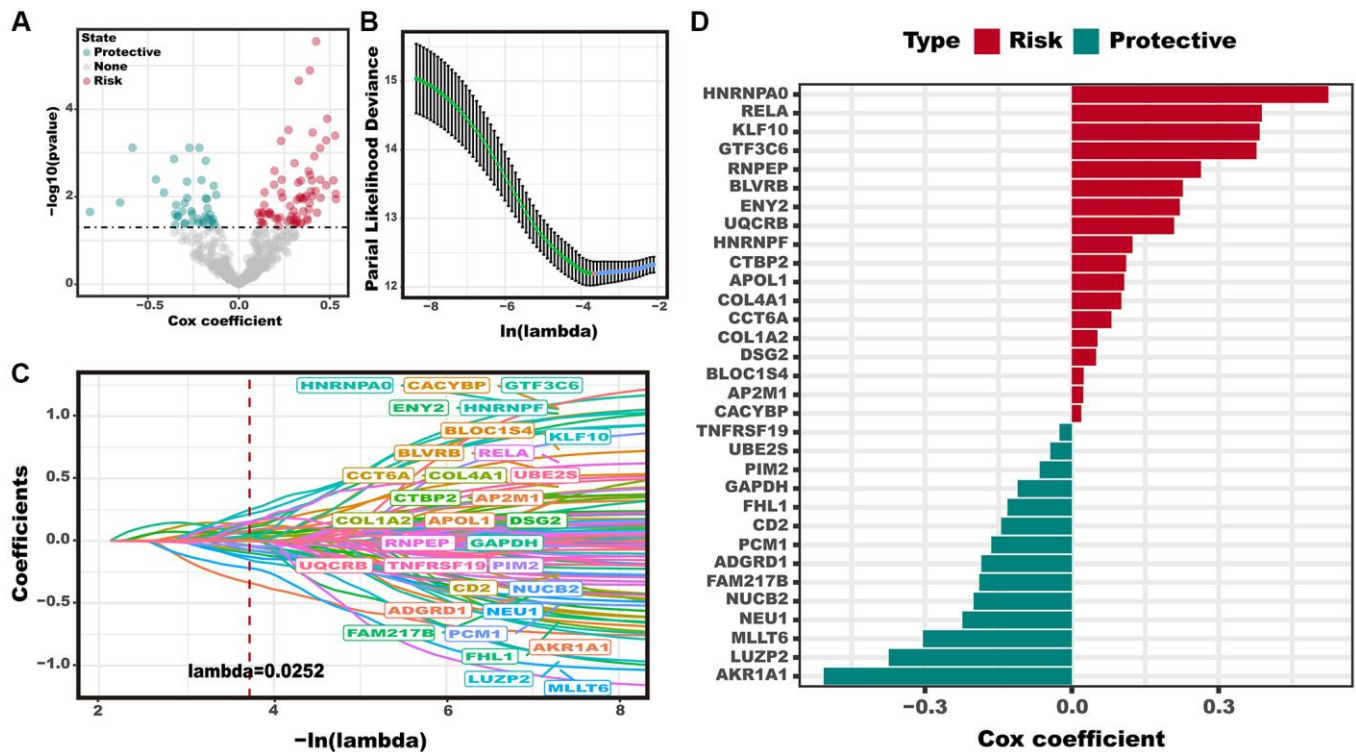
Supplementary Figures



Supplementary Figure 1. Multiple aggregophagy cell subtypes influenced the prognosis of LUAD patients. The K-M curve analysis for aggregophagy-related cell subtypes of DYNC112+CAF (A), PARK7+Mac (B), PARK7+CAF (C), and VIM+CD8+T cell (D) in the TCGA cohort; DYNLL1+CAF (E), PARK7+CAF (F), VIM+CD8+T cell (G), VIM+Mac (H), and TUBA4A+CD8+T cell (I) in the GSE68465 cohort; DYNLL1+CAF (J), PARK7+CAF (K), PARK7+Mac (L), VIM+CD8+T cell (M), VIM+Mac (N), and UBA52+CD8+T cell (O) in the GSE50081 cohort; DYNC112+CAF (P), PARK7+CAF (Q), PARK7+Mac (R), TUBA1A+CD8+T cell (S), and TUBA4A+CD8+T cell (T) in the GSE37745 cohort; PARK7+Mac (U), VIM+CD8+T cell (V), and VIM+Mac (W) in the GSE3141 cohort; DYNC112+CAF (X), PARK7+Mac (Y), TUBA1A+CD8+T cell (Z), UBA52+CD8+T cell (AA), and VIM+Mac (AB) in the GSE31210 cohort; DYNC112+CAF (AC), DYNLL1+CAF (AD), PARK7+Mac (AE), TUBA1A+CD8+T cell (AF), TUBA4A+CD8+T cell (AG), UBA52+CD8+T cell (AH), VIM+CD8+T cell (AI), and VIM+Mac (AJ) in the metaGEO cohort.



Supplementary Figure 2. Multiple aggregate cell subtypes predicted outcomes of immune checkpoint blockade (ICB) therapy. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Abbreviation: ns: not significant.



Supplementary Figure 3. The construction of ADPS. (A) Volcano plot of prognosis-related genes identified from univariate Cox regression analysis. (B) The trajectory of each independent variable with lambda. (C) Plots of the produced coefficient distributions for the logarithmic (lambda) series for parameter selection (lambda). (D) The multivariate Cox coefficients for each gene in the ADPS.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 2, 4, 5 and 8.

Supplementary Table 1. Details of data sources.

Supplementary Table 2. Feature genes of each NMF cluster in CAFs.

Supplementary Table 3. Feature genes of each NMF cluster in TAMs.

	<i>p_val</i>	<i>avg_log2FC</i>	<i>pct.1</i>	<i>pct.2</i>	<i>cluster</i>	<i>gene</i>	<i>Name/Cell type</i>
VIM	0.000211	1.411586	0.595	0.154	0	VIM	VIM+Mac-C1
MT-ND5	0.00067	1.147509	0.952	0.923	0	MT-ND5	VIM+Mac-C1
MT-ND4	0.003345	0.660753	0.952	1	0	MT-ND4	VIM+Mac-C1
MT-ATP6	0.003482	0.668498	0.929	1	0	MT-ATP6	VIM+Mac-C1
EEF2	0.005899	0.673139	0.929	0.808	0	EEF2	VIM+Mac-C1
LYZ	0.006498	0.875906	0.69	0.423	0	LYZ	VIM+Mac-C1
DSG2	0.007983	0.971472	0.238	0	0	DSG2	VIM+Mac-C1
C11orf96	0.009459	0.861673	0.643	0.462	0	C11orf96	VIM+Mac-C1
TOX4	0.009773	1.003141	0.286	0.038	0	TOX4	VIM+Mac-C1
PARK7	1.13E-05	1.208244	0.769	0.333	1	PARK7	PARK7+Mac-C2
DGCR6L	1.96E-05	0.974723	0.5	0.048	1	DGCR6L	PARK7+Mac-C2
TSTD1	0.000156	0.979324	0.692	0.167	1	TSTD1	PARK7+Mac-C2
TMEM9	0.000365	0.785661	0.346	0.024	1	TMEM9	PARK7+Mac-C2
PSMA4	0.000406	0.865812	0.654	0.214	1	PSMA4	PARK7+Mac-C2
PSMB3	0.000409	1.004955	0.538	0.143	1	PSMB3	PARK7+Mac-C2
SMIM4	0.000426	0.763388	0.577	0.143	1	SMIM4	PARK7+Mac-C2
CFAP298	0.000456	0.719377	0.269	0	1	CFAP298	PARK7+Mac-C2
PSENE1	0.000984	0.581205	0.385	0.048	1	PSENE1	PARK7+Mac-C2
CMTM6	0.001005	0.774987	0.308	0.024	1	CMTM6	PARK7+Mac-C2
MCUR1	0.001084	0.560481	0.308	0.024	1	MCUR1	PARK7+Mac-C2
TUBA4A	0.001124	1.001012	0.346	0.048	1	TUBA4A	PARK7+Mac-C2
FH	0.001169	0.570502	0.308	0.024	1	FH	PARK7+Mac-C2
NARS2	0.001286	0.700944	0.231	0	1	NARS2	PARK7+Mac-C2
CRTAP	0.001286	0.68988	0.231	0	1	CRTAP	PARK7+Mac-C2
DYNLL2	0.001286	0.654695	0.231	0	1	DYNLL2	PARK7+Mac-C2
TMEM92	0.001286	0.53154	0.231	0	1	TMEM92	PARK7+Mac-C2
PPIE	0.001693	0.693616	0.346	0.048	1	PPIE	PARK7+Mac-C2
PSMC4	0.001693	0.616486	0.346	0.048	1	PSMC4	PARK7+Mac-C2
NRGN	0.001901	0.577685	0.5	0.119	1	NRGN	PARK7+Mac-C2
YWHAQ	0.001942	0.599586	0.654	0.238	1	YWHAQ	PARK7+Mac-C2
SEMA4A	0.001944	0.820897	0.385	0.071	1	SEMA4A	PARK7+Mac-C2
LSM4	0.002023	0.843285	0.423	0.095	1	LSM4	PARK7+Mac-C2
QKI	0.00216	0.687433	0.269	0.024	1	QKI	PARK7+Mac-C2
SLC35C2	0.00216	0.658456	0.269	0.024	1	SLC35C2	PARK7+Mac-C2
SMIM19	0.002199	0.628397	0.385	0.071	1	SMIM19	PARK7+Mac-C2
DNPH1	0.002358	0.531032	0.346	0.048	1	DNPH1	PARK7+Mac-C2
HIST1H1C	0.002408	1.004697	0.423	0.095	1	HIST1H1C	PARK7+Mac-C2
CALM2	0.002578	0.654472	1	0.69	1	CALM2	PARK7+Mac-C2
SERINC3	0.002691	0.757213	0.308	0.048	1	SERINC3	PARK7+Mac-C2
NUCB2	0.002891	0.682196	0.462	0.119	1	NUCB2	PARK7+Mac-C2
GJA1	0.00291	0.67725	0.269	0.024	1	GJA1	PARK7+Mac-C2
MTIF3	0.00291	0.572991	0.269	0.024	1	MTIF3	PARK7+Mac-C2
DESI2	0.002979	0.559832	0.385	0.071	1	DESI2	PARK7+Mac-C2
NDUFB2	0.003046	0.747608	0.692	0.262	1	NDUFB2	PARK7+Mac-C2

CDK4	0.003131	0.620115	0.269	0.024	1	CDK4	PARK7+Mac-C2
CYC1	0.003243	0.797487	0.346	0.071	1	CYC1	PARK7+Mac-C2
NDUFA4	0.003264	0.769977	0.846	0.476	1	NDUFA4	PARK7+Mac-C2
UQCRB	0.003271	0.691112	0.923	0.595	1	UQCRB	PARK7+Mac-C2
GPRC5C	0.003287	0.632321	0.308	0.048	1	GPRC5C	PARK7+Mac-C2
C21orf2	0.003368	0.501911	0.269	0.024	1	C21orf2	PARK7+Mac-C2
RAB11B	0.003479	0.521213	0.538	0.167	1	RAB11B	PARK7+Mac-C2
PLPP5	0.003559	0.655067	0.385	0.071	1	PLPP5	PARK7+Mac-C2
TIMM10B	0.003567	0.658733	0.192	0	1	TIMM10B	PARK7+Mac-C2
MORN2	0.003567	0.601826	0.192	0	1	MORN2	PARK7+Mac-C2
UBLCP1	0.003567	0.551008	0.192	0	1	UBLCP1	PARK7+Mac-C2
TMEM129	0.003567	0.504385	0.192	0	1	TMEM129	PARK7+Mac-C2
RAN	0.003633	0.719248	0.615	0.238	1	RAN	PARK7+Mac-C2
EIF3F	0.003655	0.569545	0.538	0.167	1	EIF3F	PARK7+Mac-C2
TAOK3	0.003748	0.654915	0.308	0.048	1	TAOK3	PARK7+Mac-C2
RAP1A	0.004	0.606681	0.423	0.095	1	RAP1A	PARK7+Mac-C2
HMGCL	0.004001	0.526426	0.308	0.048	1	HMGCL	PARK7+Mac-C2
ATXN2	0.004268	0.657027	0.308	0.048	1	ATXN2	PARK7+Mac-C2
EIF5A	0.004314	0.643102	0.692	0.286	1	EIF5A	PARK7+Mac-C2
HADHB	0.004643	0.577399	0.385	0.095	1	HADHB	PARK7+Mac-C2
PFN1	0.004726	0.704733	0.846	0.548	1	PFN1	PARK7+Mac-C2
HMOX1	0.004852	0.508662	0.308	0.048	1	HMOX1	PARK7+Mac-C2
MRPL42	0.004942	0.606266	0.346	0.071	1	MRPL42	PARK7+Mac-C2
CDK2AP2	0.005013	0.962554	0.577	0.286	1	CDK2AP2	PARK7+Mac-C2
MAGED1	0.005241	0.534179	0.346	0.071	1	MAGED1	PARK7+Mac-C2
P3H2	0.005377	0.67345	0.5	0.143	1	P3H2	PARK7+Mac-C2
MLF2	0.005384	0.567608	0.538	0.167	1	MLF2	PARK7+Mac-C2
ARPC3	0.005721	0.589591	0.615	0.262	1	ARPC3	PARK7+Mac-C2
PSMA5	0.005791	0.563824	0.385	0.095	1	PSMA5	PARK7+Mac-C2
HDDC2	0.006115	0.58931	0.385	0.095	1	HDDC2	PARK7+Mac-C2
NT5DC1	0.006149	0.649745	0.231	0.024	1	NT5DC1	PARK7+Mac-C2
ADGRD1	0.006149	0.647112	0.231	0.024	1	ADGRD1	PARK7+Mac-C2
CAPRIN1	0.006611	0.608691	0.231	0.024	1	CAPRIN1	PARK7+Mac-C2
MT1X	0.006617	0.872285	0.423	0.119	1	MT1X	PARK7+Mac-C2
ENY2	0.0067	0.584781	0.5	0.167	1	ENY2	PARK7+Mac-C2
ANXA7	0.006791	0.748266	0.462	0.143	1	ANXA7	PARK7+Mac-C2
ZNF593	0.006964	0.519489	0.423	0.119	1	ZNF593	PARK7+Mac-C2
HIP1R	0.007105	0.550525	0.231	0.024	1	HIP1R	PARK7+Mac-C2
MRPS26	0.007215	0.622555	0.308	0.071	1	MRPS26	PARK7+Mac-C2
MRPL52	0.007583	0.58825	0.385	0.095	1	MRPL52	PARK7+Mac-C2
CIB1	0.007611	0.716517	0.615	0.31	1	CIB1	PARK7+Mac-C2
GCN1	0.007631	0.568252	0.231	0.024	1	GCN1	PARK7+Mac-C2
SNF8	0.007631	0.506375	0.231	0.024	1	SNF8	PARK7+Mac-C2
ME2	0.007631	0.503717	0.231	0.024	1	ME2	PARK7+Mac-C2
SELENOM	0.008263	0.670829	0.654	0.381	1	SELENOM	PARK7+Mac-C2
ST3GAL5	0.008292	0.575376	0.462	0.167	1	ST3GAL5	PARK7+Mac-C2
COX7A2	0.008756	0.618256	0.731	0.381	1	COX7A2	PARK7+Mac-C2
BLVRB	0.0088	0.681885	0.615	0.238	1	BLVRB	PARK7+Mac-C2
CTSD	0.00922	0.731674	0.808	0.452	1	CTSD	PARK7+Mac-C2
IFT57	0.009459	0.782298	0.769	0.452	1	IFT57	PARK7+Mac-C2
IGF2R	0.009644	0.704461	0.308	0.071	1	IGF2R	PARK7+Mac-C2
CHCHD2	0.009661	0.525212	0.846	0.524	1	CHCHD2	PARK7+Mac-C2
UBE2S	0.009769	0.59751	0.154	0	1	UBE2S	PARK7+Mac-C2
CHMP7	0.009769	0.521948	0.154	0	1	CHMP7	PARK7+Mac-C2
PIGF	0.009769	0.520695	0.154	0	1	PIGF	PARK7+Mac-C2
TRPT1	0.009769	0.507902	0.154	0	1	TRPT1	PARK7+Mac-C2

CGGBP1	0.009773	0.61168	0.346	0.095	1	CGGBP1	PARK7+Mac-C2
FNIP1	0.009823	0.619406	0.269	0.048	1	FNIP1	PARK7+Mac-C2
ZNF511	0.009823	0.590366	0.269	0.048	1	ZNF511	PARK7+Mac-C2
PSMA2	0.009857	0.590351	0.385	0.095	1	PSMA2	PARK7+Mac-C2
MGLL	0.009864	0.558292	0.615	0.286	1	MGLL	PARK7+Mac-C2

Supplementary Table 4. Feature genes of each NMF cluster in CD8+T cells.

Supplementary Table 5. Differentially expressed genes (DEGs) in each aggrephagy subtype.

Supplementary Table 6. Univariate Cox regression analysis in the TCGA, GSE3141, GSE3141, GSE37745, GSE50081, GSE68465, MetaGEO, and Meta cohorts.

TCGA				
id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	2.581799054	0.604868	11.02007	0.200201
DYNLL1+CAF-C2	4.077003236	0.404556	41.08693	0.233168
PARK7+CAF-C3	0.802670531	0.104466	6.167338	0.832667
VIM+CD8+T_cells-C1	0.183536583	0.051212	0.657767	0.009236
UBA52+CD8+T_cells-C2	3.037373724	0.253424	36.40397	0.380636
TUBA4A+CD8+T_cells-C3	2.073828361	0.069197	62.15291	0.674162
TUBA1A+CD8+T_cells-C4	2.275863571	0.255933	20.23791	0.460759
VIM+Mac-C1	1.63499456	0.235508	11.35083	0.618977
PARK7+Mac-C2	0.875312247	0.107903	7.100527	0.90077
GSE3141				
id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	2.007242198	0.16733	24.07832	0.582561
DYNLL1+CAF-C2	0.968835927	0.049885	18.81621	0.983311
PARK7+CAF-C3	1.956635837	0.080021	47.84294	0.680674
VIM+CD8+T_cells-C1	0.232823649	0.040624	1.334361	0.101808
UBA52+CD8+T_cells-C2	0.858300774	0.009532	77.28693	0.946942
TUBA4A+CD8+T_cells-C3	0.813117539	0.010904	60.63317	0.925077
TUBA1A+CD8+T_cells-C4	0.315129975	0.008938	11.11082	0.525245
VIM+Mac-C1	0.823288349	0.157964	4.29087	0.817434
PARK7+Mac-C2	12.6988644	0.041296	3905.024	0.384542
GSE31210				
id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	20.1376868	0.21446	1890.92	0.195108
DYNLL1+CAF-C2	2.927174004	0.049746	172.2432	0.605436
PARK7+CAF-C3	3.486241245	0.009597	1266.455	0.677998
VIM+CD8+T_cells-C1	1.431754564	0.074284	27.59557	0.812078
UBA52+CD8+T_cells-C2	769.3726441	0.576163	1027372	0.070325
TUBA4A+CD8+T_cells-C3	44.61804366	0.007434	267805.8	0.392181
TUBA1A+CD8+T_cells-C4	214.9467924	0.189983	243191.3	0.134392
VIM+Mac-C1	1610.479881	2.262259	1146485	0.027554
PARK7+Mac-C2	15455.07325	37.59304	6353817	0.001684

GSE37745

id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	0.128390992	0.003231	5.101943	0.274582
DYNLL1+CAF-C2	0.133755798	0.004667	3.833287	0.239963
PARK7+CAF-C3	0.009723816	8.47E-05	1.11609	0.055546
VIM+CD8+T_cells-C1	0.382683171	0.044922	3.259984	0.379507
UBA52+CD8+T_cells-C2	0.198820187	0.000245	161.6298	0.636573
TUBA4A+CD8+T_cells-C3	0.000288953	9.14E-07	0.091323	0.005521
TUBA1A+CD8+T_cells-C4	0.003347176	2.15E-05	0.520577	0.026864
VIM+Mac-C1	0.248788483	0.022933	2.699033	0.252752
PARK7+Mac-C2	0.000769389	1.10E-06	0.537409	0.031888

GSE50081

id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	0.737622411	0.036687	14.83047	0.842455
DYNLL1+CAF-C2	0.532903037	0.015622	18.17822	0.726709
PARK7+CAF-C3	0.011703745	4.57E-05	2.998346	0.115973
VIM+CD8+T_cells-C1	0.075712218	0.007643	0.750048	0.027399
UBA52+CD8+T_cells-C2	4.950905497	0.004651	5270.491	0.65287
TUBA4A+CD8+T_cells-C3	0.107274024	0.00327	3.519127	0.210032
TUBA1A+CD8+T_cells-C4	0.225455696	0.004102	12.39032	0.466177
VIM+Mac-C1	0.917315219	0.03395	24.78576	0.959077
PARK7+Mac-C2	2.65E-05	3.01E-08	0.02332	0.002317

GSE68465

id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	0.724513433	0.135067	3.886377	0.706904
DYNLL1+CAF-C2	0.228289121	0.012141	4.292474	0.323763
PARK7+CAF-C3	0.039698618	0.003022	0.521497	0.014071
VIM+CD8+T_cells-C1	0.203946722	0.072489	0.573797	0.002591
UBA52+CD8+T_cells-C2	0.477837403	0.013018	17.53924	0.687883
TUBA4A+CD8+T_cells-C3	0.183298861	0.013166	2.551969	0.206694
TUBA1A+CD8+T_cells-C4	1.211091094	0.103747	14.13772	0.878589
VIM+Mac-C1	3.310688928	0.503132	21.78487	0.212989
PARK7+Mac-C2	0.239756992	0.00661	8.695918	0.4357

MetaGEO

id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	0.152365738	0.049322	0.470686	0.001078
DYNLL1+CAF-C2	0.14848377	0.022511	0.979421	0.047528
PARK7+CAF-C3	0.825739866	0.153063	4.45469	0.823797
VIM+CD8+T_cells-C1	0.079643086	0.03838	0.165268	1.10E-11
UBA52+CD8+T_cells-C2	10.40697494	2.854632	37.94014	0.000386
TUBA4A+CD8+T_cells-C3	0.009589995	0.001341	0.068575	3.66E-06
TUBA1A+CD8+T_cells-C4	0.020902299	0.00466	0.093757	4.39E-07
VIM+Mac-C1	2.372205799	0.762971	7.37559	0.135561
PARK7+Mac-C2	368.5834853	76.27593	1781.083	1.94E-13

Meta

id	HR	HR.95L	HR.95H	p-value
DYNC1I2+CAF-C1	0.319938207	0.132078	0.774998	0.011582
DYNLL1+CAF-C2	0.238401367	0.047903	1.186461	0.079921

PARK7+CAF-C3	2.963187473	1.006161	8.726711	0.048711
VIM+CD8+T_cells-C1	0.096224073	0.051225	0.180753	3.39E-13
UBA52+CD8+T_cells-C2	13.37013396	4.244167	42.1191	9.47E-06
TUBA4A+CD8+T_cells-C3	0.015943941	0.003321	0.076536	2.33E-07
TUBA1A+CD8+T_cells-C4	0.163465178	0.082465	0.324029	2.13E-07
VIM+Mac-C1	4.621473744	1.864315	11.45623	0.000951
PARK7+Mac-C2	56.51273294	16.12425	198.0675	2.88E-10

Supplementary Table 7. Logistic regression analysis in the TCGA, GSE3141, GSE3141, GSE37745, GSE50081, GSE68465, MetaGEO, and Meta cohorts.

TCGA				
gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	2.02E-05	1.56E-06	0.000222941	1.26E-17
DYNLL1+CAF-C2	0.000147	4.94E-06	0.003793619	1.84E-07
PARK7+CAF-C3	0.001265	4.42E-05	0.0272173	4.56E-05
VIM+CD8+T_cells-C1	40.32358	7.557707	229.156878	2.10E-05
UBA52+CD8+T_cells-C2	3.943248	0.172431	98.16945486	0.394881
TUBA4A+CD8+T_cells-C3	0.008325	9.72E-05	0.671910613	0.033252
TUBA1A+CD8+T_cells-C4	4.518784	0.251759	86.37529831	0.309754
VIM+Mac-C1	0.865458	0.067049	11.55283268	0.912121
PARK7+Mac-C2	54.95842	2.77556	1353.95107	0.011093
GSE3141				
gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	0.066885	0.001432	2.721712047	0.155861
DYNLL1+CAF-C2	8.904625	0.095688	939.4348189	0.347523
PARK7+CAF-C3	0.027254	0.000181	3.088246493	0.141055
VIM+CD8+T_cells-C1	4476.675	123.5161	308868.5754	2.20E-05
UBA52+CD8+T_cells-C2	621.5397	0.656213	1128245.268	0.077061
TUBA4A+CD8+T_cells-C3	0.566985	0.001741	238.2310752	0.846678
TUBA1A+CD8+T_cells-C4	8573.647	21.15131	6205252.904	0.004536
VIM+Mac-C1	6.641665	0.483144	140.100899	0.186218
PARK7+Mac-C2	0.145823	9.51E-05	222.6762122	0.595843
GSE31210				
gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	1.57E-08	5.52E-11	2.30E-06	3.09E-11
DYNLL1+CAF-C2	8.33E-07	6.41E-09	6.45E-05	2.29E-09
PARK7+CAF-C3	5.54E-05	2.63E-07	0.008595146	0.000208
VIM+CD8+T_cells-C1	0.230344	0.01945	2.626560295	0.238935
UBA52+CD8+T_cells-C2	18.13677	0.047548	7346.64122	0.339587
TUBA4A+CD8+T_cells-C3	9.01E-05	4.59E-08	0.130091197	0.013627
TUBA1A+CD8+T_cells-C4	1.215349	0.003995	380.510701	0.946583
VIM+Mac-C1	1.68501	0.020167	157.2439218	0.818457
PARK7+Mac-C2	11112.94	81.65885	2129205.262	0.000313

GSE37745

gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	1.56E-05	9.64E-09	0.011584066	0.001776
DYNLL1+CAF-C2	0.011445	2.63E-05	3.352479758	0.131502
PARK7+CAF-C3	0.001902	7.41E-07	3.301046233	0.105386
VIM+CD8+T_cells-C1	233.1595	3.44978	25464.279	0.015854
UBA52+CD8+T_cells-C2	1329463	13.95029	4.08834E+11	0.020925
TUBA4A+CD8+T_cells-C3	995.7538	0.017683	121548210.4	0.227426
TUBA1A+CD8+T_cells-C4	440.6525	0.062058	5707936.256	0.188809
VIM+Mac-C1	3.665756	0.064324	314.2394703	0.541082
PARK7+Mac-C2	15.80532	0.000352	1051811.376	0.615967

GSE50081

gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	0.000681	7.33E-06	0.03809135	0.000771
DYNLL1+CAF-C2	0.005953	5.15E-05	0.511578096	0.02799
PARK7+CAF-C3	1.62E-07	1.97E-11	0.000440894	0.000274
VIM+CD8+T_cells-C1	55.8003	2.057267	1976.221672	0.020897
UBA52+CD8+T_cells-C2	290.5348	0.066782	2378699.731	0.197526
TUBA4A+CD8+T_cells-C3	13.8598	0.123247	2421.284465	0.292375
TUBA1A+CD8+T_cells-C4	7.610202	0.049434	1659.706252	0.440633
VIM+Mac-C1	0.120726	0.002691	4.754539676	0.259671
PARK7+Mac-C2	51058593	1595.171	6.24062E+12	0.001531

GSE68465

gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	0.000139	6.99E-06	0.002332162	1.95E-09
DYNLL1+CAF-C2	0.001069	1.22E-05	0.082101774	0.002288
PARK7+CAF-C3	0.000621	1.06E-05	0.032392375	0.000302
VIM+CD8+T_cells-C1	39.02222	7.384114	221.9286704	2.35E-05
UBA52+CD8+T_cells-C2	10.77913	0.055413	2186.920845	0.377501
TUBA4A+CD8+T_cells-C3	51.31209	0.942561	3272.378611	0.05756
TUBA1A+CD8+T_cells-C4	2.662463	0.07605	98.85240086	0.591231
VIM+Mac-C1	2.441731	0.179217	35.77694955	0.507038
PARK7+Mac-C2	0.34741	0.001827	66.31239484	0.692391

MetaGEO

gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	0.012352	0.002438	0.060578191	8.02E-08
DYNLL1+CAF-C2	1.696705	0.138144	21.00039256	0.679659
PARK7+CAF-C3	5.38E-05	4.30E-06	0.000624731	9.39E-15
VIM+CD8+T_cells-C1	28.65448	9.698947	87.3198209	2.10E-09
UBA52+CD8+T_cells-C2	18280.18	2659.294	132406.6492	6.83E-23
TUBA4A+CD8+T_cells-C3	89.62167	5.018813	1742.524109	0.002559
TUBA1A+CD8+T_cells-C4	10.70487	1.327287	88.44545704	0.026751
VIM+Mac-C1	128.4665	27.19828	648.9844333	1.93E-09
PARK7+Mac-C2	13.64775	2.197253	87.68294855	0.005405

Meta				
gene	OR	OR.95L	OR.95H	p-value
DYNC1I2+CAF-C1	0.000153	3.48E-05	0.000648301	5.30E-32
DYNLL1+CAF-C2	1.025265	0.217904	4.826020841	0.974798
PARK7+CAF-C3	0.002304	0.000485	0.010668186	1.29E-14
VIM+CD8+T_cells-C1	10.81759	4.553453	26.00137186	8.31E-08
UBA52+CD8+T_cells-C2	12.15326	3.784088	39.39289892	2.90E-05
TUBA4A+CD8+T_cells-C3	0.509344	0.086712	2.983289218	0.454582
TUBA1A+CD8+T_cells-C4	0.42945	0.221176	0.832975222	0.01243
VIM+Mac-C1	1.318582	0.750752	2.319263204	0.33623
PARK7+Mac-C2	0.206939	0.035085	1.211979935	0.081093

Supplementary Table 8. A total of 571 genes form aggrephagy-related subtypes.

Supplementary Table 9. The ADPS-related genes and corresponding coefficients.

Gene	Coef	Type
LUZP2	-0.374121637	Protective
COL4A1	0.101170202	Risk
GTF3C6	0.377254728	Risk
CTBP2	0.110815334	Risk
AKR1A1	-0.506961197	Protective
COL1A2	0.052344441	Risk
TNFRSF19	-0.025303508	Protective
APOL1	0.107021888	Risk
AP2M1	0.023375955	Risk
KLF10	0.383501703	Risk
FHL1	-0.131528	Protective
FAM217B	-0.189208623	Protective
BLOC1S4	0.024211454	Risk
UQCRB	0.208941759	Risk
HNRNPF	0.124065381	Risk
NEU1	-0.223584911	Protective
CCT6A	0.080899978	Risk
RNPEP	0.263793282	Risk
GAPDH	-0.110872233	Protective
RELA	0.388110159	Risk
NUCB2	-0.200828737	Protective
PCM1	-0.164689427	Protective
MLLT6	-0.304064015	Protective
CD2	-0.144521072	Protective
PIM2	-0.065727646	Protective
HNRNPA0	0.524133487	Risk
CACYBP	0.019093532	Risk

DSG2	0.049422674	Risk
ADGRD1	-0.18487887	Protective
ENY2	0.220663515	Risk
BLVRB	0.226676186	Risk
UBE2S	-0.044076029	Protective
